LEARNING AND PRICING WITH STRATEGIC AGENTS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Ruifan Yang

August 2025

LEARNING AND PRICING WITH STRATEGIC AGENTS

Ruifan Yang, Ph.D.

Cornell University 2025

Understanding how to design learning and pricing mechanisms in the presence of strategic agents is essential for the effective operation of societal systems. These challenges arise in domains such as transportation, supply chains, and online platforms, where individuals respond strategically to information and incentives.

In the first part, we introduce a novel learning model based on hypothesis testing, wherein agents form beliefs about their opponents' strategies and update them via a stochastic process driven by hypothesis testing and utility-based exploration. We show that in any game, this learning dynamic converges to a Nash equilibrium that maximizes the minimum utility among all players.

The second part of the thesis presents two applications of strategic pricing and intervention. In High Occupancy Toll (HOT) lane systems, we develop a game-theoretic model to design toll prices that incentivize carpooling among travelers with heterogeneous values of time and carpooling constraints. Using empirical data from California's I-880 highway, we identify Pareto-efficient tolling strategies that balance travel time reduction, revenue generation, and social welfare. In a two-tier supply chain setting, we study a commission-based pricing game in which manufacturers delegate pricing to retailers through linear commission contracts. We characterize the subgame-perfect equilibrium in closed form and identify conditions under which such delegation induces price subsidization and leads to higher retail prices compared to direct sales.

BIOGRAPHICAL SKETCH

Ruifan Yang has a BS in Mathematics and is pursuing a Ph.D. in Operations Research and Information Engineering.

To my grandmother, Yuzhu Li — in loving memory.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Professor Manxi Wu, for her patience, encouragement, and continuous support throughout my PhD journey. She has not only taught me how to read, think, and write like a researcher, but also to believe in my capabilities, articulate my ideas with clarity, and engage critically with the world around me. I am also profoundly grateful to my committee members, Professor David Goldberg and Professor Jamol Pender, for their constructive feedback, thoughtful questions, and support throughout this process. Their diverse perspectives greatly enriched the quality of my work.

I would like to thank the faculty and staff of the ORIE Department for fostering an intellectually stimulating academic environment and for providing the resources that made this research possible. I am especially thankful to the ORACL program, and Kathryn Schumacher from General Motors for the funding support and for offering a unique opportunity to bridge theoretical research with practical, real-world applications. I would like to express my sincere gratitude to my lab mates and research collaborators Zhanhao and Priya, whose contributions were instrumental to this thesis.

I would like to express my heartfelt thanks to my friends — Ailin, Fei, Hemeng, Ofer, Shangyuan, Shuyang, Tonghua, Xumei, Yanqiu, Yiyi, Yu, Yujie, and those whose names I may not have mentioned — for their friendship, encouragement, and support throughout this journey. Your presence brought joy, balance, and strength to my life during the most challenging moments of this thesis. I would like to thank my therapists Huilin and Yee-Ching, and my yoga teacher Lin, for their compassion, wisdom, and care throughout this journey.

To my parents — thank you for believing in me, for giving me the space

to grow, and for being my constant foundation. Even from afar, knowing you were always there gave me strength and comfort throughout this journey. To my partner, William — thank you for the many adventures we've shared around the world, for taking on the everyday chores that gave me space to work, and for always caring for me in every way that matters. And to my cat, Mochi — thank you for sitting by my side during countless late-night work sessions and waking me up each morning. In caring for you, I was reminded to care for myself. Lastly, I thank myself for the courage and resilience that carried me through. I will carry with me every act of kindness and every moment of growth from this journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Algorithms and data-driven decision systems are increasingly ubiquitous in various aspects of modern society, fundamentally reshaping how prices are determined, resources are allocated, and strategic interactions unfold. E-commerce companies such as Amazon and eBay employ algorithms that dynamically guide pricing based on real-time demand, competition, and inventory conditions. In urban transportation, dynamic tolling mechanisms adapt toll prices in response to live traffic conditions to manage congestion and incentivize optimal commuting choices. Ride-sharing platforms like Uber and Lyft utilize surge pricing models, algorithmically adjusting fares in real-time to balance driver supply with rider demand, ensuring availability during peak usage and optimizing network efficiency.

Despite the rapid developments of algorithmic decision tools, foundational theoretical questions about learning and pricing in the presence of strategic agents have yet to be fully addressed. In particular, most existing theoretical results about convergence properties of learning algorithms either guarantee convergence only in specific types of games, such as zero-sum games, potential games, or restrict to weaker equilibrium concepts, such as coarse correlated equilibria (CCE). A more significant issue arises when there are multiple equilibria in a game, as there is few fundamental analysis on which equilibrium the learning algorithm will converge to. This equilibrium-selection problem is critical: in games with multiple Nash equilibria, the particular equilibrium reached can dramatically affect social welfare and the distribution of benefits among players. Real-world applications also illustrate the importance of using strategic interventions to skew equilibrium outcomes in directions favorable to system designers. For instance, in supply chains, indirect sales structure or linear commission-based contracts can shift market dynamics and affect profitability across manufacturers, wholesalers, and consumers. In road-pricing contexts, toll prices strategically alter travelers' decisions regarding transportation mode and route choice. By intervening through pricing mechanisms and contracts, the designer can influence agent behaviors to steer the equilibrium outcome toward desirable states.

My thesis addresses these theoretical and practical challenges within two domains—toll pricing with HOT lanes and commission-based pricing in two-tier supply chains.

Chapter 2 introduces a new belief-based learning dynamics in which players update their strategies by combining statistical hypothesis testing with utility-driven exploration. In this dynamics, each player forms beliefs about opponents' strategies and episodically tests these beliefs using empirical observations. Beliefs are resampled either when the hypothesis test is rejected or through exploration, where the probability of exploration decreases with the player's expected utility. In general finite normal-form games, we show that the learning process converges to a set of approximate Nash equilibria and, more importantly, to a refinement that selects equilibria maximizing the minimum utility across all players. Our result establishes convergence to equilibrium in general finite games and reveals a novel mechanism for equilibrium selection induced by the structure of the learning dynamics.

Chapter 3 examines the strategic pricing mechanism in HOT lane systems. We formalize a game-theoretic model where travelers with heterogeneous values of times and carpool disutilities choose among paying a toll to use the HOT lane, forming carpools, or driving in ordinary lanes. We characterize Wardrop equilibria, and explore how equilibrium outcomes—including traffic flow, congestion levels, and toll revenues—shift in response to changes in lane capacities and toll pricing. Using empirical data from California's I-880 HOT lanes, we identify Pareto-efficient tolling strategies that optimize various objectives such as travel time reduction, social welfare enhancement, and revenue maximization. Our analysis illustrates how pricing interventions can strategically influence heterogeneous players to achieve socially desirable outcomes.

Chapter 4 examines the commission-based pricing game in a two-tier supply chain. We model a duopolistic competition of manufacturers who employ linear commission contracts with downstream dedicated retailers. These retailers subsequently engage in price competition over differentiated products. We establish the uniqueness of subgame-perfect equilibrium (SPE) and derive a closed-form SPE solution for games with general linear demand. Under regularization conditions, in a fully competitive market, strategic delegation through linear contracts typically leads manufacturers to subsidize retailers, and raises equilibrium retail prices above a direct-sales benchmark. We reveal how strategic interactions via linear contracts fundamentally alter market outcomes and influence the welfare of manufacturers, retailers, and consumers.

To maintain the flow of the main text, we have relegated most proofs to appendices A, B, and C.

CHAPTER 2

LEARNING WITH EPISODIC HYPOTHESIS TESTING IN GENERAL GAMES

2.1 Introduction

A central question in the study of learning in games is how adaptive behavior shapes long-run outcomes and under what conditions these outcomes correspond to equilibrium concepts. A rich body of work has explored dynamic models where players update their strategies based on past observations or payoff feedback. In specific classes of games like zero-sum, dominance-solvable, or potential games, belief-based models such as fictitious play converge to Nash equilibrium when players best respond to the empirical frequency of opponents' play [40, 74]. In more general games, Bayesian learning models update prior beliefs using Bayes' rule and converge to Nash equilibrium under the grain-of-truth assumption [39, 57, 82–84], while calibrated forecasting ensures convergence to correlated or approximate Nash equilibria when players best respond to forecasts [38, 56, 71]. Apart from belief-based learning, other dynamics dispense with explicit modeling of opponents actions and only adpt strategies based on payoffs. In particular, regret-based approaches ensure convergence to coarse correlated equilibrium in general finite games [37, 42, 45]. Furthermore, trial-and-error learning and log-linear learning introduce random payoff-based perturbations into strategy updates and are analyzed via Markov chain analysis characterizing the stochastic stability set, which show that learning plays (pure) Nash equilibrium for the majority of the time [6, 72, 73, 75, 76, 88, 118].

When games admit multiple equilibria, a central and pressing question is:

which equilibria are likely to emerge from the learning process? This challenge has led to research on adaptive dynamics as a mechanism for equilibrium selection. Examples include log-linear learning and its variants, which select welfare-maximizing pure Nash equilibria by introducing small random perturbations and applying resistance tree analysis [73, 75, 88]. In repeated two-player games, Jindani [55] extends belief-based hypothesis testing to achieve Pareto-efficient outcomes via structured exploration. Our work contributes to this literature by developing a belief-based hypothesis testing learning dynamic for general finite normal-form games, integrating belief revision, statistical testing, and utility-driven exploration to enable equilibrium selection based on players' exploration behavior. Our results hold without relying on specific game structures or restricting to two-player settings.

In our learning dynamic, each player maintains a belief about others' strategies, plays a smooth best response, and periodically tests whether the belief aligns with observed actions. If a statistical test rejects the current hypothesis, the player resamples a new belief. Even without rejection, players may explore alternative beliefs with a probability that decreases in expected utility, allowing dissatisfaction to guide strategic adaptation. A key feature of this dynamic is the integration of episodic hypothesis testing with utility-sensitive exploration. Hypothesis testing serves to detect whether a player's belief is statistically inconsistent with observed empirical play. If the discrepancy exceeds a tolerance threshold with high confidence, the belief is rejected and resampled. Exploration provides a complementary mechanism for belief revision, allowing players to experiment when dissatisfied with their utility, even if their belief passes the statistical test. The probability of such exploration decreases with expected utility, mapped through a transformation function that encodes the player's sensitivity to dissatisfaction. This transformed utility plays a central role in determining how likely a player is to explore at a given utility level and, in turn, influences long-run equilibrium selection.

We analyze the long-run behavior of this learning dynamics by characterizing its stochastically stable states. Formally, the joint evolution of players' beliefs and strategies induces a Markov chain over a finite state space, where each state represents a tuple of player beliefs and their corresponding smooth best response strategies. A state is stochastically stable if it retain positive stationary probability in the stationary distribution as the exploration rate becomes very small. In other words, when the exploration rate is positive but close to zero, the system spends most of its time in the stochastically stable states.

We show that the stochastically stable set lies within the set of approximate Nash equilibria—specifically, it selects those equilibria that maximize the minimum transformed utility across all players. There are two key contributions in this result. First, it guarantees that the stochastically stable set is composed of approximate equilibria in general finite games under certain parameter conditions, which implies that this learning dynamics will eventually play those approximate equilibria for the vast majority of time. Second, it identifies a principled selection criterion among equilibria: those with higher minimum transformed utility are more stable, as players with lower transformed utility are more prone to explore and destabilize the equilibrium.

The intuition behind this result follows directly from the design of the learning dynamics. States in which players hold beliefs that are inconsistent with observed actions are unlikely to persist, due to the high probability of hypothesis rejection. We show in Proposition 2 that all consistent states correspond to

ϵ-Nash equilibria under appropriate conditions on the smoothness, tolerance, and discretization parameters. Hence, all long-run stochastically stable outcomes must be approximate equilibria. For consistent states (i.e., approximate Nash equilibria), it is the least satisfied player—measured in terms of transformed utility—who is most likely to initiate exploration and leave the state. As a result, equilibria that maximize the minimum transformed utility are the most stable ones in the long run. We formalize this result in Theorem 1 using stochastic stability analysis and resistance-tree methods, showing that such equilibria correspond to recurrent classes with minimal stochastic potential.

Furthermore, our analysis reveals how the structure of the utility transformation functions governs equilibrium selection. When all players use identical transformation functions—i.e., the same sensitivity to utility in exploration—the learning dynamics select equilibria that maximize the minimum utility across agents, leading to a max-min selection. When a particular player's transformation consistently maps their utility to lower values than those of others, the stochastically stable set selects equilibria that maximize this player's utility. Corollary 1 formally characterizes this effect and shows how the choice of transformation functions directly governs which players are favored by the equilibrium selection process.

2.2 Model and Preliminaries

2.2.1 The Static Game Model

We consider a static game $G = (I, A = (A_i)_{i \in I}, (u_i)_{i \in I})$, where:

- I is a finite set of players, with $|I| = n$,
- A_i is a finite set of actions for player i, and $A := \prod_{i \in I} A_i$ denotes the set of joint action profiles. Let $A_{-i} := \prod_{j \neq i} A_j$ denote the set of joint actions of players other than i,
- $u_i : A \to \mathbb{R}$ is the payoff function for player i.

For each player i, a mixed strategy $\pi_i = (\pi_i(a_i))_{a_i \in A_i} \in \Delta_i$ is a probability distribution over A_i, where Δ_i denotes the simplex on A_i. A joint strategy profile is $\pi = (\pi_i)_{i \in I} \in \Delta := \prod_{i \in I} \Delta_i$, where Δ is the space of joint mixed strategies. The expected utility for player i given π is:

$$U_i(\pi) = \mathbb{E}_{a \sim \pi}[u_i(a)].$$

Each player $i \in I$ holds a *belief* $b_i = (b_{ij})_{j \in I \setminus \{i\}}$ of their opponents' strategies, where each $b_{ij} \in \Delta_j^M$ represents player i's belief about player j's strategy. We assume that beliefs are discretized in that each b_{ij} is in a *discretized probability simplex* over A_j with granularity parameter $M \in \mathbb{N}_+$ defined as follows: For all $i \in I,\ j \in I \setminus \{i\}$,

$$b_{ij} \in \Delta_j^M := \left\{ b_{ij} \in \mathbb{R}_{\geq 0}^{|A_j|} \;\middle|\; \sum_{a_j \in A_j} b_{ij}(a_j) = 1,\ b_{ij}(a_j) \in \left\{ \frac{m}{M} : m = 0, \ldots, M \right\} \right\}.$$

Accordingly, player i's *discretized belief space* is $\mathcal{B}_i := \prod_{j \neq i} \Delta_j^M$. We denote the joint belief vector as $b = (b_i)_{i \in I}$ and the the *joint discretized belief space* is $\mathcal{B} := \prod_{i \in I} \mathcal{B}_i$.

We next define the smooth best response function and ϵ-Nash equilibrium.

Definition 1 (Smooth Best Response)**.** *For any player $i \in I$ and temperature param-*

eter $\sigma > 0$, the smooth best response function $\mathrm{Br}_i^\sigma : \mathcal{B}_i \to \Delta_i$ is given by:

$$\mathrm{Br}_i^\sigma(a_i \mid b_i) = \frac{e^{\frac{1}{\sigma}U_i(a_i,b_i)}}{\sum_{a_i' \in A_i} e^{\frac{1}{\sigma}U_i(a_i',b_i)}}, \quad \forall a_i \in A_i, \quad \forall b_i \in \mathcal{B}_i, \tag{2.1}$$

where $U_i(a_i', b_i) = \sum_{a_{-i} \in A_{-i}} u_i(a_i, a_{-i}) b_i(a_{-i})$. The image set of the smooth best response function is $\mathrm{Im}\,(\mathrm{Br}_i^\sigma) := \{\pi_i \,|\, \pi_i = \mathrm{Br}_i^\sigma(b_i),\ b_i \in \mathcal{B}_i\}$.

We denote the joint smooth best response function as $\mathrm{Br}^\sigma(b) = (\mathrm{Br}_i^\sigma(b_i))_{i \in I}$, and the joint smooth best response function for all players other than i as $\mathrm{Br}_{-i}^\sigma(b_{-i}) = (\mathrm{Br}_j^\sigma(b_j))_{j \in I \setminus \{i\}}$.

Definition 2 (ϵ-Nash Equilibrium)**.** *For any $\epsilon > 0$, a strategy profile $\pi^* \in \Delta$ is an ϵ-Nash equilibrium if:*

$$U_i(\pi_i^*, \pi_{-i}^*) \geq U_i(\pi_i', \pi_{-i}^*) - \epsilon, \quad \forall \pi_i' \in \Delta_i, \quad \forall i \in I.$$

As $\epsilon \to 0$, π^* becomes a Nash equilibrium.

2.2.2 Preliminaries for Hypothesis Testing

A hypothesis test evaluates whether player i's belief b_i is close to the true strategy profile π_{-i} of the opponents. Given a tolerance level $\tau > 0$, we define:

- The *null hypothesis* H_0: The strategy profile π_{-i} is within τ-distance to player i's belief b_i:

$$\|\pi_{-i} - b_i\|_2 \leq \tau.$$

- The *alternative hypothesis* H_1: The strategy profile π_{-i} is not within τ-distance to player i's belief b_i:

$$\|\pi_{-i} - b_i\|_2 > \tau.$$

For each player i, let $\{a^t_{-i}\}_{t=1}^T$ be T independent and identically distributed (i.i.d.) action profiles sampled from π_{-i}. For a given significance level $\alpha \in (0, 1)$, each player i conducts a hypothesis test that rejects H_0 if and only if

$$\|\hat{\pi}_{-i} - b_i\|_2 > \tau + \sqrt{\frac{|A_{-i}|}{2T} \cdot \ln(\frac{2}{\bar{\alpha}})},$$

where $\hat{\pi}_{-i}$ is the empirical estimate of the opponents' strategy profile:

$$\hat{\pi}_{-i}(a_{-i}) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\{a^t_{-i} = a_{-i}\}, \qquad \forall a_{-i} \in A_{-i}.$$

That is, the reject region for player i's hypothesis test is given by

$$R_i(\alpha) = \left\{ (a^t_{-i})_{t=1}^T \middle| \|\hat{\pi}_{-i} - b_i\| > \tau + \sqrt{\frac{|A_{-i}|}{2T} \cdot \ln\left(\frac{2}{\alpha}\right)} \right\}. \tag{2.2}$$

The type I and type II errors of the hypothesis test are defined as follows:

- *Type I error (false positive)*: rejecting H_0 even though $\|\pi_{-i} - b_i\|_2 \le \tau$.
- *Type II error (false negative)*: failing to reject H_0 even though $\|\pi_{-i} - b_i\|_2 > \tau$.

Proposition 1. *For any $\alpha \in (0, 1)$, if*

$$T \ge T(\alpha) := \max_{\substack{i \in I,\ \pi_{-i} \in \mathrm{Im}(\mathrm{Br}^\sigma_{-i}),\\ b_i \in \mathcal{B}_i,\ \|\pi_{-i} - b_i\|_2 > \tau}} \frac{2|A_{-i}|}{(\|\pi_{-i} - b_i\|_2 - \tau)^2} \ln\left(\frac{2}{\alpha}\right), \tag{2.3}$$

then the maximum type I and type II error rates of the hypothesis test are upper bounded by α for all players, all beliefs and smooth best response strategies. That is, for any player i, any belief $b_i \in \mathcal{B}_i$ and smooth best response strategy profile $\pi_{-i} = \mathrm{Br}^\sigma_i(b_{-i})$,

$$\Pr(\textit{reject } H_0 |\ \|b_i - \pi_{-i}\|_2 \le \tau) \le \alpha, \quad \Pr(\textit{fail to reject } H_0 |\ \|b_i - \pi_{-i}\|_2 > \tau) \le \alpha.$$

Proposition 1 ensures that, if the number of samples T exceeds a certain threshold, then both the Type I and Type II error rates of the hypothesis test are uniformly bounded by the significance level α, across all players, all discretized beliefs, and all smooth best response strategy profiles. The proof of this result is provided in Appendix A.1.

Remark. *The hypothesis test introduced above is a simple test based on an ℓ_2-distance threshold and a Chernoff-type bound on the empirical estimation error. It is not necessarily the most powerful or optimal test for detecting discrepancies between beliefs and strategies. Since the focus of this paper is not on optimal test design but on the learning dynamics and convergence behavior in general games, we adopt this test for simplicity. Our convergence results remain valid under any hypothesis testing procedure that ensures small Type I and Type II error given a finite number of samples.*

2.3 Learning Dynamics and Stochastically Stable Set

In this section, we present the learning dynamics (Algorithm 1), where players periodically test their beliefs about other players' strategies and resample new beliefs of opponents based on hypothesis test result and their utility. We also present our main theorem that demonstrates the long-run outcomes of the learning dynamics.

2.3.1 Learning Dynamics

We begin by outlining the overall structure of the learning dynamics, before detailing the key quantities that govern the algorithm's dynamics. In Algorithm 1, each player i begins with an initial belief b_i^0 about others' strategies and selects a smooth best response strategy π_i^0 corresponding to their belief, as in (2.1). The learning proceeds in *epochs* ($k = 1, 2, \ldots$). In each epoch k, each player i

maintains belief b_i^k and plays the game according to a smooth best response strategy π_i^k for $T(\xi^{\bar{u}})$ rounds, collecting observations of joint action profiles. Here, $\xi \in (0,1)$ is a hyperparameter of the learning dynamics and $\bar{u}$ is a sufficiently large constant, which will be described in details later. Each player i may independently decide whether to conduct a hypothesis test against their current belief b_i^k, with probability $\gamma_i > 0$. The hypothesis test is as defined in Section 2.2.2, with $\xi^{\bar{u}}$ as the significance level and $T(\xi^{\bar{u}})$ as the sample size. If the null hypothesis is rejected (indicating a significant discrepancy between the observed actions and belief b_i^k with high probability), player i samples a new belief b_i^{k+1} uniformly from the belief space and updates their strategy to the corresponding best response $\pi_i^{k+1} = \mathrm{Br}_i^{\sigma}(b_i^{k+1})$. If the null hypothesis is not rejected, player i may still sample a new belief (referred to as exploration) with probability $\xi^{f_i(U_i(\pi_i^k, b_i^k))}$, where $f_i(\cdot)$ is an increasing function and $U_i(\pi_i^k, b_i^k)$ is player i's anticipated utility given their belief and best response. If neither testing nor exploration occurs, the player does not update their belief and strategy. Fig. 2.1 illustrates the belief update process for each player in each decision epoch, based on randomized testing, hypothesis rejection, and exploration decisions.

We now provide more details on the key quantities of the learning algorithm:

- γ_i is the probability that each player conducts a test in each epoch.
- $\xi \in (0,1)$ is a hyperparameter that affects hypothesis test significance level, epoch length, and players' probability of exploration.
- $f_i : \mathbb{R} \to \mathbb{R}_{>0}$ is an increasing function that maps player i's utility to a positive value. When player i's hypothesis test fails to reject in epoch k, the player explores (resamples a new belief) with probability $\xi^{f_i(U_i(a_i^k, b_i^k))}$, where $U_i(a_i^k, b_i^k)$ is the expected utility of player i given their belief and

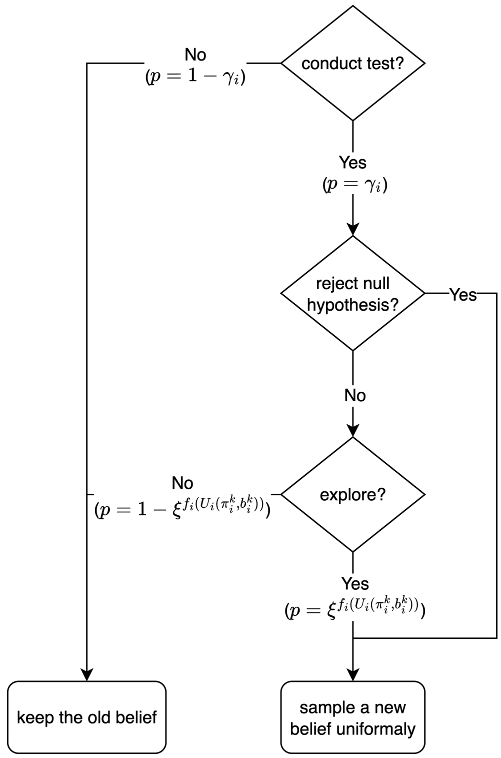


Figure 2.1: Belief update flowchart.

smoothed best response. Since $\xi \in (0,1)$ and $f_i(\cdot)$ is increasing, the exploration probability is higher when the utility $U_i(a_i^k, b_i^k)$ is low. The function $f_i(\cdot)$ can be viewed as player i's sensitivity function, governing how their exploration probability changes with their utility, and may differ for different players. We refer to $f_i(U_i(a_i^k, b_i^k))$ as player i's transformed utility in epoch k.

- $\bar{u} > \max_{\pi \in \Delta} \sum_{i \in I} f_i(U_i(\pi))$ is a constant larger than the maximum total transformed utility of all players in the game. The constant $\bar{u}$ being sufficiently large ensures that the test significance level $\xi^{\bar{u}}$ in each epoch is sufficiently small.
- $T(\xi^{\bar{u}})$ is the length of each epoch, which is also the sample size of each hypothesis test. Following from Proposition 1, $T(\xi^{\bar{u}})$ satisfies (2.3), and ensures that the hypothesis test conducted by each player has Type I and

Type II errors upper bounded by $\xi^{\bar{u}}$.

Algorithm 1: Learning with Episodic Hypothesis Testing

Input : $\xi \in (0,1)$, $\gamma_i(0,1)$ for each $i \in I$;
$f_i : \mathbb{R} \to \mathbb{R}_+$ is an increasing utility transformation function for each player $i \in I$;
$T(\xi^{\bar{u}})$ satisfies (2.3) and $\bar{u} > \max_{\pi \in \Delta} \sum_{i \in I} f_i(U_i(\pi))$;
Initalization: Each player i holds a belief $b_i^0 \in \mathcal{B}_i$ and chooses a smooth best response strategy $\pi_i^0 = \mathrm{Br}_i^\sigma(b_i^0)$.
for *epoch* $k = 1, 2, \cdots$ **do**
  Each player $i \in i$ plays the game with strategy π_i^k for $T(\xi^{\bar{u}})$ periods: $(a_i^t) \sim \pi_i^k$ for $t = 1, \cdots, T(\xi^{\bar{u}})$, and observe full action profiles $(a^t)_{t=1,\cdots,T(\xi^{\bar{u}})}$;
  for *each player* $i = 1, \cdots, |I|$ **do**
    with probability γ_i **:**
      Conduct a hypothesis test for b_i^k using observations $(a_{-i}^t)_{t=1,\cdots,T(\xi^{\bar{u}})}$;
      if $(a_{-i}^t)_{t=1}^{T(\xi^{\bar{u}})} \in R_i(\xi^{\bar{u}})$ *as in* (2.2) **then**
        Sample b_i^{k+1} uniformly from $\mathcal{B}_i$ and $\pi_i^{k+1} = \mathrm{Br}_i^\sigma(b_i^{k+1})$;
      else
        with probability $\xi^{f_i(U_i(\pi_i^k, b_i^k))}$ **:**
          Sample b_i^{k+1} uniformly from $\mathcal{B}_i$ and $\pi_i^{k+1} = \mathrm{Br}_i^\sigma(b_i^{k+1})$;
        else
          $b_i^{k+1} = b_i^k$; $\pi_i^{k+1} = \pi_i^k$;
    else
      $b_i^{k+1} = b_i^k$; $\pi_i^{k+1} = \pi_i^k$;

A key feature of the learning dynamics is the integration of hypothesis testing with utility-sensitive exploration. The hypothesis test evaluates whether player i's belief b_i^k is within τ distance to the actual strategy profile π_{-i}^k. Specifically, when the test rejects the null hypothesis with significance level $\xi^{\bar{u}}$, it indicates with probability at least $1 - \xi^{\bar{u}}$, the discrepancy $\|\pi_{-i}^k - b_i^k\|_2 > \tau$; in this case, player will resample a new belief. Conversely, when the null hypothesis fails to be rejected, it indicates that b_i^k is τ-consistent with π_{-i}^k with probability at least $1 - \xi^{\bar{u}}$. In this scenario, belief resampling occurs when the player explores with probability $\xi^{f_i(U_i(\pi_i^k, b_i^k))}$, where $f_i(\cdot)$ encodes the player's sensitivity to utility dissatisfaction. The integration of statistical hypothesis testing and utility-sensitive exploration ensures that the learning process corrects inconsistent beliefs with high probability while allowing adaptive exploration when current utility is unsatisfactory.

2.3.2 Stochastically Stable Set Characterization

In the learning dynamics, we refer to the tuple (b, π) as the *state* of the system. We note that for any belief $b_i \in \mathcal{B}_i$, each player $i \in I$ always selects their strategy $\pi_i \in \Delta_i$ via the smooth best response function $\pi_i = \mathrm{Br}_i^\sigma(b_i)$ with temperature parameter σ. Therefore, given any $\xi > 0$, the learning dynamics induces a Markov chain on the finite state space:

$$\mathcal{Z} := \{(b, \pi) \mid b_i \in \mathcal{B}_i,\ \pi_i = \mathrm{Br}_i^\sigma(b_i), \quad \forall i \in I\}.$$

Definition 3 (Consistent state)**.** *A state $z = (b, \pi) \in \mathcal{Z}$ is consistent if, for all players $i \in I$, their belief b_i is within τ-distance of the strategy profile π_{-i}, and π is a smooth best response to b. That is, the set of all consistent states is given by*

$$\mathcal{Z}^\dagger := \{z = (b, \pi) \in \mathcal{Z} \mid \|b_i - \pi_{-i}\|_2 \leq \tau, \quad \forall i \in I\}.$$

Assumption 1. *Given any $\epsilon > 0$, we assume that parameters σ, τ, M satisfy:*

$$\sigma \leq \frac{\epsilon}{2 \cdot \log(\max_{i \in I} |A_i|)}, \quad \tau \leq \frac{\epsilon \cdot \sigma}{2\sqrt{|A|} \cdot (\max_{i \in I, a \in A} u_i(a) \cdot (\max_{i \in I} |A_i| \cdot |A_{-i}|)}, \tag{2.4a}$$

$$M \geq \frac{|I| \cdot \max_{i \in I} |A_i|}{\tau} \cdot \left(1 + \frac{\sqrt{\max_{i \in I} |A_i|}}{\sigma} \cdot \max_{i \in I, a \in A} u_i(a) \cdot \max_{i \in I} |A_i||A_{-i}| \cdot |I|\right). \tag{2.4b}$$

Assumption 1 ensures that the smooth best response temperature parameter σ and the hypothesis testing tolerance level τ are sufficiently small, while the belief granularity parameter M is sufficiently large. A small σ guarantees that the smooth best response closely approximates the exact best response to any belief. A small τ ensures that, under consistency, each player's belief is sufficiently close to the actual strategy of their opponents. A large M ensures that for any strategy profile, there exists a discretized belief vector that satisfies the τ-consistency condition. The following proposition shows that under Assumption 1, each consistent state $z \in \mathcal{Z}^\dagger$ is an ϵ-Nash equilibrium. The proof of this result is provided in Appendix A.2.

Proposition 2. *For any $\epsilon > 0$, under Assumption 1, the consistent state set $\mathcal{Z}^\dagger$ is non-empty. Furthermore, for any $z^\dagger = (b^\dagger, \pi^\dagger) \in \mathcal{Z}^\dagger$, $\pi^\dagger$ is an ϵ-Nash equilibrium.*

For each epoch k, we denote $z^k = (b^k, \pi^k)$ as the state in epoch k. Given ξ, the state transition probability matrix is $P^\xi = (P^\xi_{zz'})_{zz' \in \mathcal{Z}}$, where $P^\xi_{zz'}$ is the probability of state transitioning from $z = (b, \pi)$ to $z' = (b', \pi')$ in one epoch:

$$P^\xi_{zz'} := \mathbb{P}(z^{k+1} = z' \mid z^k = z), \quad \forall z, z' \in \mathcal{Z}.$$

The following lemma shows that the state transition Markov chain is finite and has a unique stationary distribution.

Lemma 1. *The Markov process described by transition P^ξ has a finite state space and is aperiodic and irreducible for all $\xi \in (0, 1)$, and has a unique stationary distribution μ^ξ.*

A state is stochastically stable if it is in the support set of the stationary distribution $\mu^\xi(z)$ as $\xi \to 0$.

Definition 4 ([117]). *A state $z \in \mathcal{Z}$ is stochastically stable if*

$$\lim_{\xi \to 0} \mu_z^{\xi} > 0,$$

where μ_z^{ξ} is the stationary distribution of state z given P^{ξ}.

We denote the set of stochastically stable states by $\mathcal{Z}^*$. These are the states of the belief-based learning dynamics that retain positive probability in the limit stationary distribution as the perturbation parameter $\xi \to 0$. The notion of stochastic stability captures the asymptotic state distribution of the Markov process induced by the learning dynamics under vanishing ξ. Formally, for any $\delta > 0$, there exists $\bar{\xi} > 0$ such that for all $\xi \in (0, \bar{\xi})$, the stationary distribution μ^{ξ} places at least $1 - \delta$ probability mass on the stochastically stable set $\mathcal{Z}^*$. This implies that when ξ is sufficiently small, the learning dynamics visits states in $\mathcal{Z}^*$ for the majority of epochs. Therefore, states in the stochastically stable set can be viewed as high-probability long-run outcomes of the learning dynamics—that is, the beliefs and strategies most frequently observed in the evolution of the system.

Before presenting the theorem, we introduce the following assumption, which ensures that for each player i, regardless of the opponents' beliefs and strategies, there exists a belief of player i that is inconsistent with the opponents' strategies, and the corresponding smooth best response of player i is also inconsistent with the opponents' beliefs.

Assumption 2. *For any player $i \in I$, any belief profile $b_{-i} \in \prod_{j \neq i} \mathcal{B}_j$ and strategy profile $\pi_{-i} = \mathrm{Br}^{\sigma}_{-i}(b_{-i})$, there exists $\tilde{b}_i \in \Delta^M_{-i}$, and $\tilde{\pi}_i = \mathrm{Br}^{\sigma}_i(\tilde{b}_i)$ such that belief $\tilde{b}_i$ is at least τ distance away from strategy π_{-i}:*

$$\|\tilde{b}_i - \pi_{-i}\|_2 > \tau,$$

and strategy $\tilde{\pi}_i$ is at least τ distance away from belief b_{ji} for any $j \neq i$:

$$\|\tilde{\pi}_i - b_{ji}\|_2 > \tau, \quad \forall j \neq i.$$

This assumption ensures that, regardless of the system state, each player i has a belief $\tilde{b}_i$ that can, with high probability, trigger belief resampling for all players. Specifically, player i will resample their own belief with probability greater than $1 - \xi^{\bar{u}}$, because $\tilde{b}_i$ is at least τ-distant from the opponents' strategy profile π_{-i}. Furthermore, the corresponding smooth best response $\tilde{\pi}_i$ is inconsistent with the opponents' beliefs about player i, which can also prompt the opponents to resample. As $\xi \to 0$, the probability of belief resampling triggered by $\tilde{b}_i$ approaches 1. In Appendix A.4, we show in Lemma 11 that this assumption is mild and can be satisfied when the image of each player's smooth best response function contains more than $|I|$ strategies and the tolerance parameter τ is sufficiently small.

Theorem 1. *Suppose Assumptions 1 and 2 hold. The stochastically stable state set is given by*

$$\mathcal{Z}^* = \left\{ z = (b, \pi) \in \mathcal{Z}^\dagger \,\middle|\, \min_{i \in I} f_i(U_i(\pi_i, b_i)) = \max_{z'=(b',\pi') \in \mathcal{Z}^\dagger} \min_{i \in I} f_i(U_i(\pi'_i, b'_i)) \right\}.$$

Theorem 1 shows that any stochastically stable state must be a consistent state, as defined in Definition 3. Following Proposition 2, the strategy profile associated with every consistent state is an ϵ-Nash equilibrium for sufficiently small τ, σ, and large M. Hence, the long-run outcomes of the learning process lie within the ϵ-Nash equilibrium set. This property arises from the hypothesis test, which rejects states where empirical play deviates significantly from held beliefs with high probability, thereby preventing inconsistent belief-strategy tuples from attaining stochastic stability.

Furthermore, the long-run outcome of the learning dynamics selects among consistent states (i.e., approximate equilibria) those that maximize the minimum transformed utility across players. This effect arises from the structure of the exploration mechanism. The probability of leaving a consistent state depends on the likelihood that some player initiates exploration. The dominant contribution to this probability comes from the player with the highest exploration probability, $\xi^{f_i(U_i(\pi_i,b_i))}$. Since $\xi < 1$, this is governed by the player with the lowest transformed utility $f_i(U_i)$, who is most inclined to explore. As a result, the equilibrium strategy profiles that maximize $\min_{i\in I} f_i(U_i(\pi_i, b_i))$ among all equilibria are the most stable and remain in the stochastically stable set as $\xi \to 0$.

Corollary 1. *Suppose Assumptions 1 and 2 hold.*

(i) If $f_i(u) = f_j(u)$ for any $i, j \in I$ and any $u \in [\underline{u}, \bar{u}]$, where $\underline{u}$ (resp. $\bar{u}$) is the minimum (resp. maximum) utility of all players given all feasible (π, b), then

$$\mathcal{Z}^* = \left\{ z \in \mathcal{Z}^\dagger \,\middle|\, \min_{i\in I} U_i(\pi_i, b_i) = \max_{z'=(b',\pi')\in\mathcal{Z}^\dagger} \min_{i\in I} U_i(\pi'_i, b'_i) \right\}.$$

- *If there exists a player $\hat{i}$ such that $f_{\hat{i}}(u_i) \le f_j(u_j)$ for all $j \in I \setminus \{i\}$, $u_{\hat{i}} \in [\underline{u}_{\hat{i}}, \bar{u}_{\hat{i}}]$ and $u_j \in [\underline{u}_j, \bar{u}_j]$, where $\underline{u}_i$ (resp. $\bar{u}_i$) is the minimum (resp. maximum) utility of each player $i \in I$ given all feasible (π, b), then*

$$\mathcal{Z}^* = \left\{ z \in \mathcal{Z}^\dagger \,\middle|\, U_{\hat{i}}(\pi_{\hat{i}}, b_{\hat{i}}) = \max_{z'=(b',\pi')\in\mathcal{Z}^\dagger} U_{\hat{i}}(\pi'_{\hat{i}}, b'_{\hat{i}}) \right\}.$$

Corollary 1 shows that the utility transformation functions $\{f_i(\cdot)\}_{i\in I}$ govern how equilibrium refinement may favor certain players over others by modulating their exploration probabilities. When f_i is identical across all players, the learning dynamics select equilibria that maximize the minimum utility across players, corresponding to a max-min equilibrium refinement (Corollary 1 (i)).

More generally, the choice of $\{f_i(\cdot)\}_{i\in I}$ influences which player is most likely to initiate exploration at equilibrium. A player $\hat{i}$ whose function $f_{\hat{i}}(\cdot)$ consistently maps their utility to a lower value than those of others across equilibria will have the highest exploration probability at all equilibrium states. As a result, the equilibrium refinement process selects equilibria that maximize the utility of player $\hat{i}$, since such equilibria are least likely to be destabilized by their exploration (Corollary 1 (ii)).

Example 1 **(Cooperative Outcome):** Two players simultaneously choose between Stag (S) and Hare (H). The payoff matrix is

Player 1\Player 2	S	H
S	$(4,4)$	$(0,3)$
H	$(3,0)$	$(3,3)$

.

The Nash Equilibria are:

- Pure: (S,S), (H,H)
- Mixed: each player plays S with probability $\frac{3}{4}$, H with probability $\frac{1}{4}$. The expected utility is $(3,3)$.

Suppose each player i chooses identity utility transformation function $f_i(u) = u$. In the mixed NE and the pure NE (H,H), the minimum utility of all players are both 3; whereas in (S,S), the minimum utility of all players is 4. Therefore, the learning converges to the stochastically stable set with strategy profile is (S,S) – the fully cooperative outcome.

Example 2 (Transformation Function Governs Convergence): Two players simultaneously choose between Opera (O) and Football (F). The payoff matrix is

Player 1\Player 2	O	F
O	(2, 1)	(0, 0)
F	(0, 0)	(1, 2)

.

The Nash Equilibria are:

- Pure: (O, O), (F, F)
- Mixed: player 1 plays O with probability $\frac{2}{3}$ and F with probability $\frac{1}{3}$; player 2 plays O with probability $\frac{1}{3}$, and F with probability $\frac{2}{3}$. The expected utility is $(\frac{2}{3}, \frac{2}{3})$

Suppose each player i chooses the identity utility transformation function $f_i(u) = u$. In the mixed NE, the minimum utility of all players is $\frac{2}{3}$; whereas in (O, O) and (F, F), the minimum utility of all players is 1. Therefore, learning converges to the stochastically stable set with strategy profiles (O, O) and (F, F).

On the other hand, suppose player 1 chooses the identity utility transformation function $f_1(u) = u$, and player 2 chooses $f_2(u) = u - 0.1$. Then, in (O, O), the minimum utility of all players is 0.9, whereas in (F, F), the minimum utility of all players is 1. Therefore, learning converges to play the unique stochastically stable state (F, F), where player 2 receives the higher utility of 2 by increasing their exploration probability.

2.4 Proof of Theorem 1

To prove Theorem 1, we proceed in three steps. First, we show that the joint evolution of beliefs and strategies forms a regular perturbation of an unperturbed Markov chain with $\xi = 0$, where hypothesis tests are exact and exploration occurs with probability tending to zero. Second, we show that the recurrent classes of the unperturbed chain are exactly the consistent states, each forming an absorbing singleton. Moreover, every inconsistent state transitions to one of these consistent states with positive probability. Finally, we apply the resistance-tree method from [117] to compute the stochastic potential of each recurrent class and identify the stochastically stable set as the set of recurrent classes with minimum stochastic potential in a regular perturbed Markov chain. We compute the resistance of transitions between recurrent classes in our learning dynamics and show that the classes minimizing this potential correspond exactly to the set of consistent states that maximize the minimum transformed utility across players.

2.4.1 Regular Perturbation

We define the limiting process as the exploration parameter $\xi \to 0$. In this limit, the original Markov process defined by the transition matrix $P^{\xi} = (P^{\xi}_{zz'})_{z,z' \in \mathcal{Z}}$ converges to an unperturbed process governed by the transition matrix $P^0 = (P^0_{zz'})_{z,z' \in \mathcal{Z}}$. In the unperturbed process, both Type I and Type II error rates in hypothesis testing become 0, and players do not explore when they fail to reject the null hypothesis. As a result, for each player $i \in I$, if $\|b_i - \pi_{-i}\|_2 \leq \tau$, then they retain their current belief with probability 1. Conversely, if $\|b_i -$

$\pi_{-i}\|_2 > \tau$, then the player rejects the null hypothesis and updates their belief with probability 1.

We introduce the definition of a regular perturbation and show that for any $\xi > 0$, the original Markov process P^ξ is a regular perturbation of the unperturbed process P^0.

Definition 5 (Young [117])**.** *Let P^0 be the transition matrix of a stationary Markov chain defined on a finite state space $\mathcal{Z}$. Let P^ξ be the transition matrix of a family of Markov chains on $\mathcal{Z}$ perturbed from P^0, indexed by $\xi \in (0, \bar{\xi})$ for some $\bar{\xi} > 0$. The family $\{P^\xi\}_{\xi>0}$ is a* regular perturbation *of P^0 if the following conditions hold for all $z, z' \in \mathcal{Z}$:*

- *(i)* *P^ξ is aperiodic and irreducible for all $\xi \in (0, \bar{\xi})$.*
- *(ii)* $\lim_{\xi \to 0} P^\xi_{zz'} = P^0_{zz'}$.
- *(iii)* *If $P^\xi_{zz'} > 0$ for some ξ, then there exists $r_{zz'} \geq 0$ such that $0 < \lim_{\xi \to 0} P^\xi_{zz'} \cdot \xi^{-r_{zz'}} < \infty$.*

Lemma 2. *(1) For any $z = (b, \pi), z' = (b', \pi') \in \mathcal{Z}$, the transition probability of the unperturbed process P^0 satisfies:*

$$P^0_{zz'} = \prod_{i \in I} P^0_{i,zz'},$$

where $P^0_{i,zz'}$ denotes the probability of player i updating their belief from b_i to b'_i:

$$P^0_{i,zz'} = \begin{cases} 1, & \text{if } \|b_i - \pi_{-i}\|_2 \leq \tau \text{ and } b_i = b'_i, \\ 0, & \text{if } \|b_i - \pi_{-i}\|_2 \leq \tau \text{ and } b_i \neq b'_i, \\ (1-\gamma_i) + \gamma_i \cdot \frac{1}{|\mathcal{B}_i|}, & \text{if } \|b_i - \pi_{-i}\|_2 > \tau \text{ and } b_i = b'_i, \\ \gamma_i \cdot \frac{1}{|\mathcal{B}_i|}, & \text{if } \|b_i - \pi_{-i}\|_2 > \tau \text{ and } b_i \neq b'_i. \end{cases} \tag{2.5}$$

(2) The Markov process P^{ξ} is a regular perturbation of P^0. Moreover, for any $z = (b, \pi), z' = (b', \pi') \in \mathcal{Z}$,

$$0 < \lim_{\xi \to 0} P^{\xi}_{zz'} \cdot \xi^{-r_{zz'}} < \infty,$$

where

$$r_{zz'} = \sum_{\substack{i \in \{i \in I \mid b_i \neq b'_i, \\ \|b_i - \pi_{-i}\|_2 \leq \tau\}}} f_i(U_i(\pi_i, b_i)). \tag{2.6}$$

In Lemma 2, part (1) formalizes the structure of transitions in the unperturbed process: players update beliefs only when the hypothesis test rejects the null, and belief updates are uniformly random over the discretized belief space. Part (2) verifies that the belief-based learning dynamics define a *regular perturbation* of the limiting Markov process P^0 as the exploration parameter $\xi \to 0$, and shows how every nonzero transition in the perturbed process changes with ξ. This result is essential for applying the resistance-tree framework of stochastic stability.

2.4.2 Recurrent Communication Class of the Unperturbed Process

A key concept of Markov chains is the recurrent communication classes. Once the chain enters a recurrent communication class, it never leaves, and any state in the class can be reached from any other state in the class with positive probability. This is formally defined as follows:

Definition 6. *A recurrent communication class of a Markov chain is a non-empty subset of states $C \subseteq \mathcal{Z}$ such that:*

(i) For all $z, w \in C$, there exist integers $t, t' \geq 0$ such that

$$\Pr(z^t = w \mid z^0 = z) > 0, \quad \textit{and} \quad \Pr(z^{t'} = z \mid z^0 = w) > 0.$$

That is, all states in C communicate with each other.

(ii) For any $z \in C$ and any $w \notin C$, we have

$$\Pr(z^t = w \mid z^0 = z) = 0, \quad \forall t \geq 0.$$

That is, the class C is closed*: the chain cannot exit once it enters.*

In the following lemma, we show that each consistent belief state $z \in \mathcal{Z}^\dagger$ forms a recurrent communication class of the unperturbed Markov process P^0.

Lemma 3. *For each consistent belief state $z \in \mathcal{Z}^\dagger$, $\{z\}$ is a recurrent communication class of the unperturbed process P_0. Furthermore, there are no other recurrent communication class of P_0.*

Proof. We first show that every state z in $\mathcal{Z}^\dagger$ is an absorbing state. For any state $z = (b, \pi) \in \mathcal{Z}^\dagger$, every player's hypothesis is τ-consistent with the actual strategy, i.e. $\|b_i - \pi_{-i}\|_2 \leq \tau$ for all $i \in I$. From Lemma 2, we know for all $i \in I$, when $\|b_i - \pi_{-i}\|_2 \leq \tau$, we have

$$P^0_{i,zz'} = \begin{cases} 1, & \text{if } b_i = b'_i, \\ 0, & \text{if } b_i \neq b'_i. \end{cases}$$

Therefore, all players will keep their current belief and strategy, and the Markov chain remains in the same state:

$$P^0_{zz'} = \prod_{i \in I} P^0_{i,zz'} = \begin{cases} 1, & \text{if } z = z', \\ 0, & \text{if } z \neq z'. \end{cases}$$

Therefore, any consistent belief state $z \in \mathcal{Z}^\dagger$ is an absorbing state and $\{z\}$ is a recurrent communication class of the unperturbed process P^0 for all $z \in \mathcal{Z}^\dagger$.

To show that there are no other recurrent communication class in the unperturbed process, it suffices to show that for any non-consistent state $w = (b^w, \pi^w) \notin \mathcal{Z}^\dagger$, there exists a belief consistent state $z = (b^z, \pi^z) \in \mathcal{Z}^\dagger$ that is reachable from w. Since w is not a consistent belief state, there exists at least one player i whose belief is not τ-consistent, i.e.

$$\|b_i^w - \pi_{-i}^w\|_2 > \tau. \tag{2.7}$$

From Assumption 2, we know that there exists a belief $\tilde{b}_i$ that satisfies:

$$\|\tilde{b}_i - \pi_{-i}^w\|_2 > \tau, \tag{2.8}$$

and the associated smooth best respond strategy $\tilde{\pi}_i = \mathrm{Br}_i^\sigma(\tilde{b}_i)$ is not consistent with any other player's belief in w:

$$\|\tilde{\pi}_i - b_{ji}^w\|_2 > \tau, \quad \forall j \in I \setminus \{i\}. \tag{2.9}$$

We denote another state $\tilde{w} = (\tilde{b}_i, b_{-i}^w, \tilde{\pi}_i, \pi_{-i}^w)$. We show that the state transition path $w \to \tilde{w} \to z$ has positive probability in the unperturbed process.

Transition step 1: $w \to \tilde{w}$. Player i updates their belief from b_i^w to $\tilde{b}_i$ with probability $P^0_{i,w\tilde{w}}$, and all other players $j \neq i$ do not update. From (2.5) and (2.7), $P^0_{i,w\tilde{w}}$ is given by:

$$\begin{aligned} P^0_{i,w\tilde{w}} &= \begin{cases} (1-\gamma_i) + \gamma_i \cdot \frac{1}{|\mathcal{B}_i|}, & \text{if } b_i^w = \tilde{b}_i, \\ \gamma_i, \cdot \frac{1}{|\mathcal{B}_i|} & \text{if } b_i^w \neq \tilde{b}_i, \end{cases} \\ &\geq \gamma_i \cdot \frac{1}{|\mathcal{B}_i|}. \end{aligned}$$

For any $j \neq i$, player j's belief b_j^w does not change when the state transits from w to $\tilde{w}$:

$$P^0_{j,w\tilde{w}} = \begin{cases} 1 & \text{if } \|b_j^w - \pi_{-j}^w\|_2 \leq \tau \\ (1-\gamma_i) + \gamma_i \cdot \frac{1}{|\mathcal{B}_j|} & \text{if } \|b_j^w - \pi_{-j}^w\|_2 > \tau \end{cases}$$
$$\geq (1-\gamma_i) + \gamma_i \cdot \frac{1}{|\mathcal{B}_j|}.$$

Hence, the probability of state transits from w to $\tilde{w}$ is positive:

$$P^0_{w\tilde{w}} = P^0_{i,w\tilde{w}} \cdot \prod_{j\neq i} P^0_{j,w\tilde{w}} \geq (\gamma_i \cdot \frac{1}{|\mathcal{B}_i|}) \cdot ((1-\gamma_i) + \gamma_i \cdot \frac{1}{|\mathcal{B}_j|})^{n-1} > 0.$$

Transition step 2: $\tilde{w} \to z$. Player i updates their belief from $\tilde{b}_i$ to b_i^z, and every other player $j \neq i$ updates their belief from b_j^w to b_j^z. Since $\|\pi_{-j}^{\tilde{w}} - b_j^{\tilde{w}}\|_2 \geq \|\tilde{\pi}_i - b_{ji}^w\|_2 \overset{(2.9)}{>} \tau$, from (2.5), $P^0_{j,\tilde{w}\tilde{z}}$ satisfies

$$P^0_{j,\tilde{w}\tilde{z}} = \begin{cases} (1-\gamma_i) + \gamma_i \cdot \frac{1}{|\mathcal{B}_j|}, & \text{if } b_j^w = b_j^z, \\ \gamma_i \cdot \frac{1}{|\mathcal{B}_j|}, & \text{if } b_j^w \neq b_j^z, \end{cases}$$
$$\geq \gamma_i \cdot \frac{1}{|\mathcal{B}_j|}.$$

From (2.8), we have $\|\tilde{b}_i - \pi_{-i}^w\|_2 > \tau$. From (2.5),

$$P^0_{i,\tilde{w}z} = \begin{cases} (1-\gamma_i) + \gamma_i \cdot \frac{1}{|\mathcal{B}_i|}, & \text{if } \tilde{b}_i = b_i^z, \\ \gamma_i \cdot \frac{1}{|\mathcal{B}_i|}, & \text{if } \tilde{b}_i \neq b_i^z. \end{cases}$$
$$\geq \gamma_i \cdot \frac{1}{|\mathcal{B}_j|}.$$

Hence, transition from $\tilde{w}$ to $\tilde{z}$ happens with positive probability:

$$P^0_{\tilde{w}\tilde{z}} = P^0_{i,\tilde{w}\tilde{z}} \cdot \prod_{j\neq i} P^0_{j,\tilde{w}\tilde{z}} \geq (\gamma_i \cdot \frac{1}{|\mathcal{B}_i|})^n > 0.$$

Summarizing, the probability of the state transitioning from w to z in two epoches is positive:

$$\Pr(z^2 = z \mid z_0 = w) \geq P^0_{w\tilde{w}} \cdot P^0_{\tilde{w}z} > 0.$$

Hence, any non belief consistent state $w \notin \mathcal{Z}^\dagger$ cannot be in a recurrent communication class. □

2.4.3 Tree Resistance and Stochastic Potential

Having established in Steps 1 and 2 that the learning dynamics form a regular perturbation of an unperturbed Markov chain and that the recurrent classes of the unperturbed chain correspond to consistent states, we now characterize which of these classes comprise the stochastically stable set. While all consistent states are absorbing in the unperturbed dynamics, vanishing but nonzero perturbations induce transitions between them with small probabilities. To compare the relative stability of these recurrent classes under perturbation, we adopt the notion of resistance introduced in [117], which captures the leading-order exponent of the transition probability as a function of the perturbation magnitude. Formally, we present the definitions of edge and path resistance, z-tree and stochastic potential as follows:

Definition 7 (Edge resistance, Path Resistance, Resistence Tree, and Stochastic Potential [117])**.** *Consider a perturbed Markov process on a finite state space $\mathcal{Z}$, with transition matrix P^ξ. Let G be a fully connected directed graph with states $\mathcal{Z}$ being the node set.*

- ***Edge Resistance.*** *For any $z, z' \in \mathcal{Z}$, if $P^\xi_{zz'} > 0$, the resistance of the edge*

$(z \to z')$ is the unique number $r_{zz'} \geq 0$ such that

$$0 < \lim_{\xi \to 0} P^{\xi}_{zz'} \cdot \xi^{-r_{zz'}} < \infty.$$

If $P^{\xi}_{zz'} = 0$, we define $r_{zz'} = \infty$.

- ***Path Resistance.*** *For a transition path $\rho_{z_1 \to z_n} = z_1 \to z_2 \to \cdots \to z_n$, the resistance of the path is the sum of its edge resistances:*

$$R(\rho_{z_1 \to z_n}) = \sum_{j=1}^{n-1} r_{z_j z_{j+1}}.$$

- ***Minimum Path Resistance.*** *For any $z, z' \in \mathcal{Z}$, let $\hat{r}_{wz}$ be the smallest path resistance over all paths that start in w and end in z:*

$$\hat{r}_{wz} = \min_{\rho_{w \to z}} R(\rho_{w \to z}).$$

Let $\mathcal{G}$ be a fully connected directed graph with each node being a recurrent class of the Markov chain. The edge weight between any two recurrent classes (nodes) $C_i, C_j \in \mathcal{C}$ is defined as the minimum path resistance $\hat{r}_{ij}$ across all paths that begin in C_i and end in C_j in the original graph G. The following concepts are defined on the graph $\mathcal{G}$:

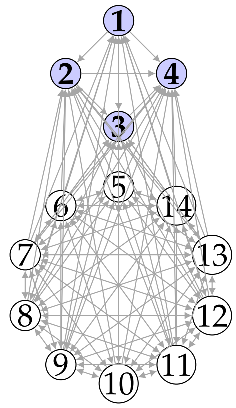
- *z-**Tree.*** *For any node $C_j \in \mathcal{C}$, a j-tree Γ is a spanning directed tree over $\mathcal{C}$ such that for every node $C_i \neq C_j$, there is a unique directed path from C_i to C_j. The weight of a j-tree is the sum of the weight of its constituent edges.*
- ***Stochastic Potential.*** *The stochastic potential $\phi(j)$ of a recurrent class $C_j \in \mathcal{C}$ is defined as the minimum weight over all possible j-trees:*

$$\phi(j) = \min_{\Gamma \text{ is a } j\text{-tree}} \sum_{(C_j \to C_{j'}) \in \Gamma} \hat{r}_{jj'}.$$

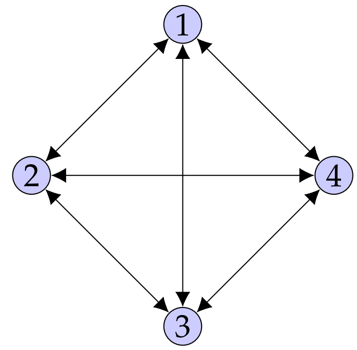
In Definition 7, the edge resistance $r_{zz'}$ captures the leading-order exponent of the one-epoch state transition probability $P^{\xi}_{zz'}$ in ξ; transitions with higher resistance become exponentially less likely as ξ decreases. From Lemma 3, we know that each consistent state z induces a singleton recurrent class $\{z\}$. A path resistance then aggregates these one-step state transition resistances over sequences of transitions. The notion of a z-tree captures how the state transition process can reach a particular consistent state z from all other consistent states through the unique paths determined by the tree, and its total resistance reflects the leading-order exponent of the total probability of reaching that state through the paths in the tree. The z-tree with the minimum potential contains paths such that reaching a consistent state z through those paths has the highest probability, and the minimum resistance is defined as the stochastic potential $\phi(z)$ of consistent state z. The stochastic potential $\phi(z)$ quantifies how the maximum probability of reaching a consistent state z changes with ξ in the perturbed process: the probability of reaching states with lower stochastic potential decays more slowly than for other consistent states with higher stochastic potential as $\xi \to 0$. Indeed, the following lemma from Young [117] shows that the stochastically stable states are those belonging to a recurrent communication class that minimizes its stochastic potential.

Lemma 4 (Young [117])**.** *Let P^0 be a stationary Markov process on the finite state space $\mathcal{Z}$ with recurrent communication classes $C_1, \ldots, C_J$. Let P^ξ be a regular perturbation of P^0. Then, a state z is stochastically stable if and only if z is contained in a recurrent communication class C_j that minimizes its stochastic potential.*

To complete the proof of Theorem 1, it remains to identify which of the recurrent classes of the unperturbed process P^0 have the minimum stochastic potential. Recall from Lemma 3, each consistent state in $\mathcal{Z}^\dagger$ forms a singleton recurrent class. The following lemma characterizes the edge and minimum path resistances between any pair of consistent states, which will be used for com-



(a) Graph G



(b) Graph $\mathcal{G}$

Figure 2.2: Graph G represents the fully connected Markov Process graph in which the blue nodes represent consistent states, and white nodes represent inconsistent states. Graph $\mathcal{G}$ is the fully connected reduced graph where each node represents a consistent state.

puting the weight of any z-tree.

Lemma 5 (Minimum resistance between consistent states). *Let $z = (b, \pi)$, $z' = (b', \pi') \in \mathcal{Z}$. The edge resistance between z and z' is*

$$r_{zz'} = \sum_{i \in I} r_{i,zz'}, \qquad r_{i,zz'} = \begin{cases} f_i(U_i(\pi_i, b_i)), & \text{if } \|b_i - \pi_{-i}\|_2 \leq \tau \text{ and } b_i \neq b'_i, \\ 0, & \text{otherwise.} \end{cases} \tag{2.10}$$

For any $z = (b, \pi) \in Z^\dagger$ and any $w \in Z^\dagger$, the edge weight $\hat{r}_{zw}$ is

$$\hat{r}_{zw} = \min_{i \in I} f_i(U_i(\pi_i, b_i)). \tag{2.11}$$

Proof of Lemma 5. From part (2) of Lemma 2, for any $z = (b, \pi), z' \in \mathcal{Z}$, we have that

$$0 < \lim_{\xi \to 0} P^{\xi}_{zz'} \cdot \xi^{-r_{zz'}} < \infty,$$

for

$$r_{zz'} = \sum_{\substack{i \in \{i \in I \mid b_i \neq b'_i, \\ \|b_i - \pi_{-i}\|_2 \leq \tau\}}} f_i(U_i(\pi_i, b_i)).$$

Therefore, the edge resistance between z and z' satisfies

$$r_{zz'} = \sum_{i \in I} r_{i,zz'}, \qquad r_{i,zz'} = \begin{cases} f_i\big(U_i(\pi_i, b_i)\big), & \text{if } \|b_i - \pi_{-i}\|_2 \leq \tau \text{ and } b_i \neq b'_i, \\ 0, & \text{otherwise.} \end{cases}$$

We next prove that the edge weight is given by (2.11). Consider any belief consistent states $z = (b^z, \pi^z), w = (b^w, \pi^w) \in \mathcal{Z}^\dagger$. We first prove that $\hat{r}_{zw} \leq \min_{i \in I} f_i(U_i(\pi_i, b_i))$. Recall that the edge weight $\hat{r}_{zw}$ is defined as the minimum resistance over all possible paths between z and w. It suffices to present a transition path from z to w with resistance $\min_{i \in I} f_i(U_i(\pi_i, b_i))$. Let player $i^\dagger$ be a player with the smallest $f_i(U_i(\pi_i, b_i))$ among all players. That is,

$$f_{i^\dagger}(U_{i^\dagger}(\pi_i, b_i)) = \min_{j \in I} f(U_j(\pi_i, b_i)).$$

From Assumption 2, there exists a belief $\tilde{b}_{i^\dagger} \in \mathcal{B}_{i^\dagger}$ for player $i^\dagger$ that is at least τ distance away from strategy $\pi^z_{-i^\dagger}$:

$$\|\tilde{b}_{i^\dagger} - \pi^z_{-i^\dagger}\|_2 > \tau,$$

and its σ-smooth strategy $\tilde{\pi}_{i^\dagger} = \mathrm{Br}^\sigma_{i^\dagger}(\tilde{b}_{i^\dagger})$ is far from all other players' beliefs in b^z:

$$\|\tilde{\pi}_{i^\dagger} - b^z_{ji^\dagger}\|_2 > \tau, \quad \forall j \in I \setminus \{i^\dagger\}.$$

We define $\tilde{z} = (\tilde{b}_{i^\dagger}, b^z_{-i^\dagger}, \tilde{\pi}_{i^\dagger}, \pi^z_{-i^\dagger})$. We compute the resistance of state transition path: $z \to \tilde{z} \to w$.

Transition step $z \to \tilde{z}$. Player $i^\dagger$ updates their belief to $\tilde{b}_{i^\dagger}$, and all other players $j \neq i^\dagger$ keep their belief b^z_j. Since $z \in \mathcal{Z}^\dagger$ is a belief consistent state, we have $\|b^z_j - \pi^z_{-j}\|_2 \leq \tau$ for all $j \in I$. From (2.10), we have $r_{i^\dagger,z\tilde{z}} = f_{i^\dagger}(U_{i^\dagger}(\pi_{i^\dagger}, b_{i^\dagger}))$ and $r_{j,z\tilde{z}} = 0$ for all $j \neq i^\dagger$, and the edge resistance $r_{z\tilde{z}}$ between z and $\tilde{z}$ is

$$r_{z\tilde{z}} = \sum_{j\in I} r_{j,z\tilde{z}} = f_{i^\dagger}(U_{i^\dagger}(\pi_{i^\dagger}, b_{i^\dagger})).$$

Transition step $\tilde{z} \to w$. Recall that $\tilde{z} = (\tilde{b}_{i^\dagger}, b^z_{-i^\dagger}, \tilde{\pi}_{i^\dagger} = \mathrm{Br}^\sigma_{i^\dagger}(\tilde{b}_{i^\dagger}), \pi^z_{-i^\dagger})$, and $\|\tilde{b}_{i^\dagger} - \pi^z_{-i^\dagger}\|_2 > \tau$. Hence from (2.10), we have

$$r_{i^\dagger,\tilde{z}w} = 0.$$

Since for all $j \neq i^\dagger$, we have $\|\tilde{\pi}_{i^\dagger} - b^z_{ji^\dagger}\|_2 > \tau$. Then,

$$\|\pi^{\tilde{z}}_{-j} - b^{\tilde{z}}_j\|_2 \geq \|\tilde{\pi}_{i^\dagger} - b^z_{ji^\dagger}\|_2 > \tau.$$

From (2.10), we have

$$r_{j,\tilde{z}w} = 0, \quad \forall j \in I \setminus \{i^\dagger\}.$$

Hence, the edge resistance $r_{\tilde{z}w}$ between $\tilde{z}$ and w is

$$r_{\tilde{z}w} = \sum_{j\in I} r_{j,\tilde{z}w} = 0.$$

The path resistance of path $\rho_{z\to w} = z \to \tilde{z} \to w$ is

$$R(\rho_{z\to w}) = r_{z\tilde{z}} + r_{\tilde{z}w} = f_{i^\dagger}(U_{i^\dagger}(\pi_{i^\dagger}, b_{i^\dagger})).$$

Since the edge weight $\hat{r}_{zw}$ is defined as the smallest resistance over all possible paths between z and w, we have

$$\hat{r}_{zw} \leq R(\rho_{z\to w}) = f_{i^\dagger}(U_{i^\dagger}(\pi_{i^\dagger}, b_{i^\dagger}) = \min_{j\in I} f(U_j(\pi_i, b_i)). \tag{2.12}$$

Next, we show that the edge weight $\hat{r}_{zw}$ is at least $\min_{i\in I} f_i(U_i(\pi_i, b_i))$. We argue that any transition path that starts at $z = (b^z, \pi^z) \in \mathcal{Z}^\dagger$ and ends at state $w \in \mathcal{Z}^\dagger$ has resistance larger than or equal to $\min_{i\in I} f_i(U_i(\pi_i, b_i))$. Consider any transition path from z to w. Since $z \neq w$, there must exists $z' = (b', \pi') \neq z$ such that $z \to z'$ is on the path. Since $z' \neq z$, we have $b^z \neq b'$. Hence, there must exist at least one player $i \in I$ who changes their belief, i.e. $b_i^z \neq b_i'$. Since $\|b_i^z - \pi_{-i}^z\|_2 \leq \tau$, from (2.10), we have

$$r_{i,zz'} = f_i(U_i(\pi_i, b_i)).$$

Since $r_{j,zz'} \geq 0$ for all $j \neq i$, the edge resistance $r_{zz'}$ between z and z' satisfies

$$r_{zz'} = \sum_{j\in I} r_{j,zz'} \geq r_{i,zz'} = f_i(U_i(\pi_i, b_i)).$$

Since the resistance of any edge is nonnegative, for any path $\rho_{z\to w} = z \to z' \to \cdots \to w$ where $z' = (b', \pi')$ is the first state visited on the path other than z, the path resistance of $\rho_{z\to w}$ must satisfy:

$$R(\rho_{z\to w}) \geq r_{zz'} = f_i(U_i(\pi_i, b_i)), \quad \forall i \text{ such that } b_i^z \neq b_i'.$$

Since the edge weight $\hat{r}_{zw}$ is defined as the smallest resistance over all $\rho_{z\to w} \in \mathcal{P}_{z\to w}$, it satisfies:

$$\hat{r}_{zw} = \min_{\rho_{z\to w}\in\mathcal{P}_{z\to w}} R(\rho_{z\to w}) \geq \min_{i\in I} f_i(U_i(\pi_i, b_i)). \tag{2.13}$$

Combing (2.12) and (2.13), we have

$$\hat{r}_{zw} = \min_{i\in I} f_i(U_i(\pi_i, b_i)), \quad \forall z = (b, \pi), w \in \mathcal{Z}^\dagger.$$

□

With the edge weight between consistent states characterized in Lemma 5, we are now ready to complete the proof of Theorem 1. Specifically, we show that the consistent states that maximize the minimum transformed utility across all players also minimize stochastic potential, and thus constitute the stochastically stable set.

Proof of Theorem 1. Consider any $z \in \mathcal{Z}^\dagger$ and the associated z-tree. For any $w \neq z \in \mathcal{Z}^\dagger$, there exists exactly one edge that starts from w in the z-tree. That is, there exists exactly one $w' \in \mathcal{Z}^\dagger$ such that (w, w') is in the z-tree. From Lemma 5, for any $w = (b^w, \pi^w) \neq z \in \mathcal{Z}^\dagger$, the edge weight is

$$\hat{r}_{ww'} = \min_{i \in I} f(U_i(\pi_i^w, b_i^w)).$$

Notice that the edge weight $\hat{r}_{ww'}$ only depends on w. Any z-tree must have the same total weight given by

$$\sum_{\substack{w=(b^w,\pi^w) \\ w \in \mathcal{Z}^\dagger \setminus \{z\}}} \min_{i \in I} f(U_i(\pi_i^w, b_i^w)) = \sum_{\substack{w=(b^w,\pi^w) \\ w \in \mathcal{Z}^\dagger}} \min_{i \in I} f(U_i(\pi_i^w, b_i^w)) - \min_{i \in I} f(U_i(\pi_i^z, b_i^z)).$$

Therefore, the stochastic potential of $z \in \mathcal{Z}^\dagger$ is

$$\begin{aligned} \phi(z) &= \sum_{\substack{w=(b^w,\pi^w) \\ w \in \mathcal{Z}^\dagger \setminus \{z\}}} \min_{i \in I} f_i(U_i(\pi_i^w, b_i^w)) \\ &= \underbrace{\sum_{w=(b^w,\pi^w) \in \mathcal{Z}^\dagger} \min_{i \in I} f(U_i(\pi_i^w, b_i^w))}_{\text{constant}} - \min_{i \in I} f(U_i(\pi_i^z, b_i^z)). \end{aligned}$$

From Proposition 4, we know that a state z is stochastically stable if and only if z is in a recurrent communication class and z has the smallest stochastic potential. Therefore, a state $z^* = (b^*, \pi^*)$ is stochastically stable if and only if $\min_{i \in I} f(U_i(\pi_i^*, b_i^*))$ is the largest among all states $z \in \mathcal{Z}^\dagger$. That is,

$$\min_{i \in I} f(U_i(\pi_i^*, b_i^*)) = \max_{z=(b^z,\pi^z) \in \mathcal{Z}^\dagger} \min_{i \in I} f(U_i(\pi_i^z, b_i^z)).$$

□

2.5 Concluding Remarks

This work develops a new learning dynamics that integrates hypothesis testing and utility-sensitive exploration, and analyzes how players adapt in general strategic environments. Beyond proving that the learning dynamics converge to approximate Nash equilibria in general normal form games, we show that the dynamics further select equilibria that maximize the minimum transformed utility across all players. This refinement is endogenously induced by the exploration mechanism adopted by agents in the learning process. Our analysis relies on interpreting the dynamics as a regular perturbation of a Markov process and characterizing the stochastically stable states using resistance-tree methods. Future directions include exploring equilibrium refinements induced by alternative exploration functions and applying this learning model to stochastic games and extensive-form games.

CHAPTER 3

DESIGNING HIGH-OCCUPANCY TOLL LANES: A GAME-THEORETIC ANALYSIS

3.1 Introduction

High Occupancy Toll (HOT) lanes are traffic lanes or roadways that are open to vehicles satisfying a minimum occupancy requirement but also offer access to other vehicles with a toll price. In practice, HOT lanes have been implemented on several interstate highways in California, Texas, and Washington states. With the proper design of lane capacity and toll price, HOT lanes can effectively mitigate traffic congestion through incentivizing carpooling and transit use, while also generating revenue to support transportation infrastructure through toll collection.

The goal of our work is to study the optimal design of HOT lane systems and its impact on traffic congestion, social welfare, and revenue generated from toll collection. In our model, a traffic authority designs the HOT lane systems by choosing the road capacity of HOT lanes, and the toll price. Given the design of HOT, we develop a game-theoretic model to analyze the strategic decisions made by travelers who have the action set of paying or carpooling to use the HOT lane, or using the ordinary lane. Travelers are modeled as a population of nonatomic agents with a continuous distribution of value of time and carpool disutility. Both the HOT lanes and the ordinary lanes are congestible in that the travel time of each lane increases with the aggregate flow induced by agents' decisions. The outcomes of the system in terms of agent travel time cost and toll collection are jointly determined by agents' equilibrium strategies and the design of the traffic authority.

In the first part of our work (Sec. 3.2-3.3), we consider highway segments with a single entrance and exit node. We provide a complete characterization of Wardrop equilibrium in this game. In particular, we identify two qualitatively distinct equilibrium regimes that depend on the traffic authority's design of lane capacity and toll price. In the first equilibrium regime, all agents who take the HOT lane form carpools, and no one pays the toll due to the relatively high toll price. In the second equilibrium regime, a fraction of agents with high carpool disutilities and high value of times make toll payments to take the HOT lanes, while the rest either form carpools or take the ordinary lanes. In both regimes, agents are split between taking the HOT lanes and the ordinary lanes.

The equilibrium characterization provides the system designer with insights on how the equilibrium flows and travel time costs of both the HOT lanes and the ordinary lanes depend on the system parameters that include travel time cost functions, capacity allocation and toll price. Moreover, we present comparative static analysis on how the equilibrium flow and costs change with the fraction of capacity that is allocated to the HOT lanes and the toll price of HOT lanes. We find that if we increase the HOT capacity while holding the toll price fixed, the latency difference between ordinary lanes and HOT lanes increases. Moreover, more agents use the HOT lanes by paying the toll price or carpooling and fewer agents use ordinary lanes. On the other hand, increasing the toll price while holding the HOT capacity fixed will lead to an increase in the latency difference between ordinary lanes and HOT lanes, an increase in carpooling agents, and a decrease in toll-paying agents. The number of agents using ordinary lanes can change in either direction.

In the second part of our work (Sec. 3.4-3.5), we generalize our model to highways with multiple segments separated by entrance and exit nodes and a carpool system with multiple occupancy levels. The toll price of each segment set by the system designer varies with the vehicle's occupancy level. The population consists of agents with different entry and exit points. Agents choose their carpool occupancy levels before entering the highway, and can switch between ordinary lanes and HOT lanes for different road segments. We generalize our equilibrium concept to this model extension, and prove equilibrium existence. We also provide a generic sufficient condition under which the equilibrium is unique.

We apply our model and equilibrium analysis in the numerical study using the data collected on the California I-880 HOT lane system, from Dixon Landing Road to Lewelling Boulevard. We calibrate the latency function using vehicle travel time and flow data provided by the Caltrans Performance Measurement System (PeMS). To compute the equilibrium strategy distribution, we need to estimate the demand for each entry and exit pair and the distribution of preference parameters among the population. To ensure tractability in estimating demand and preference distribution, we partition the preference parameter vector space into equally sized subsets, and estimate the demand of agents with preference in each subset building on the idea of inverse optimization. That is, given the data on toll prices and driving time, we compute the equilibrium strategy profile of agents with all preference parameters and entrance and exit nodes. This allows us to map the demand estimate of each preference set to induced vehicle flows on ordinary lanes and HOT lanes at equilibrium. We formulate a convex program to estimate the demand volumes as the one that minimizes the difference between the equilibrium vehicle flows and the observed flows on each lane and each segment.

Next, we compute the equilibrium strategy profile for a set of discretized design parameters, including capacity allocation and toll prices. We consider four objective functions for the HOT design: (i) the total agent travel time, (ii) the total vehicle driving time, (iii) the total revenue measures the toll prices collected, and (iv) total cost of all agents taking into account their driving time, toll payments, and carpool disutilities. We select a time interval (5-6 pm) during the evening rush hour to compute the Pareto front for the design of HOT lanes under various toll prices and HOT capacities. This analysis illustrates the system authority's trade-off between reducing traffic congestion, enhancing social welfare, and maximizing total toll revenue. Charging a high toll price on road segments with higher demand is effective in incentivizing agents to carpool and reduce both agent travel time and vehicle driving time. On the other hand, for revenue maximization, a lower toll price is optimal to increase the fraction of toll-paying agents.

Moreover, we compute the optimal toll design under the current HOT capacity across all operating hours of a workday. Since demand volume is lower during the morning hours but higher in the afternoon, setting a high toll price in the afternoon is more effective for reducing agent travel time, vehicle driving time, and costs. However, for revenue maximization, it is optimal to set a low toll price throughout the day. By adjusting the current toll price to the optimal toll prices for each of these four objectives, we can achieve significant improvements in the corresponding objective.

Related Literature. Our model and analysis build on the rich literature of congestion games that includes the equilibrium analysis of routing strategies made by atomic agents [80, 91] and nonatomic agents [97] in networks, and the analysis on the price of anarchy [24, 87, 92, 95]. Most of the classical results in congestion games have focused on the settings where all agents have homogeneous preferences. The papers [77, 79] extended these results to study the equilibrium existence and efficiency with player-specific costs. Previous literature has also examined the optimal design of tolling mechanisms that minimize the social cost of nonatomic agents with homogeneous preferences [19, 93, 94]. The optimal toll design with heterogeneous values of time has also been studied. For example, studies such as [10–12, 51, 58, 65, 66, 90, 102] consider a finite number of agent classes, where the value of time is the same among agents within the same class but differs across different classes. Furthermore, other works examine the setting where the value of time of agents follows continuous distribution. These works do not incorporate carpooling into their models [23, 36, 54, 68, 85, 116].

There are extensive works modeling travelers' decisions regarding HOT/HOV lanes. The first stream of work examined the static user equilibrium of travel mode and/or route choices on a single-segment, multi-lane road. Yang and Huang [115] studied the user equilibrium with and without HOV lanes and discussed the optimal congestion pricing. Zang et al., Chu et al., and Hughes and Kaffine [20, 52, 122] extended the HOV system with various schemes and settings. These works focus on HOV systems, which do not allow single occupancy vehicles to use the express lane. Moreover, they assume that all commuters have a homogeneous value of time and carpool disutility, except for Hughes and Kaffine [52], which considered commuters with heterogeneous carpool disutilities but homogeneous value of time. Considering the HOT systems that allow single occupancy vehicles to drive on the express lanes by paying a toll, Jang et al., Lou et al.,and Tan and Gao [53, 67, 99] discussed the pricing schemes where only single occupancy vehicles make decisions between HOT lanes and ordinary lanes while the number of high occupancy vehicles remains constant. Konishi and Mun [59], allowing travelers to choose both their travel modes and routes simultaneously, compared the road efficiency for HOT and HOV systems, and explored the optimal congestion pricing scheme for both ordinary lanes and HOV lanes on a multi-lane highway. They also considered commuters with heterogeneous carpool disutilities but a homogeneous value of time. Yuan et al. [121] studied the impact of ride-sourcing vehicles on both HOV and HOT systems, considering commuters with heterogeneous values of time but homogeneous carpool disutility. Additionally, serval other works [18, 28, 29, 62, 63, 112–114] extended these models to consider ridesharing user equilibrium (RUE) with HOT lanes. None of the above works considers commuters with both heterogeneous values of time and carpool disutilities, and they only considered homogeneous carpool occupancy levels.

More broadly, the second stream of works studied the impact of carpooling in a dynamic setting that extends the bottleneck model by Vickrey [104], and examined departure time equilibrium and/or travel mode choice equilibrium with HOT/HOV lanes [61]. Wu et al., and Qian and Zhang [89, 108] studied the HOV/HOT system with three traveling modes: transit, driving alone, and carpooling. They derived the departure time equilibrium for each traveling mode and how different factors affect the mode shares and network performance. A few other papers [64, 105, 106, 109, 110, 123] extended the same problem to include parking space constraints or ride-sharing compensation. However, all the above works considered only homogeneous commuters. Yu et al. [119] studied departure time equilibrium with heterogeneous users in the preference for cost of carpooling, values of time, and values of schedule delay. However, they modeled travelers' mode choice and route choice separately: although travelers achieve departure time equilibrium within each mode, the shares among the modes are determined by a nested logit model. Also, they only considered a single-lane scenario, which does not incorporate separate HOV/HOT lanes. Moreover, all the dynamic works above either assume a constant carpool occupancy level ([111], [110], [89], [106], [120]), or carpool occupancy level that is a continuous variable as in [123]. The paper [107] extended the temporal capacity allocation scheme with discrete heterogeneous occupancy level and heterogeneous carpool inconvenience costs, but considered a homogeneous value of time.

Our work contributes to the modeling and analysis of HOT lane systems in static settings. In particular, our model incorporates agents with heterogeneous values of time and heterogeneous carpool disutilities, where both parameters are continuously distributed. This dual-dimensional heterogeneity is essential for optimal HOT lane design since choices between ordinary and HOT lanes, and whether to pay or carpool, depend on both parameters. We prove equilibrium existence, uniqueness, and fully characterize the equilibrium structure with general latency functions and continuous preference distributions. Our results identify two distinct equilibrium regimes and lead to comparative statics that provide insights into how toll prices and HOT lane capacity affect lane usage and carpooling ratios. Furthermore, we extend our basic model to multi-segment settings with multiple occupancy levels and differentiated pricing. We prove equilibrium existence and uniqueness. These results support practical HOT lane design with multiple segments and tiered toll prices.

Moreover, our case study of the California I-880 highway contributes to the literature of numerical/empirical analysis on HOT design and carpooling. Ekström et al. [30] investigated the impact of toll design on social welfare through a case study of Stockholm. Fan [31, 32] demonstrated the improvement of revenue and social welfare in the Sioux Falls network by adopting optimal tolling mechanisms. Michalaka et al., and He et al. [47, 78] used simulation to compute optimal toll prices across multiple objectives that incorporate time-of-day pricing, drivers' lane choice behaviors in the presence of tolls, and different toll structures across various road segments of HOT lane facilities. Cohen et al. [21, 22] designed field experiments and found a positive impact of HOV lanes on both carpooling intent and adoption. Additionally, Finkleman et al. [35] surveyed 250+ drivers, and find relationships between willingness to pay and the improvement in travel speeds in HOT lanes, the length of the trip, and the urgency of on-time arrival.

Our numerical study contributes to the above literature by bridging the game theory analysis of HOT lane system with data collected on the HOT project along the California I-880 highway. In particular, we inversely estimated the distribution of travelers' preferences using the traffic sensor data and data requested from Caltrans on the aggregate lane choice and carpool ratios. We designed optimal tools for four proposed objectives, building on the equilibrium analysis and preference distribution estimates. Our results provide the optimal hourly toll pricing with the objectives of improving traffic congestion mitigation and toll revenue under different HOT capacities, and demonstrate the trade-off between different objectives.

3.2 The basic model

Consider a highway segment consisting of *ordinary* and *high occupancy toll* (HOT) lanes. An ordinary lane is toll-free and open to all vehicles. A high occupancy toll lane is accessible to vehicles that either pay the toll price $\tau \in \mathbb{R}_{\geq 0}$ or meet the minimum occupancy requirement with passenger size that is an integer $A \geq 2$. A central planner (e.g. transportation authority) determines the toll price τ, the minimum occupancy requirement A, and the allocation of road capacity between HOT lanes and ordinary lanes. In particular, we denote the fraction of capacity allocated to HOT lanes as $\rho \in [0, 1]$, and the remaining $(1 - \rho)$-fraction of capacity is allocated to the ordinary lanes. The capacity allocation affects the travel time cost (i.e. latency function) of the two types of lanes. We denote the latency function of the HOT lanes as $\ell_{\mathrm{h}}(x_{\mathrm{h}}, \rho)$, and the latency function of the ordinary lanes as $\ell_{\mathrm{o}}(x_{\mathrm{o}}, 1 - \rho)$, where x_{h} (resp. x_{o}) is the flow of vehicles using the HOT lanes (resp. the ordinary lanes). We assume that the latency functions satisfy the following assumption:

Assumption 3.

(a) The latency function $\ell_{\mathrm{o}}(x_{\mathrm{o}}, 1 - \rho)$ (resp. $\ell_{\mathrm{h}}(x_{\mathrm{h}}, \rho)$) is increasing in the flow x_{o} (resp. x_{h}), and increasing (resp. decreasing) in the capacity ratio ρ.

(b) $\ell_{\mathrm{o}}(0, 1 - \rho) = \ell_{\mathrm{h}}(0, \rho)$ for any $\rho \in [0, 1]$.

Assumption 3(a) indicates that both lanes are congestible in that the latency increases as the flow increases.[1] Additionally, the latency decreases in one type of lanes as the allocated capacity of that lane increases. Assumption 3(b) implies

[1]This assumption is supported by queueing based models [69, 70] and empirical validations [46].

that the free flow travel time, defined as the latency when the flow is zero, is the same for both types of lanes. This is a reasonable assumption since the free flow travel time is determined by the length of the highway and the speed limit.

We model travelers as non-atomic agents with a total demand of $D > 0$. The action set of each agent is $A = \{\text{toll}, \text{pool}, \text{o}\}$, where toll (resp. pool) is the action of taking the HOT lanes by paying the toll price (resp. by meeting occupancy requirement), and o is to take the ordinary lane. Agents have heterogeneous preferences about the travel time cost (relative to the monetary payment) as well as the disutility of forming carpool groups. We model the heterogeneous preferences of agents by the parameter of value of time (i.e. the amount of monetary cost that is equivalent to one unit time cost), denoted as $\beta \in \mathrm{B} = [0, \bar{\beta}]$, and the carpool disutility, denoted as $\gamma \in \Gamma = [0, \bar{\gamma}]$. The distribution of agents' preference parameters (β, γ) is represented by the probability density function $f : \mathrm{B} \times \Gamma \to \mathbb{R}$ such that $f(\beta, \gamma) > 0$ for all (β, γ) and $\int_{\mathrm{B}\times\Gamma} f(\beta, \gamma) d\beta d\gamma = 1$.

We define the strategy of an agent as a mapping from their preference parameters (β, γ) to a pure strategy in action set A, denoted as $s : \mathrm{B} \times \Gamma \to A$. The set of agents who choose each action a, denoted by R_a, is given by:

$$R_a = \{\mathrm{B} \times \Gamma | s(\beta, \gamma) = a\}, \quad \forall a \in A.$$

We represent the strategy distribution of the agent population as $\sigma = (\sigma_a)_{a\in A}$, where

$$\sigma_a = \frac{1}{D} \int_{R_a} f(\beta, \gamma) d\beta d\gamma \tag{3.1}$$

is the fraction of agents who choose each action $a \in A$, and $\sigma_{\text{toll}} + \sigma_{\text{pool}} + \sigma_{\text{o}} = 1$. Here, both R_a and σ_a for each a depend on s. We drop the dependence from the notation for simplicity. The flow on each type of lanes induced by σ is as follows:

$$x_{\text{h}} = \left(\sigma_{\text{toll}} + \frac{\sigma_{\text{pool}}}{A}\right) D, \quad x_{\text{o}} = \sigma_{\text{o}} D. \tag{3.2}$$

The cost of each agent with preference parameters (β, γ) for choosing actions toll, pool, o is given by:

$$C_{\text{toll}}(\sigma, \beta, \gamma) = \beta \cdot \ell_{\text{h}}(x_{\text{h}}, \rho) + \tau, \tag{3.3a}$$

$$C_{\text{pool}}(\sigma, \beta, \gamma) = \beta \cdot \ell_{\text{h}}(x_{\text{h}}, \rho) + \gamma, \tag{3.3b}$$

$$C_{\text{o}}(\sigma, \beta, \gamma) = \beta \cdot \ell_{\text{o}}(x_{\text{o}}, 1-\rho), \tag{3.3c}$$

where $\beta \cdot \ell_{\text{h}}(x_{\text{h}}, \rho)$ (resp. $\beta \cdot \ell_{\text{o}}(x_{\text{o}}, 1-\rho)$) represents the cost of enduring the latency on the HOT lanes (resp. the ordinary lanes), and the cost of toll payment or the carpool disutility is added for action toll and pool, respectively.

A strategy profile s^* is a Wardrop equilibrium if no agent has incentive to deviate:

Definition 8. *A strategy profile $s^* : \mathrm{B} \times \Gamma \to A$ is a Wardrop equilibrium if*

$$\begin{aligned} & s^*(\beta, \gamma) = a, \\ \Rightarrow \quad & C_a(\sigma^*, \beta, \gamma) = \operatorname*{argmin}_{a' \in A} C_{a'}(\sigma^*, \beta, \gamma), \\ & \forall (\beta, \gamma) \in \mathrm{B} \times \Gamma, \end{aligned}$$

and σ^* is the associated equilibrium strategy distribution given by (3.1).

That is, the action chosen by an agent with parameters (β, γ) in equilibrium minimizes their own cost compared to choosing the other two actions. Given equilibrium strategy distribution σ^*, we denote the induced equilibrium flow on the HOT lanes and the ordinary lanes by x_{h}^* and x_{o}^*, respectively.

3.3 Equilibrium characterization and comparative statics

3.3.1 Equilibrium characterization

In this section, we characterize the Wardrop equilibrium of the game. For ease of exposition, we define $\ell_\delta(\sigma, \rho)$ as the difference of the latency between the ordinary lanes and the HOT lanes given the strategy distribution σ and the capacity allocation ρ:

$$\ell_\delta(\sigma, \rho) := \ell_{\mathrm{o}}(x_{\mathrm{o}}, 1-\rho) - \ell_{\mathrm{h}}(x_{\mathrm{h}}, \rho), \tag{3.4}$$

where x_{o} and x_{h} are derived from σ as in (3.2).

We first show that the latency difference between the ordinary lanes and the HOT lanes is always non-negative. Furthermore, when the toll price is strictly positive, there will always be some agents taking the ordinary lane or carpooling. The proof is deferred to Appendix B.1.

Lemma 6. *If $\tau > 0$, then $\ell_\delta(\sigma^*, \rho) > 0$, $\sigma^*_{\mathrm{pool}} > 0$, and $\sigma^*_{\mathrm{o}} > 0$.*

We next characterize the best response strategies of agents for a given strategy distribution σ. In particular, given σ, the best response of an agent with parameter (β, γ) is the action that minimizes the associated cost as in (3.3). We denote the best response as $\mathrm{BR}(\sigma, \beta, \gamma) \in A$. Then, we can separate the parameter set $\mathrm{B} \times \Gamma$ into three regions $(\Lambda_a(\sigma))_{a \in A}$, where $\Lambda_a(\sigma) := \{\mathrm{B} \times \Gamma | \mathrm{BR}(\sigma, \beta, \gamma) = a\}$. The following lemma characterizes the three regions with respect to the latency difference $\ell_\delta(\sigma, \rho)$ induced by σ and the toll price τ. [2]

[2] We do not consider the boundary cases where the inequalities in Lemma 7 hold with equality. Agents with preference parameters on the boundaries of each region are indifferent between two or even all three actions. Their tie-breaking rule does not affect the equilibrium analysis, as agents are nonatomic, and the demand from agents with boundary preference parameters is effectively zero.

Lemma 7. *Given σ,*

$$\Lambda_{\text{toll}}(\sigma) = \{B \times \Gamma | \beta \ell_\delta(\sigma, \rho) \geq \tau, \ \gamma \geq \tau\},$$
$$\Lambda_{\text{pool}}(\sigma) = \{B \times \Gamma | \beta \ell_\delta(\sigma, \rho) \geq \gamma, \ \gamma \leq \tau\},$$
$$\Lambda_{\text{o}}(\sigma) = \{B \times \Gamma | \beta \ell_\delta(\sigma, \rho) \leq \min\{\tau, \gamma\}\}.$$

A detailed proof can be found in the appendix B.1. We illustrate the three regions in Figure 3.1. We note that $\Lambda_{\text{toll}}(\sigma)$ includes agents with both high value of time β and high carpool disutility γ. Such agent prefers to take the HOT lanes via paying rather than taking the ordinary lanes due to their high value for time (i.e. the cost saving given $\ell_\delta(\sigma, \rho)$ is no less than the toll price τ), and also prefers to pay the toll price over carpooling due to their high carpool disutility (i.e. γ is no less than the toll price). Similarly, agents in $\Lambda_{\text{pool}}(\sigma)$ have carpool disutility at most τ, and thus prefer to carpool than paying the toll price. Their value of time β is high relative to the carpool disutility γ so that the cost saving given $\ell_\delta(\sigma, \rho)$ is no less than the carpool disutility, i.e. they prefer to take the HOT lanes by carpool rather than taking the ordinary lane. Finally, $\Lambda_{\text{o}}(\sigma)$ includes agents whose value of time is low relative to both their carpool disutility and toll price, and hence they prefer to take the ordinary lanes compared to taking the HOT lanes via carpool or toll payment.

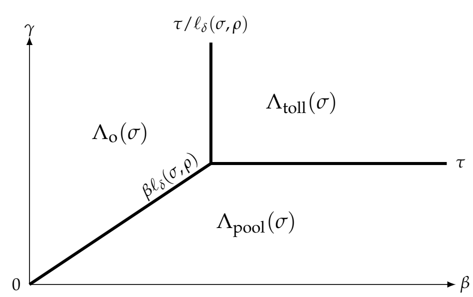


Figure 3.1: Characterization of best response strategies.

We are now ready to present equilibrium characterization of this game. We define the following latency difference threshold value $\ell_\delta^\dagger$ that will be used for separating different equilibrium regimes:

$$\ell_\delta^\dagger := \ell_\delta(\sigma^\dagger, \rho), \tag{3.5}$$

where $\sigma^\dagger$ is the threshold strategy distribution given by:

$$\sigma_{\text{toll}}^\dagger = 0, \tag{3.6a}$$

$$\sigma_{\text{pool}}^\dagger = \int_0^{\bar{\beta}} \int_0^{\min\{\tau, \overline{\gamma}\}\beta/\bar{\beta}} f(\beta, \gamma) d\gamma d\beta, \tag{3.6b}$$

$$\sigma_{\text{o}}^\dagger = 1 - \sigma_{\text{pool}}^\dagger. \tag{3.6c}$$

The following theorem shows that the game has a unique equilibrium that falls into either one of the two regimes depending on the game parameters. In each regime, the equilibrium strategy profile can be computed by solving a fixed point equation.

Theorem 2. *The game has a unique Wardrop equilibrium.*

Regime A: The toll price τ is relatively high, i.e. $\tau \geq \min\left\{\bar{\gamma}, \bar{\beta}\ell_\delta^\dagger\right\}$. No agent takes HOT lanes by paying the toll, i.e. $\sigma_{\text{toll}}^* = 0$. Furthermore,

(A-1) If $\bar{\beta}\ell_\delta^\dagger \leq \overline{\gamma}$, then σ_{pool}^* is the unique solution that satisfies the following equation:

$$\sigma_{\text{pool}}^* = \int_0^{\bar{\beta}} \int_0^{\ell_\delta(0, \sigma_{\text{pool}}^*, 1-\sigma_{\text{pool}}^*, \rho)\beta} f(\beta, \gamma) d\gamma d\beta, \tag{3.7}$$

and $\sigma_{\text{o}}^* = 1 - \sigma_{\text{pool}}^*$.

(A-2) If $\bar{\beta}\ell_\delta^\dagger > \overline{\gamma}$, then σ_{pool}^* is the unique solution of the following equation:

$$1 - \sigma_{\text{pool}}^* = \int_0^{\overline{\gamma}} \int_0^{\gamma/\ell_\delta(0, \sigma_{\text{pool}}^*, 1-\sigma_{\text{pool}}^*, \rho)} f(\beta, \gamma) d\beta d\gamma, \tag{3.8}$$

and $\sigma_{\text{o}}^* = 1 - \sigma_{\text{pool}}^*$.

Regime B: The toll price τ is relatively low, i.e. $0 < \tau < \min\{\bar{\gamma}, \bar{\beta}\ell_\delta^\dagger\}$. All three actions are taken by agents in equilibrium, and σ^* is the unique solution that satisfies the following equations:

$$\sigma^*_{\text{toll}} = \int_\tau^{\overline{\gamma}} \int_{\tau/\ell_\delta(\sigma^*,\rho)}^{\overline{\beta}} f(\beta,\gamma)d\beta d\gamma, \tag{3.9a}$$

$$\sigma^*_{\text{pool}} = \int_0^{\tau} \int_{\gamma/\ell_\delta(\sigma^*,\rho)}^{\overline{\beta}} f(\beta,\gamma)d\beta d\gamma, \tag{3.9b}$$

$$\sigma^*_{\text{o}} = 1 - (\sigma^*_{\text{toll}} + \sigma^*_{\text{pool}}). \tag{3.9c}$$

We provide the proof intuition of Theorem 2 in this section. The complete proof is in Appendix B.1. Our equilibrium characterization builds on the two lemmas 6 – 7 introduced before. In particular, Lemma 6 shows that in equilibrium both lanes are used, and either (A) all agents who take the HOT lanes choose to carpool, or (B) a positive fraction of agents who take the HOT lanes pay the toll τ. Indeed, (A) and (B) are each associated with equilibrium regimes A and B, respectively.

Furthermore, following Definition 8, an equilibrium strategy distribution σ^* must satisfy

$$\sigma^*_a = \iint_{\Lambda_a(\sigma^*)} f(\beta,\gamma)d\beta d\gamma, \quad a \in A, \tag{3.10}$$

where $\Lambda_a(\sigma^*)$ is the best response region characterized in Lemma 7. In particular, using the best response characterization in Lemma 7, the equilibrium distribution of carpool σ^*_{pool} can be written as:

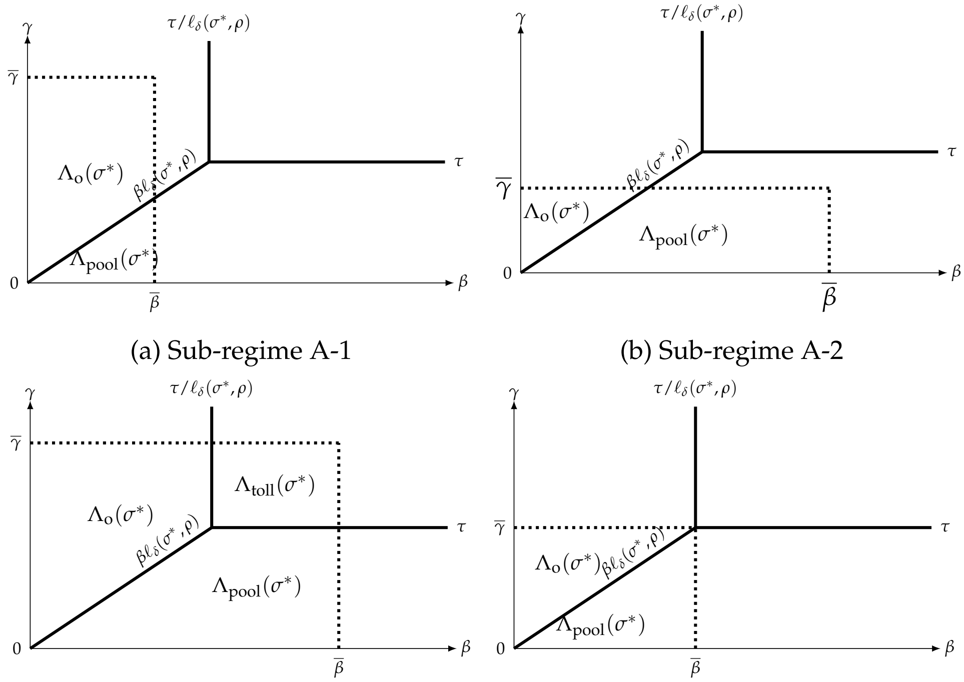
$$\sigma^*_{\text{pool}} = \int_0^{\overline{\beta}} \int_0^{\min\{\ell_\delta(\sigma^*,\rho)\beta,\tau,\overline{\gamma}\}} f(\beta,\gamma)d\gamma d\beta.$$

The two sub-regimes, (A-1) and (A-2), and regime B each corresponds to the scenario where one of the three elements, $\bar{\beta}\ell_\delta(\sigma^*,\rho)$, τ, or $\overline{\gamma}$, is the smallest. In particular, sub-regime A-1 (resp. A-2) corresponds to the case where $\bar{\beta}\ell_\delta(\sigma^*,\rho)$ (resp. $\overline{\gamma}$) is the smallest among three elements. Thus, no agents use the HOT lane by paying the toll price since the toll price τ is either larger than the value of the time saved by taking the HOT lane $\bar{\beta}\ell_\delta(\sigma^*,\rho)$ or larger than the maximum carpool disutility $\bar{\gamma}$. Figures 3.2a – 3.2b illustrate that, in the equilibrium of sub-regimes A-1 and A-2, agents only choose to take the ordinary lane or carpool to take the HOT lane. This is because the preference parameter set does not intersect with the set corresponding to choosing to pay the toll to take the HOT lane as the best response strategy. In regime B, τ is the smallest element of the three (i.e. $0 < \tau < \min\{\bar{\gamma}, \bar{\beta}\ell_\delta(\sigma^*,\rho)\}$ as shown in Figure 3.2c), and the strategy distributions for all three actions are positive in equilibrium.

Finally, the threshold strategy distribution $\sigma^\dagger$ as in (3.6) and the threshold latency cost difference $\ell_\delta^\dagger$ as in (3.5) are derived from the boundary case $\bar{\beta}\ell_\delta(\sigma^*,\rho) = \bar{\gamma} = \tau$, see Fig. 3.2d for the illustration. In this threshold case, $\sigma^* = \sigma^\dagger$ and the latency difference $\ell_\delta(\sigma^*,\rho) = \ell_\delta^\dagger$. We can verify that this boundary case indeed marks the transition between different regimes. For example, in sub-regime A-1,

$$\begin{aligned}\sigma^*_{\text{pool}} &= \int_0^{\overline{\beta}} \int_0^{\min\{\ell_\delta(\sigma^*,\rho)\beta,\tau,\overline{\gamma}\}} f(\beta,\gamma)d\gamma d\beta \\ &\overset{(a)}{=} \int_0^{\overline{\beta}} \int_0^{\ell_\delta(\sigma^*,\rho)\beta} f(\beta,\gamma)d\gamma d\beta \\ &\overset{(b)}{\leq} \int_0^{\overline{\beta}} \int_0^{\min\{\tau,\overline{\gamma}\}\beta/\bar{\beta}} f(\beta,\gamma)d\gamma d\beta = \sigma^\dagger_{\text{pool}},\end{aligned}$$

and $\ell_\delta(\sigma^*,\rho) \leq \ell_\delta^\dagger$, where both (a) and (b) are due to the sub-regime A-1 conditions. Therefore, the sub-regime A-1 boundary characterization $\overline{\gamma} \geq \bar{\beta}\ell_\delta^\dagger$ guarantees that the sub-regime equilibrium condition $\overline{\gamma} \geq \bar{\beta}\ell_\delta(\sigma^*,\rho)$ holds when problem instance parameters are in A-1. Similarly, we can show that $\sigma^*_{\text{pool}} \geq \sigma^\dagger_{\text{pool}}$ and $\ell_\delta(\sigma^*,\rho) \geq \ell_\delta^\dagger$ in sub-regime A-2 and regime B. As a result, the sub-regime A-2 (resp. regime B) characterization $\bar{\gamma} \leq \bar{\beta}\ell_\delta^\dagger$ (resp. $0 < \tau < \min\{\bar{\gamma}, \bar{\beta}\ell_\delta^\dagger\}$) guarantees that the condition $\bar{\gamma} \leq \bar{\beta}\ell_\delta^*(\sigma^*,\rho)$ (resp. $0 < \tau < \min\{\bar{\gamma}, \bar{\beta}\ell_\delta(\sigma^*,\rho)\}$) is satisfied when problem instance parameters are in A-2 (resp. regime B). The detailed description of these conditions of each regime is in the theorem proof in Appendix B.1.



(a) Sub-regime A-1

(b) Sub-regime A-2

(c) Regime B

(d) Threshold case $\bar{\beta}\ell_\delta(\sigma^*,\rho)=\bar{\gamma}=\tau$, $\sigma^*=\sigma^\dagger$

Figure 3.2: Equilibrium outcomes in each regime.

3.3.2 Comparative statics

We analyze the change of equilibrium strategy distribution with HOT capacity ρ and toll price τ.

Theorem 3. *The comparative statics of equilibrium strategy distribution σ^* with respect to (ρ, τ) are summarized in Table 3.1.*

Intuitively, if we fix the toll price τ and increase the HOT capacity ρ, some agents will switch from ordinary lanes to HOT lanes by either paying the toll price or carpooling. Therefore, σ^*_{toll} increases, σ^*_{pool} increases, and σ^*_{o} decreases.

	Fix τ increase ρ	Fix ρ increase τ
σ_{o}^*	Decreasing	Either Direction
σ_{toll}^*	Increasing	Non-Increasing
σ_{pool}^*	Increasing	Non-Decreasing
$\ell_\delta(\sigma^*, \rho)$	Increasing	Non-Decreasing

Table 3.1: Comparative statics.

From analyzing the regime boundary change, we can also conclude that the latency difference between the ordinary lanes and the HOT lanes increases. See Appendix B.1 for the complete proof.

On the other hand, if we fix the HOT capacity ρ and increase the toll price τ, some agents will deviate from paying the toll to carpooling or taking the ordinary lane. Hence, the latency in the ordinary lanes will increase and in the HOT lanes will decrease, which means the latency difference between the two lanes $\ell_\delta(\sigma^*_, \rho)$ will increase. Additionally, σ_{toll}^* decreases and σ_{pool}^* increases. As for the ordinary lanes, agents with high value of time and low carpool disutilities will switch to carpools as HOT lane becomes less congested. Depending on the preference distribution and the toll prices before and after the change, we can either have more agents switch from ordinary lanes to carpool or more agents switch from toll paying to ordinary lanes. Hence, σ_{o}^* can either increase or decrease. We remark that when toll price is already higher than the threshold where no agents pay the toll (regime A in Theorem 2), further increasing the toll price has no impact on the strategy distributions and therefore does not affect the latency difference.

3.4 Extensions to multiple segments and occupancy levels

Consider a highway partitioned into multiple segments by separation nodes, which represent the locations where vehicles get on and off the highway. We number the segments sequentially from upstream to downstream as $e \in [E] := 1, \ldots, E$, where E is the total number of road segments. Same as the basic model, the central planner divides the capacity of the highway into the HOT lane and the ordinary lane, and $\rho \in [0, 1]$ is the fraction of capacity allocated to the HOT lane. Agents can form carpools with different occupancy levels $m \in [M] := 1, \ldots, M$, where $m = 1$ indicates that the agent does not carpool with others, and M is the maximum carpool size. For each segment $e \in [E]$, a toll price $\tau_{e,m} \geq 0$ is charged on each vehicle using the HOT lane with occupancy level m on segment e. Agents split the toll price evenly, i.e. a agent pays $\tau_{e,m}/m$ when carpooling with $m - 1$ other agents on the HOT lane of segment e. The latency function on each road segment e is given by $\ell_{\mathrm{o},e}(x_{\mathrm{o},e}, 1-\rho)$ for ordinary lanes and $\ell_{\mathrm{h},e}(x_{\mathrm{h},e}, \rho)$ for HOT lanes, where $x_{\mathrm{o},e}$ and $x_{\mathrm{h},e}$ are the vehicle flows on ordinary lanes and HOT lanes of road segment e, respectively. Same as the basic model, we assume that the latency functions on all road segments satisfy Assumption 3.

For each pair of $i \leq j \in [E]$, a population of agents with demand D^{ij} enters the highway from the entrance node (the beginning) of segment i and leaves the highway from the exit node (the ending) of segment j. We refer the population who traverses segments $[i : j] := i, \ldots, j$ as the population (i, j). Agents in each population (i, j) decides their carpool size $m \in [M]$ and whether to take the HOT lane or the ordinary lane in each segment $e \in [i : j]$. We denote an action of population (i, j) as $a = (a_{\mathrm{occu}}, (a_e)_{e \in [i,j]})$, where $a_{\mathrm{occu}} \in [M]$ is the occupancy level, and $a_e \in \{\mathrm{o}, \mathrm{h}\}$ is to take the ordinary lane or the HOT lane for each segment $e \in [i, j]$. Thus, the action set of the population (i, j) is $A^{ij} = [M] \times \{\mathrm{o}, \mathrm{h}\}^{[i:j]}$. We note that agents must select a single carpool size for traversing all the segments but they can switch between the ordinary lanes and the HOT lanes at the separation nodes in between segments based on the HOT toll price and latency of the next segment.

Analogous to the basic model, we represent the heterogeneous preference of agents of using value of time $\beta \in [0, \overline{\beta}]$ and carpool disutilities $\gamma := (\gamma_m)_{m \in [M]}$, where $\gamma_m \in [0, \overline{\gamma}_m]$ denotes the disutilities for choosing occupancy level m. We set the disutility of single occupancy γ_1 to 0. The distribution of agents' preference parameters (β, γ) for population (i, j) is represented by the probability density function $f^{ij} : \mathrm{B} \times \Gamma \rightarrow \mathbb{R}$ such that $f^{ij}(\beta, \gamma) > 0$ for all (β, γ) and $\int_{\mathrm{B} \times \Gamma} f^{ij}(\beta, \gamma) d\beta d\gamma = 1$.

We define the strategy of an agent in population (i, j) as a mapping from their preference parameters (β, γ) to a pure strategy in action set A^{ij}, denoted as $s^{ij} : \mathrm{B} \times \Gamma \rightarrow A^{ij}$. The set of agents of the population (i, j) who choose each action $a \in A^{ij}$, denoted by R^{ij}_a, is given by:

$$R^{ij}_a = \left\{ \mathrm{B} \times \Gamma | s^{ij}(\beta, \gamma) = a \right\}, \quad \forall a \in A^{ij}. \tag{3.11}$$

We represent the strategy distribution of the population (i, j) as $\sigma^{ij} = (\sigma^{ij}_a)_{a \in A^{ij}}$, where

$$\sigma^{ij}_a = \frac{1}{D^{ij}} \int_{R^{ij}_a} f^{ij}(\beta, \gamma) d\beta d\gamma \tag{3.12}$$

is the fraction of agents who choose each action $a \in A$, and $\sum_{a \in A^{ij}} \sigma^{ij}_a = 1$. On each road segment e, the vehicle flow on ordinary lanes $x_{\mathrm{o},e}$ and on HOT lanes $x_{\mathrm{h},e}$ are induced by all agents of population (i, j) that enters the highway on or before e and exits on or after e, i.e. $i \leq e \leq j$. In particular,

$$x_{\text{o},e} = \sum_{i=1}^{e} \sum_{j=e}^{E} D^{ij} \left(\sum_{a \in A^{ij}} \frac{1}{a_{\text{occu}}} \sigma_a^{ij} \mathbb{1}_{a_e = \text{o}} \right), \quad \forall e \in [E], \tag{3.13a}$$

$$x_{\text{h},e} = \sum_{i=1}^{e} \sum_{j=e}^{E} D^{ij} \left(\sum_{a \in A^{ij}} \frac{1}{a_{\text{occu}}} \sigma_a^{ij} \mathbb{1}_{a_e = \text{h}} \right), \quad \forall e \in [E]. \tag{3.13b}$$

The cost of each agent of population (i, j) with preference parameters (β, γ) for choosing action $a \in A^{ij}$ is given by

$$\begin{aligned} C_{\text{a}}(\sigma, \beta, \gamma) &= \gamma_{a_{\text{occu}}} \\ &+ \sum_{e=i}^{j} \beta \cdot \ell_{\text{o},e}(x_{\text{o},e}, 1-\rho) \mathbb{1}_{a_e = \text{o}} \\ &+ \sum_{e=i}^{j} \left(\beta \cdot \ell_{\text{h},e}(x_{\text{h},e}, \rho) + \frac{1}{a_{\text{occu}}} \tau_{e, a_{\text{occu}}} \right) \mathbb{1}_{a_e = \text{h}}, \end{aligned} \tag{3.14}$$

where $\beta \cdot \ell_{\text{o},e}(x_{\text{o},e}, 1-\rho)$ (resp. $\beta \cdot \ell_{\text{h},e}(x_{\text{h},e}, \rho)$) represents the cost of enduring the latency on ordinary lanes (resp. HOT lanes) on segment e. Additionally, the carpool disutility $\gamma_{a_{\text{occu}}}$ is added according to the associated occupancy level of the action, and toll payment $\frac{1}{a_{\text{occu}}} \tau_{e, a_{\text{occu}}}$ is added based on whether the action takes the HOT lane on each segment. Analogous to the basic model, we define the Wardrop equilibrium as:

Definition 9. *A strategy profile $s^* : \mathrm{B} \times \Gamma \to A$ is a Wardrop equilibrium if*

$$\begin{aligned} & s^{*ij}(\beta, \gamma) = a, \\ & \Rightarrow \quad C_a(\sigma^{*ij}, \beta, \gamma) = \operatorname*{argmin}_{a' \in A^{ij}} C_{a'}(\sigma^{*ij}, \beta, \gamma), \\ & \forall (\beta, \gamma) \in \mathrm{B} \times \Gamma, \ \forall i \leq j \in [E], \end{aligned}$$

and σ^{*ij} is equilibrium strategy distribution of population (i, j) induced by s^* given by (3.11) – (3.12).

We next provide conditions under which equilibrium is unique in the multi-segment setting. Before presenting this result, we note that an agent's equilibrium strategy $s^{*ij}(\beta, \gamma)$ depends on the segment latency of each lane, but only through the difference between them, not their individual values. We define $\delta = (\delta_e)_{e \in E}$ as the vector of latency difference, where δ_e is the latency of the ordinary lane exceeding that of the HOT lane in segment e. Given δ and the toll price vector τ, the best response of agent in population (i, j) with preference parameter (β, γ) is uniquely determined. In particular, when an agent with preference parameter (β, γ) chooses occupancy level m, they will select the ordinary lane (resp. HOT lane) of segment e if $\tau_{e,m}/m > \beta\delta_e$ (resp. $\tau_{e,m}/m < \beta\delta_e$). The agent's best response occupancy level is $\text{argmin}_{m \in [M]} \{\sum_{e \in [i,j]} \min\{\tau_{e,m}/m,\ \beta\delta_e\} + \gamma_m\}$. With slight abuse of notation, we denote $x(\delta)$ as the lane flow vector induced by all agents taking their best response to the cost difference vector δ according to (3.12)-(3.13). We define $\Phi_e(\delta) = \ell_{o,e}(x_{o,e}(\delta)) - \ell_{h,e}(x_{h,e}(\delta))$ as the latency cost difference between the two lanes of segment e induced by all agents' best response given δ, and the vector function is $\Phi(\delta) = (\Phi_e(\delta))_{e \in E}$.

We note that the set of fixed point solution of $\Phi(\delta) = \delta$ is the vector of latency cost difference in equilibrium δ^*. This is because when the latency cost difference induced by agents' best response $\Phi(\delta^*)$ is consistent with the actual latency cost difference δ^*, no agent has incentive to deviate. Moreover, since agents have unique best response for any δ, there is a one-to-one correspondence between an equilibrium strategy profile s^* and an equilibrium latency cost difference vector δ^*. The following theorem shows that in the multi-segment model, equilibrium exists, and is unique when the function $\Phi(\cdot)$ satisfies certain condition.

Theorem 4. *Given any toll price vector τ, Wardrop equilibrium s^* exists. Moreover, s^* is unique if the Jacobian matrix $\nabla\Phi(\delta)$ does not have 1 as its eigenvalue for any $\delta \in \prod_{e \in E}[\underline{\delta}_e, \bar{\delta}_e]$, where $\underline{\delta}_e = \ell_{o,e}(0) - \ell_{\mathrm{h},e}\left(\sum_{i=1}^{e}\sum_{j=e}^{E} D^{ij}\right)$ and $\bar{\delta}_e = \ell_{\mathrm{o},e}\left(\sum_{i=1}^{e}\sum_{j=e}^{E} D^{ij}\right) - \ell_{\mathrm{h},e}(0)$.*

In Theorem 4, the equilibrium existence result is established using the Kakutani's fixed point theorem, relying on the boundedness of the latency cost difference vector and the continuity of the function $\Phi(\cdot)$. The equilibrium uniqueness result follows from the one-to-one correspondence between equilibrium and fixed point solution of $\Phi(\delta) = \delta$, and the mean value theorem. We note that the sufficient condition holds generically, meaning that equilibrium is generically unique. Additionally, in the single-segment setting, we can verify that $\Phi(\delta)$ is monotone in δ and that $\nabla\Phi(\delta)$ does not have an eigenvalue of 1. As demonstrated in Theorem 2, equilibrium is indeed unique in the single segment setting. The complete proof of Theorem 4 can be found in Appendix B.1.

3.5 Design HOT on California I-880

In this section, we redesign the high-occupancy toll (HOT) lane on California's I-880. Using data collected from the HOT operations on I-880 in 2021, we calibrate a multi-segment model. The primary challenge in this calibration is estimating the distribution of agents' preferences from aggregate lane choice data. To address this, we apply inverse optimization to estimate the preference distribution such that the resulting equilibrium flow closely matches the observed data. With the calibrated model, we compute the Wardrop equilibrium and determine the optimal toll price to achieve various policy goals. We then compare the computed toll price with the actual prices from the 2021 operations.

3.5.1 Data description

The Metropolitan Transportation Commission in California started the conversion of the existing HOV lanes to HOT lanes on the I-880 highway in 2019. The HOT lanes run from Hegenberger Road to Dixon Landing Road in the southbound direction and from Dixon Landing Road to Lewelling Boulevard in the northbound direction. This stretch of highway is partitioned into multiple segments. The toll price is charged for using the HOT lane on each segment from 5 am to 8 pm on each workday, and the toll is updated every 5 min. Vehicles with carpool size of 3 can use the HOT lane for free, vehicles with carpool size of 2 pay half of the toll price, and vehicles with a single person pay the full price.

We calibrate our multi-segment model using data collected from the Northbound of I-880 between the Dixon Landing Rd and Lewelling Blvd. The total distance is 22 miles and the highway has three ordinary lanes and one HOT lane. The highway is partitioned into five segments with separation nodes named as the Auto Mall Pkwy, Mowry Ave, Decoto Rd, Whipple Rd, and Hesperian Blvd, see Fig. 3.3. The distance of these segments are 5.75 miles, 3.17 miles, 3.46 miles, 2.11 miles, and 7.16 miles, respectively.

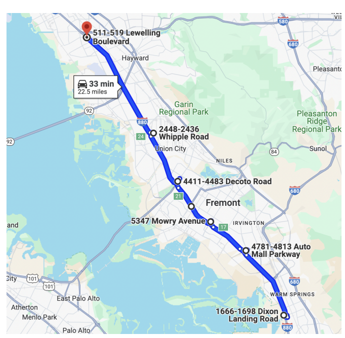


Figure 3.3: Interstate 880 (I-880) Highway from Dixon Landing Rd to Lewelling Blvd (Highlighted in Blue). The nodes separating the segments are marked in bold text.

The California Department of Transportation have installed hundreds of sensors along the I-880 highway. Each sensor measures the vehicle flow data and average speed data at the 5-minute level. We obtain the data from [8] for all workdays between March 1st 2021 and August 31st 2021. We number these workdays sequentially as $n \in [N] := 1, \ldots, N$, where N is the total number of workdays. For each workday n, we aggregate those 5-minute vehicle flows into per-hour vehicle flows of each sensor. We take the average of vehicle speed across all 5-minute intervals of each hour to infer the per-hour average speed. We divide the distance between two adjacent sensors with the average speed to compute the average travel time between each pair of adjacent sensors during each hour of the day n. We number the hours from 5 am to 8 pm as $t \in [T] := 1, \ldots, T$, where $T = 15$ is the total number of HOT operation hours of each workday.

For each road segment e between Auto Mall Pkwy and Hesperian Blvd, we identify the list of all sensors covering this road segment. We sum up the average travel time across all these sensors to obtain the latency of ordinary lanes $\hat{\ell}^{t,n}_{\mathrm{o},e}$ and HOT lanes $\hat{\ell}^{t,n}_{\mathrm{h},e}$ of this road segment for each hour t of each day n. Additionally, we take an average of vehicle flows across all these sensors to obtain the observed vehicle flows for ordinary lanes $\hat{x}^{t,n}_{\mathrm{o},e}$ and HOT lanes $\hat{x}^{t,n}_{\mathrm{h},e}$ of this road segment for each hour t of each day n.

Figure 3.4 illustrates the travel time from Auto Mall Pkwy to Hesperian Blvd in different hours of a day. The dotted lines are the mean travel time on ordinary lanes and HOT lanes in each hour of the day averaged across all days, while the shaded regions are the corresponding 95% confidence intervals for each hour. Ordinary lanes have a uniformly higher travel time than HOT lanes in all hours. Particularly, in the afternoon hours (i.e. 2-6 pm), ordinary lanes can reach 33%

higher travel time compared to the ordinary lanes on average.

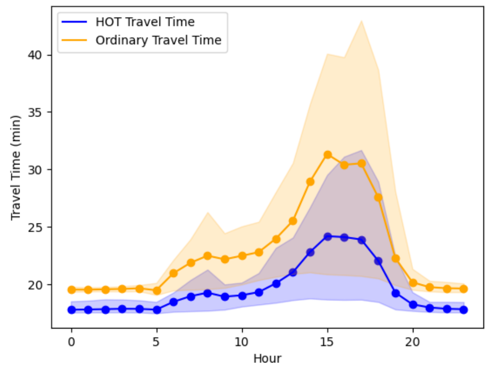


Figure 3.4: Average travel time (min) and 95% confidence interval of the HOT and ordinary lanes from Auto Mall Pkwy to Hesperian Blvd in each hour of a workday.

We requested from Caltrans the toll price data for each road segment for each 5 min time interval, and the daily number of vehicles at each occupancy level using the HOT lanes for the entire I-880 highway.[3] Using this data, we compute the hourly averaged toll price of each segment and the fraction of vehicles taking the HOT lanes with occupancy level m on each day n, denoted as r_m^n.

3.5.2 Model calibration

Latency functions. We estimate the latency function of the ordinary and HOT lanes of each road segment based on the Bureau of Public Roads (BPR) function [101]. Given the vehicle flow x and the HOT capacity $\rho = 0.25$ (one lane out of 4 is HOT in the current design), the BPR function can be written as follows:

$$\ell_o(x) = T_f \cdot \left[1 + \left(\eta \cdot \frac{x_o}{(1-\rho)V}\right)^b\right],$$
$$\ell_h(x) = T_f \cdot \left[1 + \left(\eta \cdot \frac{x_h}{\rho V}\right)^b\right],$$

[3]Due to privacy concern, the occupancy level data shared with us is aggregated across all segments and all HOT operation hours for each day.

where η and b are the BPR coefficients, T_f is the free flow travel time, and V is the road capacity. We set $b = 4.0$ following [7], and estimate parameters T_f and $\frac{\eta}{V}$ of each segment using the flow data $\left\{\hat{x}^{t,n}_{\text{o},e}, \hat{x}^{t,n}_{\text{h},e}\right\}_{e\in[E],t\in[T],n\in[N]}$ and driving time $\left\{\hat{\ell}^{t,n}_{\text{o},e}, \hat{\ell}^{t,n}_{\text{h},e}\right\}_{e\in[E],t\in[T],n\in[N]}$ via linear regression.

Calibration of preference distribution. To compute the equilibrium strategy distribution, we need to estimate the population demand for each entrance and exit pair, and estimate the agents' preference distribution. To ensure the tractability of the demand and preference distribution estimates, we create a grid for each preference parameter β and $\{\gamma\}_{m\in[M]}$ with evenly spaced intervals. Then, the entire preference parameter vector space is partitioned into equally sized subsets. We denote the set of all partitioned preference sets as K with generic member k, and assume that the preference distribution in each subset k is uniform. This can be viewed as an approximation of the original probability density function of the preference parameters: As the partition of preference parameter space becomes finer, the approximation becomes closer to the original density function.

For each pair of entrance and exit nodes (i, j) and each hour t, we estimate $d^{ij,t}_k$ as the mass of agents with preference parameters in each subset k. This estimate corresponds to the multiplication of the total agent demand for the population (i, j) at time t and the fraction of agents with preference parameters in subset k. Our estimate captures the variation of agent demand and preference distribution across different entrance and exit nodes and times of the day. We assume the agents' demand for each hour is the same across all workdays.

We estimate $d := \left(d^{ij,t}_k\right)_{i\leq j\in[E],t\in[T],k\in[K]}$ using inverse optimization. Given data on toll prices $\left\{\tau_{e,m}^{t,n}\right\}_{e\in[E]}$ and driving time $\left\{\hat{\ell}_{\mathrm{o},e}^{t,n}, \hat{\ell}_{\mathrm{h},e}^{t,n}\right\}_{e\in[E]}$ for each hour t of each day n, we compute the best response strategy profile $s^{*t,n}$ of all agents for hour t and day n. With the estimated demand $\left(d_k^{ij,t}\right)_{i\leq j\in[E],k\in[K]}$ for hour t, we obtain the best response strategy distribution $\left(\sigma_a^{*t,n}(d)\right)_{a\in A}$ following analysis in Sec. 3.4, and the induced vehicle flow $x^{*t,n}(d) := \left\{x_{\mathrm{o},e}^{*t,n}(d), x_{\mathrm{h},e}^{*t,n}(d)\right\}_{e\in[E]}$ based on (3.13). Additionally, we compute the induced best response vehicle flow of each occupancy level $\bar{x}^{*n}(d) := (\bar{x}_m^{*n}(d))_{m\in[M]}$ as follows: $\forall m \in [M]$,

$$\bar{x}_m^{*n}(d) = \sum_{t\in[T]} \sum_{i\leq j\in[E]} D^{ij,t} \cdot \left(\sum_{a\in A^{ij}} \frac{1}{a_{\text{occu}}} \sigma_a^{*ij,t,n}(d) \mathbb{1}_{a_e=\mathrm{h}, a_{\text{occu}}=m}\right). \tag{3.15}$$

We estimate d to be the demand vector such that the induced best response vehicle flow is close to the observed flows $\{\hat{x}_{\mathrm{o},e}^{t,n}, \hat{x}_{\mathrm{h},e}^{t,n}\}_{e\in[E],t\in[T],n\in[N]}$ and the induced fraction of vehicles taking each occupancy level is close to $(r_m^n)_{m\in[M],n\in[N]}$. We formulate the estimation problem as the following convex optimization program:

$$\min_d \sum_{n\in[N]} \sum_{n\in[N]} \sum_{m\in[M]} \left(r_m^n \left(\sum_{m'\in[M]} \bar{x}_{m'}^{*n}(d)\right) - \bar{x}_m^{*n}(d)\right)^2 + \sum_{t\in[T]} \sum_{e\in[E]} \left(\left(x_{\mathrm{o},e}^{*t,n}(d) - \hat{x}_{\mathrm{o},e}^{t,n}\right)^2 + \left(x_{\mathrm{h},e}^{*t,n}(d) - \hat{x}_{\mathrm{h},e}^{t,n}\right)^2\right).$$

We use the calibrated demand to compute the number of agents taking each occupancy level and compare it with the actual numbers. In Fig. 3.5, the dotted lines show the observed daily fraction of agents on HOT lanes for each occupancy level, while the curved lines show the equilibrium daily ratio of agents for each occupancy level based on calibrated demand and induced equilibrium. The close alignment of curved lines and dotted lines demonstrates that our calibrated demand closely matches the actual demand.

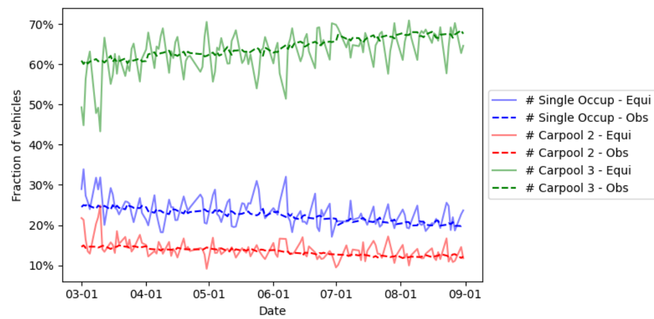


Figure 3.5: Fractions of agents on HOT lanes taking each occupancy level per day.

3.5.3 Optimal design of toll price and HOT capacity at 5-6pm

Based on the calibrated model, we compute the optimal toll price and HOT capacity allocation for 5-6 pm. We consider each of the following four objectives in equilibrium:

1. The total agent travel time.

$$C(\tau, \rho) := \sum_{i \le j \in [E]} D^{ij} \left[\sum_{a \in A^{ij}} \sigma_a^{*ij} \cdot \left(\sum_{e=i}^{j} \ell_{\text{o},e}(x^*_{\text{o},e}, 1-\rho) \mathbb{1}_{a_e = \text{o}} + \ell_{\text{h},e}(x^*_{\text{h},e}, \rho) \mathbb{1}_{a_e = \text{h}} \right) \right]. \tag{3.16}$$

2. The total vehicle driving time. The total vehicle driving time differs from the total agent travel time. Specifically, for every minute a vehicle with $m \in [M]$ agents spends on the highway, the total vehicle driving time counts this as 1 minute, whereas the agent travel time counts it as m minutes.

$$E(\tau, \rho) := \sum_{i \le j \in [E]} D^{ij} \left[\sum_{a \in A^{ij}} \sigma_a^{*ij} \frac{1}{a_{\text{occu}}} \cdot \left(\sum_{e=i}^{j} \ell_{\text{o},e}(x^*_{\text{o},e}, 1-\rho) \mathbb{1}_{a_e = \text{o}} + \ell_{\text{h},e}(x^*_{\text{h},e}, \rho) \mathbb{1}_{a_e = \text{h}} \right) \right]. \tag{3.17}$$

3. Total revenue: The total toll prices paid by all agents.

$$R(\tau, \rho) := \sum_{i \le j \in [E]} D^{ij} \left[\sum_{a \in A^{ij}} \sigma_a^{*ij} \frac{1}{a_{\text{occu}}} \left(\sum_{e=i}^{j} \tau_{e, a_{\text{occu}}} \mathbb{1}_{a_e = \text{h}} \right) \right]. \tag{3.18}$$

4. Total cost: The total cost (driving time, toll price and carpool disutility) incurred by all agents incorporating their heterogeneous preferences.

$$U(\tau,\rho) := \sum_{i \leq j \in [E]} D^{ij} \left[\sum_{a \in A^{ij}} \int_{\gamma} \int_{\beta} C_a(\sigma^*, \beta, \gamma) \right.$$
$$\left. \cdot \, \mathbb{1}_{s^{*ij}(\beta,\gamma)=a} f^{ij}(\beta,\gamma) d\beta d\gamma \right]. \tag{3.19}$$

The problems (3.16) – (3.19) compute the HOT toll price for each segment that optimizes the objective function based on the induced equilibrium strategy. As such, these are mathematical programs with equilibrium constraints (MPEC). Furthermore, since the toll is only applied to the HOT lane, these problems can be seen as a generalization of optimal tolling with support constraints, which has been proven to be NP-hard ([43, 49]). We adopt the enumeration algorithm for the optimal design of toll price and HOT capacity. In particular, we set the toll price on each road segment within the range of 0 and 7 dollars discretized by \$0.5. We also choose the set $\rho \in \{\frac{1}{4}, \frac{2}{4}, \frac{3}{4}\}$ since the highway has four lanes. For each pair of (τ, ρ), we compute the equilibrium strategies, and the corresponding objective function value. We choose the toll price and capacity fraction with the optimal value.

Table 3.2 summarizes our result. The first row shows the current HOT capacity and the average toll price. The remaining rows in Table 3.2 shows the optimal toll prices on each road segment for agent travel time minimization, vehicle driving time minimization, revenue maximization, and cost minimization, under different HOT capacities.

When the HOT capacity is 0.25 (the setting in practice), the optimal toll prices for agent travel time, vehicle driving time, and cost minimization are lower than the current prices on Auto Mall Pkwy, Mowry Ave, Decoto Rd, and Whipple

Rd. On Hesperian Blvd, they match the current average toll price. For revenue maximization, the optimal toll prices are lower than the current average prices on all five road segments. As HOT capacity increases, the optimal toll prices for these four objectives may vary, either increasing or decreasing, depending on each road segment.

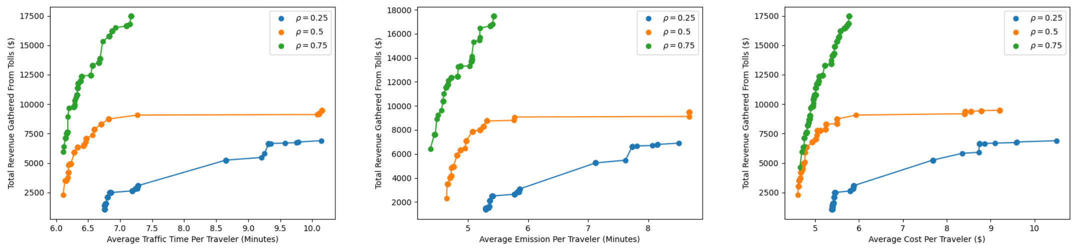
Note that the optimal toll prices for agent travel time, vehicle driving time, and cost minimization are higher on Hesperian Blvd, and lower on other segments. We remark that it aligns with the fact that the demand volume of agents is also higher on Hesperian Blvd and lower on other segments. Given the convex nature of the latency function, a large demand volume on a road segment means that the reduction in travel time achieved by incentivizing agents to carpool becomes more significant compared to segments with smaller demand volumes. Therefore, setting a high toll price on road segments with large demand is more effective in incentivizing carpool, which in turn reduces the agent travel time, vehicle driving time, and cost of agents.

However, charging a high toll price can lead agents to either carpool or take the ordinary lane, leaving fewer agents willing to pay the toll. Consequently, the optimal toll price that maximizes the revenue is often lower than the ones associated with the other three objectives. This creates a trade-off between revenue maximization and the minimization of agent travel time, vehicle driving time, and costs. This tradeoff is illustrated in the Pareto front in Fig. 3.6. The blue, orange, and green curves in Fig. 3.6 show the maximum attainable revenue at each value of agent travel time, vehicle driving time, and cost for HOT capacity of 0.25, 0.5, and 0.75, respectively.[4]

[4]While the optimal toll design of a high HOT capacity dominates the optimal toll design of lower HOT capacities, it does not mean that setting more lanes as the HOT lane is better in reality. We need to take other aspects into considerations, for example the accessibility and

HOT Capacity	Objective	Toll Prices				
		Auto Mall	Mowry	Decoto	Whipple	Hesperian
0.25	Current Prices	$1.1	$2.2	$2.5	$4.0	$5.0
0.25	Agent Time Minimization	$0.5	$1.5	$0.5	$0.5	$5.0
	Vehicle Time Minimization	$0.5	$1.5	$0.5	$0.5	$4.0
	Revenue Minimization	$0.0	$1.0	$1.0	$0.5	$1.5
	Cost Minimization	$0.5	$1.5	$0.5	$0.5	$5.0
0.50	Agent Time Minimization	$0.0	$0.5	$0.0	$0.0	$5.0
	Vehicle Time Minimization	$0.0	$0.5	$0.0	$0.0	$5.0
	Revenue Minimization	$0.5	$1.5	$1.5	$1.5	$1.0
	Cost Minimization	$0.0	$0.5	$0.0	$0.0	$5.0
0.75	Agent Time Minimization	$0.0	$0.5	$0.0	$0.0	$5.0
	Vehicle Time Minimization	$0.0	$1.5	$0.0	$0.0	$5.0
	Revenue Minimization	$0.5	$2.5	$4.5	$1.5	$4.5
	Cost Minimization	$0.0	$0.0	$0.0	$0.0	$5.0

Table 3.2: Optimal toll prices and HOT capacity design for 5-6 pm on all five road segments.



(a) Agent Time (b) Vehicle Time (c) Cost

Figure 3.6: Pareto fronts of agent travel time, vehicle driving time, and cost w.r.t revenue for 5-6 pm on all road segments. The blue, orange, and green curves show the maximum attainable revenue at each value of agent travel time, vehicle driving time, and cost for HOT capacity of 0.25, 0.5, and 0.75 respectively.

3.5.4 Hourly optimal toll pricing

In this section, we compute the optimal hourly toll price from 5am to 8pm for each one of the four objectives. We set the HOT capacity as 0.25 to match the current HOT capacity on I-880. In Figure 3.7, red curves represent the optimal toll prices, blue dots represent the average current toll price recorded in the data, and the blue shaded regions represent the 95% confidence interval of the current toll price.

For Auto Mall Parkway, the optimal toll prices for minimizing agent travel time, vehicle driving time, and costs are lower than the current prices for most hours of the day. The optimal toll prices for revenue maximization are higher in the early morning but lower in the afternoon (Figures 3.7a-3.7d). On Mowry Ave (Figures 3.7e-3.7h), the optimal toll prices for agent travel time minimization, vehicle driving time minimization, and cost minimization are lower than the current prices in the morning (before 12 pm) but higher during evening rush hours (around 5 pm). The optimal toll prices for revenue maximization are lower than the current prices during evening rush hours but similar in other

equity of agents with different incomes.

hours. For Decoto Rd (Figures 3.7i-3.7l) and Whipple Rd (Figures 3.7m-3.7p), the optimal toll prices for all four objectives are lower than the current prices during evening rush hours and similar to the current prices at other times. Lastly, for Hesperian Blvd (Figures 3.7q-3.7t), the optimal toll prices for agent travel time and vehicle driving time minimization are lower than the current prices during morning rush hours (6-8 am), while the optimal toll prices for revenue maximization are lower during evening rush hours.

On each road segment, the optimal toll prices for agent travel time minimization, vehicle driving time minimization, and cost minimization are almost identical at each hour. However, the optimal toll prices for revenue maximization are similar during the morning hours (before 12 pm) but significantly lower than the optimal toll prices for the other three objectives during the evening rush hours (around 5 pm), especially on Hesperian Blvd.

In the morning, when the agent demand volume is low, setting a high toll price to incentivize carpooling does not lead to a significant reduction in HOT latency. Therefore, the optimal toll prices for both the minimization of agent travel time, vehicle driving time, and costs, as well as for revenue maximization, are low during the morning hours.

On the other hand, in the evening hours, when agent demand is higher, setting a high toll price can lead to a substantial reduction in HOT latency, particularly on road segments with high demand. However, this also discourages most people from paying the toll. As a result, the optimal toll prices for revenue maximization are lower than those for minimizing agent travel time, vehicle driving time, and costs. This difference is more pronounced on road segments with high agent demand, such as Hesperian Blvd.

Figure 3.8 illustrates the improvements achieved by implementing optimal toll prices in terms of agent travel time, vehicle driving time, revenue, and cost, compared to the current toll prices. Blue bars represent the percentage improvements, while red curves depict the numerical improvements. By using the optimal toll prices, we can achieve reductions of up to 30-40% in agent travel time, vehicle driving time, and costs, and an increase in revenue of up to 2500%. Specifically, optimal toll prices can lead to reductions of up to 300,000 minutes in total agent travel time, 350,000 minutes in total vehicle driving time, and $250,000 in total costs, while increasing total revenue by up to $40,000. The largest numerical improvements for all four objectives occur during the afternoon hours, when travel demand is high.

3.6 Concluding remarks

In this article, we examine a game-theoretic model that analyzes the lane choice of travelers with heterogeneous values of time and carpool disutilites on highways equipped with HOT lanes. For highways with a single road segment, we characterize the equilibrium strategies, and identify two qualitatively distinct equilibrium regimes that depends on the HOT lane capacity and toll price. We discuss how equilibrium strategies and latency difference of ordinary lanes and HOT lanes will change by increasing the HOT capacity or toll price. Additionally, we extend our model to highways with multiple entrance and exit nodes. We calibrate our model using the data of California Interstate highway 880 and determine the optimal capacity allocation and toll design. As a future direction of research, we will investigate the equilibrium property in a fully generalized network, and the design of HOT systems with multiple combined objectives.

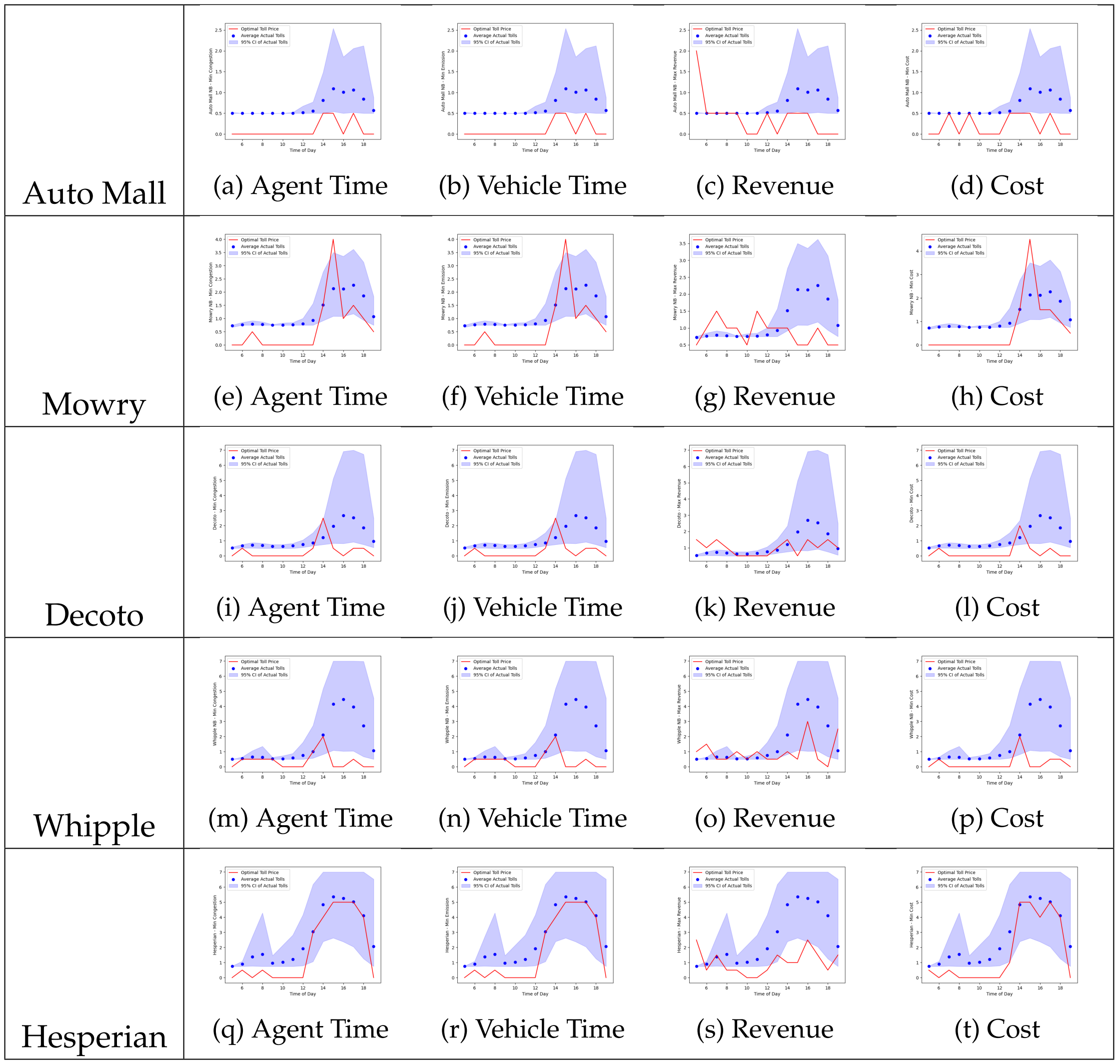
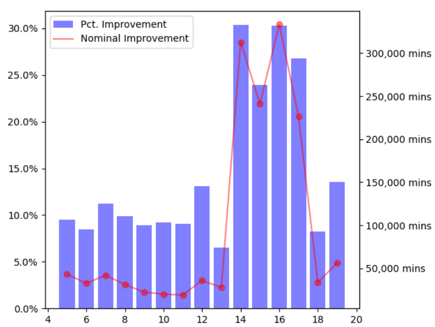
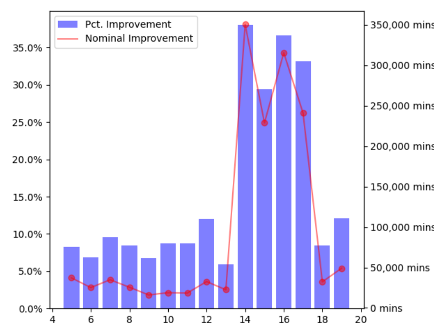


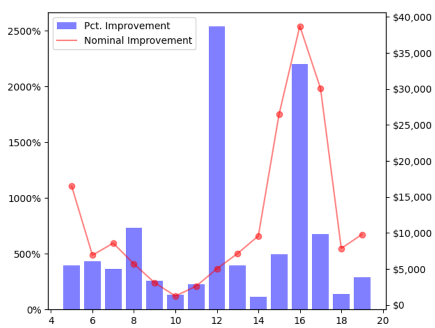
Figure 3.7: Optimal toll prices for agent travel time minimization, vehicle driving time minimization, revenue maximization, and cost minimization in each hour from 5 am to 8 pm of a workday. Red curves represent the optimal toll prices, blue dots represent the average current toll prices at each hour, while the blue shaded regions represent the 95% confidence regions of the current toll prices, across all workdays from March 1st 2021 to August 31st 2021.



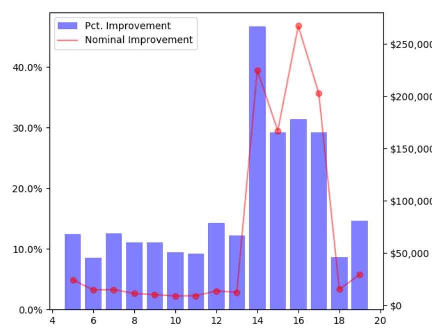
(a) Agent Time



(b) Vehicle Time



(c) Revenue



(d) Cost

Figure 3.8: Improvement of objectives by optimal toll price design in each hour from 5 am to 8 pm of a workday.

CHAPTER 4

GAME-THEORETIC ANALYSIS OF COMMISSION-BASED PRICING IN TWO-TIER SUPPLY CHAINS

4.1 Introduction

In many supply chain settings, manufacturers rely on intermediaries—such as retailers or retailers—to deliver their products to consumers. These two-tier structures raise important questions about how upstream manufacturers can incentivize downstream agents to set prices in ways that align with their strategic interests. In this paper, we study a game-theoretic model of such a two-tier supply chain involving manufacturers and their associated retailers. Each manufacturer produces a set of differentiated products and delegates pricing decisions to a dedicated retailer by offering a linear contract that scales linearly with the product prices and sale quantities. We seek to understand how these contracts affect equilibrium outcomes, and how they compare to a baseline one-tier market where manufacturers sell directly to consumers.

We model the supply chain as a two-stage game involving two competing manufacturers and their respective retailers. In the first stage, each manufacturer selects a linear contract scheme for their retailer that consists of a proportional rate on the retailer's revenue and an additive offset on each product's sales price. In the second stage, given the linear contracts, the two retailers simultaneously set product prices in a downstream pricing game to maximize their individual payoffs that equal the total revenue minus contract payments. The market demand for each product is determined by a linear function of all product prices that captures substitutability and complementarity across products. The solution concept of the two-stage game is subgame-perfect equilibrium (SPE) that characterizes the equilibrium contract parameter design from the manufacturers and the equilibrium prices induced by the competing retailers.

The first step in our analysis shows that the two-stage game admits a natural simplification. Specifically, we identify a reduced game where each manufacturer's strategic decision reduces to only selecting the additive parameters for their products, while the proportional parameters are determined to ensure the retailer earns a fixed minimum utility. This reduction significantly simplifies the equilibrium analysis.

To illustrate the economic implications of this structure, we begin with a motivating example where each manufacturer produces a single product, and the two products are substitutes. In this case, we show that manufacturers set strictly positive additive parameters at equilibrium that induce higher prices compared to the direct-sale benchmark. This suggests that strategic contract design can result in inflated prices even in competitive settings in the two-tier supply chain competition.

We then formalize this insight under general linear demand functions with multiple products. We derive closed-form expressions for the equilibrium price and contract parameters profiles, and provide sufficient conditions under which manufacturers always choose positive additive parameters at SPE. In particular, when all products are substitutes and the price sensitivity matrix is diagonally dominant, the resulting retail prices under the two-tier structure are strictly higher than that in the direct-sale game. These results highlight how strategic contract design can systematically distort market prices. Having conducted extensive numerical searches without uncovering any counterexamples, we further conjecture that both the positivity of the equilibrium additive parameters and the price-comparison outcome continue to hold even without the diagonal-dominance requirement on price sensitivity matrices.

Finally, we extend our analysis to cases where products can be complementary. We illustrates with a simple example that manufacturers may optimally choose negative additive parameters—that is, pay their retailers to lower prices on certain products to boost the sales of another complementary product. As a result, the equilibrium price resulted from the two-tier competition can be higher for some products while lower on other products compared to the price resulted from competition in direct sales.

4.1.1 Literature Review

Our work contributes to three lines of research: strategic delegation, network games, and competition in two-tier supply chains.

Strategic Delegation. There is a rich literature on strategic delegation in vertical structures, most of which considers a simplified setting where each manufacturer-retailer pair handles a single product. For a comprehensive overview, see the book chapter of Kopel and Pezzino [60]. These works model the competition and strategic delegation as a two-stage game: in the first stage, manufacturers choose incentive schemes for their retailers; in the second stage, retailers compete in the market by setting quantities or prices. Foundational works Vickers [103], Fershtman and Judd [34], and Sklivas [98] show that manufacturers can influence the market outcomes by delegating decision-making to retailers with specific incentive structures. In Cournot competition, manufacturers incentivize retailers to set higher product quantities, and force rivals to also increase their production. Such delegation increases the total quantity, lowers the prices, and reduces profits in equilibrium in comparison to direct sale competition. In Bertrand competition, manufacturers incentivizes retailers to set higher prices, which softens the competition and can raise equilibrium prices above the price in direct competition. These articles also demonstrated that manufacturers can effectively achieve their profit-maximizing goals by incentivizing retailers to pursue a combined objective of manufacturers' profit and revenue.

Most of this literature focuses on single-product manufacturers except for a few exceptions. Moner-Colonques et al. [81] examine the case where a manufacturer with two products competes with a manufacturer with one product in a Cournot competition. The authors found that the manufacturer with two products may prefer delegating the sales of one product while selling the other directly. Alles and Datar [1] explore a Bertrand competition, where two manufacturers both producing the same two products with different product costs. Both manufacturers strategically incentivize the their retailers to increase the price of the product with lower cost, and reduces the price of the product with higher cost. None of the aforementioned studies accommodates manufacturers with an arbitrary number of products.

Furthermore, prior studies have also investigated strategic delegation through shared retailers. Choi [17] explores a similar setting where two manufacturers both sell through two common retailers for a single product. Federgruen and Hu [33] later extend Choi's model by incorporating multiple manufacturers, multiple differentiated products, and multiple common retailers.

Our work extends the literature on strategic delegation with dedicated retailers by moving beyond the single- or two-product settings. We model strategic delegation as a two-stage game in which each manufacturer-retailer pair manages multiple products. We fully characterize the subgame perfect equilibrium of this game under linear demand functions. Interestingly, our results generalize the core insight of Fershtman and Judd [34] and Sklivas [98] and provide sufficient conditions under which strategic delegation can soften downstream competition and increase product prices even in a multi-product setting. Moreover, we show that a manufacturer may subsidize certain products to boost the sales of complementary products—an effect that does not arise in single-product models.

Network Games. There is extensive research on how network structure shapes strategic behavior and market outcomes. Foundational work by Ballester et al. [3] shows that in linear-quadratic network games, players' equilibrium actions align with their Bonacich centrality. Subsequent studies extend this insight to pricing settings where firms set personalized prices for consumers embedded in a network and exhibiting positive consumption externalities. For example, Bloch and Quérou [5] and Candogan et al. [9] examine a monopolist optimizing prices under equilibrium consumption constraints; later works (e.g., Bimpikis et al. [4], Chen et al. [14]) explore competitive pricing under symmetric or asymmetric network structures, with varying assumptions on consumer behavior, product differentiation, and information availability.

While these studies provide important insights into network externalities and price-setting behavior, their settings differ fundamentally from ours in two key respects. First, most prior work models externalities between consumers that arise from social interactions within a network, while we model externalities between products, where the price of one product influences the demand for others. This distinction shifts the focus from consumer-level peer effects to product-level demand interactions in a two-tier supply chain. Second, prior models typically assume that firms directly set consumer prices, while we analyze a two-stage structure in which upstream manufacturers delegate pricing decisions to retailers by designing linear contracts. This delegation structure introduces a new layer of strategic behavior. Our modeling framework reveals that optimal linear contract design depends explicitly on the nature of cross-product externalities—whether products are substitutes or complements.

Additional literature on two-tier supply chain competition. Apart from the articles summarized in the strategic delegation literature, prior research has also examined other forms of incentive design and pricing in two-tier supply chain settings. For example, Demirag et al. [27] and Chen et al. [13] study settings where manufacturers use lump-sum commissions to incentivize retailers and analyze the resulting pricing outcomes. Taylor [100], Chen and Riordan [16], Arya and Mittendorf [2], and Demirag et al. [26] explore rebate mechanisms—such as linear or target-based rebates—in supply chains without strategic competition between upstream firms. These papers investigate the role of different contractual forms in shaping downstream pricing but typically do not model multiple products or strategic upstream interactions. In addition, a separate line of work has studied a different type of two-tier supply chain competition, where downstream firms purchase goods from upstream producers and compete in serving end consumers. These articles studied models of both Cournot and Bertrand competition across vertically related firms Salinger [96], Ordover et al. [86], Hart et al. [44], Chen [15], and Chen and Riordan [16]. While our model also features vertical competition and a two-stage game structure, it differs significantly in that manufacturers delegate pricing to retailers through commissions, rather than selling intermediate goods to competing downstream firms.

In contrast to these studies, our paper introduces a comprehensive framework that integrates strategic commission contract design, a two-tier supply chain with exclusive delegation, and competition between duopolistic manufacturers. We derived closed-form equilibrium outcomes under general linear demand, demonstrating how intermediation and linear contract structures jointly distort prices in competitive environments.

4.2 Model Setup

Two-tier supply chain. Consider two manufacturers $k \in \{1, 2\}$, each producing a set of differentiated products, denoted as I_k. The set of all products is $I = I_1 \cup I_2$. Each manufacturer k sells their products in the market through a corresponding retailer k who sets the price p_i for each product $i \in I_k$. Let $p := (p_i)_{i \in I}$ be the price vector for all products, and $p_{I_k} = (p_i)_{i \in I_k}$ be the price vector chosen by each retailer k. The demand for each product $i \in I$ is

$$\pi_i(p) = 1 + \sum_{j \in I} d_{ij} p_j,$$

where each d_{ij} is the price sensitivity parameter that measures the impact of product j price on the demand of product i. We assume that the demand $\pi_i(p)$ for product i decreases with product i's own price p_i (i.e. $d_{ii} < 0$), and may decrease or increase with other products' prices p_j for $j \neq i$. If $d_{ij} > 0$, then product j is a *substitute* of product i; if $d_{ij} < 0$, then product j is a *complement* of product i. Let $D = (d_{ij})_{i,j \in I}$ denote the price sensitivity matrix.

For each product $i \in I_k$, the retailer k receives utility modeled by a linear contract: $h_i(p_i) = \alpha_k(p_i - c_i)$ for selling a unit of product i. The proportional parameters $\alpha_k \in [0,1]$ and additive parameters $c_i \in \mathbb{R}$ are determined by manufacturer k, where α_k represents the part of the utility that is proportional to the product price, and c_i is a product-specific additive contract parameter. Let $\alpha = (\alpha_k)_{k=1,2}$ and $c = (c_i)_{i \in I}$ be vectors of all contract parameters. We denote the the contract parameters set by manufacturer k as $c_{I_k} = (c_i)_{i \in I_k}$. We formally define the two-stage game as follows:

Stage 1 (manufacturer game): Each manufacturer k simultaneously chooses contract parameters $(\alpha_k, c_{I_k}) \in [0,1] \times \mathbb{R}^{|I_K|}$.

Stage 2 (retailer game): Observing the contract parameters (α, c), each retailer k simultaneously chooses product prices $p_{I_k} \in \mathbb{R}^{|I_K|}$. We assume that there is a lower bound LB on the retailer's utility that is guaranteed by the manufacturer to ensure that retailers have sufficient incentive to provide the service. Given any price vector p and contract parameters (α, c), each retailer k's utility can be written as:

$$U_k^w(p; \alpha, c) = \max\left\{\sum_{i \in I_k} \pi_i(p)\alpha_k(p_i - c_i),\ \text{LB}\right\}, \quad \forall k = 1, 2. \tag{4.1}$$

Manufacturer k receives all the utility from the sales of the products $i \in I_k$ minus the contract utility paid to retailer k. Given price vector p and contract parameters (α, c), each manufacturer k's utility is given by:

$$U_k^m(p;\alpha,c) = \sum_{i\in I_k} p_i \pi_i(p) - U_k^w(p;\alpha,c), \quad \forall k = 1,2. \tag{4.2}$$

Now we define subgame-perfect equilibrium (SPE) of the two-stage game.

Definition 10. *A strategy profile (α^*, c^*, p^*) is an SPE of the two-stage game if*

1. *The price vector $p^*(\alpha, c) \in \mathbb{R}^{|I|}$ is a Nash equilibrium of the retailer game for any contract parameters $(\alpha, c) \in [0,1]^2 \times \mathbb{R}^{|I|}$: for all $k \in \{1,2\}$,*

$$p^*_{I_k}(\alpha,c) \in \operatorname*{argmax}_{p_{I_k}\in\mathbb{R}^{|I_k|}} U_k^w((p_{I_k}, p^*_{I_{-k}}(\alpha,c));\alpha,c).$$

2. *The manufacturer's strategy (α^*, c^*) is a Nash equilibrium of the manufacturer game given that the retailers choose an equilibrium price function $p^*(\alpha, c)$ in the subgame: for all $k \in \{1,2\}$,*

$$(\alpha_k^*, c^*_{I_k}) \in \operatorname*{argmax}_{\alpha_k\in[0,1], c_{I_k}\in\mathbb{R}^{|I_k|}} U_k^m(p^*((\alpha_k,\alpha^*_{-k}),(c_{I_k}, c^*_{I_{-k}}));(\alpha_k,\alpha^*_{-k}),(c_{I_k}, c^*_{I_{-k}})).$$

Equilibrium selection in the retailer game. For the retailer game, we only consider the set of price equilibria $P^*(\alpha, c)$ that maximize the actual share of revenue $\sum_{i\in I_k} \pi_i(p)\alpha_k(p_i - c_i)$ without considering the retailer's utility lower bound. Formally, $p^*(\alpha, c) \in P^*(\alpha, c)$ if and only

$$p^*_{I_k}(\alpha,c) \in \operatorname*{argmax}_{p_{I_k}\in\mathbb{R}^{|I_k|}} \sum_{i\in I_k} \pi_i((p_{I_k}, p^*_{I_{-k}}(\alpha,c)))\alpha_k(p_i - c_i), \quad \forall k \in \{1,2\}. \tag{4.3}$$

For each k, the objective $\sum_{i\in I_k} \pi_i((p_{I_k}, p^*_{I_{-k}}(\alpha,c)))\alpha_k(p_i - c_i)$ is jointly continuous in $(p_{I_k}, \alpha, c, p^*_{I_{-k}})$ and tends to $-\infty$ as $\|p_{I_k}\| \to \infty$, so any maximizer must lie in a compact region. By the extended Berge Maximum Theorem for sup-compact objectives, $P^*(\alpha, c)$ is nonempty. When the actual share of revenue induced by retailer competition is lower than LB, the retailers may gain the same utility LB from choosing other price vectors not in $P^*(\alpha, c)$. Those price vectors are all Nash equilibria of the sub-game, but impose varying impacts on the manufacturers' utilities. By focusing on the equilibria that maximize the actual share of revenue, we ensure that the lower bound constraint does not skew the price equilibrium of the retailer game. We show that $P^*(\alpha, c)$ is a subset of all equilibria in the retailer subgame in Lemma 12 in Appendix C.1. For the rest of the paper, we only consider the subset of SPE that is in $P^*(\alpha, c)$.

Lemma 8. *For any c, α and α', $P^*(\alpha, c) = P^*(\alpha', c)$.*

Proof. Given (α, c), let $p^* \in P^*(\alpha, c)$ be a Nash equilibrium. We show that p_{I_k} is still a best response to $p^*_{I_{-k}}$ for both k with α':

$$\begin{aligned} p^*_{I_k} &\in \arg\max_{p_{I_k}} \sum_{i \in I_k} \alpha_k \pi_i(p_{I_k}, p^*_{I_{-k}})(p_i - c_i) \\ &= \arg\max_{p_{I_k}} \alpha_k \left(\sum_{i \in I_k} \pi_i(p_{I_k}, p^*_{I_{-k}})(p_i - c_i) \right) \\ &= \arg\max_{p_{I_k}} \left(\sum_{i \in I_k} \pi_i(p_{I_k}, p^*_{I_{-k}})(p_i - c_i) \right) \\ &= \arg\max_{p_{I_k}} \alpha'_k \left(\sum_{i \in I_k} \pi_i(p_{I_k}, p^*_{I_{-k}})(p_i - c_i) \right). \end{aligned} \tag{4.4}$$

Therefore, $p^*_{I_k}$ is a best response to $p^*_{I_{-k}}$ given (α', c) for both $k = 1, 2$. Consequently, $p^* \in P^*(\alpha', c)$. □

Since the set of Nash equilibrium in the lower level game does not depend on α, we simplify the notation, and denote the equilibrium price vector in the lower level game as $p^*(c) \in P^*(c)$.

Proposition 3. *If the tuple (α^*, c^*) that satisfies the following conditions exists, then it constitutes a subgame perfect equilibrium (SPE) strategy for the manufacturers:*

$$c_k^* \in \arg\max_{c_k} \sum_{i \in I_k} \pi_i(p^*(c_k, c_{-k}^*)) p_i^*(c_k, c_{-k}^*). \tag{4.5}$$

and

$$\alpha_k^* = \min\left\{ \frac{LB}{\sum_{i \in I_k} \pi_i(p^*(c^*))(p_i^*(c^*) - c_i^*)}, \ 1 \right\}. \tag{4.6}$$

Consequently, the equilibrium utility of the manufacturer is given by:

$$U_k^m(\alpha^*, c^*) = \sum_{i \in I_k} \pi_i(p^*(c^*)) p_i^*(c^*) - LB,$$

where $p^*(c^*) \in P^*(c)$ is an equilibrium of the retailer game, and the equilibrium utility of the retailer is LB for both $k \in \{1, 2\}$.

Proof. We show that (α_k^*, c_k^*) is a best response strategy to $(\alpha_{-k}^*, c_{-k}^*)$:

$$\begin{aligned}
& U_k^m(\alpha_k, c_k, \alpha_{-k}^*, c_{-k}^*) \\
=& \sum_{i \in I_k} \pi_i(p^*(c_k, c_{-k}^*)) p_i^*(c_k, c_{-k}^*) \\
& \quad - \max\left\{ \alpha_k \left(\sum_{i \in I_k} \pi_i(p^*(c_k, c_{-k}^*))(p_i^*(c_k, c_{-k}^*) - c_i^*) \right), \mathrm{LB} \right\} \\
\leq& \max_{c_k} \sum_{i \in I_k} \pi_i(p^*(c_k, c_{-k}^*)) p_i^*(c_k, c_{-k}^*) - \mathrm{LB} \\
=& U_k^m(\alpha_k^*, c_k^*, \alpha_{-k}^*, c_{-k}^*).
\end{aligned}$$

□

Based on the above proposition, the set of SPE can be obtained by solving a reduced two-stage game. In this reduced game, each manufacturer $k \in \{1, 2\}$

selects an additive parameter c_k and receives utility given by

$$R_k^m(c) = \sum_{i \in I_k} \pi_i(p^*(c)) p_i^*(c),$$

where $p^*(c)$ represents a Nash equilibrium of the reduced retailer game for any given additive parameters c. In the reduced retailer game, each retailer k selects a price p_{I_k} and gains utility expressed as

$$R_k^w(p; c) = \sum_{i \in I_k} \pi_i(p)(p_i - c_i).$$

A strategy profile (α^*, p^*, c^*) is an SPE of the original game if and only if (p^*, c^*) is an SPE of the reduced game and α^* satisfies (4.6).

Direct sale. We also define a direct sale game, where manufacturers set prices and sell directly to the consumers. Each manufacturer $k \in \{1, 2\}$ sets prices $\tilde{p}_{I_k}$ for all products $i \in I_k$ and aims to maximize the total utility from selling all products:

$$\tilde{R}_k(\tilde{p}) := \sum_{i \in I_k} \pi_i(\tilde{p}) \tilde{p}_i.$$

The Nash equilibrium of direct sale game is defined as follow:

Definition 11. *A price vector $\tilde{p}^* \in \mathbb{R}^{|I|}$ is a Nash equilibrium of the direct sale game if $\tilde{p}^*_{I_k}$ maximizes the utility for manufacturer k given their opponent manufacturer sets the prices to be $\tilde{p}^*_{I_{-k}}$. That is, for all $k \in \{1, 2\}$,*

$$\tilde{p}^*_{I_k} \in \operatorname*{argmax}_{tp_{I_k} \in \mathbb{R}^{|I_k|}} \tilde{R}_k((\tilde{p}_{I_k}, \tilde{p}^*_{I_{-k}})).$$

To understand the impact of the two-tier supply chain structure, we compare the equilibrium price $p^*(c^*)$ from the subgame-perfect equilibrium (SPE) of the two-tier supply chain with the equilibrium price $p^*(0)$ from the direct sale game. This comparison will reveal how the two-tier supply chain competition affects the final market prices.

4.3 Motivating Example

We provide a motivating example showing that in an equilibrium of the two stage game, the additive parameters c^* are in general not zero. The example studies a simple setting with two products, each owned by one manufacturer. The additive parameters c^* at equilibrium are always positive, indicating that manufacturers sets contract parameters strategically to induce higher pricing in the retailer competition compared that in direct sale.

Example 1 Suppose each manufacturer has one product. $I_1 = \{1\}, I_2 = \{2\}$. Let the demand functions be linear:

$$\pi_1(p) = 1 - p_1 + dp_2,$$
$$\pi_2(p) = 1 - p_2 + dp_1,$$

for some price sensitivity parameter $d \in (0, 1)$. Given any additive parameters c, the price equilibrium $p^*(c)$ of the retailer game satisfies:

$$\begin{aligned} p_k^*(c) &= \operatorname*{argmax}_{p_k} R_k^w((p_k, p_{-k}^*(c)); c) \\ &= \operatorname*{argmax}_{p_k} \pi_k((p_k, p_{-k}^*(c)))(p_k - c_k) \\ &= \operatorname*{argmax}_{p_k} (1 - p_k + dp_{-k}^*(c))(p_k - c_k), \quad \forall k \in \{1, 2\}, \end{aligned}$$

The first order condition for the retailer game is for all $k \in \{1, 2\}$,

$$\left. \frac{\partial R_k^w(p_k, p_{-k}^*(c); c)}{\partial p_k} \right|_{p_k = p_k^*(c)} = 1 - p_k^*(c) + dp_{-k}^*(c) - (p_k^*(c) - c_k) = 0.$$

and the second-order derivative is negative:

$$\frac{\partial^2 R_1^w((p_1, p_2; c)}{\partial p_1^2} = \frac{\partial^2 R_2^w((p_1, p_2); c)}{\partial p_2^2} = -2.$$

This implies:

$$2p_1^*(c) - dp_2^*(c) = c_1 + 1,$$
$$2p_2^*(c) - dp_1^*(c) = c_2 + 1.$$

Therefore, the equilibrium price function is

$$p_1^*(c) = \frac{2\,(c_1+1) + d\,(c_2+1)}{4-d^2}$$
$$p_2^*(c) = \frac{d\,(c_1+1) + 2\,(c_2+1)}{4-d^2}.$$

Note that for each $k \in \{1,2\}$, $p_k^*(c)$ is is increasing in both c_k and c_{-k}. We can write the manufacturer k's utility as a function of c:

$$\begin{aligned}
R_k^m(c) =& \pi_k(p^*(c))p_k^*(c) \\
=& \left(1 - \frac{2\,(c_k+1) + d\,(c_{-k}+1)}{4-d^2} + d \cdot \frac{d\,(c_k+1) + 2\,(c_{-k}+1)}{4-d^2}\right) \\
& \cdot \frac{2\,(c_k+1) + d\,(c_{-k}+1)}{4-d^2} \\
=& \frac{2\,(c_k+1) + d\,(c_{-k}+1)}{(4-d^2)^2} \\
& \cdot \Big((4-d^2) - 2\,(c_k+1) - d\,(c_{-k}+1) + d^2\,(c_k+1) + 2d\,(c_{-k}+1)\Big) \\
=& \frac{2\,(c_k+1) + d\,(c_{-k}+1)}{(4-d^2)^2} \\
& \cdot \Big((4-d^2) + (d^2-2)\,(c_k+1) + d\,(c_{-k}+1)\Big)
\end{aligned}$$

The first order conditions for the manufacturer game are: for all $k \in \{1,2\}$,

$$\begin{aligned}
& \frac{\partial R_k^m(c)}{\partial c_k} \\
=& \frac{(d^2-2)\Big(2\,(c_k+1) + d\,(c_{-k}+1)\Big)}{(4-d^2)^2} \\
& + \frac{2\,\big((4-d^2) + (d^2-2)\,(c_k+1) + d\,(c_{-k}+1)\big)}{(4-d^2)^2} \\
=& \frac{1}{(4-d^2)^2}\left(2(4-d^2) + \Big(4(d^2-2)\,(c_k+1) + (d^3\,(c_{-k}+1)\Big)\right) = 0.
\end{aligned}$$

The second-order derivative is negative:

$$\frac{\partial^2 R_1^m(c)}{\partial c_1^2} = \frac{\partial^2 R_2^m(c)}{\partial c_2^2} = \frac{4(d^2-2)}{(4-d^2)^2} < 0.$$

Therefore, the first order conditions give the best response function for each manufacturer k. Simplifying the first order condition, we have

$$2(4-d^2) + \Big(4(d^2-2)\,(c_k+1) + d^3\,(c_{-k}+1)\Big) = 0,$$
$$2(4-d^2) + \Big(4(d^2-2)\,(c_k+1) + d^3\,(c_{-k}+1)\Big) = 0, \quad \forall k \in \{1,2\}.$$

Let $\text{BR}_k(c_{-k})$ be the best response for manufacturer k when the opponent manufacturer choose c_{-k}. We have

$$\text{BR}_k(c_{-k}) = -\frac{(4-d^2)}{2(d^2-2)} - \frac{d^3}{4(d^2-2)}\,(c_{-k}+1) - 1, \quad \forall k \in \{1,2\}.$$

The intersection of the best response functions is the additive contract parameter equilibrium. Therefore, the SPE of the two tier game is

$$c_1^* = c_2^* = -\frac{2(4-d^2)}{d^3+4d^2-8} - 1.$$
$$p_1^*(c^*) = p_2^*(c^*) = -\frac{2(d+2)}{d^3+4d^2-8}.$$

Notice that $c_1^*, c_2^* > 0$ for any $d \in (0,1)$, and increases with d. Since $p^*(c)$ increases with c, the Nash equilibrium of the direct sale game $\tilde{p}^* = p^*(0)$ is smaller than the price equilibrium $p^*(c^*)$ in the two tier game:

$$\tilde{p}_1^* = p_1^*(\vec{0}) < p_1^*(c^*)$$
$$\tilde{p}_2^* = p_2^*(\vec{0}) < p_2^*(c^*).$$

In this motivating example, the additive parameters c^* at equilibrium are always positive. With the two-tier structure, manufacturers utilize the linear contract to incentivize their retailers to set a higher equilibrium price compared to the direct sale game.

4.4 Main Result

In this section, we analyze the equilibrium of the two-tier supply chain game with linear demand functions, and compare the equilibrium price with that of the direct sale game. For conciseness, we define $N \in \{0,1\}^{|I|\times|I|}$ be the matrix that indicates whether two products belong to the same manufacturer. That is,

$$N_{ij} = \begin{cases} 1 & \text{if } i,j \in I_k \text{ for some } k \in \{1,2\} \\ 0 & \text{if } i \in I_k, j \in I_{-k} \text{ for some } k \in \{1,2\}. \end{cases}$$

Notice that N is a symmetric square matrix that consists of two block matrices on the diagonal with all elements being 1, and the remaining two off-diagonal block matrices have all elements being 0. Additionally, we define the submatrices D_1, D_2, D_3, D_4 of the sensitivity matrix D:

$$D = \begin{bmatrix} D_1 & D_2 \\ D_3 & D_4 \end{bmatrix}, \tag{4.7}$$

where

$$D_1 = (D_{ij})_{i,j\in I_1}, \quad D_2 = (D_{ij})_{i\in I_1, j\in I_2}, \quad D_3 = (D_{ij})_{i\in I_2, j\in I_1}, \quad D_4 = (D_{ij})_{i,j\in I_2}.$$

Here, D_1 and D_4 capture the sensitivity parameters of products within the first and second manufacturer's product sets, I_1 and I_2, respectively, and D_2 and D_3 represent the sensitivities between products belong to two different manufacturers.

Other standard notation. For a matrix $U \in \mathbb{R}^{m\times n}$, we write $U > 0$ to indicate that the inequality holds element-wise, i.e., $U_{ij} > 0$ for all $i = 1, \dots, m, j = 1, \dots, n$. Similarly, $U \geq 0$, $U < 0$, and $U \leq 0$ are also interpreted element-wise.

For any two matrices U, V with the same dimension, the Hadamard product (element-wise multiplication) is defined as $W = U \circ V$, where $w_{ij} = u_{ij}v_{ij}$. For a vector $z \in \mathbb{R}^{|I|}$, we use $z_{I_k} := (z_i)_{i \in I_k}$ to represent the entries z_i corresponding to $i \in I_k$. Similarly, for a matrix $U \in \mathbb{R}^{|I| \times |I|}$, we use $U_{I_k} := (U_{ij})_{i,j \in I_k}$ to represent the sub-matrix block corresponding to the rows and columns in I_k. We use J_n to represent an all-one matrix with size n, and I_n to represent an identity matrix with size n. We use $\mathrm{Diag}(z)$ to represent a diagonal matrix with z on the diagonal.

We make the following assumptions to ensure that our problems and solutions are well-defined. Assumption 5 ensures that certain matrices are invertible. Assumptions 6 and 7 ensure that the utility functions of the manufacturer and retailers are concave in their strategies so that equilibrium exists. Assumption 8 restricts our discussion to settings in which all equilibria lie in the interior of the feasible region. That is, the equilibrium price $p^*(c^*)$ and the demand $\pi_i(p_j^*(c^*))$ of each product in equilibrium are positive.

Assumption 4 (Block-Level Invertibility)**.** *For each $k \in (1, 2)$, the matrices D_{I_k}, $D_{I_k} + D_{I_k}^T$ are invertible.*

For notational brevity, we introduce the following two matrices:

$$\Gamma = D + (N \circ D^T), \tag{4.8}$$

$$A = \begin{bmatrix} D_3^T (D_4 + D_4^T)^{-1} D_2^T & 0 \\ 0 & D_2^T (D_1 + D_1^T)^{-1} D_3^T \end{bmatrix}. \tag{4.9}$$

Assumption 5 (Full-Rank and Invertibility)**.** *The following matrices are invertible:*

$$D, \quad \Gamma, \quad \Gamma - A, \quad \textit{and} \quad D + \left(N \circ \Gamma^{-T}\right)^{-1} (N \circ \Gamma^{-1} D^T).$$

Assumption 6 (Concavity of Retailer Payoffs)**.** *For any k, the matrix $(D_{I_k} + D_{I_k}^T)$ is negative definite.*

Assumption 7 (Concavity of Manufacturer Payoffs)**.** *For any k, the matrix $((\Gamma^{-1})^T D^T)_{I_k}(\Gamma^{-1})_{I_k} + (\Gamma^{-1})_{I_k}^T (D(\Gamma^{-1}))_{I_k}$ is negative definite.*

Assumption 8 (Interior Equilibrium Condition)**.** *The matrix $-(\Gamma - A)^{-1}\vec{1} > 0$, and $1 - D(\Gamma - A)^{-1}\vec{1} > 0$.*

We first show that given any additive parameters $c \in \mathbb{R}^{|I|}$, the equilibrium price vector $p^*(c)$ in the retailer game is unique and has a closed-form solution.

Proposition 4. *Under Assumptions 4, 5, 8, and 6, for any c, the equilibrium price vector $p^*(c)$ in the retailer game is unique, and is given by:*

$$p^*(c) = (\Gamma)^{-1}((N \circ D^T)c - \vec{1}), \tag{4.10}$$

where Γ is given by (4.8). Furthermore, the price equilibrium $\tilde{p}$ of the direct sale game is

$$\tilde{p}^* = -(\Gamma)^{-1}\vec{1}.$$

Proof. In the retailer game, retailer k's utility is

$$\begin{aligned} R_k^w(p; c) &= \sum_{i \in I_k} \pi_i(p)(p_i - c_i) \\ &= \sum_{i \in I_k} (1 + \sum_{j \in I} d_{ij} p_j)(p_i - c_i) \\ &= (\vec{1} + Dp)_{I_k}^T (p - c)_{I_k}. \end{aligned}$$

Take the partial derivative with respect to p_{I_k}, we have

$$\begin{aligned}\frac{\partial R_k^w(p;c)}{\partial p_{I_k}} &= (\vec{1}+Dp)_{I_k} + \frac{\partial (Dp)_{I_k}}{\partial p_{I_k}}(p-c)_{I_k} \\ &= (\vec{1}+Dp)_{I_k} + ((N\circ \frac{\partial (Dp)}{\partial p})(p-c))_{I_k} \\ &= (\vec{1}+Dp)_{I_k} + ((N\circ D^T)(p-c))_{I_k}.\end{aligned}$$

At equilibrium, each retailer k will choose the best response $p^*_{I_k}(c)$ such that

$$\frac{\partial R_k^w(p;c)}{\partial p_{I_k}} = 0.$$

Combining the first order conditions for both $k=1,2$, we have

$$\begin{aligned}&\vec{1} + Dp^*(c) + (N\circ D^T)(p^*(c)-c) = 0 \\ \Rightarrow &\vec{1} + Dp^*(c) + (N\circ D^T)p^*(c) - (N\circ D^T)c = 0 \\ \Rightarrow &p^*(c) = (D + (N\circ D^T))^{-1}((N\circ D^T)c - \vec{1}).\end{aligned}$$

Now we verify the second order condition. For any $k\in\{1,2\}$,

$$\frac{\partial^2 R_k^w(p;c)}{\partial p_{I_k}^2} = D_{I_k}^T + D_{I_k}.$$

Assumption 6 ensures that $D_{I_k}^T + D_{I_k}$ is negative definite for any k. Therefore, the solution that satisfies the first order conditions is the price equilibrium of the retailer game.

Notice for each $k\in\{1,2\}$, the utility function of the direct sale game equals the utility function of the retailer game when the additive parameters $c=0$:

$$\tilde{R}_k(p) = R_k^w(p;\vec{0}).$$

Consequentially, price equilibrium $\tilde{p}$ of the direct sale game is given by

$$\tilde{p}^* = p^*(0) = -(\Gamma)^{-1}\vec{1}.$$

□

Given price equilibrium function p^*, the equilibrium contract parameters c^* and the equilibrium price $p^*(c^*)$ of the two stage game are also unique and has closed form expression.

Proposition 5. *Suppose that assumptions 4, 5, 8, 6, 7 holds, and let* Γ *and* A *be defined as in* (4.8) *and* (4.9)*. Equilibrium contract parameters* c^* *in equilibrium is unique and can be written as:*

$$c^* = -(N \circ D^T)^{-1} A (\Gamma - A)^{-1} \vec{1}. \tag{4.11}$$

Consequently, the equilibrium price $p^*(c^*)$ *can be written as:*

$$p^*(c^*) = -(\Gamma - A)^{-1} \vec{1}. \tag{4.12}$$

Proof of proposition 5. From proposition 4, price equilibrium $p^*(c) = (\Gamma)^{-1}((N \circ D^T)c - \vec{1})$ depends on additive parameters c only through the quantity

$$\lambda := (N \circ D^T)c.$$

That is,

$$p^*(\lambda) = \Gamma^{-1}(\lambda - \vec{1}). \tag{4.13}$$

Due to the block structure of $N \circ D^T$, λ_{I_k} only depends on c_{I_k} for each $k \in \{1, 2\}$. From assumption 4, since the matrix $N \circ D^T = \begin{bmatrix} D_1^T & 0 \\ 0 & D_4^T \end{bmatrix}$ is invertible, there exists a one-to-one correspondence between λ_{I_k} and c_{I_k}. Therefore, we can view λ_{I_k} as the decision variables for each manufacturer k. Each manufacturer k's utility can be written as a function of λ:

$$\begin{aligned} R_k^m(\lambda) &= \sum_{i \in I_k} p_i^*(\lambda) \pi_i(p^*(\lambda)) \\ &= (\vec{1} + Dp^*(\lambda))_{I_k} \cdot (p^*(\lambda))_{I_k}. \end{aligned}$$

Take derivative with respect to λ on both sides of equation (4.13), we have

$$\frac{\partial p^*(\lambda)}{\partial \lambda} = (\Gamma^{-1})^T.$$

Taking the first order derivative of $R_k^m(p^*(\lambda);\lambda)$ with respect to $(\lambda)_{I_k}$, we have for all $k \in \{1,2\}$,

$$\begin{aligned}
\frac{\partial R_k^m(p^*(\lambda);\lambda)}{\partial(\lambda)_{I_k}} &= \frac{\partial(p^*(\lambda))_{I_k}}{\partial(\lambda)_{I_k}}(\vec{1} + Dp^*(\lambda))_{I_k} + \frac{\partial(Dp^*(\lambda))_{I_k}}{\partial(\lambda)_{I_k}}(p^*(\lambda))_{I_k} \\
&= \left((N \circ \frac{\partial p^*(\lambda)}{\partial \lambda})(\vec{1} + Dp^*(\lambda))\right)_{I_k} + \left((N \circ \frac{\partial Dp^*(\lambda)}{\partial \lambda})p^*(\lambda)\right)_{I_k} \\
&= \left((N \circ (\Gamma^{-1})^T)(\vec{1} + Dp^*(\lambda))\right)_{I_k} + \left((N \circ (\Gamma^{-1})^T D^T)p^*(\lambda)\right)_{I_k}.
\end{aligned}$$

We can verify the second order condition:

$$\begin{aligned}
\frac{\partial^2 R_k^m(p^*(\lambda);\lambda)}{\partial(\lambda)_{I_k}^2} &= \frac{\partial(Dp^*(\lambda))_{I_k}}{\partial(\lambda)_{I_k}}(\frac{\partial(p^*(\lambda))_{I_k}}{\partial(\lambda)_{I_k}})^T + \frac{\partial(p^*(\lambda))_{I_k}}{\partial(\lambda)_{I_k}}(\frac{\partial(Dp^*(\lambda))_{I_k}}{\partial(\lambda)_{I_k}})^T \\
&= ((\Gamma^{-1})^T D^T)_{I_k}(\Gamma^{-1})_{I_k} + (\Gamma^{-1})_{I_k}^T(D(\Gamma^{-1}))_{I_k}.
\end{aligned}$$

By assumption 7, $((\Gamma^{-1})^T D^T)_{I_k}(\Gamma^{-1})_{I_k} + (\Gamma^{-1})_{I_k}^T(D(\Gamma^{-1}))_{I_k}$ is negative definite. Therefore, the solution for the first order condition is the maximizer of retailer k's utility. In equilibrium, each manufacturer k will choose the best response λ_k^* such that

$$\frac{\partial R_k^w(p^*(\lambda^*);\lambda^*)}{\partial(\lambda)_{I_k}} = 0, \quad \forall k \in \{1,2\}.$$

Combining the first order conditions for all manufacturer k, we have

$$(N \circ (\Gamma^{-1})^T)(\vec{1} + Dp^*(\lambda^*)) + (N \circ (\Gamma^{-1})^T D^T)p^*(\lambda^*) = 0.$$

From the first order condition, we can solve for $p^*(\lambda^*)$:

$$\begin{aligned}
& p^*(\lambda^*) \\
&= -\left((N \circ (\Gamma^{-1})^T)D + (N \circ (\Gamma^{-1})^T D^T)\right)^{-1}(N \circ (\Gamma^{-1})^T)\vec{1} \\
&= -\left((N \circ (\Gamma^{-1})^T)^{-1}(N \circ (\Gamma^{-1})^T)D + (N \circ (\Gamma^{-1})^T)^{-1}(N \circ (\Gamma^{-1})^T D^T)\right)^{-1}\vec{1} \\
&= -\left(D + (N \circ (\Gamma^{-1})^T)^{-1}(N \circ (\Gamma^{-1})^T D^T)\right)^{-1}\vec{1}.
\end{aligned}$$

The second equality is acquired from rearranging the $(N \circ (\Gamma^{-1})^T)$ term into the matrix inverse. We can plug in equation (4.13) to the left hand side and solve for λ^*:

$$\lambda^* = \vec{1} - \Gamma \left(D + (N \circ (\Gamma^{-1})^T)^{-1} (N \circ (\Gamma^T)^{-1} D^T) \right)^{-1} \vec{1}.$$

Therefore, the equilibrium additive parameter c^* can be written as:

$$\begin{aligned} c^* &= (N \circ D^T)^{-1} \lambda^* \\ &= (N \circ D^T)^{-1} \vec{1} - (N \circ D^T)^{-1} \Gamma \left(D + (N \circ (\Gamma^{-1})^T)^{-1} (N \circ (\Gamma^T)^{-1} D^T) \right)^{-1} \vec{1}. \end{aligned}$$

Now, we will try to further simplify the expression for $p^*(c^*)$:

$$p^*(c^*) = - \left(D + (N \circ (\Gamma^{-1})^T)^{-1} (N \circ (\Gamma^T)^{-1} D^T) \right)^{-1} \vec{1}. \tag{4.14}$$

Let the inverse matrix Γ^{-1} be partitioned into four submatrices as follows:

$$\Gamma^{-1} = \begin{bmatrix} (\Gamma^{-1})_1 & (\Gamma^{-1})_2 \\ (\Gamma^{-1})_3 & (\Gamma^{-1})_4 \end{bmatrix},$$

where $(\Gamma^{-1})_1 \in \mathbb{R}^{|I_1| \times |I_1|}$, $(\Gamma^{-1})_2 \in \mathbb{R}^{|I_1| \times |I_2|}$, $(\Gamma^{-1})_3 \in \mathbb{R}^{|I_2| \times |I_1|}$, $(\Gamma^{-1})_4 \in \mathbb{R}^{|I_2| \times |I_2|}$. From matrix inversion lemma 13, we have

$$\begin{aligned} (\Gamma^{-1})_2 &= -(D_1 + D_1^T)^{-1} D_2 (\Gamma^{-1})_4, \\ (\Gamma^{-1})_3 &= -(D_4 + D_4^T)^{-1} D_3 (\Gamma^{-1})_1. \end{aligned} \tag{4.15}$$

We now simplify each terms in (4.14). We first consider the term $N \circ$

$((\Gamma^T)^{-1}D^T)$.

$$
\begin{aligned}
&N \circ ((\Gamma^T)^{-1}D^T) \\
&= N \circ \begin{bmatrix} (\Gamma^{-1})_1^T & (\Gamma^{-1})_3^T \\ (\Gamma^{-1})_2^T & (\Gamma^{-1})_4^T \end{bmatrix} \begin{bmatrix} D_1^T & D_3^T \\ D_2^T & D_4^T \end{bmatrix} \\
&= \begin{bmatrix} (\Gamma^{-1})_1^T D_1^T + (\Gamma^{-1})_3^T D_2^T & 0 \\ 0 & (\Gamma^{-1})_2^T D_3^T + (\Gamma^{-1})_4^T D_4^T \end{bmatrix} \\
&\overset{(4.15)}{=} \begin{bmatrix} (\Gamma^{-1})_1^T D_1^T - (\Gamma^{-1})_1^T D_3^T (D_4 + D_4^T)^{-1} D_2^T & 0 \\ 0 & -(\Gamma^{-1})_4^T D_2^T (D_1 + D_1^T)^{-1} D_3^T + (\Gamma^{-1})_4^T D_4^T \end{bmatrix} \\
&= \begin{bmatrix} (\Gamma^{-1})_1^T (D_1^T - D_3^T (D_4 + D_4^T)^{-1} D_2^T) & 0 \\ 0 & (\Gamma^{-1})_4^T (D_4^T - D_2^T (D_1 + D_1^T)^{-1} D_3^T) \end{bmatrix}.
\end{aligned} \tag{4.16}
$$

Using the above equality, we can compute the expression $(N \circ (\Gamma^{-1})^T)^{-1}(N \circ (\Gamma^T)^{-1}D^T)$.

$$
\begin{aligned}
&(N \circ (\Gamma^{-1})^T)^{-1}(N \circ (\Gamma^T)^{-1}D^T) \\
&\overset{(4.16)}{=} \begin{bmatrix} ((\Gamma^{-1})_1^T)^{-1} & 0 \\ 0 & ((\Gamma^{-1})_4^T)^{-1} \end{bmatrix} \\
&\quad \begin{bmatrix} (\Gamma^{-1})_1^T (D_1^T - D_3^T (D_4 + D_4^T)^{-1} D_2^T) & 0 \\ 0 & (\Gamma^{-1})_4^T (D_4^T - D_2^T (D_1 + D_1^T)^{-1} D_3^T) \end{bmatrix} \\
&= \begin{bmatrix} (D_1^T - D_3^T (D_4 + D_4^T)^{-1} D_2^T) & 0 \\ 0 & (D_4^T - D_2^T (D_1 + D_1^T)^{-1} D_3^T) \end{bmatrix}.
\end{aligned} \tag{4.17}
$$

Lastly, we can simplify the term $D + (N \circ (\Gamma^{-1})^T)^{-1}(N \circ (\Gamma^T)^{-1} D^T)$.

$$
\begin{aligned}
& D + (N \circ (\Gamma^{-1})^T)^{-1}(N \circ (\Gamma^T)^{-1} D^T) \\
\overset{(4.17)}{=}\ & D + N \circ D^T - \begin{bmatrix} D_3^T (D_4 + D_4^T)^{-1} D_2^T & 0 \\ 0 & D_2^T (D_1 + D_1^T)^{-1} D_3^T \end{bmatrix} \\
=\ & \Gamma - A.
\end{aligned}
$$

Following equation (4.14), the equilibrium price at SPE can be written as

$$
p^*(c^*) = -(\Gamma - A)^{-1}\vec{1}.
$$

From proposition 4, we know that $p^*(c) = (\Gamma)^{-1}((N \circ D^T)c - \vec{1})$. Therefore, we can solve for contract parameter equilibrium c^*:

$$
(\Gamma)^{-1}((N \circ D^T)c^* - \vec{1}) = -(\Gamma - A)^{-1}\vec{1}.
$$

By rearranging, we obtain

$$
\begin{aligned}
(N \circ D)c^* &= \vec{1} - \Gamma(\Gamma - A)^{-1}\vec{1} \\
&= \vec{1} - (I - A\Gamma^{-1})^{-1}\vec{1} \\
&= \vec{1} - (I + A(\Gamma - A)^{-1})\vec{1} \\
&= -A(\Gamma - A)^{-1}\vec{1},
\end{aligned}
$$

where the second equality is acquired by moving Γ to the inside of the inverse term $(\Gamma - A)^{-1}$, and the third equality follows from the Lemma 14. Therefore, the additive parameter equilibrium c^* in SPE can be written as:

$$
c^* = -(N \circ D)^{-1} A(\Gamma - A)^{-1}\vec{1}.
$$

□

In Example 1, we have shown that in a competitive market where each manufacturer has one product, the additive parameters c^* in equilibrium is positive. We generalize this observation in the following proposition to provide sufficient conditions under which $c^* > 0$ in SPE, and consequently, the equilibrium price $p^*(c^*)$ is higher than the price $\tilde{p}^*$ of the direct sale game.

Definition 12. *A square matrix $U = (u_{ij})_{i,j\in[n]}$ is diagonally dominant if for every row i, the absolute value of the diagonal entry is greater than or equal to the sum of the absolute values of the other entries in that row:*

$$|u_{ii}| \geq \sum_{\substack{j=1 \\ j\neq i}}^{n} |u_{ij}|.$$

Proposition 6. *Under assumptions 4, 5, 8, 6, 7, and assume further*

1. *The price sensitivity matrix $D = (d_{ij})_{i,j\in I}$ satisfies:*
   - $d_{ii} < 0$ *for all* $i \in I$.
   - $d_{ij} > 0$ *for all* $i, j \in I, i \neq j$.
2. *The matrices D_1, D_4, and Γ as defined in (4.8) are diagonally dominant.*

Then, the equilibrium additive parameters in SPE is positive

$$c^* > 0.$$

Furthermore, the equilibrium price $p^*(c^*)$ in SPE is higher than the price $\tilde{p}^*$ of the direct sale game:

$$p^*(c^*) > \tilde{p}^*.$$

We first present a technical lemma showing that under the stated assumptions, matrices $(N \circ D^T)^{-1}$, A, and Γ^{-1} are elementwise negative.

Lemma 9. *Under the assumptions stated in Proposition 6, the following matrices are elementwise negative:*

- $\Gamma^{-1} < 0$,
- $D_3^T(D_4 + D_4^T)^{-1}D_2^T < 0$,
- $D_2^T(D_1 + D_1^T)^{-1}D_3^T < 0$,
- $(D_1^T)^{-1} < 0$,
- $(D_4^T)^{-1} < 0$.

Proof. Under Assumption 1 , matrix Γ is diagonally dominant. From assumption 2, matrix D satisfies $d_{ii} < 0$ for all $i \in I$ and $d_{ij} > 0$ for all $i, j \in I, i \neq j$. From equation (4.8), we have

$$\Gamma_{ii} = 2d_{ii} < 0,$$
$$\Gamma_{ij} = \begin{cases} d_{ij} + d_{ji} > 0 & \text{if } i, j \in I_k, \text{ for some } k \in \{1, 2\}, \\ d_{ij} > 0 & \text{if } i \in I_k, j \in I_{-k}, \text{ for some } k \in \{1, 2\}. \end{cases} \tag{4.18}$$

From lemma 15, we have

$$\Gamma^{-1} < 0.$$

Since Γ is diagonally dominant and $\Gamma_{ij} > 0$ for all $i, j \in I$, we have for any $k \in \{1, 2\}, i \in I_k$,

$$|\Gamma_{ii}| \geq \sum_{j \neq i} |\Gamma_{ij}| > \sum_{\substack{j \neq i \\ j \in I_k}} |\Gamma_{ij}|.$$

From equation (4.18), we have for any $k \in \{1, 2\}, i \in I_k$,

$$|2d_{ii}| \geq \sum_{\substack{j \neq i \\ j \in I_k}} |d_{ij} + d_{ji}|.$$

By definition 12, matrices $D_1 + D_1^T, D_4 + D_4^T$ are diagonally dominant. Hence, matrices $D_1, D_4, (D_1 + D_1^T), (D_4 + D_4^T)$ and Γ are all diagonally dominant, have negative diagonal entries and positive non-diagonal entries. From lemma 15, we have

$$D_1^{-1} < 0, \quad D_4^{-1} < 0, \quad (D_1 + D_1^T)^{-1} < 0, \quad (D_4 + D_4^T)^{-1} < 0.$$

Since also $D_2 > 0, D_3 > 0$, we have

$$D_3^T(D_4 + D_4^T)^{-1}D_2^T < 0, \quad \text{and } D_2^T(D_1 + D_1^T)^{-1}D_3^T < 0.$$

Also,

$$(D_1^T)^{-1} = (D_1^{-1})^T < 0, \quad (D_4^T)^{-1} = (D_4^{-1})^T < 0.$$

□

Proof. Let A be defined as in equation (4.9). From proposition 5 and assumption 8, the equilibrium price in SPE is

$$p^*(c^*) = -(\Gamma - A)^{-1}\vec{1} > 0,$$

and the additive parameter c^* in SPE can be written as:

$$\begin{aligned}
c^* &= -(N \circ D^T)^{-1}A(\Gamma - A)^{-1}\vec{1} \\
&= (N \circ D^T)^{-1}Ap^*(c^*) \\
&= \begin{bmatrix} (D_1^T)^{-1} & 0 \\ 0 & (D_4^T)^{-1} \end{bmatrix} \begin{bmatrix} D_3^T(D_4 + D_4^T)^{-1}D_2^T & 0 \\ 0 & D_2^T(D_1 + D_1^T)^{-1}D_3^T \end{bmatrix} p^*(c^*) \\
&= \begin{bmatrix} (D_1^T)^{-1}D_3^T(D_4 + D_4^T)^{-1}D_2^T & 0 \\ 0 & (D_4^T)^{-1}D_2^T(D_1 + D_1^T)^{-1}D_3^T \end{bmatrix} p^*(c^*)
\end{aligned}$$

From lemma 9, we have $D_3^T(D_4 + D_4^T)^{-1}D_2^T < 0$, $D_2^T(D_1 + D_1^T)^{-1}D_3^T < 0$, $(D_1^T)^{-1} < 0$,and $(D_4^T)^{-1} < 0$. From assumption 8, we know $p^*(c^*) > 0$, then contract parameter equilibrium satisfies

$$c^* = \begin{bmatrix} (D_1^T)^{-1}D_3^T(D_4 + D_4^T)^{-1}D_2^T & 0 \\ 0 & (D_4^T)^{-1}D_2^T(D_1 + D_1^T)^{-1}D_3^T \end{bmatrix} p^*(c^*) > 0.$$

From proposition 4, we know that the price equilibrium $\tilde{p}^*$ of the direct sale game is

$$\tilde{p}^* = -\Gamma^{-1}\vec{1}.$$

From proposition 5 and assumption 8, the equilibrium price $p^*(c^*)$ in SPE is:

$$p^*(c^*) = -(\Gamma - A)^{-1}\vec{1} > 0.$$

We can expand the term $(\Gamma - A)^{-1}$ in $p^*(c^*)$ using Lemma 14 (in the appendix):

$$\begin{aligned} p^*(c^*) &= -(\Gamma - A)^{-1}\vec{1} \\ &= -\left(\Gamma^{-1} + \Gamma^{-1}A(I^{-1} - \Gamma^{-1}A)^{-1}\Gamma^{-1}\right)\vec{1} \\ &= -\Gamma^{-1}\vec{1} - \Gamma^{-1}A(\Gamma - A)^{-1}\vec{1} \\ &= \tilde{p}^* + \Gamma^{-1}Ap^*(c^*). \end{aligned}$$

From lemma 9, we have $\Gamma^{-1} < 0$, $D_3^T(D_4 + D_4^T)^{-1}D_2^T < 0$, and $D_2^T(D_1 + D_1^T)^{-1}D_3^T < 0$. From assumption 8, we have $p^*(c^*) > 0$. Hence, $\Gamma^{-1}Ap^*(c^*) > 0$. Therefore,

$$p^*(c^*) = \tilde{p}^* + \Gamma^{-1}Ap^*(c^*) > \tilde{p}^*.$$

□

Proposition 6 provides a condition where the equilibrium contract parameters are all positive at SPE. The following example demonstrates a scenario where the conditions of Proposition 6 are not satisfied, resulting in negative contract parameters at SPE. When c_i^* is negative, the corresponding manufacturer subsidizes product i and incentivizes their retailer to set a lower, more competitive price for product i.

Example 2 Suppose there are three products in the market: $I_1 = \{1,2\}, I_2 = \{3\}$. Let $d \in (-0.5, 0.5)$. Consider the demand sensitivity matrix:

$$D = \begin{bmatrix} -1 & d & 0.5 \\ d & -1 & 0 \\ 0.5 & 0 & -1 \end{bmatrix}.$$

At SPE, the additive parameter $c_1^* > 0$ and $c_3^* > 0$ for all $d \in (-0.5, 0.5)$. Also, $c_2^* > 0$ if $d \in (0.5, 0)$, and $c_2^* < 0$ if $d \in (0, 0.5)$. We can compute the SPE using proposition 5. For the sake of a continuous flow in this example, we postpone the verification of assumptions to the end of this example. From equation (4.8) and (4.9), we have

$$\Gamma = \begin{bmatrix} -2 & 2d & 0.5 \\ 2d & -2 & 0 \\ 0.5 & 0 & -2 \end{bmatrix}, \quad A = \begin{bmatrix} D_3^T(D_4 + D_4^T)^{-1}D_2^T & 0 \\ 0 & D_2^T(D_1 + D_1^T)^{-1}D_3^T \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{1}{8} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{8}\frac{1}{1-d^2} \end{bmatrix}.$$

From equation (4.12), the equilibrium price $p^*(c^*)$ can be computed as:

$$
\begin{aligned}
p^*(c^*) &= -(\Gamma - A)^{-1}\vec{1} \\
&= -\begin{bmatrix} -\frac{15}{8} & 2d & 0.5 \\ 2d & -2 & 0 \\ 0.5 & 0 & -2+\frac{1}{8}\frac{1}{1-d^2} \end{bmatrix}^{-1} \vec{1} \\
&= -\frac{1}{-7+8d^2+\frac{15}{32}\frac{1}{1-d^2}-\frac{d^2}{2(1-d^2)}} \\
&\quad \cdot \begin{bmatrix} 4-\frac{1}{4}\frac{1}{1-d^2} & -2d\left(-2+\frac{1}{8}\frac{1}{1-d^2}\right) & 1 \\ -2d\left(-2+\frac{1}{8}\frac{1}{1-d^2}\right) & \frac{7}{2}-\frac{15}{64}\frac{1}{1-d^2} & d \\ 1 & d & \frac{15}{4}-4d^2 \end{bmatrix} \vec{1} \\
&= -\frac{1}{-7+8d^2+\frac{15}{32}\frac{1}{1-d^2}-\frac{d^2}{2(1-d^2)}} \begin{bmatrix} 5-\frac{1}{4}\frac{1+d}{1-d^2}+4d \\ 5d-\frac{1}{4}\frac{d+\frac{15}{16}}{1-d^2}+\frac{7}{2} \\ d+\frac{19}{4}-4d^2 \end{bmatrix}. \qquad (4.19)
\end{aligned}
$$

From equation (4.11), additive parameter c^* at SPE can be computed as:

$$
\begin{aligned}
c^* &= \begin{bmatrix} D_1^{-1} & 0 \\ 0 & D_4^{-1} \end{bmatrix} A p^*(c^*) \\
&= \begin{bmatrix} -\frac{1}{1-d^2} & -\frac{d}{1-d^2} & 0 \\ -\frac{d}{1-d^2} & -\frac{1}{1-d^2} & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} -\frac{1}{8} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{8}\frac{1}{1-d^2} \end{bmatrix} p^*(c^*) \\
&= \begin{bmatrix} \frac{1}{8}\frac{1}{1-d^2} & 0 & 0 \\ \frac{1}{8}\frac{d}{1-d^2} & 0 & 0 \\ 0 & 0 & \frac{1}{8}\frac{1}{1-d^2} \end{bmatrix} p^*(c^*) \\
&= \frac{1}{8}\frac{1}{1-d^2} \begin{bmatrix} p_1^*(c^*) \\ d \cdot p_1^*(c^*) \\ p_3^*(c^*) \end{bmatrix}.
\end{aligned}
$$

For all $d \in (-0.5, 0.5)$, we have $p^*(c^*) > 0$ and $1 - d^2 > 0$. Hence, $c_1^* > 0$ and $c_3^* > 0$ for all $d \in (-0.5, 0.5)$. Furthermore, we have $c_2^* > 0$ if $d \in (0.5, 0)$, and $c_2^* < 0$ if $d \in (0, 0.5)$.

Now we check that all the assumptions are satisfied.

Assumption 4, 5: Since $I_1 = \{1, 2\}$, we have

$$D_{I_1} = \begin{bmatrix} -1 & d \\ d & -1 \end{bmatrix}, \quad D_{I_1} + D_{I_1}^T = \begin{bmatrix} -2 & 2d \\ 2d & -2 \end{bmatrix}.$$

Both matrices are invertible as their determinant equals $1 - d^2$, which is strictly positive for all $d \in (-0.5, 0.5)$. For $I_2 = \{3\}$, the submatrix $D_{I_2} = [-1]$ is trivially invertible. Hence, all matrices D_{I_k} and $D_{I_k} + D_{I_k}^T$ are invertible. The determinant of D is $-1 + d^2$, which is nonzero for $d \in (-0.5, 0.5)$, thus D is invertible. The determinant of Γ is $-8(1 - d^2)$, which is strictly negative for all $d \in (-0.5, 0.5)$. Therefore, Γ is invertible. The determinant of matrix $\Gamma - A$ is $-7 + 8d^2 + \frac{15}{32}\frac{1}{1-d^2} - \frac{d^2}{2(1-d^2)}$ which is nonzero for $d \in (-0.5, 0.5)$, thus $\Gamma - A$ is invertible.

Assumption 8: It is easy to verify that for $d \in (-0.5, 0.5)$, we have $p^*(c^*)$ as computed in equation (4.19) is positive and $\pi^*(p^*) = 1 + Dp^*(c^*)$ is positive.

Assumption 6: The matrices $D_1 = \begin{bmatrix} -1 & d \\ d & -1 \end{bmatrix}$ and $D_4 = \begin{bmatrix} -1 \end{bmatrix}$ are negative definite for $d \in (-0.5, 0.5)$.

Assumption 7: For $k = 1$, the matrix $((\Gamma^{-1})^T D^T)_{I_k} (\Gamma^{-1})_{I_k} + (\Gamma^{-1})_{I_k}^T (D(\Gamma^{-1}))_{I_k}$

can be written as

$$
\begin{aligned}
&((\Gamma^{-1})^T D^T)_{I_1}(\Gamma^{-1})_{I_1} + (\Gamma^{-1})_{I_1}^T (D(\Gamma^{-1}))_{I_1} \\
=&\frac{1}{(8d^2-7.5)^2}\begin{bmatrix} 4d^2-3.5 & 0 \\ 0.25d & 4d^2-3.75 \end{bmatrix}\begin{bmatrix} 4 & 4d \\ 4d & 3.75 \end{bmatrix} \\
&+\frac{1}{(8d^2-7.5)^2}\begin{bmatrix} 4 & 4d \\ 4d & 3.75 \end{bmatrix}\begin{bmatrix} 4d^2-3.5 & 0.25d \\ 0 & 4d^2-3.75 \end{bmatrix} \\
=&\frac{1}{(8d^2-7.5)^2}\left(\begin{bmatrix} 4(4d^2-3.5) & 4d(4d^2-3.5) \\ d+4d(4d^2-3.75) & d^2+3.75(4d^2-3.75) \end{bmatrix}\right. \\
&\left.+\begin{bmatrix} 4(4d^2-3.5) & d+4d(4d^2-3.75) \\ 4d(4d^2-3.5) & d^2+3.75(4d^2-3.75) \end{bmatrix}\right) \\
=&\frac{1}{(8d^2-7.5)^2} \\
&\cdot\begin{bmatrix} 2*4(4d^2-3.5) & 4d(8d^2-3.5-3.75)+d \\ 4d(8d^2-3.5-3.75)+d+4d(4d^2-3.75) & 2d^2+2*3.75(4d^2-3.75) \end{bmatrix}.
\end{aligned}
$$

We can check that for all $d \in (-0.5, 0.5)$, it is negative definite. For $k = 2$, the matrix $((\Gamma^{-1})^T D^T)_{I_k}(\Gamma^{-1})_{I_k} + (\Gamma^{-1})_{I_k}^T (D(\Gamma^{-1}))_{I_k}$ can be written as:

$$((\Gamma^{-1})^T D^T)_{I_2}(\Gamma^{-1})_{I_2} + (\Gamma^{-1})_{I_2}^T (D(\Gamma^{-1}))_{I_2} = \frac{8}{(8d^2-7.5)^2}(1-d^2)(4d^2-3.5).$$

We can check that it is negative for all $d \in (-0.5, 0.5)$.

In the above example, if product 2 is complementary to product 1, the first manufacturer has an incentive to subsidize product 2. By setting a negative contract parameter $c_2^* < 0$, the retailer is incentivized to reduce the price of product 2, which will boost sales of both product 1 and 2.

Although Proposition 6 is proved under a diagonal-dominance assumption, we believe the same positivity and price-comparison results hold even when

D_1, D_4, and Γ are not diagonally dominant; extensive numerical searches have uncovered no counter-examples, indicating that the assumption is likely superfluous.

Conjecture 1. *Under Assumptions 4, 5, 8, 6, and 7, and assuming the price-sensitivity matrix $D = (d_{ij})_{i,j \in I}$ satisfies*

- $d_{ii} < 0$ *for all* $i \in I$,
- $d_{ij} > 0$ *for all* $i \neq j$,

then the unique SPE additive parameters satisfy

$$c^* \;>\; 0,$$

and the corresponding equilibrium price exceeds that of the direct sale game,

$$p^*(c^*) \;>\; \widetilde{p}^*.$$

4.5 Concluding Remarks

In this paper, we study strategic pricing in a two-tier supply chain where two competing manufacturers sell multiple products through their dedicated retailers. The products may be substitutes or complements, and market demand of each product is linear in product prices. In the first stage, manufacturers design linear contracts, composed of additive and proportional parameters; in the second stage, retailers set product prices to maximize their utilities.

We first analyze the retailer game and study how equilibrium prices change with additive parameters in linear contract: when the additive parameter is positive, manufacturers incentivize the retailers to set higher prices, hence reduce competition. When the additive parameter is negative, manufacturers subsidize the retailers to set lower prices, hence intensify the competition. We show that the equilibrium contract parameters and product prices are unique, and we provide a complete characterization of the subgame perfect equilibrium. We compare the equilibrium prices of the two-stage game to a benchmark direct-sale game, where manufacturers sell directly to consumers without retailers. We show that when products are substitutes, manufacturers set positive additive parameters for all products, and incentivize retailers to set higher prices. We also provide an example in which, at equilibrium, the manufacturer offers negative additive parameters—effectively subsidizing the retailer—to incentivize a lower price for a product that complements another. This subsidy boosts the sales of both the subsidized product and its complement, thereby increasing the manufacturer's total utility. One future direction is to systematically characterize the conditions under which manufacturers strategically choose to subsidize their retailers in equilibrium. Also, we aim to establish a rigorous proof of Conjecture 1, thereby confirming that the positive additive parameter results hold without the diagonal-dominance assumption.

APPENDIX A

CHAPTER 2 APPENDIX

A.1 Proof of Proposition 1

Consider any $i \in I, b_i \in \Delta^M_{-i}, \pi_{-i} \in \Delta_{-i}$ and any T observations $\{a^t_{-i}\}^T_{t=1}$ sampled from π_{-i}. We have

$$\mathbb{E}[\hat{\pi}_{-i}(a_{-i})] = \pi_{-i}(a_{-i}), \quad \forall a_{-i} \in A_{-i}.$$

Since each indicator $\mathbf{1}\{a^t_{-i} = a_{-i}\} \in \{0, 1\}$, by Hoeffding's inequality, for any $\delta > 0$,

$$\begin{aligned}
&P\Big(\Big|\hat{\pi}_{-i}(a_{-i}) - \pi_{-i}(a_{-i})\Big| \geq \delta\Big) \\
&= P\Big(\Big|\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}\{a^t_{-i} = a_{-i}\} - \frac{1}{T}\,\mathbb{E}\Big[\sum_{t=1}^{T}\mathbf{1}\{a^t_{-i} = a_{-i}\}\Big]\Big| \geq \delta\Big) \\
&= P\Big(\Big|\sum_{t=1}^{T}\mathbf{1}\{a^t_{-i} = a_{-i}\} - \mathbb{E}\Big[\sum_{t=1}^{T}\mathbf{1}\{a^t_{-i} = a_{-i}\}\Big]\Big| \geq \delta \cdot T\Big) \\
&\leq 2\exp\Big(-\frac{2\delta^2 \cdot T^2}{T}\Big) \\
&= 2\exp\Big(-2T \cdot \delta^2\Big).
\end{aligned}$$

By Cauchy–Schwarz,

$$\begin{aligned}
\|\hat{\pi}_{-i} - \pi_{-i}\|_2 &= \sqrt{\sum_{a_{-i} \in A_{-i}} (\hat{\pi}_{-i}(a_{-i}) - \pi_{-i}(a_{-i}))^2} \\
&\leq \sqrt{|A_{-i}|} \cdot \max_{a_{-i} \in A_{-i}} |\hat{\pi}_{-i}(a_{-i}) - \pi_{-i}(a_{-i})|.
\end{aligned}$$

Hence, for any $\delta > 0$,

$$\begin{aligned}
\Pr\left(\|\hat{\pi} - \pi\|_2 \geq \delta\right) &\leq \Pr\Big(\max_{a_{-i} \in A_{-i}} |\hat{\pi}_{-i}(a_{-i}) - \pi_{-i}(a_{-i})| \geq \tfrac{\delta}{\sqrt{|A_{-i}|}}\Big) \\
&\leq 2\exp\big(-2T\,(\tfrac{\delta}{\sqrt{|A_{-i}|}})^2\big) \\
&= 2\exp\left(-\frac{2T}{|A_{-i}|} \cdot \delta^2\right). \qquad \text{(A.1)}
\end{aligned}$$

Type-I error analysis. Suppose that $\|\pi_{-i} - b_i\| \leq \tau$. Let

$$\delta_T = \sqrt{\frac{|A_{-i}|}{2T} \cdot \ln\left(\tfrac{2}{\bar{\alpha}}\right)}.$$

If H_0 is rejected:

$$\|\hat{\pi}_{-i} - b_i\|_2 > \tau + \delta_T,$$

then from the triangle inequality, we have

$$\|\hat{\pi}_{-i} - \pi_{-i}\|_2 \geq \|\hat{\pi}_{-i} - b_i\|_2 - \|b_i - \pi_{-i}\|_2 > (\tau + \delta_T) - \tau = \delta_T.$$

Therefore,

$$\begin{aligned}
\Pr\left(\|\hat{\pi}_{-i} - b_i\|_2 > \tau + \delta_T\right) &\leq \Pr\left(\|\hat{\pi}_{-i} - \pi_{-i}\|_2 > \delta_T\right) \\
&\overset{\text{(A.1)}}{\leq} 2\exp\left(-\frac{2T}{|A_{-i}|} \cdot \left(\delta_T\right)^2\right) \\
&= \bar{\alpha}.
\end{aligned}$$

Type-II error analysis Suppose that $d := \|\pi_{-i} - b_i\|_2 > \tau$. When player i fails to reject H_0, we have

$$\|\hat{\pi}_{-i} - b_i\|_2 \leq \tau + \delta_T.$$

From the triangle inequality, we have

$$\|\hat{\pi}_{-i} - \pi_{-i}\|_2 \geq \|\pi_{-i} - b_i\|_2 - \|\hat{\pi}_{-i} - b_i\|_2 \geq d - (\tau + \delta_T).$$

Therefore,

$$\Pr\Big(\|\hat{\pi}_{-i} - \pi_{-i}\|_2 \le \tau + \delta_T\Big) \le \Pr\Big(\|\hat{\pi}_{-i} - \pi_{-i}\|_2 \ge d - (\tau + \delta_T)\Big) \overset{(A.1)}{\le} 2\exp\left(-\frac{2T}{|A_{-i}|}\cdot(d - \tau - \delta_T)^2\right).$$

Suppose

$$T(\bar{\alpha}) \ge \max_{\substack{i \in I,\, b_i \in \mathcal{B}_i, \\ \pi_{-i} \in \mathrm{Im}(\mathrm{Br}_i^\sigma), \\ \|\pi_{-i} - b_i\|_2 > \tau}} \frac{2|A_{-i}|}{(\|\pi_{-i} - b_i\|_2 - \tau)^2} \ln\left(\frac{2}{\bar{\alpha}}\right) \ge \frac{2|A_{-i}|}{(d-\tau)^2} \ln\left(\frac{2}{\bar{\alpha}}\right). \tag{A.2}$$

Then,

$$\delta_{T(\bar{\alpha})} = \sqrt{\frac{|A_{-i}|}{2T(\bar{\alpha})}\cdot \ln\left(\tfrac{2}{\bar{\alpha}}\right)} \le \frac{d - \tau}{2}. \tag{A.3}$$

Since function $\exp(-x)$ is decreasing in x,

$$\Pr\Big(\|\hat{\pi}_{-i} - \pi_{-i}\|_2 \le \tau + \delta_T\Big) \le 2\exp\left(-\frac{2T(\bar{\alpha})}{|A_{-i}|}\cdot\Big(d - \tau - \delta_{T(\bar{\alpha})}\Big)^2\right) \overset{(A.3)}{\le} 2\exp\left(-\frac{2T(\bar{\alpha})}{|A_{-i}|}\cdot\left(\frac{d-\tau}{2}\right)^2\right) \overset{(A.2)}{\le} \bar{\alpha}.$$

□

A.2 Proof of Proposition 2

Lemma 10 (Gao 2018). *For any $\sigma > 0$, $i \in I$, the smooth response function with temperature parameter σ is Lipschitz in $U_i(\cdot, b_i) \in \mathbb{R}^{|A_i|}$ with Lipschitz parameter $1/\sigma$. That is, for any $\sigma > 0$, $i \in I$, $b_i, b_i' \in \Delta_{-i}$,*

$$\|\mathrm{Br}_i^\sigma(b_i) - \mathrm{Br}_i^\sigma(b_i')\|_2 \le \frac{1}{\sigma}\|U_i(\cdot, b_i) - U_i(\cdot, b_i')\|_2. \tag{A.4}$$

Proof. First, for any $\sigma > 0$, define $\pi^* \in \Delta$ as a fixed point of the smooth best response mapping. That is,

$$\pi_i^*(a_i) = \frac{\exp\left(\frac{1}{\sigma} U_i(a_i, \pi_{-i}^*)\right)}{\sum_{a_i' \in A_i} \exp\left(\frac{1}{\sigma} U_i(a_i', \pi_{-i}^*)\right)}, \quad \forall a_i \in A_i,\ \forall i \in I. \tag{A.5}$$

Such a fixed point π^* exists because the smooth best response function on the right-hand side of (A.5) is continuous in π, and the joint strategy space Δ is a nonempty, compact, and convex subset of a Euclidean space. Existence then follows from Brouwer's fixed point theorem.

We now consider a belief vector in the discretized belief space $b^\dagger = (b_i^\dagger)_{i \in I} \in \mathcal{B} = \Delta^M$ such that each $b_i^\dagger$ has the minimum ℓ_2 distance with the strategy profile π_{-i}^*. We also define the strategy profile $\pi^\dagger = \text{Br}^\sigma(b^\dagger)$.

Part 1: We show that $(b^\dagger, \pi^\dagger)$ satisfies the following conditions under Assumption 1:

$$\|\pi_{-i}^* - b_i^\dagger\|_2 \leq \tau, \quad \forall i \in I, \tag{A.6a}$$

$$\|\text{Br}_i^\sigma(b_i^*) - \text{Br}_i^\sigma(b_i^\dagger)\|_2 \leq \tau, \quad \forall i \in I, \tag{A.6b}$$

$$\|b_i^\dagger - \pi_{-i}^\dagger\|_2 \leq \tau, \quad \forall i \in I. \tag{A.6c}$$

In particular, (A.6c) concludes that $(b^\dagger, \pi^\dagger)$ is τ-consistent.

Proof of Part 1: To prove (A.6a), we note that since $b^\dagger$ is the belief in the discretized set $\mathcal{B}$ that is closest to π^*, we must have

$$|\pi_j^*(a_j) - b_{ij}(a_j)| \leq \frac{1}{M}, \quad \forall a_j \in A_j, \quad \forall i, j \in I.$$

Hence, the ℓ_1 distance between b_{ij} and π_j^* can be bounded as follows:

$$\|\pi_j^* - b_{ij}\|_1 = \sum_{a_j \in A_j} |\pi_j^*(a_j) - b_{ij}(a_j)| \leq \sum_{a_j \in A_j} \frac{1}{M} \leq \frac{\max_{i' \in I} |A_{i'}|}{M}, \quad \forall i, j \in I. \tag{A.7}$$

Moreover,

$$\|\pi_{-i}^* - b_i\|_2 \leq \|\pi_{-i}^* - b_i\|_1 = \sum_{j \neq i} \|\pi_j^* - b_{ij}\|_1 \overset{(A.7)}{\leq} \frac{\max_{i' \in I} |A_{i'}| \cdot |I|}{M}$$
$$\overset{(2.4b)}{\leq} \frac{\tau}{\left(1 + \sqrt{\max_{i' \in I} |A_{i'}|} \cdot \max_{i' \in I, a \in A} u_{i'}(a) \cdot \max_{i' \in I}\{|A_{i'}| \cdot |A_{-i'}|\} \cdot |I|/\sigma\right)}$$
$$\leq \tau.$$

To prove (A.6b), we first show that

$$\begin{aligned}
&\|U_i(\cdot, b_i^*) - U_i(\cdot, b_i^\dagger)\|_2 \\
&= \left(\sum_{a_i \in A_i} \left[\sum_{a_{-i} \in A_{-i}} (b_i^*(a_{-i}) - b_i^\dagger(a_{-i})) u_i(a_i, a_{-i}) \right]^2 \right)^{1/2} \\
&\leq \left(\sum_{a_i \in A_i} \left[\|b_i^* - b_i^\dagger\|_2 \cdot \|u_i(a_i, \cdot)\|_2 \right]^2 \right)^{1/2} \qquad \text{(Cauchy–Schwarz)} \\
&= \|b_i^* - b_i^\dagger\|_2 \cdot \left(\sum_{a_i \in A_i} \|u_i(a_i, \cdot)\|_2^2 \right)^{1/2} \\
&\leq \|b_i^* - b_i^\dagger\|_2 \cdot \left(\sum_{a_i \in A_i} \sum_{a_{-i} \in A_{-i}} u_i(a_i, a_{-i})^2 \right)^{1/2} \\
&\leq \|b_i^* - b_i^\dagger\|_2 \cdot \left(|A_i| \cdot |A_{-i}| \cdot \max_{i' \in I, a \in A} u_{i'}(a)^2 \right)^{1/2} \\
&= \max_{i' \in I, a \in A} u_{i'}(a) \cdot \sqrt{|A_i| \cdot |A_{-i}|} \cdot \|b_i^* - b_i^\dagger\|_2 \\
&\leq \max_{i' \in I, a \in A} u_{i'}(a) \cdot \max_{i' \in I}\{|A_{i'}| \cdot |A_{-i'}|\} \cdot \|b_i^* - b_i^\dagger\|_2, \quad \forall i \in I.
\end{aligned} \tag{A.8}$$

Then, following Lemma 10, we have

$$\begin{aligned}
&\|\mathrm{Br}_i^\sigma(b_i^*) - \mathrm{Br}_i^\sigma(b_i^\dagger)\|_2 \\
\overset{(A.4)}{\leq}\ & \frac{1}{\sigma}\|U_i(\cdot, b^*) - U_i(\cdot, b^\dagger)\|_2 \\
\overset{(A.8)}{\leq}\ & \frac{\max_{i'\in I, a\in A} u_{i'}(a) \cdot \max_{i'\in I}\{|A_{i'}|\cdot|A_{-i'}|\}}{\sigma}\|b_i^* - b_i^\dagger\|_2.
\end{aligned} \tag{A.9}$$

We next show that (A.6c) holds:

$$\begin{aligned}
&\|b_i^\dagger - \pi_{-i}^\dagger\|_2 \leq \|b_i^\dagger - \pi_{-i}^\dagger\|_1 \leq \sum_{j\neq i}\|b_{ij}^\dagger - \pi_j^\dagger\|_1 \\
&\leq \sum_{j\neq i}\|b_{ij}^\dagger - \pi_j^* + \pi_j^* - \mathrm{Br}_j^\sigma(b_j^*) + \mathrm{Br}_j^\sigma(b_j^*) - \pi_j^\dagger\|_1 \\
&\leq \sum_{j\neq i}\left(\|b_{ij}^\dagger - \pi_j^*\|_1 + \|\pi_j^* - \mathrm{Br}_j^\sigma(b_j^*)\|_1 + \|\mathrm{Br}_j^\sigma(b_j^*) - \pi_j^\dagger\|_1\right) \\
&\leq \sum_{j\neq i}\left(\|b_{ij}^\dagger - \pi_j^*\|_1 + \|\mathrm{Br}_j^\sigma(b_j^*) - \pi_j^\dagger\|_1\right) \\
&\leq \sum_{j\neq i}\left(\|b_{ij}^\dagger - \pi_j^*\|_1 + \sqrt{|A_j|}\|\mathrm{Br}_j^\sigma(b_j^*) - \mathrm{Br}_j^\sigma(b_j^\dagger)\|_2\right) \\
&\underset{(A.9)}{\overset{(A.7)}{\leq}} \sum_{j\neq i}\left(\frac{\max_{i'\in I}|A_{i'}|}{M} + \right. \\
&\qquad \left.\sqrt{\max_{i'\in I}|A_{i'}|}\cdot\frac{\max_{i'\in I, a\in A} u_{i'}(a)\cdot\max_{i'\in I}\{|A_{i'}|\cdot|A_{-i'}|\}}{\sigma}\cdot\frac{\max_{i'\in I}|A_{i'}|\cdot|I|}{M}\right) \\
&\leq |I|\cdot\frac{\max_{i'\in I}|A_{i'}|}{M}\cdot \\
&\qquad \left(1 + \sqrt{\max_{i'\in I}|A_{i'}|}\cdot\frac{\max_{i'\in I, a\in A} u_{i'}(a)\cdot\max_{i'\in I}\{|A_{i'}|\cdot|A_{-i'}|\}}{\sigma}\cdot|I|\right) \\
&\overset{(2.4b)}{\leq}\ \tau.
\end{aligned}$$

Part 2: We further show that $\pi^\dagger$ is an ϵ-Nash equilibrium.

Proof of Part 2: For any $i \in I$, we note that

$$\begin{aligned}
\|\pi_i^\dagger - \mathrm{Br}_i^\sigma(\pi_{-i}^\dagger)\|_2 &= \|\mathrm{Br}_i^\sigma(b_i^\dagger) - \mathrm{Br}_i^\sigma(\pi_{-i}^\dagger)\|_2 \\
&\overset{(A.9)}{\leq} \frac{\max_{i'\in I, a\in A} u_{i'}(a) \cdot \max_{i'\in I}\{|A_{i'}| \cdot |A_{-i'}|\}}{\sigma} \cdot \|b_i^\dagger - \pi_{-i}^\dagger\|_2 \\
&\overset{(A.6c)}{\leq} \frac{\max_{i'\in I, a\in A} u_{i'}(a) \cdot \max_{i'\in I}\{|A_{i'}| \cdot |A_{-i'}|\} \cdot \tau}{\sigma} \\
&\overset{(2.4a)}{\leq} \frac{\epsilon}{2\sqrt{|A|} \cdot \max_{i'\in I, a\in A} u_{i'}(a)}.
\end{aligned} \tag{A.10}$$

We use the triangle inequality to bound the difference in expected utilities: for any $i \in I$,

$$\begin{aligned}
&|U_i(\pi_i^\dagger, \pi_{-i}^\dagger) - \max_{\pi_i' \in \Delta_i} U_i(\pi_i', \pi_{-i}^\dagger)| \\
\leq &|U_i(\pi_i^\dagger, \pi_{-i}^\dagger) - U_i(\mathrm{Br}_i^\sigma(\pi_{-i}^\dagger), \pi_{-i}^\dagger)| + |U_i(\mathrm{Br}_i^\sigma(\pi_{-i}^\dagger), \pi_{-i}^\dagger) - \max_{\pi_i' \in \Delta_i} U_i(\pi_i', \pi_{-i}^\dagger)|.
\end{aligned} \tag{A.11}$$

For any $i \in I$, $\pi, \pi' \in \Delta_i$, we note that

$$\begin{aligned}
\Big|U_i(\pi) - U_i(\pi')\Big| &= \left|\sum_{a\in A} (\pi(a) - \pi'(a)) u_i(a)\right| \\
&\leq \|\pi - \pi'\|_2 \cdot \|u_i\|_2 \\
&\leq \|\pi - \pi'\|_2 \cdot \sqrt{|A|} \max_{i\in I, a\in A} u_i(a) \\
&= \sqrt{|A|} \cdot \max_{i'\in I, a\in A} u_{i'}(a) \cdot \|\pi - \pi'\|_2.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\Big|U_i(\pi_i^\dagger, \pi_{-i}^\dagger) - U_i(\mathrm{Br}_i^\sigma(\pi_{-i}^\dagger), \pi_{-i}^\dagger)\Big| \\
\leq &\sqrt{|A|} \cdot \max_{i'\in I, a\in A} u_{i'}(a) \cdot \left\|\pi^\dagger - \left(\mathrm{Br}_i^\sigma(\pi_{-i}^\dagger), \pi_{-i}^\dagger\right)\right\|_2.
\end{aligned} \tag{A.12}$$

Following from (A.12) and (A.10), we have

$$|U_i(\pi_i^\dagger, \pi_{-i}^\dagger) - U_i(\mathrm{Br}_i^\sigma(\pi_{-i}^\dagger), \pi_{-i}^\dagger)| \leq \frac{\epsilon}{2}. \tag{A.13}$$

Additionally, we know that

$$\mathrm{Br}_i^\sigma(b_i^\dagger) = \arg\max_{\pi_i \in \Delta_i} U_i(\pi_i, b_i) + \sigma H(\pi)_i,$$

where $H(\pi_i) = -\sum_{a_i} \pi_i(a_i) \log \pi_i(a_i)$ is the entropy function of π_i. Therefore,

$$U_i(\mathrm{Br}_i^\sigma(b_i^\dagger), b_i^\dagger) + \sigma H(\mathrm{Br}_i^\sigma(b_i^\dagger)) \geq U_i(\pi_i^*, b_i^\dagger) + \sigma H(\pi_i^*).$$

Since $\mathrm{Br}_i^\sigma(b_i) \in \Delta_i$ is a distribution over $|A_i|$ elements, its entropy satisfies $H(\mathrm{Br}_i^\sigma(b_i)) \leq \log(|A_i|)$ [25]. Since entropy is non-negative for any distribution, we have $H(\pi_i^*) \geq 0$. Therefore, we have

$$\begin{aligned} U_i(\mathrm{Br}_i^\sigma(b_i^\dagger), b_i^\dagger) + \sigma \cdot \log(|A_i|) &\geq U_i(\mathrm{Br}_i^\sigma(b_i^\dagger), b_i^\dagger) + \sigma \cdot H(\mathrm{Br}_i^\sigma(b_i^\dagger)) \\ &\geq U_i(\pi_i^*, b_i^\dagger) + \sigma H(\pi_i^*) \geq \max_{\pi_i \in \Delta_i} U_i(\pi_i, b_i^\dagger). \end{aligned}$$

Rearranging, we have

$$|\max_{\pi_i \in \Delta_i} U_i(\pi_i, b_i^\dagger) - U_i(\mathrm{Br}_i^\sigma(b_i^\dagger), b_i^\dagger)| \leq \sigma \cdot \log(|A_i|) \leq \sigma \cdot \log(\max_{i \in I} |A_i|) \overset{(2.4a)}{\leq} \frac{\epsilon}{2}. \tag{A.14}$$

Combining (A.11), (A.13) and (A.14), we have

$$|U_i(\pi_i^\dagger, \pi_{-i}^\dagger) - \max_{\pi_i' \in \Delta_i} U_i(\pi_i', \pi_{-i}^\dagger)| \leq \epsilon,$$

which concludes that $\pi^\dagger$ is an ϵ-Nash equilibrium. □

A.3 Proof of Lemma 1

Proof. Recall that the state space $\mathcal{Z}$ is

$$\mathcal{Z} := \{(b, \pi) \mid b_i \in \mathcal{B}_i,\ \pi_i = \mathrm{Br}_i^\sigma(b_i), \quad \forall i \in I\}.$$

Since $\mathcal{B}_i = \Delta_i^M$ is finite for all $i \in I$, the state space $\mathcal{Z}$ is finite. For any $\xi \in (0,1)$, $z = (b, \pi), z' = (b', \pi') \in \mathcal{Z}$, the transition probability satisfies

$$
\begin{aligned}
& P^{\xi}_{zz'} \\
\geq\ & \Pr\Big(\text{all players test, fail to reject, or reject and explore,} \\
& \qquad \text{then update to belief specified in } b'\Big) \\
=\ & \Pr\Big(\text{all players test}\Big) \cdot \Big(\sum_{I_r \subseteq I} \Pr(I_r \text{ reject}, I \setminus I_r \text{ fail to reject }) \cdot \Pr(I \setminus I_r \text{ explore})\Big) \\
& \cdot \Pr(\text{all players update belief to } b') \\
=\ & \Pr\Big(\text{all players test}\Big) \cdot \Big(\sum_{I_r \subseteq I} \Pr(I_r \text{ reject}, I \setminus I_r \text{ fail to reject }) \cdot \prod_{i \in I \setminus I_r} \xi^{f_i(U_i(\pi_i, b_i))}\Big) \\
& \cdot \Pr(\text{all players update belief to } b') \\
\geq\ & \Pr\Big(\text{all players test}\Big) \cdot \Big(\sum_{I_r \subseteq I} \Pr(I_r \text{ reject}, I \setminus I_r \text{ fail to reject })\Big) \\
& \cdot \min_{I_r \subset I} \prod_{i \in I \setminus I_r} \xi^{f_i(U_i(\pi_i, b_i))} \cdot \Pr\Big(\text{all update belief to } b'\Big) \\
=\ & \gamma_i^{|I|} \cdot 1 \cdot \min_{I_r \subset I} \prod_{i \in I \setminus I_r} \xi^{f_i(U_i(\pi_i, b_i))} \cdot \prod_{i \in I} \frac{1}{|\mathcal{B}_i|} > 0.
\end{aligned}
$$

Since transition $P^{\xi}_{zz'}$ is positive for any $z, z' \in Z$, the Markov chain P^{ξ} is aperiodic and irreducible, and the Markov process has a unique stationary distribution.

□

A.4 Sufficient Condition for Assumption 2

Lemma 11. *Suppose*

$$|\text{Im}(\text{Br}_i^{\sigma})| > |I|, \quad \forall i \in I,$$

and

$$\tau < \min_{b_{-i}\in\mathcal{B}_{-i}, i\in I} \max_{\tilde{b}_i\in\mathcal{B}_i} \min\left\{\|\tilde{b}_i - \mathrm{Br}^{\sigma}_{-i}(b_{-i})\|_2,\ \min_{j\in I\setminus\{i\}}\{\|\mathrm{Br}^{\sigma}_i(\tilde{b}_i) - b_{ji}\|_2\}\right\}.$$

Then, Assumption 2 holds.

Proof. We first show that

$$\min_{b_{-i}\in\mathcal{B}_{-i}, i\in I} \max_{\tilde{b}_i\in\mathcal{B}_i} \min\left\{\|\tilde{b}_i - \mathrm{Br}^{\sigma}_{-i}(b_{-i})\|_2,\ \min_{j\in I\setminus\{i\}}\{\|\mathrm{Br}^{\sigma}_i(\tilde{b}_i) - b_{ji}\|_2\}\right\} > 0.$$

Consider any $i \in I$, any $b_{-i} \in \mathcal{B}_{-i}$. Since b_{ji} for $j \in I \setminus \{i\}$ are $|I| - 1$ points in Δ_i, and $\mathrm{Im}(\mathrm{Br}^{\sigma}_i) \subset \Delta_i$, by Pigeonhole principle, there exists at least two points $\pi^{\dagger}, \pi' \in \mathrm{Im}(\mathrm{Br}^{\sigma}_i)$ such that

$$\pi' \neq \pi^{\dagger} \neq b_{ji}, \quad \forall j \in I \setminus \{i\}.$$

Let $b'_i, b^{\dagger}_i \in \mathcal{B}_i$ be such that

$$\mathrm{Br}^{\sigma}_i(b'_i) = \pi', \quad \mathrm{Br}^{\sigma}_i(b^{\dagger}_i) = \pi^{\dagger}.$$

Since $\mathrm{Br}^{\sigma}_{-i}(b_{-i})$ is one point in $\mathcal{B}_i$, by Pigeonhole principle, at least one of b'_i or $b^{\dagger}_i$ is different from $\mathrm{Br}^{\sigma}_{-i}(b_{-i})$. WLOG, suppose $b^{\dagger}_i \neq \mathrm{Br}^{\sigma}_{-i}(b_{-i})$. Then,

$$\|b^{\dagger}_i - \mathrm{Br}^{\sigma}_{-i}(b_{-i})\|_2 > 0.$$

Since $\pi^{\dagger} \neq b_{ji}$ for all $j \in I \setminus \{i\}$,

$$\min_{j\in I\setminus\{i\}}\{\|\mathrm{Br}^{\sigma}_i(b^{\dagger}_i) - b_{ji}\|_2\} > 0.$$

Therefore,

$$\max_{\tilde{b}_i\in\mathcal{B}_i} \min\left\{\|\tilde{b}_i - \mathrm{Br}^{\sigma}_{-i}(b_{-i})\|_2,\ \min_{j\in I\setminus\{i\}}\{\|\mathrm{Br}^{\sigma}_i(\tilde{b}_i) - b_{ji}\|_2\}\right\} > 0, \quad \forall i \in I, b_{-i} \in \mathcal{B}_{-i}.$$

Consider any $i \in I$, any $b_{-i} \in \mathcal{B}_{-i}$. Let $\pi_{-i} = \mathrm{Br}^{\sigma}_{-i}(b_{-i})$, and

$$b^{\dagger}_i = \operatorname*{argmax}_{\tilde{b}_i\in\mathcal{B}_i} \min\left\{\|\tilde{b}_i - \mathrm{Br}^{\sigma}_{-i}(b_{-i})\|_2,\ \min_{j\in I\setminus\{i\}}\{\|\mathrm{Br}^{\sigma}_i(\tilde{b}_i) - b_{ji}\|_2\}\right\}.$$

Since

$$\tau < \max_{\tilde{b}_i \in \mathcal{B}_i} \min \left\{ \|\tilde{b}_i - \mathrm{Br}^{\sigma}_{-i}(b_{-i})\|_2, \min_{j \in I \setminus \{i\}} \{\|\mathrm{Br}^{\sigma}_i(\tilde{b}_i) - b_{ji}\|_2\} \right\},$$

we have

$$\tau < \min \left\{ \|b^{\dagger}_i - \mathrm{Br}^{\sigma}_{-i}(b_{-i})\|_2, \min_{j \in I \setminus \{i\}} \{\|\mathrm{Br}^{\sigma}_i(b^{\dagger}_i) - b_{ji}\|_2\} \right\}.$$

Therefore, $b^{\dagger}_i$ satisfies

$$\|\tilde{b}_i - \pi_{-i}\|_2 = \|b^{\dagger}_i - \mathrm{Br}^{\sigma}_{-i}(b_{-i})\|_2 > \tau,$$

and

$$\|\tilde{\pi}_i - b_{ji}\|_2 > \tau, \quad \forall j \neq i.$$

□

A.5 Proof of Lemma 2

(1) For any $z = (b, \pi) \in \mathcal{Z}, z' = (b', \pi) \in \mathcal{Z}$, let I_c be the set of players whose beliefs are τ-consistent in z, and I_{nc} be the set of players whose beliefs are not τ-consistent in z:

$$I_c = \{i \in I \mid \|b_i - \pi_{-i}\|_2 \leq \tau\},$$
$$I_{nc} = \{i \in I \mid \|b_i - \pi_{-i}\|_2 > \tau\}.$$

Let I_d be the set of players whose belief changes from z to z', and I_{nd} be set of players whose belief does not change from z to z'

$$I_d = \{i \in I \mid b_i \neq b'_i\},$$
$$I_{nd} = \{i \in I \mid b_i = b'_i\}.$$

Let $(a^t)_{t=1}^T$ be sequence of realized action profiles in one episode. From Proposition (1),

$$\Pr\left((a^t)_{t=1}^T \in R_i(\xi)\right) \leq \xi^{\bar{u}}, \quad \forall i \in I_c,$$
$$\Pr\left((a^t)_{t=1}^T \notin R_i(\xi)\right) \leq \xi^{\bar{u}}, \quad \forall i \in I_{nc}.$$

We first show that as ξ goes to zero, all players with inconsistent beliefs reject their null hypothesis with probability 1, and all players with τ-consistent beliefs fail to reject their null hypothesis with probability 1 (if they choose to conduct a test):

$$\begin{aligned}
1 &\geq \Pr(i \in I_c \text{ all fail to reject, } i \in I_{nc} \text{ all reject}) \\
&= \Pr\left((a^t)_{t=1}^T \notin \bigcup_{i \in I_c} R_i(\xi), (a^t)_{t=1}^T \in \bigcap_{i \in I_{nc}} R_i(\xi)\right) \\
&= 1 - \Pr\left((a^t)_{t=1}^T \in \bigcup_{i \in I_c} R_i(\xi) \text{ or } (a^t)_{t=1}^T \notin \bigcap_{i \in I_{nc}} R_i(\xi)\right) \\
&\geq 1 - \sum_{i \in I_c} \Pr\left((a^t)_{t=1}^T \in R_i(\xi)\right) - \sum_{i \in I_{nc}} \Pr\left((a^t)_{t=1}^T \notin R_i(\xi)\right) \\
&\geq 1 - |I| \cdot \xi^{\bar{u}}. \qquad \text{(A.16)}
\end{aligned}$$

Therefore,

$$\lim_{\xi \to 0} \Pr(i \in I_c \text{ all fail to reject}, i \in I_{nc} \text{ all reject}) = 1, \qquad \text{(A.17)}$$

$$\lim_{\xi \to 0} \Pr(\neg(i \in I_c \text{ all fail to reject}, i \in I_{nc} \text{ all reject})) = 0. \qquad \text{(A.18)}$$

For any $z, z' \in Z$, the transition probability $P^0_{zz'}$ from z to z' as $\xi \to 0$ satisfies:

$$
\begin{aligned}
P^0_{zz'} &= \lim_{\xi \to 0} \Pr(z \to z') \\
&= \lim_{\xi \to 0} \Pr(z \to z' \text{ and } (i \in I_c \text{ all fail to reject}, i \in I_{nc} \text{ all reject})) \\
&\quad + \lim_{\xi \to 0} \Pr(z \to z' \text{ and } \neg(i \in I_c \text{ all fail to reject}, i \in I_{nc} \text{ all reject})) \\
&\overset{(A.18)}{=} \lim_{\xi \to 0} \Pr(z \to z' \text{ and } (i \in I_c \text{ fail to reject}, i \in I_{nc} \text{ all reject})) \\
&= \lim_{\xi \to 0} \Pr(z \to z' \mid (i \in I_c \text{ fail to reject}, i \in I_{nc} \text{ all reject})) \\
&\qquad \cdot \Pr(i \in I_c \text{ fail to reject}, i \in I_{nc} \text{ all reject}) \\
&\overset{(A.17)}{=} \lim_{\xi \to 0} \Pr(z \to z' \mid (i \in I_c \text{ fail to reject}, i \in I_{nc} \text{ all reject})).
\end{aligned}
$$

Conditioned on the event that all players $i \in I_c$ fail to reject the null hypothesis and all players $i \in I_{nc}$ reject the null hypothesis, the probability of transitioning from state z to state z' depends only on: (i) the exploration probabilities of players in I_c, and (ii) the probabilities of resampling the specific beliefs in z' for those players who resample. These probabilities are independent across players. Therefore,

$$
\begin{aligned}
&\lim_{\xi \to 0} \Pr(z \to z' \mid (i \in I_c \text{ fail to reject}, i \in I_{nc} \text{ all reject})) \\
&= \lim_{\xi \to 0} \prod_{i \in I} \Pr(z_i \to z'_i \mid (i \in I_c \text{ fail to reject}, i \in I_{nc} \text{ all reject})) \\
&= \prod_{i \in I} \lim_{\xi \to 0} \Pr(z_i \to z'_i \mid (i \in I_c \text{ fail to reject}, i \in I_{nc} \text{ all reject})).
\end{aligned}
$$

Analogous to the above argument, the the transition probability $P^0_{i,zz'}$ can be written as follows:

$$
P^0_{i,zz'} = \lim_{\xi \to 0} \Pr(z_i \to z'_i \mid (i \in I_c \text{ fail to reject}, i \in I_{nc} \text{ all reject})).
$$

Therefore,

$$P^0_{zz'} = \prod_{i \in I} P^0_{i,zz'},$$

We next compute the transition probability $P^0_{i,zz'}$ for each player i. As $\xi \to 0$, the state transition probability of player $i \in I_c$ is derived from the following event in sequence:

- Player i starts a test with probability γ_i;
- After a test is started, player i fails to reject the null hypothesis with probability 1 following from (A.17);
- Player i explores with probability 0 since $\lim_{\xi \to 0} \xi^{f_i(U_i(\pi_i, b_i))} = 0$, and uniformly chooses any new belief with probability $1/\mathcal{B}_i$.

Therefore,

$$P^0_{i,zz'} = \begin{cases} 1, & \text{if } \|b_i - \pi_{-i}\|_2 \leq \tau,\ b_i = b'_i, \\ 0, & \text{if } \|b_i - \pi_{-i}\|_2 \leq \tau,\ b_i \neq b'_i. \end{cases}$$

Similarly,

- Each player $i \in I_{nc}$ starts a test with probability γ_i;
- If a test is started, player i rejects the null hypothesis with probability 1 following (A.17) and uniformly chooses any new belief with probability $1/\mathcal{B}_i$.

That is,

$$P^0_{i,zz'} = \begin{cases} (1-\gamma_i) + \gamma_i \cdot \frac{1}{|\mathcal{B}_i|}, & \text{if } \|b_i - \pi_{-i}\|_2 > \tau,\ b_i = b'_i, \\ \gamma_i \cdot \frac{1}{|\mathcal{B}_i|}, & \text{if } \|b_i - \pi_{-i}\|_2 > \tau,\ b_i \neq b'_i. \end{cases}$$

Summarizing, we have

$$P^0_{i,zz'} = \begin{cases} 1, & \text{if } \|b_i - \pi_{-i}\|_2 \le \tau,\ b_i = b'_i, \\ 0, & \text{if } \|b_i - \pi_{-i}\|_2 \le \tau,\ b_i \ne b'_i, \\ (1-\gamma_i) + \gamma_i \cdot \frac{1}{|\mathcal{B}_i|}, & \text{if } \|b_i - \pi_{-i}\|_2 > \tau,\ b_i = b'_i, \\ \gamma_i \cdot \frac{1}{|\mathcal{B}_i|}, & \text{if } \|b_i - \pi_{-i}\|_2 > \tau,\ b_i \ne b'_i. \end{cases}$$

(2) Condition (i) follows from Lemma 1, and condition (ii) follows by the definition. To verify condition (iii), we need to show that $0 < \lim_{\xi \to 0} P^{\xi}_{zz'} \cdot \xi^{-r_{zz'}} < \infty$, where $r_{zz'}$ is defined in (2.6) and can be rewritten as follows:

$$r_{zz'} = \sum_{i \in I_d \cap I_c} f_i(U_i(\pi_i, b_i)) \ge 0, \quad \forall z = (b, \pi), z' \in \mathcal{Z}.$$

We first compute the lower bound of $\lim_{\xi \to 0} P^{\xi}_{zz'} \cdot \xi^{-r_{zz'}}$. We consider a particular event that leads the state transition from z to z':

- All players in I_{nd} (whose beliefs do not change from z to z') do not start hypothesis tests.
- All players in $I_c \cap I_d$ (whose beliefs are τ-consistent and do change):
  (a) start hypothesis tests,
  (b) fail to reject their null hypothesis,
  (c) explore and resample new beliefs as specified in b'.
- All players in $I_{nc} \cap I_d$ (whose beliefs are not τ-consistent and do change):
  (a) start hypothesis tests,
  (b) reject their null hypothesis,
  (c) update beliefs as specified in b'.

We can lower bound the transition probability $P^{\xi}_{zz'}$ as the probability of the above event:

$$
\begin{aligned}
P^{\xi}_{zz'} \geq \Pr\Big(& i \in I_{nd} \text{ does not test}, \\
& i \in I_d \bigcap I_c \text{ tests, fails to reject, and explores to new belief } b_i', \\
& i \in I_d \bigcap I_{nc} \text{ tests, rejects, and changes to new belief } b_i' \Big) \\
= & \prod_{i \in I_{nd}} \Pr\left(i \text{ does not test}\right) \cdot \prod_{i \in I_d} \Pr\left(i \text{ tests}\right) \\
& \cdot \Pr\left(i \in I_d \bigcap I_c \text{ all fail to reject}, i \in I_d \bigcap I_{nc} \text{ all reject}\right) \\
& \cdot \prod_{i \in I_d \bigcap I_c} \Pr\left(i \text{ explores}\right) \cdot \Pr\left(\text{all } i \in I_d \text{ choose } b_i'\right) \qquad (*) \\
\geq & \prod_{i \in I_{nd}} \Pr\left(i \text{ does not test}\right) \cdot \prod_{i \in I_d} \Pr\left(i \text{ tests}\right) \cdot \Pr\left(\text{all } i \in I_d \text{ choose } b_i'\right) \\
& \cdot \Pr\left(i \in I_c \text{ all fail to reject}, i \in I_{nc} \text{ all reject}\right) \cdot \prod_{i \in I_d \bigcap I_c} \Pr\left(i \text{ explores}\right) \\
\overset{(A.16)}{\geq} & (1-\gamma_i)^{|I_{nd}|} \cdot \gamma_i^{|I_d|} \cdot \left(\frac{1}{\mathcal{B}_i}\right)^{|I_d|} \cdot \left(1 - |I| \cdot \xi^{\bar{u}}\right) \cdot \prod_{i \in I_d \cap I_c} \xi^{f_i(U_i(\pi_i, b_i))} \\
= & (1-\gamma_i)^{|I_{nd}|} \cdot \gamma_i^{|I_d|} \cdot \left(\frac{1}{\mathcal{B}_i}\right)^{|I_d|} \cdot \left(1 - |I| \cdot \xi^{\bar{u}}\right) \cdot \xi^{\sum_{i \in I_d \cap I_c} f_i(U_i(\pi_i, b_i))},
\end{aligned}
$$

where (*) is due to the fact that players independently start a test or explore. Consequently,

$$
\begin{aligned}
& \lim_{\xi \to 0} P^{\xi}_{zz'} \cdot \xi^{-r_{zz'}} \\
\geq & \lim_{\xi \to 0} (1-\gamma_i)^{|I_{nd}|} \cdot \left(\frac{1}{\mathcal{B}_i}\right)^{|I_d|} \cdot \gamma_i^{|I_d|} \cdot \left(1 - |I| \cdot \xi^{\bar{u}}\right) \cdot \xi^{\sum_{i \in I_d \cap I_c} f_i(U_i(\pi_i, b_i))} \\
& \cdot \xi^{-\sum_{i \in I_d \cap I_c} f_i(U_i(\pi_i, b_i))} \\
= & (1-\gamma_i)^{|I_{nd}|} \cdot \gamma_i^{|I_d|} \cdot \left(\frac{1}{\mathcal{B}_i}\right)^{|I_d|} > 0.
\end{aligned}
$$

Next, we compute an upper bound of $\lim_{\xi \to 0} P^{\xi}_{zz'} \cdot \xi^{-r_{zz'}}$. For any $z, z' \in Z$, the state transition probability $P^{\xi}_{zz'}$ satisfies

$$
\begin{aligned}
& P^{\xi}_{zz'} \\
&= \Pr\left(\text{all } i \in I \text{ transit from } b_i \text{ to } b'_i\right) \\
&\leq \Pr\left(\text{all } i \in I_d \text{ transit from } b_i \text{ to } b'_i\right) \\
&= \Pr\left(\text{all } i \in I_d \text{ test, either reject or fail to reject and explore,}\right. \qquad \text{(A.20)} \\
&\qquad \left.\text{then transit from } b_i \text{ to } b'_i\right) \\
&= \Pr\left(\text{all } i \in I_d \text{ test}\right) \cdot \Pr\left(\text{all } i \in I_d \text{ choose } b'_i\right) \\
&\quad \cdot \sum_{I_r \subseteq I_d} \Big(\Pr\left(\text{all } i \in I_d \setminus I_r \text{ fail to reject, all } i \in I_r \text{ reject}\right) \\
&\qquad \cdot \Pr\left(\text{all } i \in I_d \setminus I_r \text{ choose to explore}\right)\Big) \\
&= \gamma_i^{|I_d|} \cdot \left(\frac{1}{\mathcal{B}_i}\right)^{|I_d|} \\
&\quad \cdot \sum_{I_r \subseteq I_d} \Big(\Pr\left(\text{all } i \in I_d \setminus I_r \text{ fail to reject, all } i \in I_r \text{ reject}\right) \\
&\qquad \cdot \Pr\left(\text{all } i \in I_d \setminus I_r \text{ choose to explore}\right)\Big). \qquad \text{(A.21)}
\end{aligned}
$$

We consider a specific set $I_r^{\dagger} = I_d \bigcap I_{nc}$. We have $I_d \setminus I_r^{\dagger} = I_d \setminus (I_d \bigcap I_{nc}) = I_d \bigcap I_c$. Then,

$$
\begin{aligned}
& \Pr\left(\text{all } i \in I_d \setminus I_r^{\dagger} \text{ fail to reject, all } i \in I_r^{\dagger} \text{ reject}\right) \\
=& \Pr\left(\text{all } i \in I_d \bigcap I_c \text{ fail to reject, all } i \in I_d \bigcap I_{nc} \text{ reject}\right) \\
\geq& \Pr\left(\text{all } i \in I_c \text{ fail to reject, all } i \in I_{nc} \text{ reject}\right) \\
\overset{\text{(A.16)}}{\geq}& 1 - |I| \cdot \xi^{\bar{u}}.
\end{aligned}
$$

Therefore,

$$\sum_{\substack{I_r \neq I_r^\dagger \\ I_r \subseteq I_d}} \Pr\left(\text{all } i \in I \setminus I_r \text{ fail to reject, all } i \in I_r \text{ reject}\right)$$
$$\leq 1 - \Pr\left(\text{all } i \in I \setminus I_r^\dagger \text{ fail to reject, all } i \in I_r^\dagger \text{ reject}\right)$$
$$\leq |I| \cdot \xi^{\bar{u}}. \tag{A.22}$$

Hence,

$$P_{zz'}^{\xi}$$
$$\overset{(A.21)}{\leq} \gamma_i^{|I_d|} \cdot \left(\frac{1}{\mathcal{B}_i}\right)^{|I_d|} \cdot$$
$$\left(\Pr\left(i \in I_d \bigcap I_c \text{ fail to reject, } i \in I_d \bigcap I_{nc} \text{ reject}\right) \cdot \Pr\left(i \in I_d \bigcap I_c \text{ explore}\right)\right.$$
$$\left. + \sum_{\substack{I_r \neq I_i^\dagger \\ I_r \subseteq I_d}} \Pr\left(i \in I_d \setminus I_r \text{ fail to reject, } i \in I_r \text{ reject}\right) \cdot \Pr\left(i \in I_d \setminus I_r \text{ explore}\right)\right)$$
$$\overset{(A.22)}{\leq} \gamma_i^{|I_d|} \cdot \left(\frac{1}{\mathcal{B}_i}\right)^{|I_d|} \cdot \left(1 \cdot \Pr\left(\text{all } i \in I_d \bigcap I_c \text{ choose to explore}\right) + |I| \cdot \xi^{\bar{u}} \cdot 1\right) \tag{A.23a}$$
$$= \gamma_i^{|I_d|} \cdot \left(\frac{1}{\mathcal{B}_i}\right)^{|I_d|} \cdot \left(\prod_{i \in I_d \cap I_c} \xi^{f_i(U_i(\pi_i, b_i))} + |I| \cdot \xi^{\bar{u}}\right)$$
$$= \gamma_i^{|I_d|} \cdot \left(\frac{1}{\mathcal{B}_i}\right)^{|I_d|} \cdot \left(\xi^{\sum_{i \in I_d \cap I_c} f_i(U_i(\pi_i, b_i))} + |I| \cdot \xi^{\bar{u}}\right). \tag{A.23b}$$

Inequality (A.23a) follows from inequality (A.22), and the fact that all probabilities are smaller than or equal to 1. Then,

$$\lim_{\xi \to 0} P_{zz'}^{\xi} \cdot \xi^{-r_{zz'}}$$
$$\overset{(A.23b)}{\leq} \lim_{\xi \to 0} \gamma_i^{|I_d|} \cdot \left(\frac{1}{\mathcal{B}_i}\right)^{|I_d|} \cdot \left(\xi^{\sum_{i \in I_d \cap I_c} f_i(U_i(\pi_i, b_i))} + |I| \cdot \xi^{\bar{u}}\right) \cdot \xi^{-\sum_{i \in I_d \cap I_c} f_i(U_i(\pi_i, b_i))}$$
$$= \lim_{\xi \to 0} \gamma_i^{|I_d|} \cdot \left(\frac{1}{\mathcal{B}_i}\right)^{|I_d|} \cdot \left(1 + |I| \cdot \xi^{\bar{u} - \sum_{i \in I_d \cap I_c} f_i(U_i(\pi_i, b_i))}\right)$$
$$= \gamma_i^{|I_d|} \cdot \left(\frac{1}{\mathcal{B}_i}\right)^{|I_d|} < \infty,$$

where the last equality follows from

$$\bar{u} - \sum_{i \in I_d \cap I_c} f_i(U_i(\pi_i, b_i)) > \max_{\pi \in \Delta} \sum_{i \in I} f_i(U_i(\pi)) - \sum_{i \in I_d \cap I_c} f_i(U_i(\pi_i, b_i)) \geq 0.$$

Therefore, we can conclude that given $r_{zz'} = \sum_{i \in I_d \cap I_c} f_i(U_i(\pi_i, b_i))$, we have

$$0 < \lim_{\xi \to 0} P^{\xi}_{zz'} \cdot \xi^{-r_{zz'}} < \infty.$$

□

APPENDIX B

CHAPTER 3 APPENDIX

B.1 Proof of Results

Proof of Lemma 6. We first show that $\ell_\delta(\sigma^*, \rho) > 0$. Assume that $\ell_\delta(\sigma^*, \rho) \leq 0$ for the sake of contradiction. Then, for any (β, γ) with $\gamma > 0$, we have

$$C_{\mathrm{o}}(\sigma^*, \beta, \gamma) - C_{\mathrm{toll}}(\sigma^*, \beta, \gamma) = \beta\ell_\delta(\sigma^*, \rho) - \tau < 0,$$
$$C_{\mathrm{o}}(\sigma^*, \beta, \gamma) - C_{\mathrm{pool}}(\sigma^*, \beta, \gamma) = \beta\ell_\delta(\sigma^*, \rho) - \gamma < 0.$$

That is, all agent will choose to take the ordinary lane, and thus

$$\begin{aligned}\ell_\delta(\sigma^*, \rho) &= \ell_{\mathrm{o}}(D, 1-\rho) - \ell_{\mathrm{h}}(0, \rho)\\ &= \ell_{\mathrm{o}}(D, 1-\rho) - \ell_{\mathrm{o}}(0, 1-\rho) > 0,\end{aligned}$$

where the second equality is due to Assumption 3 (b) and the last inequality is due to Assumption 3 (a). We obtain a contradiction. Hence, $\ell_\delta(\sigma^*, \rho) > 0$.

Next, we prove that $\sigma^*_{\mathrm{pool}} > 0$. Consider agents whose value of time satisfies $\beta \in [\bar{\beta}/2, \bar{\beta}]$ and carpool disutility satisfies $\gamma \in \left[0, \frac{1}{3}\min\left\{\bar{\beta}\ell_\delta(\sigma^*, \rho), \tau\right\}\right]$. For those agents, we have $\gamma \leq \frac{1}{3}\min\left\{\bar{\beta}\ell_\delta(\sigma^*, \rho), \tau\right\} < \tau$, which leads to

$$\begin{aligned}C_{\mathrm{pool}}(\sigma^*, \beta, \gamma) &= \beta\ell_{\mathrm{h}}(\sigma^*, \rho) + \gamma\\ &< \beta\ell_{\mathrm{h}}(\sigma^*, \rho) + \tau = C_{\mathrm{toll}}(\sigma^*, \beta, \gamma).\end{aligned}$$

Additionally, we have

$$\begin{aligned}
C_{\text{pool}}(\sigma^*,\beta,\gamma) =&\beta\ell_{\text{h}}(\sigma^*,\rho)+\gamma\\
\leq&\beta\ell_{\text{h}}(\sigma^*,\rho)+\frac{1}{3}\bar{\beta}\ell_{\delta}(\sigma^*,\rho)\\
<&\beta\ell_{\text{h}}(\sigma^*,\rho)+\beta\ell_{\delta}(\sigma^*,\rho)\\
=&\beta\ell_{\text{o}(}\sigma^*,1-\rho)=C_{\text{o}}(\sigma^*,\beta,\gamma),
\end{aligned}$$

where the last inequality is due to $\beta \geq \frac{1}{2}\bar{\beta} > \frac{1}{3}\bar{\beta}$. Hence, for those agents, we have $C_{\text{pool}}(\sigma^*,\beta,\gamma) < C_{\text{o}}(\sigma^*,\beta,\gamma)$ and $C_{\text{pool}}(\sigma^*,\beta,\gamma) < C_{\text{toll}}(\sigma^*,\beta,\gamma)$. That is, all agents with preference parameters in this set choose to carpool. Therefore,

$$\sigma^*_{\text{pool}} \geq \int_{\bar{\beta}/2}^{\bar{\beta}}\int_0^{\frac{1}{3}\min\left\{\bar{\beta}\ell_{\delta}(\sigma^*,\rho),\tau\right\}} f(\beta,\gamma)d\gamma d\beta > 0.$$

Finally, we are going to argue that $\sigma^*_{\text{o}} > 0$. Assume for the sake of contradiction that $\sigma^*_{\text{o}} = 0$, then

$$\ell_{\delta}(\sigma^*,\rho) = \ell_{\text{o}}(0,1-\rho) - \ell_{\text{h}}(\sigma^*,\rho) < 0,$$

which is a contradiction with the fact that $\ell_{\delta}(\sigma^*,\rho) > 0$. □

Proof of Lemma 7 Recall from (3.3), agents whose best response is toll satisfy $C_{\text{toll}}(\sigma,\beta,\gamma) \leq C_{\text{pool}}(\sigma,\beta,\gamma)$ and $C_{\text{toll}}(\sigma,\beta,\gamma) \leq C_{\text{o}}(\sigma,\beta,\gamma)$, i.e. the cost of paying the toll is smaller than or equal to the cost of any other two actions. From $C_{\text{toll}}(\sigma,\beta,\gamma) \leq C_{\text{pool}}(\sigma,\beta,\gamma)$, we obtain $\beta\ell_{\text{h}}(\sigma,\rho)+\tau \leq \beta\ell_{\text{h}}(\sigma,\rho)+\gamma$, which yields $\tau \leq \gamma$. Moreover, from $C_{\text{toll}}(\sigma,\beta,\gamma) \leq C_{\text{o}}(\sigma,\beta,\gamma)$, we obtain $\beta\ell_{\text{h}}(\sigma,\rho)+\tau \leq \beta\ell_{\text{o}}(\sigma,1-\rho)$, which yields $\tau \leq \beta\ell_{\delta}(\sigma,\rho)$. Thus, we obtain $\Lambda_{\text{toll}}(\sigma)$.

Similarly, agents whose best response is pool satisfy $C_{\text{pool}}(\sigma,\beta,\gamma) = \beta\ell_{\text{h}}(\sigma,\rho)+\gamma \leq C_{\text{toll}}(\sigma,\beta,\gamma) = \ell_{\text{h}}(\sigma,\rho)+\tau$ and $C_{\text{pool}}(\sigma,\beta,\gamma) = \beta\ell_{\text{h}}(\sigma,\rho)+\gamma \leq$

$C_{\rm o}(\sigma,\beta,\gamma) = \beta\ell_{\rm o}(\sigma,1-\rho)$. It yields $\gamma \le \tau$ and $\gamma \le \beta\ell_\delta(\sigma,\rho)$, and thus we obtain $\Lambda_{\rm pool}(\sigma)$. Lastly, agents whose best response is o (ordinary lane) satisfy $C_{\rm o}(\sigma,\beta,\gamma) = \beta\ell_{\rm o}(\sigma,1-\rho) \le C_{\rm toll}(\sigma,\beta,\gamma) = \beta\ell_{\rm h}(\sigma,\rho)+\tau$ and $C_{\rm o}(\sigma,\beta,\gamma) = \beta\ell_{\rm o}(\sigma,1-\rho) \le C_{\rm pool}(\sigma,\beta,\gamma) = \beta\ell_{\rm h}(\sigma,\rho)+\gamma$. It yields $\beta\ell_\delta(\sigma,\rho) \le \min\{\tau,\gamma\}$, and thus we obtain $\Lambda_{\rm o}(\sigma)$. $\square$

Proof of Theorem 2: In each regime, we first show that under the regime condition, the equations described in the theorem has a unique fixed point. Then, we verify that $(\sigma^*_{\rm toll},\sigma^*_{\rm pool},\sigma^*_{\rm o})$ provided for each regime indeed satisfies the equilibrium definition.

Regime A:

(A-1) $\bar\beta\ell^\dagger_\delta \le \overline{\gamma}$.

Consider a function

$$q(z) := \frac{z}{\int_0^{\overline{\beta}}\int_0^{\ell_\delta(0,z,1-z,\rho)\beta} f(\beta,\gamma)\,d\gamma d\beta}.$$

Recall that under regime A, we have $\tau \ge \min\{\overline{\beta}\ell^\dagger_\delta,\overline{\gamma}\}$. Therefore, we have $\overline{\beta}\ell^\dagger_\delta \le \min\{\tau,\overline{\gamma}\}$. Thus, for any $\beta > 0$, we have $\frac{\min\{\tau,\overline{\gamma}\}}{\overline{\beta}}\beta \ge \ell^\dagger_\delta\beta = \ell_\delta(0,\sigma^\dagger_{\rm pool},1-\sigma^\dagger_{\rm pool})\beta$. Therefore,

$$\sigma^\dagger_{\rm pool} = \int_0^{\overline{\beta}}\int_0^{\frac{\min\{\tau,\overline{\gamma}\}}{\overline{\beta}}\beta} f(\beta,\gamma)\,d\gamma d\beta$$
$$\ge \int_0^{\overline{\beta}}\int_0^{\ell_\delta(0,\sigma^\dagger_{\rm pool},1-\sigma^\dagger_{\rm pool})\beta} f(\beta,\gamma)\,d\gamma d\beta.$$

Hence, $q(\sigma^\dagger_{\rm pool}) \ge 1$. Additionally, it is easy to see that $q(0)=0$.

Since $\ell_\delta(0,z,1-z,\rho)$ is monotonically decreasing with z, $q(z)$ is continuous and monotonically increasing in z. Therefore there must be a unique point $\sigma^*_{\text{pool}} \in [0, \sigma^\dagger_{\text{pool}}]$ such that $q(\sigma^*_{\text{pool}}) = 1$, which is the unique solution of equation (3.7):

$$\sigma^*_{\text{pool}} = \int_0^{\overline{\beta}} \int_0^{\ell_\delta(0,\sigma^*_{\text{pool}},1-\sigma^*_{\text{pool}})\beta} f(\beta,\gamma)\, d\gamma d\beta.$$

Now we want to show that $\sigma^*_{\text{toll}} = 0, \sigma^*_{\text{pool}}, \sigma^*_{\text{o}} = 1 - \sigma^*_{\text{pool}}$ satisfies the equilibrium condition (3.10).

First, we argue that at equilibrium, $\sigma^*_{\text{toll}} = 0$. From previous argument, we know that $\sigma^*_{\text{pool}} \leq \sigma^\dagger_{\text{pool}}$. Plugging in the definitions of strategy distribution and $\sigma^\dagger_{\text{pool}}$, we obtain $\int_0^{\overline{\beta}} \int_0^{\ell_\delta(0,\sigma^*_{\text{pool}},1-\sigma^*_{\text{pool}})\beta} f(\beta,\gamma)\, d\gamma d\beta \leq \int_0^{\overline{\beta}} \int_0^{\frac{\min\{\tau,\overline{\gamma}\}}{\overline{\beta}}\beta} f(\beta,\gamma)\, d\gamma d\beta$, which implies $\ell_\delta(0,\sigma^*_{\text{pool}}, 1-\sigma^*_{\text{pool}}) \leq \frac{\min\{\tau,\overline{\gamma}\}}{\overline{\beta}}$. Therefore, for any (β,γ), we have $\beta\ell_\delta(0,\sigma^*_{\text{pool}}, 1-\sigma^*_{\text{pool}}) \leq \min\{\tau,\overline{\gamma}\}$. Hence, for all agents, we have $C_{\text{o}}(\sigma,\beta,\gamma) \leq C_{\text{toll}}(\sigma,\beta,\gamma)$. Thus, no agents want to deviate to pay the toll and we have $\sigma^*_{\text{toll}} = 0$ at equilibrium.

Next, we argue that at equilibrium, the strategy distribution of carpooling agents is given by σ^*_{pool}. An agent (β,γ) will choose pool over o if $C_{\text{pool}}(\sigma^*,\beta,\gamma) < C_{\text{o}}(\sigma^*,\beta,\gamma)$, which is equivalent to $\gamma < \beta\ell_\delta(0,\sigma^*_{\text{pool}}, 1 - \sigma^*_{\text{pool}})$. By lemma 7, the set of agents whose best response is to carpool under regime A-1 is then given by $\Lambda_{\text{pool}}(\sigma^*) = \{(\beta,\gamma) : 0 < \beta < \overline{\beta}, 0 < \gamma < \beta\ell_\delta(0,\sigma^*_{\text{pool}}, 1-\sigma^*_{\text{pool}})\}$. Integrating over all agents in $\Lambda_{\text{pool}}(\sigma^*)$ yields the σ^*_{pool} given above. Lastly, because we have argued that $\sigma^*_{\text{toll}} = 0$ and σ^*_{pool} satisfy the equilibrium condition (3.10), we can obtain that $\sigma^*_{\text{o}} = 1 - \sigma^*_{\text{pool}}$ must also satisfy the equilibrium condition. Hence, we have proved that $(\sigma^*_{\text{toll}}, \sigma^*_{\text{pool}}, \sigma^*_{\text{o}})$ provided for regime A-1 satisfies the equilibrium condition.

(A-2) $\overline{\beta}\ell^\dagger_\delta > \overline{\gamma}$.

Consider the following function g:

$$g(z) := \frac{1-z}{\int_0^{\overline{\gamma}} \int_0^{\frac{\gamma}{\ell_\delta(0,z,1-z,\rho)}} f(\beta,\gamma) d\beta \, d\gamma}.$$

We first argue that $g(\sigma^\dagger_{\text{pool}}) > 1$. From the regime A condition such that $\tau \geq \min\{\overline{\gamma}, \bar{\beta}\ell^\dagger_\delta\}$, we obtain $\bar{\beta}\ell^\dagger_\delta > \min\{\tau, \overline{\gamma}\}$ and $\tau \geq \overline{\gamma}$. Then, we have

$$\begin{aligned}
1 - \sigma^\dagger_{\text{pool}} =& 1 - \int_0^{\overline{\beta}} \int_0^{\frac{\min\{\tau,\overline{\gamma}\}}{\overline{\beta}}\beta} f(\beta,\gamma) \, d\gamma d\beta \\
\overset{(a)}{>}& 1 - \int_0^{\overline{\beta}} \int_0^{\beta\ell_\delta(0,\sigma^\dagger_{\text{pool}},1-\sigma^\dagger_{\text{pool}})} f(\beta,\gamma) \, d\gamma d\beta \\
=& \int_0^{\overline{\beta}} \int_{\beta\ell_\delta(0,\sigma^\dagger_{\text{pool}},1-\sigma^\dagger_{\text{pool}})}^{\overline{\gamma}} f(\beta,\gamma) \, d\gamma d\beta \\
\overset{(b)}{=}& \int_0^{\overline{\gamma}} \int_0^{\frac{\gamma}{\ell_\delta(0,\sigma^\dagger_{\text{pool}},1-\sigma^\dagger_{\text{pool}},\rho)}} f(\beta,\gamma) d\beta \, d\gamma
\end{aligned}$$

(a) is due to the inequality $\min\{\tau,\overline{\gamma}\} < \bar{\beta}\ell^\dagger_\delta = \bar{\beta}\ell_\delta(0,\sigma^\dagger_{\text{pool}}, 1-\sigma^\dagger_{\text{pool}})$, and (b) is obtained by changing the integration order. Therefore, we obtain $g(\sigma^\dagger_{\text{pool}}) > 1$. Additionally, it is easy to see that $g(1) = 0$. Since $\ell_\delta(0,z,1-z,\rho)$ is monotonically decreasing with z, we know that $g(z)$ is monotonically decreasing with z. Therefore, there exist a unique $\sigma^*_{\text{pool}} > \sigma^\dagger_{\text{pool}}$ such that $g(\sigma^*_{\text{pool}}) = 1$, which means that it is the unique solution of equation (3.8):

$$1 - \sigma^*_{\text{pool}} = \int_0^{\overline{\gamma}} \int_0^{\frac{\gamma}{\ell_\delta(0,\sigma^*_{\text{pool}},1-\sigma^*_{\text{pool}})}} f(\beta,\gamma) d\beta d\gamma.$$

Next, we are going to show that $\sigma^*_{\text{toll}} = 0, \sigma^*_{\text{pool}}, \sigma^*_{\text{o}} = 1 - \sigma^*_{\text{pool}}$ satisfies the equilibrium condition (3.10).

Due to the condition of regime A-2 such that $\tau \geq \overline{\gamma}$, we obtain that $C_{\text{pool}}(\sigma,\beta,\gamma) \leq C_{\text{toll}}(\sigma,\beta,\gamma)$ holds for all agents. Therefore, no agents want to deviate to toll paying and thus $\sigma^*_{\text{toll}} = 0$ at equilibrium. Additionally, by lemma 7, the set of agents using ordinary lanes under regime A-2 is given by $\Lambda_o(\sigma^*) = \{(\beta, \gamma) : \beta \leq \gamma/\ell_\delta(\sigma^*)\}$. Hence, under equilibrium, we obtain

$$\begin{aligned}\sigma_o^* &= \iint_{\Lambda_o(\sigma^*)} f(\beta,\gamma) d\beta d\gamma \\ &= \int_0^{\overline{\gamma}} \int_0^{\frac{\gamma}{\ell_\delta(0,\sigma^*_{\text{pool}},1-\sigma^*_{\text{pool}})}} f(\beta,\gamma) d\beta d\gamma = 1 - \sigma^*_{\text{pool}},\end{aligned}$$

which is equivalent to equation (3.7). Therefore, σ^*_{pool} and $\sigma^*_o = 1 - \sigma^*_{\text{pool}}$ satisfy the equilibrium condition 3.10 as well.

Regime B:

To show that the system of equations (3.9) has a unique solution, we first represent the strategy distributions using the latency difference between ordinary lanes and HOT lanes, which we denote as δ. Consider function $g : [\tau/\bar{\beta}, \infty) \to [0, \bar{\sigma}_{\text{toll}})$, where $\bar{\sigma}_{\text{toll}}$ is defined as $\bar{\sigma}_{\text{toll}} := \int_\tau^{\overline{\gamma}} \int_0^{\bar{\beta}} f(\beta,\gamma)\, d\beta d\gamma$,

$$g(y) = \int_\tau^{\overline{\gamma}} \int_{\tau/y}^{\bar{\beta}} f(\beta,\gamma)\, d\beta d\gamma,$$

and function h with the domain $[\tau/\bar{\beta}, \infty)$:

$$h(y) = \int_0^{\tau} \int_{\gamma/y}^{\bar{\beta}} f(\beta,\gamma)\, d\beta d\gamma.$$

We note that both g and h are continuous and increase monotonically with y. Moreover, we note that the range of h is $[\sigma^\dagger_{\text{pool}}, 1 - \bar{\sigma}_{\text{toll}})$, because

$$\begin{aligned}h(\tau/\bar{\beta}) &= \int_0^{\tau} \int_{\bar{\beta}\gamma/\tau}^{\bar{\beta}} f(\beta,\gamma) d\beta d\gamma \\ &= \int_0^{\bar{\beta}} \int_0^{\beta\tau/\bar{\beta}} f(\beta,\gamma) d\gamma d\beta \\ &= \int_0^{\bar{\beta}} \int_0^{\min\{\tau,\overline{\gamma}\}\beta/\bar{\beta}} f(\beta,\gamma) d\gamma d\beta = \sigma^\dagger_{\text{pool}},\end{aligned}$$

and

$$\lim_{\delta \to \infty} h(\delta) = \int_0^{\tau} \int_0^{\bar{\beta}} f(\beta, \gamma) d\beta d\gamma$$
$$= 1 - \int_{\tau}^{\overline{\gamma}} \int_0^{\bar{\beta}} f(\beta, \gamma) d\beta d\gamma = 1 - \bar{\sigma}_{\text{toll}}.$$

Let δ^* be the induced latency difference at equilibrium. By equations (3.9), we obtain that $\sigma^*_{\text{toll}} = g(\delta^*)$, $\sigma^*_{\text{pool}} = h(\delta^*)$, and $\sigma^*_{\text{o}} = 1 - g(\delta^*) - h(\delta^*)$. Hence, in order to show that (3.9) has a unique solution, it is suffice to show that there exists a solution to the following fixed point equation:

$$\delta = \ell_\delta(g(\delta), h(\delta), 1 - g(\delta) - h(\delta)). \tag{B.1}$$

Define the function q with the domain $[\tau/\bar{\beta}, \infty)$:

$$q(\delta) = \frac{\ell_\delta(g(\delta), h(\delta), 1 - g(\delta) - h(\delta))}{\delta}.$$

We remark that q is continuous and mononotically decreasing in δ. First, we have

$$q(\tau/\bar{\beta}) = \frac{\ell_\delta(g(\tau/\bar{\beta}), h(\tau/\bar{\beta}), 1 - g(\tau/\bar{\beta}) - h(\tau/\bar{\beta}))}{\tau/\bar{\beta}}$$
$$\overset{(a)}{=} \frac{\bar{\beta}\ell^\dagger_\delta}{\tau} \overset{(b)}{>} 1,$$

where (a) is due to $g(\tau/\bar{\beta}) = 0$ and $h(\tau/\bar{\beta}) = \sigma^\dagger_{\text{pool}}$; (b) is due to the regime condition $\tau < \bar{\beta}\ell^\dagger_\delta$. Furthermore, it is easy to see that as $\delta \uparrow \infty$, $q(\delta) \to 0$. Since q is continuous monotonically decreasing, there exists a unique $\delta^* \in [\tau/\bar{\beta}, \infty)$ such that $q(\delta^*) = 1$, which implies that there exists a unique solution to (B.1). Hence, there exists a unique solution to (3.9).

Given the strategy distribution σ^*, integrating over the best response region for each of the three actions given in lemma 7 yields exactly the system of equations (3.9). Therefore, the solution obtained in (3.9) is an equilibrium. □

Proof of Theorem 3:

Fix τ and increase ρ:

We first show that $\ell_\delta(\sigma^*, \rho)$ increases. We assume for the sake of contradiction that $\ell_\delta(\sigma^*, \rho)$ decreases. Then, from lemma 7, more agents will join the ordinary lanes, and fewer will carpool or pay the toll price. That is, σ_{o}^* increases, σ_{pool}^* decreases, and σ_{toll}^* is non-increasing. Recall from assumption 3 that the latency function for ordinary lanes (resp. HOT lanes) increases (resp. decreases) with ρ. Therefore, $\ell_{\mathrm{o}}(\sigma_{\mathrm{o}}^* D, 1-\rho)$ increases as both σ_{o}^* and ρ increase. Moreover, $\ell_{\mathrm{h}}\left(\left(\sigma_{\mathrm{toll}}^* + \frac{\sigma_{\mathrm{pool}}^*}{A}\right) D, \rho\right)$ decreases because the term $\left(\sigma_{\mathrm{toll}}^* + \frac{\sigma_{\mathrm{pool}}^*}{A}\right)$ decreases and ρ increases. Hence, $\ell_\delta(\sigma^*, \rho)$ increases, which is a contradiction.

Next, we assume for the sake of contradiction that $\ell_\delta(\sigma^*, \rho)$ does not change. Using lemma 7, we obtain that the strategy distributions for all three actions should hold fixed. Nevertheless, given the same strategy distributions, increasing ρ leads to an increase in $\ell_{\mathrm{o}}(\sigma_{\mathrm{o}}^* D, 1-\rho)$ and a decrease in $\ell_{\mathrm{h}}\left(\left(\sigma_{\mathrm{toll}}^* + \frac{\sigma_{\mathrm{pool}}^*}{A}\right) D, \rho\right)$. This yields an increase in $\ell_\delta(\sigma^*, \rho)$, which is a contradiction.

Since $\ell_\delta(\sigma^*, \rho)$ increases, from lemma 7, we obtain that the strategy distributions σ_{o}^* decreases, σ_{pool}^* increases, and σ_{toll}^* increases.

Fix ρ and increase τ:

We first show that at equilibrium $\ell_\delta(\sigma^*, \rho)$ is non-decreasing. Assume for the sake of contradiction that $\ell_\delta(\sigma^*, \rho)$ decreases. Since τ increases, from lemma 7, we obtain that the population paying toll price decreases, the population taking the ordinary lanes is increases, and the population that carpools decreases. Therefore, $\ell_\delta(\sigma^*, \rho)$ increases, which is a contradiction.

Last but not least, we argue that σ_o^* can change in either direction when τ increases. When the toll price is relatively large, i.e. in regime A in theorem 2, further increasing the τ has no impact on the strategy distributions, and hence σ_o^* does not chance. When the toll price is relatively low, i.e. in regime B in theorem 2, an increase in τ can either lead to an increase or decrease of the term $\frac{\tau}{\ell_\delta(\sigma^*,\rho)}$ in (3.9a), which depends on the preference distribution. Therefore, σ_o^* can also change in either direction given the increase in τ. □

Proof of Theorem 4. We note that for each segment $e \in [E]$, $\delta_e \in [\underline{\delta}_e, \bar{\delta}_e]$, where the lower bound $\underline{\delta}_e = \ell_{o,e}(0) - \ell_{\mathrm{h},e}\left(\sum_{i=1}^{e}\sum_{j=e}^{E} D^{ij}\right)$ is achieved when all agents choose the HOT lane without pooling, and the upper bound $\bar{\delta}_e = \ell_{\mathrm{o},e}\left(\sum_{i=1}^{e}\sum_{j=e}^{E} D^{ij}\right) - \ell_{e,h}(0)$ is achieved when all agents take the ordinary lane. Moreover, since the preference distribution function $f(\beta,\gamma)$ is continuous, the function $\Phi : \prod_{e\in E}[\underline{\delta}_e, \bar{\delta}_e] \to \prod_{e\in E}[\underline{\delta}_e, \bar{\delta}_e]$ is continuous. From the Kakutani's fixed-point theorem, we know that the set of fixed-points of $\Phi(\delta) = \delta$ is non-empty. Furthermore, for any such fixed point δ^*, the associated best-response strategy s^* is unique, and satisfies the condition in Definition 9, and thus is a Wardrop equilibrium. On the other hand, for any Nash equilibrium s^*, the induced equilibrium latency cost difference must satisfy $\Phi(\delta^*) = \delta^*$. Therefore, the fixed-point set of $\Phi(\delta) = \delta$ and the set of Wardrop equilibrium has one-to-one correspondence. Since the fixed-point set is non-empty, we can conclude that Wardrop equilibrium exists.

Additionally, equilibrium is unique if and only if $\Phi(\delta) = \delta$ has a unique solution. If there exist two fixed points δ, δ' satisfying $\Phi(\delta) = \delta$ and $\Phi(\delta') = \delta'$. Let v_i denote the $|E|$-dimensional vector with i-th index being 1 and all other indices being 0. Then, by the mean value theorem, there exists z on the line segment connecting δ and δ' such that

$$v_e \cdot [\nabla\Phi(z)(\delta - \delta')] = v_e \cdot [\Phi(\delta) - \Phi(\delta')], \quad \forall e \in E.$$

Combining the equations for all $e \in E$, we obtain

$$\nabla\Phi(z)(\delta - \delta') = \Phi(\delta) - \Phi(\delta') = \delta - \delta'. \tag{B.2}$$

If the Jacobian $\nabla\Phi(z)$ does not have eigenvalue 1 for any z, then there do not exist δ, δ' that satisfies (B.2), and thus equilibrium is unique. $\square$

APPENDIX C

CHAPTER 4 APPENDIX

C.1 Technical Lemmas

Lemma 12. *Given linear contract parameters $\alpha \in \mathbb{R}^2, c \in \mathbb{R}^{|I|\times|I|}$, $P^*(\alpha, c)$ as defined in (4.3) is a subset of the price equilibria of the retailer game given α, c.*

Proof. Let $p^*(\alpha, c) \in P^*(\alpha, c)$. We want to show that p^* is a price equilibrium of the retailer game. For any $k \in \{1, 2\}$,

$$\begin{aligned} p^*_{I_k}(\alpha, c) &\in \operatorname*{argmax}_{p_{I_k} \in \mathbb{R}^{|I_k|}} \sum_{i \in I_k} \pi_i((p_{I_k}, p^*_{I_{-k}}(\alpha, c)))\alpha_k(p_i - c_i) \\ &\subseteq \operatorname*{argmax}_{p_{I_k} \in \mathbb{R}^{|I_k|}} \max\left\{ \sum_{i \in I_k} \pi_i((p_{I_k}, p^*_{I_{-k}}(\alpha, c)))\alpha_k(p_i - c_i), \text{LB} \right\} \\ &= \operatorname*{argmax}_{p_{I_k} \in \mathbb{R}^{|I_k|}} R^w_k((p_{I_k}, p^*_{I_{-k}}(\alpha, c)); \alpha, c). \end{aligned}$$

□

Lemma 13 (Matrix Inversion Lemma, [41])**.** *Let M be an invertible matrix with $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, where A and D are invertible square matrix. Then, the inverse of M can be written as $M^{-1} = \begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$, where*

$$\begin{aligned} W &= (A - BD^{-1}C)^{-1}, \\ Z &= (D - CA^{-1}B)^{-1}, \\ X &= -A^{-1}BZ = -WBD^{-1}, \\ Y &= -D^{-1}CW = -ZCA^{-1}. \end{aligned}$$

Lemma 14 (Woodbury Formula, [48])**.** *Let $A \in \mathbb{R}^{n \times n}$ be an invertible matrix, and let $U \in \mathbb{R}^{n \times k}$, $V \in \mathbb{R}^{k \times n}$, and $C \in \mathbb{R}^{k \times k}$ be matrices. If C is invertible, then the inverse of $A + UCV$ is given by:*

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}.$$

Lemma 15. *Suppose that a diagonally dominant matrix M has negative diagonal entries and positive non-diagonal entries. Then,*

$$M^{-1} \leq 0.$$

Proof. Let $M = [m_{ij}]$ be a diagonally dominant matrix with negative diagonal entries and positive non-diagonal entries. Since M is diagonally dominant, for each row i, we have:

$$|m_{ii}| \geq \sum_{j \neq i} |m_{ij}|$$

Given $m_{ii} < 0$ and $m_{ij} > 0$ for $i \neq j$, this implies:

$$m_{ii} \leq -\sum_{j \neq i} m_{ij}$$

Consider the matrix $N = -M$. The matrix N has positive diagonal entries and negative non-diagonal entries:

$$n_{ii} = -m_{ii} > 0, \quad n_{ij} = -m_{ij} \leq 0 \text{ for } i \neq j$$

The matrix N is strictly diagonally dominant because:

$$n_{ii} = -m_{ii} > \sum_{j \neq i} (-m_{ij}) = \sum_{j \neq i} n_{ij}$$

Since N is strictly diagonally dominant with non-positive off-diagonal entries, it is an M-matrix. It is a known property of M-matrices that if N is an M-matrix [50], then $N^{-1} \geq 0$. Since $N = -M$, we have:

$$N^{-1} = (-M)^{-1} = -M^{-1}$$

Therefore, $-M^{-1} \geq 0$, which implies $M^{-1} \leq 0$. □

BIBLIOGRAPHY

[1] Michael Alles and Srikant Datar. Strategic transfer pricing. *Management Science*, 44:451–461, 1998.

[2] Anil Arya and Brian Mittendorf. Managing strategic inventories via manufacturer-to-consumer rebates. *Management Science*, 59:813–818, 10 2012. doi: 10.1287/mnsc.1120.1626.

[3] Coralio Ballester, Antoni Calvó-Armengol, and Yves Zenou. Who's who in networks. wanted: The key player. *Econometrica*, 74:1403–1417, 2006.

[4] Kostas Bimpikis, Asuman Ozdaglar, and Ercan Yildiz. Competitive targeted advertising over networks. *Operations Research*, 64:705–720, 5 2016.

[5] Francis Bloch and Nicolas Quérou. Pricing in social networks. *Games and Economic Behavior*, 80:243–261, 2013.

[6] Holly Borowski, Jason Marden, and Jeff Shamma. Learning to play efficient coarse correlated equilibria. *Dynamic Games and Applications*, 9, 03 2019.

[7] David Branston. Link capacity functions: A review. *Transportation research*, 10(4):223–236, 1976.

[8] Caltrans. Pems: Freeway performance measurement system. https://pems.dot.ca.gov/, 2024. Accessed: 2024-06-16.

[9] Ozan Candogan, Kostas Bimpikis, and Asuman Ozdaglar. Optimal pricing in networks with externalities. *Operations Research*, 60:883–905, 2012.

[10] Hongyu Chen, Yang Liu, and Yu Marco Nie. Solving the step-tolled bottleneck model with general user heterogeneity. *Transportation Research Part B: Methodological*, 81:210–229, 2015.

[11] Hongyu Chen, Yu Marco Nie, and Yafeng Yin. Optimal multi-step toll design under general user heterogeneity. *Transportation Research Part B: Methodological*, 81:775–793, 2015.

[12] Mei Chen and David H Bernstein. Solving the toll design problem with multiple user groups. *Transportation Research Part B: Methodological*, 38(1):61–79, 2004.

[13] Xin Chen, Qiong Wang, and Juan Xu. Manufacturers' assortment planning and pricing in a competitive two-tier supply chain. Technical report, Department of Industrial and Systems Engineering, University of Illinois at Urbana-Champaign, 2021.

[14] Ying-Ju Chen, Yves Zenou, and Junjie Zhou. Competitive pricing strategies in social networks. *The RAND Journal of Economics*, 49:672–705, 2018.

[15] Yongmin Chen. On vertical mergers and their competitive effects. *The RAND Journal of Economics*, 32:667–685, 2001.

[16] Yongmin Chen and Michael H Riordan. Vertical integration, exclusive dealing, and ex post cartelization. *The RAND Journal of Economics*, 38:1–21, 2007.

[17] S.Chan Choi. Price competition in a duopoly common retailer channel. *Journal of Retailing*, 72:117–134, 1996.

[18] Pengyun Chong, Min Lv, Hao Zhu, and Dong Ding. A scheme to improve network performance based on traffic restriction and pricing in the pres-

ence of carpooling. *Mathematical Problems in Engineering*, 2022:9360726, 1 2022.

[19] George Christodoulou and Elias Koutsoupias. The price of anarchy of finite congestion games. In *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 67–73, 2005.

[20] Chih Peng Chu, Jyh Fa Tsai, and Shou Ren Hu. Optimal starting location of an hov lane for a linear monocentric urban area. *Transportation Research Part A: Policy and Practice*, 46:457–466, 3 2012.

[21] Maxime C Cohen, Michael-David Fiszer, Avia Ratzon, and Roy Sasson. Incentivizing commuters to carpool: A large field experiment with waze. *Manufacturing & Service Operations Management*, 25:1263–1284, 2023.

[22] Maxime C. Cohen, Alexandre Jacquillat, Avia Ratzon, and Roy Sasson. The impact of high-occupancy vehicle lanes on carpooling. *Transportation Research Part A: Policy and Practice*, 165:186–206, 11 2022.

[23] Richard Cole, Yevgeniy Dodis, and Tim Roughgarden. Pricing network edges for heterogeneous selfish users. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 521–530, 2003.

[24] José R Correa, Andreas S Schulz, and Nicolás E Stier-Moses. Selfish routing in capacitated networks. *Mathematics of Operations Research*, 29(4):961–976, 2004.

[25] Thomas M Cover. *Elements of information theory*. John Wiley & Sons, 1999.

[26] Ozgun Caliskan Demirag, Ozgul Baysar, Pinar Keskinocak, and Julie L Swann. The effects of customer rebates and retailer incentives on a man-

ufacturer's profits and sales. *Naval Research Logistics (NRL)*, 57:88–108, 2 2010.

[27] Ozgun Caliskan Demirag, Pinar Keskinocak, and Julie Swann. Customer rebates and retailer incentives in the presence of competition and price discrimination. *European Journal of Operational Research*, 215:268–280, 2011.

[28] Xuan Di, Henry X Liu, Xuegang (Jeff) Ban, and Hai Yang. Ridesharing user equilibrium and its implications for high-occupancy toll lane pricing. *Transportation Research Record*, 2667:39–50, 2017.

[29] Xuan Di, Rui Ma, Henry X. Liu, and Xuegang (Jeff) Ban. A link-node reformulation of ridesharing user equilibrium with network design. *Transportation Research Part B: Methodological*, 112:230–255, 6 2018.

[30] Joakim Ekström, Leonid Engelson, and Clas Rydergren. Optimal toll locations and toll levels in congestion pricing schemes: a case study of Stockholm. *Transportation Planning and Technology*, 37(4):333–353, 2014.

[31] Wei Fan. Optimal congestion pricing toll design for revenue maximization: comprehensive numerical results and implications. *Canadian Journal of Civil Engineering*, 42(8):544–551, 2015.

[32] Wei Fan. Social welfare maximization by optimal toll design for congestion management: models and comprehensive numerical results. *Transportation Letters*, 9(2):81–89, 2017.

[33] Awi Federgruen and Ming Hu. Technical note - sequential multiproduct price competition in supply chain networks. *Operations Research*, 64:135–149, 2016.

[34] Chaim Fershtman and Kenneth L Judd. Equilibrium incentives in oligopoly. *The American Economic Review*, 77:927–940, 1987.

[35] Jeremy Finkleman, Jeffrey Casello, and Liping Fu. Empirical evidence from the greater toronto area on the acceptability and impacts of hot lanes. *Transport Policy*, 18:814–824, 11 2011.

[36] Lisa Fleischer, Kamal Jain, and Mohammad Mahdian. Tolls for heterogeneous selfish users in multicommodity networks and generalized congestion games. In *45th Annual IEEE Symposium on Foundations of Computer Science*, pages 277–285. IEEE, 2004.

[37] Dean Foster and H Young. Regret testing: Learning to play nash equilibrium without knowing you have an opponent. *Theoretical Economics*, 1:341–367, 10 2006.

[38] Dean P. Foster and Rakesh V. Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21(1):40–55, 1997.

[39] Dean P. Foster and H. Peyton Young. Learning, hypothesis testing, and nash equilibrium. *Games and Economic Behavior*, 45:73–96, 10 2003.

[40] Drew Fudenberg and David K Levine. *The theory of learning in games*, volume 2. MIT press, 1998.

[41] Jean Gallier. *Geometric Methods and Applications: For Computer Science and Engineering*, volume 38 of *Texts in Applied Mathematics*. Springer, 2011.

[42] Fabrizio Germano and Gábor Lugosi. Global nash convergence of foster and young's regret testing. *Games and Economic Behavior*, 60:135–154, 7 2007.

[43] Tobias Harks, Ingo Kleinert, Max Klimm, and Rolf H Möhring. Computing network tolls with support constraints. *Networks*, 65(3):262–285, 2015.

[44] Oliver Hart, Jean Tirole, Dennis W Carlton, and Oliver E Williamson. Vertical integration and market foreclosure. *Brookings Papers on Economic Activity. Microeconomics*, 1990:205–286, 1990.

[45] Sergiu Hart and Andreu Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.

[46] HCM. *Highway capacity manual: Practical applications of research*. US Department of Commerce, Bureau of Public Roads, 1950.

[47] Xiang He, Xiqun Chen, Chenfeng Xiong, Zheng Zhu, and Lei Zhang. Optimal time-varying pricing for toll roads under multiple objectives: a simulation-based optimization approach. *Transportation Science*, 51(2):412–426, 2017.

[48] Nicholas J. Higham. *Accuracy and Stability of Numerical Algorithms*. Society for Industrial and Applied Mathematics, second edition, 2002.

[49] Martin Hoefer, Lars Olbrich, and Alexander Skopalik. Taxing subnetworks. In *International workshop on internet and network economics*, pages 286–294. Springer, 2008.

[50] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[51] Alejandro Ortega Hortelano, Jose Manuel Vassallo, and Juan Ignacio Pérez. Optimal welfare price for a road corridor with heterogeneous users. *Transport*, 34(3):318–329, 2019.

[52] Jonathan E Hughes and Daniel Kaffine. When should drivers be encouraged to carpool in hov lanes? *Economic Inquiry*, 57:667–684, 1 2019.

[53] Kitae Jang, Myoung Kyun Song, Keechoo Choi, and Dong-Kyu Kim. A bi-level framework for pricing of high-occupancy toll lanes. *Transport*, 29(3):317–325, 2014.

[54] Lan Jiang and Hani S Mahmassani. Toll pricing: computational tests for capturing heterogeneity of user preferences. *Transportation research record*, 2343(1):105–115, 2013.

[55] Sam Jindani. Learning efficient equilibria in repeated games. *Journal of Economic Theory*, 205:105551, 10 2022.

[56] Sham M. Kakade and Dean P. Foster. Deterministic calibration and nash equilibrium. *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, 3120:33–48, 2004.

[57] Ehud Kalai and Ehud Lehrer. Rational learning leads to nash equilibrium. *Econometrica*, 61:1019, 9 1993.

[58] George Karakostas and Stavros G Kolliopoulos. Edge pricing of multicommodity networks for heterogeneous selfish users. In *FOCS*, volume 4, pages 268–276, 2004.

[59] Hideo Konishi and Se il Mun. Carpooling and congestion pricing: Hov and hot lanes. *Regional Science and Urban Economics*, 40:173–186, 7 2010.

[60] Michael Kopel and Mario Pezzino. 10. strategic delegation in oligopoly. *Handbook of game theory and industrial organization*, 2:248–285, 2018.

[61] Jorge A Laval, Hyun W Cho, Juan C Muñoz, and Yafeng Yin. Real-time congestion pricing strategies for toll facilities. *Transportation Research Part B: Methodological*, 71:19–31, 2015.

[62] Meng Li, Xuan Di, Henry X Liu, and Hai-Jun Huang. A restricted path-based ridesharing user equilibrium. *Journal of Intelligent Transportation Systems*, 24:383–403, 7 2020. doi: 10.1080/15472450.2019.1658525.

[63] Xingyuan Li and Jing Bai. A ridesharing choice behavioral equilibrium model with users of heterogeneous values of time. *International Journal of Environmental Research and Public Health*, 18, 2021.

[64] Yang Liu and Yuanyuan Li. Pricing scheme design of ridesharing program in morning commute problem. *Transportation Research Part C: Emerging Technologies*, 79:156–177, 6 2017.

[65] Yang Liu and Yu Nie. Welfare effects of congestion pricing and transit services in multiclass multimodal networks. *Transportation research record*, 2283(1):34–43, 2012.

[66] Yang Liu and Yu Nie. A credit-based congestion management scheme in general two-mode networks with multiclass users. *Networks and Spatial Economics*, 17:681–711, 2017.

[67] Yingyan Lou, Yafeng Yin, and Jorge A Laval. Optimal dynamic pricing strategies for high-occupancy/toll lanes. *Transportation Research Part C: Emerging Technologies*, 19(1):64–74, 2011.

[68] Chung-Cheng Lu, Xuesong Zhou, and Hani S Mahmassani. Variable toll pricing and heterogeneous users: Model and solution algorithm for bicri-

terion dynamic traffic assignment problem. *Transportation Research Record*, 1964(1):19–26, 2006.

[69] Jing Lu and Carolina Osorio. A probabilistic traffic-theoretic network loading model suitable for large-scale network analysis. *Transportation Science*, 52(6):1509–1530, 2018.

[70] Jing Lu and Carolina Osorio. On the analytical probabilistic modeling of flow transmission across nodes in transportation networks. *Transportation research record*, 2676(12):209–225, 2022.

[71] Shie Mannor, Jeff S. Shamma, and Gürdal Arslan. Online calibrated forecasts: Memory efficiency versus universality for learning in games. *Machine Learning*, 67:77–115, 5 2007.

[72] Jason R. Marden. Selecting efficient correlated equilibria through distributed learning. *Games and Economic Behavior*, 106:114–133, 2017.

[73] Jason R. Marden, Gürdal Arslan, and Jeff S. Shamma. Cooperative control and potential games. *Trans. Sys. Man Cyber. Part B*, 39(6):1393–1407, December 2009a.

[74] Jason R. Marden, GÜrdal Arslan, and Jeff S. Shamma. Joint strategy fictitious play with inertia for potential games. *IEEE Transactions on Automatic Control*, 54(2):208–220, 2009c.

[75] Jason R. Marden and Jeff S. Shamma. Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation. *Games and Economic Behavior*, 75:788–808, 7 2012.

[76] Jason R. Marden, H. Peyton Young, and Lucy Y. Pao. Achieving pareto

optimality through distributed learning. *Proceedings of the IEEE Conference on Decision and Control*, pages 7419–7424, 2012.

[77] Marios Mavronicolas, Igal Milchtaich, Burkhard Monien, and Karsten Tiemann. Congestion games with player-specific constants. In *Mathematical Foundations of Computer Science 2007: 32nd International Symposium, MFCS 2007 Český Krumlov, Czech Republic, August 26-31, 2007 Proceedings 32*, pages 633–644. Springer, 2007.

[78] Dimitra Michalaka, Yafeng Yin, and David Hale. Simulating high-occupancy toll lane operations. *Transportation research record*, 2396(1):124–132, 2013.

[79] Igal Milchtaich. Congestion games with player-specific payoff functions. *Games and economic behavior*, 13(1):111–124, 1996.

[80] Dov Monderer and Lloyd S Shapley. Potential games. *Games and economic behavior*, 14(1):124–143, 1996.

[81] Rafael Moner-Colonques, José J Sempere-Monerris, and Amparo Urbano. Strategic delegation with multiproduct firms. *Journal of Economics & Management Strategy*, 13:405–427, 9 2004.

[82] John H. Nachbar. Prediction, optimization, and learning in repeated games. *Econometrica*, 65:275, 3 1997.

[83] John H. Nachbar. Bayesian learning in repeated games of incomplete information. *Social Choice and Welfare*, 18:303–326, 2001.

[84] John H. Nachbar. Beliefs in repeated games. *Econometrica*, 73:459–480, 3 2005.

[85] Yu Marco Nie and Yang Liu. Existence of self-financing and pareto-improving congestion pricing: Impact of value of time distribution. *Transportation Research Part A: Policy and Practice*, 44(1):39–51, 2010.

[86] Janusz A Ordover, Garth Saloner, and Steven C Salop. Equilibrium vertical foreclosure. *The American Economic Review*, 80:127–142, 1990.

[87] Dario Paccagnan, Rahul Chandan, Bryce L Ferguson, and Jason R Marden. Incentivizing efficient use of shared infrastructure: Optimal tolls in congestion games. *arXiv preprint arXiv:1911.09806*, 2019.

[88] Bary S.R. Pradelski and H. Peyton Young. Learning efficient nash equilibria in distributed systems. *Games and Economic Behavior*, 75:882–897, 7 2012.

[89] Zhen Sean Qian and H. Michael Zhang. Modeling multi-modal morning commute in a one-to-one corridor network. *Transportation Research Part C: Emerging Technologies*, 19:254–269, 4 2011.

[90] Gabriel de O Ramos, Roxana Rădulescu, Ann Nowé, and Anderson R Tavares. Toll-based learning for minimising congestion under heterogeneous preferences. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, pages 1098–1106, 2020.

[91] Robert W Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 2:65–67, 1973.

[92] Tim Roughgarden. *Selfish routing and the price of anarchy*. MIT press, 2005.

[93] Tim Roughgarden. Algorithmic game theory. *Communications of the ACM*, 53(7):78–86, 2010.

[94] Tim Roughgarden and Éva Tardos. How bad is selfish routing? *Journal of the ACM (JACM)*, 49(2):236–259, 2002.

[95] Tim Roughgarden and Éva Tardos. Bounding the inefficiency of equilibria in nonatomic congestion games. *Games and economic behavior*, 47(2):389–403, 2004.

[96] Michael A Salinger. Vertical mergers and market foreclosure. *The Quarterly Journal of Economics*, 103:345–356, 1988.

[97] William H Sandholm. Potential games with continuous player sets. *Journal of Economic theory*, 97(1):81–108, 2001.

[98] Steven D Sklivas. The strategic choice of managerial incentives. *The RAND Journal of Economics*, 18:452–458, 1987.

[99] Zhen Tan and H Oliver Gao. Hybrid model predictive control based dynamic pricing of managed lanes with multiple accesses. *Transportation Research Part B: Methodological*, 112:113–131, 2018.

[100] Terry A Taylor. Supply chain coordination under channel rebates with sales effort effects. *Management Science*, 48:992–1007, 8 2002. doi: 10.1287/mnsc.48.8.992.168.

[101] US Bureau of Public Roads. Traffic assignment manual for application with a large, high speed computer, 1964.

[102] Vincent AC van den Berg. Self-financing roads under coarse tolling and preference heterogeneity. *Transportation Research Part B: Methodological*, 182:102909, 2024.

[103] John Vickers. Delegation and the theory of the firm. *The Economic Journal*, 95:138–147, 1985.

[104] William S. Vickrey. Congestion theory and transport investment. *The American Economic Review*, 59(2):251–260, 1969.

[105] Jing Wang, Hua Wang, and Xiaohua Yu. Parking permit management of morning commuting considering carpooling with parking restraint. *Journal of Transportation Engineering, Part A: Systems*, 146:04020052, 7 2020. doi: 10.1061/JTEPBS.0000368.

[106] Jing Peng Wang, Xuegang (Jeff) Ban, and Hai Jun Huang. Dynamic ridesharing with variable-ratio charging-compensation scheme for morning commute. *Transportation Research Part B: Methodological*, 122:390–415, 4 2019.

[107] Bangyang Wei, Xiang Zhang, Wei Liu, Meead Saberi, and S. Travis Waller. Capacity allocation and tolling-rewarding schemes for the morning commute with carpooling. *Transportation Research Part C: Emerging Technologies*, 142:103789, 9 2022.

[108] Di Wu, Yafeng Yin, and Siriphong Lawphongpanich. Pareto-improving congestion pricing on multimodal transportation networks. *European Journal of Operational Research*, 210(3):660–669, 2011.

[109] Ling Ling Xiao, Tian Liang Liu, and Hai Jun Huang. On the morning commute problem with carpooling behavior under parking space constraint. *Transportation Research Part B: Methodological*, 91:383–407, 9 2016.

[110] Ling-Ling Xiao, Tian-Liang Liu, and Hai-Jun Huang. Tradable permit

schemes for managing morning commute with carpool under parking space constraint. *Transportation*, 48:1563–1586, 2021.

[111] Ling Ling Xiao, Tian Liang Liu, Hai Jun Huang, and Ronghui Liu. Temporal-spatial allocation of bottleneck capacity for managing morning commute with carpool. *Transportation Research Part B: Methodological*, 143:177–200, 1 2021.

[112] Huayu Xu, Fernando Ordóñez, and Maged Dessouky. A traffic assignment model for a ridesharing transportation market. *Journal of Advanced Transportation*, 49:793–816, 11 2015.

[113] Huayu Xu, Jong Shi Pang, Fernando Ordóñez, and Maged Dessouky. Complementarity models for traffic equilibrium with ridesharing. *Transportation Research Part B: Methodological*, 81:161–182, 11 2015.

[114] Chen-Yang Yan, Mao-Bin Hu, Rui Jiang, Jiancheng Long, Jin-Yong Chen, and Hao-Xiang Liu. Stochastic ridesharing user equilibrium in transport networks. *Networks and Spatial Economics*, 19:1007–1030, 2019.

[115] Hai Yang and Hai-Jun Huang. Carpooling and congestion pricing in a multilane highway with high-occupancy-vehicle lanes. *Transportation Research Part A: Policy and Practice*, 33(2):139–155, 1999.

[116] Hai Yang and Hai-Jun Huang. *Mathematical and economic theory of road pricing*. Emerald Group Publishing Limited, 2005.

[117] H. Peyton Young. The evolution of conventions. *Econometrica*, 61:57, 1 1993.

[118] H. Peyton Young. Learning by trial and error. *Games and Economic Behavior*, 65:626–643, 3 2009.

[119] Xiaojuan Yu, Vincent A.C. van den Berg, and Erik T. Verhoef. Carpooling with heterogeneous users in the bottleneck model. *Transportation Research Part B: Methodological*, 127:178–200, 9 2019.

[120] Xiaojuan Yu, Vincent A.C. van den Berg, and Erik T. Verhoef. Carpooling with heterogeneous users in the bottleneck model. *Transportation Research Part B: Methodological*, 127:178–200, 9 2019.

[121] Fangfang Yuan, Xiaolei Wang, and Zhibin Chen. Assessing the impact of ride-sourcing vehicles on hov-lane efficacy and management strategies. *Transport Policy*, 150:35–52, 5 2024.

[122] Guangzhi Zang, Meng Xu, and Ziyou Gao. High-occupancy vehicle lanes and tradable credits scheme for traffic congestion management: A bilevel programming approach. *PROMET - Traffic & Transportation*, 30:1, 2 2018.

[123] Lin Zhong, Kenan Zhang, Yu (Marco) Nie, and Jiuping Xu. Dynamic carpool in morning commute: Role of high-occupancy-vehicle (hov) and high-occupancy-toll (hot) lanes. *Transportation Research Part B: Methodological*, 135:98–119, 5 2020.