Analysis and Design of Information Transmission in Networks of Strategic Agents

by

Nasimeh Heydaribeni

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical and Computer Engineering)
in the University of Michigan
2022

Doctoral Committee:

Associate Professor Achilleas Anastasopoulos, Chair
Professor Tilman Borgers
Associate Professor Ketan Savla
Professor Lei Ying

Nasimeh Heydaribeni

heydari@umich.edu

ORCID iD: 0000-0001-8097-9885

To My Family

Acknowledgments

I am thankful for having the chance to complete this dissertation under the supervision of professor Achilleas Anastasopoulos. I received his continuous support and invaluable guidance throughout my Ph.D., for which I express my deepest gratitude.

I would also like to thank my thesis committee, professor Tilman Borgers, professor Lei Ying, and professor Ketan Savla for their time and valuable feedback and suggestions.

I am also lucky to have had the opportunity of working with my collaborators, Ilai Bistritz and professor Ketan Savla, who helped broaden my research horizon.

Finally, I would like to thank my family: my parents, my sister, Nargess, and my husband, Mahdi, who were my main source of emotional support during the challenging years of my Ph.D. I am also grateful for having the opportunity of making wonderful friends in Ann Arbor who were there for me during this whole journey.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF TABLES

TABLE

LIST OF APPENDICES

ABSTRACT

Most of today's systems consist of strategic/selfish agents with some private information and uncertainty towards others' information and system states. Transmission and exchange of information in such networks have been the focus of many interesting research areas such as mechanism design, information design, and Bayesian learning. The information is directly exchanged in mechanism design and information design, and the goal is to steer agents' actions towards a desirable direction by putting incentives in place (mechanism design) or designing the appropriate information structure (information design). Information can also be spread more indirectly by agents who observe each other's actions (Bayesian learning).

In this thesis, we follow two main directions of "Analysis" and "Design" to investigate the spread of information in networks of strategic agents. Specifically, we analyze dynamic systems with asymmetric information and characterize their equilibria and study the spread of information induced by these equilibrium behaviors. Furthermore, we study how incentives or information structures can be designed to shape the equilibrium behavior of agents.

In part I of this thesis (analysis part), we study structured perfect Bayesian equilibria (structured PBE) in dynamic games with asymmetric information. While there is no general framework to characterize such equilibria, we can study them for some specific information structures. Specifically, we consider games with conditionally independent types. As an example of such games, we study a setting where there is a marketplace with a product that has an unknown value and privately informed agents coming to the market to decide on buying or not buying the product. The agents get multiple chances to enter the market, and in this sense, they act non-myopically. Characterization of structured PBE in this game enables us to analyze informational cascades and suggest settings that avoid such outcomes.

In part II of this thesis (design part), we design distributed mechanisms for efficient resource allocation in networks. The message transmission is done locally in our mechanisms, and we investigate how appropriate information is propagated throughout the network so that the equilibrium outcome is efficient. We also study a joint information and mechanism design problem where agents with private types arrive at a queue with an unobservable backlog. We study how a planner that

observes the queue backlog can design taxes and type-dependent admission signals for the agents to gain the most revenue. We further analyze an information design problem for a non-atomic service scheduling game. We investigate how a planner can give suggestions to users about the time to join a queue for a service with an unknown start time to minimize the social cost.

CHAPTER 1

Introduction

1.1 Motivation and Background

In the last decades, there has been significant research into understanding how agents behave in communication, transportation, energy, economic and societal networks. It has been realized that assumptions such as fully-informed or fully-compliant agents are untenable in vastly decentralized networks. Indeed, most of today's systems comprise of strategic/selfish agents with private information and uncertainty towards others' information and system states. Transmission and exchange of information in such networks have been the focus of many interesting research areas. In "Mechanism Design", appropriate incentives are designed for the agents to incentivize them to directly share some part of their private information in order to achieve a socially optimal objective (e.g., efficient resource allocation). In "Information Design", we study how an agent can affect others' actions by directly sharing some part of her private information with them. In both mechanism design and information design frameworks, the goal is to "steer" agents' actions towards a desirable direction, either by incentivizing them via taxes/subsidies (mechanism design) or by providing appropriate information to them (information design). Another set of related problems is collectively known as "Bayesian Learning". In this research area, we study how information is spread indirectly between agents through observing each others' actions. In all such problems, the interaction between players is characterized through notions of equilibria. Therefore, studying the equilibria in games will enable us to further analyze the players' behavior.

In this thesis, we have utilized tools from these frameworks to gain a better understanding of strategic agents' behavior and the spread of information in complex systems, which are modeled as dynamic games with asymmetric information. Examples of such systems can be found in network problems such as resource allocation in networks, queuing systems, informational cascades, etc. We present two main approaches to study such problems. One is an analytical approach where we study and analyze the spread of information in different settings. That is, we study and analyze equilibria and the resulted behavior of agents in such problems. The other approach is from a designing point of view, where we utilize mechanism design and information design techniques to study such problems. Therefore, the problems discussed in this thesis can be categorized into two main parts, namely “Analysis” and “Design”.

1.2 Analysis of Equilibria in Dynamic Games with asymmetric information

In order to investigate information transmission and learning in strategic environments, one needs to study the strategic interaction between players, which is usually formalized with notions of equilibria. In this thesis, we study dynamic games with asymmetric information, for which the appropriate notion of equilibrium is perfect Bayesian equilibrium (PBE) [1, 2, 3]. There is no unified framework to characterize PBE due to its complexity in general. Hence, we restrict our focus to structured equilibria (which are akin to Markov policies in stochastic control) [4, 5]. The main challenge in characterizing structured PBE in dynamic games is finding the appropriate summary according to which players make decisions. These summaries usually include some type of belief over the unknown states of the game.

Specifically, in games with asymmetric information, finding such summaries is more complex due to the emergence of private beliefs and the need to form beliefs over beliefs, which could also be a private quantity and therefore, create an infinite hierarchy of private beliefs. Some specific information structures enable us to avoid the infinite hierarchy of beliefs. In this thesis, we study structured PBE for dynamic games with asymmetric information where agents’ types are conditionally independent given an unknown state of the world. We show summaries including a private belief and a joint public belief over the private beliefs are valid summaries for structured PBE. Essentially, we show that the hierarchy of beliefs stops at the second step due to the defined belief over (private) beliefs being public. In addition, we specialize our results to dynamic linear quadratic Gaussian (LQG) games where instead of beliefs, we form estimates over unknown states of the game due to the beliefs being Gaussian.

1.3 Analysis of Bayesian Learning and Informational Cascades

In large decentralized networks, there is usually some unknown (possibly dynamically changing) state of the world. While the agents have their own private information about the unknown state, they can also become more informed by observing others’ actions in the network. Consider a market with a new product with unknown quality. Each person has his/her observation of the quality of the product. It is intuitively clear that if one gathers all the people’s information about this product, one can get a better idea about its quality. In a strategic setting where agents may not be willing to reveal their private information, people can get information and learn about others' observations through their actions (whether they have bought the product or not). Learning the unknown state of the world can be beneficial to the agents since it enables them to make better decisions. However, there may be situations where this learning stops due to the fact that nobody has an incentive to act based on his/her private information; these situations are referred to as informational cascades. Avoiding these situations may result in more beneficial outcomes for the whole network. Sequential learning has been extensively explored in the literature, with a special focus on informational cascades [6, 7, 8, 9, 10]. In two seminal papers [6, 7] the authors investigated the occurrence of fads in a social network, which was later generalized in [8]. In this thesis, we study Bayesian learning in a dynamic game where players have multiple opportunities to decide on buying a product with unknown quality. In this sense, we allow the players to act non-myopically and investigate the occurrence of information cascades. We present results showing that while information cascades seem to be unavoidable in most situations, the non-myopic behavior of the players can help avoid them for some specific parameters of the game.

1.4 Distributed Mechanism Design

Achieving a socially optimal objective is often formulated as a centralized optimization problem where a central authority allocates scarce resources so that the sum of users' utilities (social welfare) is maximized. Versions of this problem have been extensively studied in the literature in the last few decades. Technological advances have enabled us to solve such optimization problems involving thousands of users. There is, however, a fundamental difficulty in solving such problems that cannot be overcome by technology: in large networks with heterogeneous and strategic agents with privacy constraints, agents may not be willing to reveal their private information (utility), which is related to the optimization problem. This fundamental problem has been addressed by the theory of "Mechanism Design" [11, 12]. Mechanism design is the design of appropriate incentives that, once in place, incentivize agents to share some of their private information in a way that enables achievement of the socially optimal objective. Mechanism design has been widely utilized in such areas of research as market allocations [13, 14, 15], rate and resource allocations [16, Sec. 2.7][17, 18, 19, 20, 21], spectrum sharing [22, 23, 24], data security [25], power allocation in wireless networks [26, 27], demand-side management in the power grid [28, 29, 30], etc.

In the standard mechanism design framework (Hurwicz-Reiter [1]), agents transmit messages to a central authority. The central authority, upon receiving all these messages, determines allocation and taxes/subsidies for the agents of the network. Clearly, this arrangement may result in a significant communication overhead due to the message transmission of agents to the central authority. In this thesis, we study the more realistic scenario where such message transmission to a central authority cannot take place due to network communication constraints. To investigate this problem, we consider a setting in which agents are only allowed to transmit their messages to their local neighborhood. Consequently, each agent can determine her allocation and tax based on the messages she hears locally, and therefore, there is no need for a central authority to evaluate these functions. This additional restriction is the focus of a research direction that we call "Distributed Mechanism Design" (DMD).

A complementary view of DMD stems from the literature of "Distributed Optimization" (e.g., [31, 32, 33, 34, 35, 36, 37]) where agents do exchange local messages in order to solve a centralized allocation problem. In distributed optimization, however, it is assumed that agents are not strategic. In fact, they are automata and execute a predefined message exchange algorithm. Our work thus aims at enriching the theory of Distributed Optimization by incorporating agents that are strategic and have privacy constraints. In this sense, the results presented in this thesis constitute a step towards the ultimate goal to combine two areas of research, Mechanism Design and Distributed Optimization, into a unified theory of Distributed Mechanism Design.

1.5 Information Design

In information design, there is one agent, the information designer, who has some valuable information about an unknown payoff-related state of the world. Other agents in the system are interested to know the information about the state of the world to be able to make more informed decisions and earn higher payoffs. However, the information provider investigates how he can "steer" other agents' actions towards his own interests by wisely providing some information about the state of the world to other agents. The goal of the information designer is to align the preferences of other agents with his own preferences as much as possible by designing the information structure for the agents. Information design problems with one sender and one receiver are referred to as "Bayesian persuasion" introduced in [38], where authors present a geometric form of interpreting information design and when it is profitable for the designer not to share some part of the information. However, when there are multiple receivers, the information design problem becomes more complex, and notions of equilibria must be introduced to analyze the game.

In this thesis, we study two information design problems with multiple receivers. We focus on a different aspect of information design in each of the problems studied. In the first problem, we have multiple receivers, and they have private types. Therefore, we study a joint mechanism design and information design problem similar to what is done in [39]. In our model, all agents have some private information: the planner (information designer) has private information about the state of the world, and the network agents possess private types related to their preferences. The information provided to the agents can be considered as a resource that is to be allocated to the agents. Therefore, mechanism design tools can be used to design tax functions to incentivize agents to reveal their type. As a result of this design, the information provider can provide different amounts of information for different types of agents to maximize his revenue. In the second problem, we have a continuum of receivers who need a service with an unknown start time. An informed planner will make suggestions about the time to join the queue to minimize the social cost. The continuous aspect of the receivers makes studying the information design problem more complex and challenging but gives more insight into the real-world population of information receivers.

1.6 Thesis Overview

In this thesis, we consider five different problems, each described in a separate chapter. Chapters 2 and 3 constitute the analysis part of the thesis (part I), and chapters 4, 5, and 6 are the three chapters under the design approach (part II).

In chapter 2, we present our work "Structured equilibria for dynamic games with asymmetric information and dependent types" [40, 41, 42]. In this work, we have studied a dynamic game with asymmetric information. There is a state of the world with an unknown value, $V \in \mathbb{R}$. At each time, all players get noisy private observations of V that are independent conditioned on V or, in other words, dependent to each other through V. We characterize structured perfect Bayesian equilibrium for this game. In order to find the structure of the equilibrium, we need to define summaries, or sufficient statistics, for the history of the game. A quantity commonly used as a sufficient statistic is a belief over the unknown state of the world V. The main challenge in this context is the emergence of private beliefs in sufficient statistics, i.e., the fact that different agents in the system have different (private) observations of V and, therefore, form private beliefs over it. One way to avoid this problem is to consider models in which private beliefs either do not exist (symmetric information games, or asymmetric but independent observations [4, 43, 44]) or, if they exist, they are not taken into account in agents' strategies (see for example the concept of "public perfect equilibrium" [45]). In order to intuitively explain the conceptual difficulty arising from having private beliefs in sufficient statistics, consider the following thought process. If a player acts according to her private belief ξ_t^i of a hidden variable and she expects other players to behave in the same way, she needs to form a belief over other players' beliefs to interpret and predict their actions, and she has to take that belief into account when acting. In other words, she has to form a belief over (at least) ξ_t^j for all other users $j \neq i$. This is a belief on beliefs which is also private information of user i, and it has to be taken into account in her strategies. Due to the symmetry of the information structure, all other players should do the same. Hence, it is clear that user i needs to form beliefs over beliefs over beliefs of other players. This chain continues as long as these beliefs are private. It stops whenever the beliefs in one step become public or public functions of previous step beliefs. In chapter 2, we show that for a game with conditionally independent types, this chain of beliefs stops at the second step. Using this result, we characterize structured PBE for the considered game. We then specialize our results to a dynamic LQG game where the aforementioned beliefs are Gaussian, and therefore, players only keep track of estimates and covariance matrices instead of beliefs.

In chapter 3, we present our work "Bayesian learning with non-myopic agents" [46, 47, 48]. This work was done in collaboration with Ilai Bistritz, and the modeling, problem formulation, and derivation of some of the results were mainly done by him [49]. In this work, we consider a setting where there is a product with an unknown value, $V \in \{0, 1\}$. There is a finite number of players that enter a market one at a time according to a random exogenous process. The player that has entered the market has to decide to buy or not buy the product, but he might get other chances in the future to decide on buying or not buying. In this sense, the players in our model are non-myopic, and they take the future into account for their decisions. Each player has a noisy private observation of V, which can be revealed to others when taking an action. The private observations are generated independently conditioned on V. In other words, they are dependent to each other through V. Players update their beliefs on V once someone reveals her private information. We characterize equilibria for this game which is a dynamic game with asymmetric information, for which an appropriate notion of equilibrium is the PBE. In this framework, we investigate whether informational cascades can happen or not. In other words, we investigate whether non-myopic behavior can/cannot avoid informational cascades. We have shown that although informational cascades are inevitable for most of the game parameters, they can still be avoided for some of the model parameters due to the non-myopic behavior of agents.

In chapter 4, we present our work "Distributed mechanism design for network resource allocation problems" [50, 51, 52]. In this work, we consider two network resource allocation problems, formulated as network utility maximization (NUM) problems. In particular, we consider two concrete examples of resource allocation, namely, data rate allocation for unicast transmission networks and for multicast transmission networks. The NUM problem associated with the unicast protocol is as follows.

$$\max_{x} \sum_{i \in \mathcal{N}} v_i(x_i) \tag{1.1a}$$

$$\text{s.t.} \quad x_i \geq 0 \quad \forall i \in \mathcal{N} \tag{1.1b}$$

$$\text{and} \quad \sum_{j \in \mathcal{N}^l} x_j \leq c^l \quad \forall l \in \mathcal{L}. \tag{1.1c}$$

Utility function $v_i(\cdot)$ is private information of player i, and therefore, one can not solve the above maximization problem directly. Therefore, we propose a distributed mechanism to achieve efficient rate allocation for each of the mentioned protocols. To model the distributed aspect of our mechanism, we utilize a message transmission network through which agents' messages will be exchanged. Therefore, we have two layers of networks, one for data transmission and one for message transmission related to the distributed mechanism, as it is shown in Fig. 1.1.

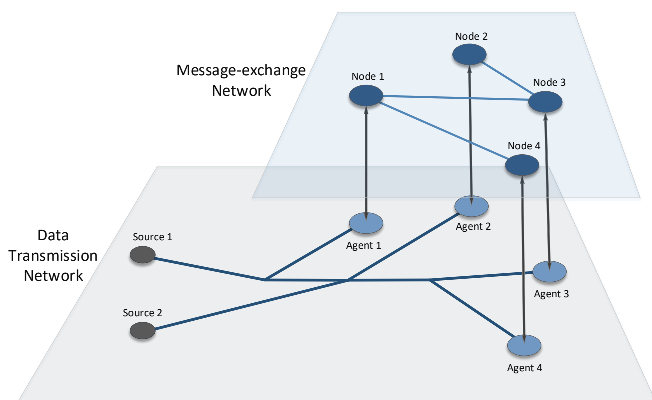


Figure 1.1: Message-exchange network vs. Data transmission network

Chapter 5 presents our work "Joint Information and Mechanism Design for Queues with Heterogeneous Users" [53]. We consider a queue with an unknown backlog to the incoming traffic. There is a planner that observes the queue backlog and sends admission signals to the agents arriving at the queue. The agents have private types that affect how much they value being serviced in the queue and how much they lose by waiting in the queue. We utilize mechanism design techniques to design tax functions to incentivize agents to report their types truthfully. The planner then optimizes the information he shares with each type of agent through the admission signal to maximize his revenue. The order of the actions is as follows. An agent arrives at the queue. She decides to either join the mechanism, which means she will be admitted by the planner, or choose the outside option, which means she has to decide to join or leave the queue without any information about the queue backlog. If she chooses to join the mechanism, she has to report her type, and then she will receive the admission signal (according to which she is either admitted to the queue or not). Also, she has to pay a tax. We characterize tax functions, investigate whether the planner prefers all types of users to join the mechanism or not, and study the structural properties of the optimal admission policy.

The extended form of this game is depicted in the figure below.

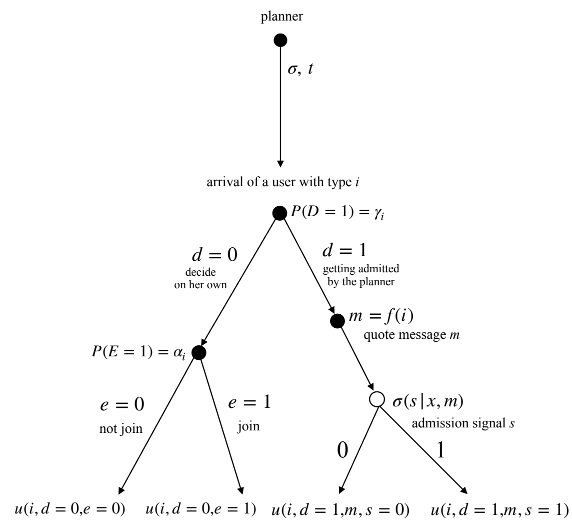


Figure 1.2: Extended form of the game faced by each user at the arrival time.

In chapter 6 we present our work "Information Design for a Non-atomic Service Scheduling Game" [54]. We study an information design problem where we have a continuum of user population as our receivers. A service scheduling game is analyzed where the users do not know the service start time, but there is an informed planner who sends suggestions to the users about when to join the queue. The objective of the planner is to minimize the social cost of users among all obedient direct signals. The cost of a user contains the time she spends in the queue and the difference between when she gets serviced and the service start time.

The proofs and complimentary discussions related to each chapter are relegated to the Appendices at the end of this thesis.

1.7 Contributions

The contributions of this thesis are as follows.

- In chapter 2, we characterize structured PBE for a dynamic game with asymmetric information and dependent types. The main contribution of this work is to show that, due to the conditional independence of the private signals given V, the private belief chain stops at the second step, and players' beliefs over others' beliefs are public functions of their own beliefs (the first step beliefs). In the LQG model, we further show that the beliefs are Gaussian and hence, are characterized by their mean and covariance matrix. Furthermore, the players' estimations over others' estimations are public linear functions of their own estimations. We hypothesize (and eventually prove) structured PBE with strategies for user i being linear in $\widehat{V}_t^i$, the private estimate of V by user i, generated by a (private) Kalman filter. We show that the equilibrium strategies can be characterized by an appropriate backward sequential decomposition algorithm akin to dynamic programming. In the LQG model, the main difference of our work from the standard stochastic control LQG framework is that the forward recursion that evaluates covariance matrices cannot be performed separately as it depends on the equilibrium strategies. A unique feature of our development is the requirement to update in a forward manner additional quantities that are observation dependent (public actions). This precludes off-line evaluation of these forward-updated quantities and necessitates their inclusion as part of the state of the above mentioned backward sequential decomposition.

- In chapter 3, we study and characterize the equilibrium of the proposed dynamic game with asymmetric information. We provide existence results and some structural properties for the equilibrium of the game. Further, we show that there exists an equilibrium with infinitely patient players, in which we can avoid bad informational cascades in some states of the world. It is shown in [49] that for not infinitely patient players, the probability of a cascade approaches one as the number of players, N, approaches infinity. Moreover, it is shown that the number of players who have revealed their information before the cascade occurs is small - which formalizes their inefficiency. These results show that in order to avoid informational cascades, we may not only allow the players to be non-myopic, but we should also allow them to be infinitely patient.

- In chapter 4, we design two distributed mechanism design for efficient data rate allocation in unicast and multicast transmission networks. The mechanisms are distributed, and there is no need for a central authority to collect the messages and determine allocation and taxes. They also fully implement the efficient allocation in their Nash equilibria, i.e., there are no extraneous non-efficient equilibria in the induced game. They are individually rational and weak budget balance. Further, the message space grows linearly with the number of agents in the network. Although we present our method for two examples of unicast transmission protocol and multicast multirate transmission protocol, the mechanisms can be used in many network resource allocation problems.

- In chapter 5, we propose a joint information and mechanism design technique for queues with heterogeneous users. We formulate the mechanism to ensure dominant strategy incentive compatibility. We also investigate whether the planner prefers all types of users to join the mechanism or not. We further characterize the planner's optimization problem that he uses to maximize his revenue. We show some structural results for the optimal admission policy of the planner through analytical reasoning that are also supported by numerical analysis. We observe how the planner discriminates between different types of players in providing information for them in order to gain more revenue. We also study the two extreme cases of full and no information strategies.

- In chapter 6, we formulate an information design problem for a service scheduling game consisting of non-atomic agents. We characterize the equilibrium in full information and no information extremes. We show some results on when the planner can do no better than revealing the full information to the agents. We impose some assumptions on our model that will allow us to express the information design problem as a generalized problem of moments (GPM) [55]. We use the computation tools for these problems such as Gloptipoly [56] to numerically solve the information design problem.

Part I

Analysis

CHAPTER 2

Structured Equilibria for Dynamic Games with Asymmetric Information and Dependent Types

2.1 Introduction

Dynamic games with asymmetric information play an important role in decision and control problems, yet there is no general framework to study such games in a tractable manner. The appropriate solution concept for these games is some notion of equilibrium such as Bayesian Nash equilibrium, perfect Bayesian equilibrium (PBE), sequential equilibrium, etc. [1, 2, 3]. Due to the dynamic nature of such games, the players' histories expand with time and therefore the corresponding strategies have an expanding domain. To mitigate this problem, researchers have introduced equilibrium concepts that summarize the time expanding histories into sufficient statistics. For symmetric information games, Markov perfect equilibria [57] have been introduced, in which the players' strategies depend only on payoff-relevant past events and not the whole history. For asymmetric information games or control problems, finding the appropriate sufficient statistic is a challenging task and various information structures and corresponding statistics have been considered in the literature [58, 4, 59, 60, 61, 62, 63].

A quantity commonly used as a sufficient statistic, is a belief over some unknown part of the system. The main challenge in this context is the emergence of private beliefs in the sufficient statistics, i.e., the fact that different agents in the system may have different (private) observations about the same quantity. One way to avoid this problem is to consider models in which private beliefs either do not exist (symmetric information games, or asymmetric but independent observations [4, 59, 64]) or, if they exist, they are not taken into account in agents' strategies (see for example the concept of "public perfect equilibrium" [45]). In order to intuitively explain the conceptual difficulty arising from having private beliefs in the sufficient statistics, consider the following thought process. If a player i acts according to her private belief ξ_t^i of a hidden variable and she expects other players to behave in the same way, she needs to form a belief over other players' beliefs to interpret and predict their actions and she has to take that belief into account when acting. In other words, she has to form a belief over (at least) ξ_t^j for all other users $j \neq i$. This is a belief on beliefs which is also a private information of user i and it has to be taken into account in her strategies. Due to symmetry of the information structure, all other players should do the same. But now, it is clear that user i needs to form beliefs over beliefs over beliefs of other players. This chain continues as long as this hierarchy of beliefs are private. It stops whenever the beliefs in one step are public or public functions of previous step beliefs.

In this chapter, we study a dynamic game with asymmetric information. We consider a model with an unknown state of the world V, where each player i has a private noisy observation X_t^i of it at each time t. The private observations of players are conditionally independent given V. We then specialize this setting to the case of a Linear Quadratic Gaussian (LQG) non-zero-sum game where V is a Gaussian random variable and players' observations are generated through a linear Gaussian model from V. Our LQG model closely follows that of [64] with one important difference: the private observations of players in [64] are independent where in our case, they are dependent through V; in particular they are conditionally independent given V.

The intuitive reason behind studying a model with dependent private observations is that in today's complex networks, agents are well connected to each other and each agent in the network is affected by other agents and also by many unknown system quantities. Although the independence assumption simplifies model analysis, it cannot always capture the ever connected aspect of today's networks. Our model can also be thought of as a generalization of the one in [65] where V models the value of a product (or a technology) and agents receive a noisy private signal about it and decide whether to adopt it or not, with the important difference that we allow multiple agents to act simultaneously and, unlike [65], we also allow them to return to the marketplace at each time instance and receive a new observation on V. A real-world application of such a model can be seen in product promotions in social networks where there is a product with unknown quality, V, and the users obtain private noisy observations of the product value (e.g., by receiving free samples or asking around about the product). Players' actions relate to how much they want to promote the product (e.g., by advertising it in social networks or writing online reviews, etc.). Depending on the reward functions, we can have different types of players. For instance, some of them may work for a competing company and have malicious intentions towards that product, while others may have the intention to help the community make more informed decisions and they promote what they think has good quality. Another example can be a security game, where V is the unknown security status of the network, and users get private observations about V by privately "poking" the system. Players act by trying to use the system based on their knowledge of its security status (e.g., requesting services or launching attacks), while at the same time learning about the security status. Similarly, we can model both malicious and not malicious players by defining appropriate reward functions (e.g., non-malicious users utilize the system more if they think it is secure, while malicious ones utilize it more if they think it is not secure).

One of the contributions of this work is to show that, due to the conditional independence of the private signals given V, the private belief chain stops at the second step and players beliefs over others' beliefs are public functions of their own beliefs (the first step beliefs). In the LQG model, we further show that the beliefs are Gaussian and hence, are characterized by their mean and covariance matrix. Furthermore, the players' estimation over others' estimations are public linear functions of their own estimations. We hypothesize (and eventually prove) structured PBE with strategies for user i being linear in $\widehat{V}_t^i$, the private estimate of V by user i, generated by a (private) Kalman filter. This is the second contribution of this work.

We show that the equilibrium strategies can be characterized by an appropriate backward sequential decomposition algorithm akin to dynamic programming. In the LQG model, the main difference of our work from the standard stochastic control LQG framework is that the forward recursion that evaluates covariance matrices cannot be performed separately as it depends on the equilibrium strategies. This was also the case in [64]. A unique feature of our development is the requirement to update in a forward manner additional quantities that are observation dependent (public actions). This precludes off-line evaluation of these forward-updated quantities and necessitates their inclusion as part of the state of the above mentioned backward sequential decomposition. This is the third contribution of this work.

2.1.1 Literature Review

In this section we give an overview of the related literature with a focus on the information structures. In [66], a framework, called precedence diagram, was introduced to characterize the information structures in team problems with asymmetric information. The evolving (dynamic) information of the decision makers is modeled by a different (new) controller making a decision at each time with the specific information corresponding to that time available to her. The authors have also provided some examples of the dynamic team problems, one of which is LQG team problem with nested information structure and have proved optimality of linear controllers. The specific information structure considered, nested information, allows the authors to form an equivalent static team problem for the dynamic model considered and hence, avoiding further challenges of dealing with dynamic models.

LQG models have been studied extensively for decision and control problems. In the simplest instance of a single centralized controller it is well known that there is separation of estimation and control, posterior beliefs of the state are Gaussian, a sufficient statistic for control is the state estimate evaluated by the Kalman filter, the optimal control is linear in the state estimate, and the required covariance matrices can be calculated offline [67]. Although it is known that, in general, linear controllers are not optimal in LQG team problems [68], as we mentioned, some information structures have been identified for which linear controllers are shown to be optimal such as the works with nested information structure [66], stochastically nested information structure [61] and partial history sharing information structure [60]. Private beliefs do not emerge in these models because of the specific information structure considered. In the nested information structure, there is no need to form beliefs to interpret the action of the predecessors because the decision maker already knows their information. In the model considered in [60], the decision makers have local memory (not perfect memory) and the authors have not defined any summaries for the history and therefore, beliefs and hence, private beliefs are not introduced.

In order to capture the strategic behavior of agents, dynamic decision problems have also been considered in the context of dynamic games and there is extensive literature on dynamic games with asymmetric information. In [69], the author considers a delayed observation sharing model where all of the previous private observations are shared with all of the players and the asymmetry of the information is only due to the private observations at current time. This specific information structure avoids the private beliefs in the sufficient statistics because they can be formed by augmenting the public belief by the current private observation. One-step delayed information sharing is also used in [70]. Similarly, in [65, 49, 46, 47], there is a public belief that can be augmented by the players' static private signals, to form the private beliefs.

Authors in [71] have used the common information approach, which breaks the history into the common and private parts and similarly, two partial strategies are introduced. One is applied to the private part of the history and the other one generates the first one based on the public part of the history. Finding the strategy that is generated based on the public part of the history does not have the challenges of asymmetric games because the public part of the history is common between all players. The solution concept used is called common information based Markov perfect equilibria. Note that in [71], the private part of the history is not summarized into any other quantity, and therefore, no private beliefs had to be defined. A similar approach is used in [4].

In [64], authors have considered a multi-stage LQG game and characterized a signaling equilibrium which is linear in agents' private observations. In addition, a backward sequential decomposition was presented for the construction of the equilibrium, based on the general development in [59]. In this work, the private observations are independent across agents and therefore there are no private beliefs in the game. This is because a player's belief over others' private observations is independent of her private observation and hence, the belief is public.

A number of works consider LQG games where information available to some players is affected by the decision of others. The works of [72] on strategic information transmission, and [73] on Gaussian cheap talk consider two-stage games and focus on Bayesian Nash equilibria. These works, however, consider games that are not dynamic. This implies that there is no need to search for the sufficient statistics and no private belief will be defined. The classic work on Bayesian persuasion [38], and the related one on strategic deception [74] consider two-stage and multi-stage games, respectively, and focus on (sender preferred) subgame perfect equilibria owing to the fact that strategies (as opposed to only the actions) of the sender are observed. Although the authors of [74] consider a dynamic game, they do not summarize the history into time invariant quantities and they search for the strategies over the whole time horizon. Therefore, although the problem becomes intractable for large time horizons, the issue of private beliefs does not appear.

The unique feature of this work is that we consider dependent private observations (specifically, conditionally independent on a hidden state of the world) between agents, in conjunctions with strategies with time-invariant domains, and so sufficient statistics (beliefs) are defined. As a result, we are forced to deal with private beliefs and the aforementioned issue of the infinite sequence of beliefs on beliefs has to be resolved. This is what makes the considered model interesting and more challenging compared to the previous works.

Games with asymmetric information are also studied in the context of hypergames [75]. In hypergames, players play different games (in an 1-level hypergame), and they have different perceptions towards each others' games (in a 2-level hypergame) and so on. This is similar to the private belief hierarchy that we study in this chapter. However, we study a Bayesian game where although players have different perceptions and uncertainty towards other players preferences, they are playing the same game and we deal with the uncertainty by considering average utility maximizing players. Furthermore, we do not impose a fixed level on beliefs over beliefs that each player can have, as opposed to hypergames where the level of the game is a fixed quantity.

The remaining part of the chapter is structured as follows. In section 2.2 the general model is described. section 2.3 is a review of the solution concept that we have considered in this chapter. We develop our main results in section 2.4. In section 2.5, we describe the special case of the model that is an LQG game, followed by the development of a concrete example in section 2.6 together with numerical results. We discuss some extensions for the model studied in this chapter in section 2.7 and we conclude in section 2.8. Most of the proofs of theorems and lemmas are relegated to Appendix A at the end of this thesis.

2.1.2 Notation

We use upper case letters for scalar and vector random variables and lower case letters for their realizations. We use the notation $\mathbb{P}(a|b)$ to denote the probability $\mathbb{P}(A = a|B = b)$ for discrete random variables and to denote $f_{A|B}(a|b)$, i.e., the probability density function of A at a given $B = b$ for continuous random variables. The superscripts in the probability distributions and expectations such as $\mathbb{P}^g$ and $\mathbb{E}^g$ indicate the strategy according to which the probability distributions are defined. We also use subscripts in the expectation operator as $\mathbb{E}^g_\mu$ to indicate the belief according to which the expectation is calculated. Bold upper case letters are used to denote matrices. Subscripts denote time indices and superscripts represent player identities. The notation $-i$ denotes the set of all players except i. All vectors are column vectors. The transpose of a matrix $\mathbf{A}$ (or vector) is denoted by $\mathbf{A}'$. We use semicolons "; " for vertical concatenation of matrices (or vectors). For any vector (or matrix) with time and player indices, a^i_t (or $\mathbf{A}^i_t$), a^{-i}_t denotes the vertical concatenation of vectors (or matrices) $a^1_t, a^2_t, \ldots, a^{i-1}_t, a^{i+1}_t, \ldots$. Further, $a^i_{1:t}$ means $(a^i_1, a^i_2, \ldots, a^i_t)$. In general, for any vector with time and player indices, a^i_t, we remove the superscript to show the vertical concatenation of the whole vectors and we remove the subscript to show the set of all vectors for all times. The matrix of all zeros with appropriate dimensions is denoted by $\mathbf{0}$ and the identity matrix of appropriate dimensions is denoted by $\mathbf{I}$. For two matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathfrak{D}(\mathbf{A}, \mathbf{B})$ represents the block diagonal concatenation of these matrices, i.e., $\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$ (it applies for any number of matrices). By $\mathfrak{D}(\mathbf{A}^{-i})$, we mean the block diagonal concatenation of matrices $\mathbf{A}^j$ for $j \in -i$. Further, $\mathrm{qd}(A; B)$ represents $B'AB$. For the equation $\begin{bmatrix} \tilde{a}\,; & \tilde{b}\,; & \tilde{c} \end{bmatrix} = \mathbf{A} \begin{bmatrix} a\,; & b\,; & c \end{bmatrix}$, the notation $(\mathbf{A})_{\tilde{a},b}$ denotes the intersection of the rows of $\mathbf{A}$ corresponding to $\tilde{a}$ and the columns that are multiplied by b. Note that both $\tilde{a}$ and b are row vectors. We use " : " for either of the row or column subscripts to indicate the whole rows or columns, e.g., $(\mathbf{A})_{:,b}$ denotes the columns of $\mathbf{A}$ that are multiplied by b. The trace of matrix $\mathbf{A}$ is denoted by $\mathrm{tr}(\mathbf{A})$. We use $\delta(\cdot)$ for the Dirac delta function. We denote the normal distribution with mean vector m and covariance matrix $\mathbf{\Sigma}$ by $\mathrm{N}(m, \mathbf{\Sigma})$. We use square brackets for mappings that produce functions, e.g., $F[a]$ is a mapping that takes a as its input and produces a function. For any Euclidean set $\mathcal{S}$, $\Delta(\mathcal{S})$ represents the space of all probability measures on $\mathcal{S}$. We use $Supp(\sigma)$ to denote the support of the probability distribution σ. To keep the expressions of integrals compact, we drop the infinitesimal variables and only present the integral variables in the integral signs.

2.2 Model

We consider a discrete time dynamic system with N strategic players in the set $\mathcal{N} = \{1, 2, ..., N\}$ over a finite time horizon $\mathcal{T} = \{1, 2, ..., T\}$. There is a static unknown state of the world $V \sim Q_V(\cdot)$. Each player has a private noisy observation X_t^i of V at every time step $t \in \mathcal{T}$. At time t, player i takes action $a_t^i \in \mathcal{A}^i$ which is observed publicly by all players. The private observations are generated according to the kernel $X_t^i \sim Q_X^i(\cdot|V, A_{t-1})$ and they are independent across agents given V and A_{t-1}, i.e.,

$$\mathbb{P}(X_t|V, A_{1:t-1}, X_{1:t-1}) = \mathbb{P}(X_t|V, A_{t-1}) = \prod_{i \in \mathcal{N}} Q_X^i(X_t^i|V, A_{t-1}) \tag{2.1}$$

The kernels Q_V and Q_X^i are known to all of the players. We assume that players have perfect recall and we can construct the history of the system at time t as $h_t = (v, x_{1:t}, a_{1:t-1}) \in \mathcal{H}_t$ and the information set of player i at time t as $h_t^i = (x_{1:t}^i, a_{1:t-1}) \in \mathcal{H}_t^i$. At the end of time step t, each player i receives the reward $r_t^i(v, a_t)$. We assume that the reward functions are known to all players, but the value of the rewards are not observed by the players until the end of the time horizon.

Let $g^i = (g_t^i)_{t \in \mathcal{T}}$ be a probabilistic strategy of player i, where $g_t^i : \mathcal{H}_t^i \to \Delta(\mathcal{A}^i)$, meaning that player i's action at time t is generated according to the distribution $A_t^i \sim g_t^i(\cdot|h_t^i)$. The strategy profile of all players is denoted by g. For the strategy profile g, player i's total expected reward is

$$J^{i,g} := \mathbb{E}^g \left\{ \sum_{t=1}^{T} r_t^i(V, A_t) \right\}, \tag{2.2}$$

and her objective is to maximize her total expected reward.

2.3 Solution concept

We can model this system as a dynamic game with asymmetric information and an appropriate solution concept for such games is Perfect Bayesian Equilibrium (PBE). A PBE consists of a pair (β, μ) (an assessment) of strategy profile $\beta = (\beta_t^i)_{t \in \mathcal{T}, i \in \mathcal{N}}$ and belief system $\mu = (\mu_t^i)_{t \in \mathcal{T}, i \in \mathcal{N}}$ where $\mu_t^i : \mathcal{H}_t^i \to \Delta(\mathcal{H}_t)$ satisfies Bayesian updating and sequential rationality holds. Bayesian updating includes both on- and off-equilibrium histories[1]. This condition requires the beliefs to be Bayesian updated, if possible, given any history, whether that history is on equilibrium or off equilibrium [76, 3]. To be more specific, given an information set h_t^i, which could be on

[1]On-equilibrium histories are the ones that have positive probability of occurrence under equilibrium strategies and similarly, off-equilibrium histories are the ones with zero probability of occurrence under equilibrium strategies [76].

or off-equilibrium, and for each h^i_{t+1}, the beliefs should be updated according to Bayes rule if $\mathbb{P}^g(h^i_{t+1}|h^i_t) > 0$, where we define $\mathbb{P}^g(h^i_{t+1}|h^i_t)$ as follows.

$$\mathbb{P}^g(h^i_{t+1}|h^i_t) = \int_{(h_{t+1}:h^i_{t+1})\backslash h^i_{t+1}} \int_{h_t} P^{g_t}(h_{t+1}|h_t)\mu^i_t(h_t|h^i_t), \tag{2.3}$$

and we define P^{g_t} to be the kernel that describes the transition probability from h_t to h_{t+1}. We also define $h_{t+1} : h^i_{t+1}$ as all histories h_{t+1} that are consistent with h^i_{t+1}, and the notation $(h_{t+1} : h^i_{t+1})\backslash h^i_{t+1}$ means we have excluded the h^i_{t+1} part from those h_{t+1} that are consistent with h^i_{t+1}. The beliefs are updated arbitrarily if $\mathbb{P}^g(h^i_{t+1}|h^i_t) = 0$.

The Bayesian updating of the belief $\mu^i_{t+1}(h_{t+1}|h^i_{t+1})$ is described below. For those h_{t+1} that are consistent with h^i_{t+1}, we write

$$\mu^i_{t+1}(h_{t+1}|h^i_{t+1}) = \mathbb{P}^g(h_{t+1}|h^i_{t+1}) = \frac{\int_{h_t} P^{g_t}(h_{t+1}|h_t)\mu^i_t(h_t|h^i_t)}{\int_{h_{t+1}:h^i_{t+1}} \int_{h_t} P^{g_t}(h_{t+1}|h_t)\mu^i_t(h_t|h^i_t)}, \tag{2.4}$$

and if h_{t+1} is not consistent with h^i_{t+1}, we have $\mu^i_{t+1}(h_{t+1}|h^i_{t+1}) = 0$.

For any $i \in \mathcal{N},\ t \in \mathcal{T}, h^i_t \in \mathcal{H}^i_t, \tilde{\beta}^i$, sequential rationality imposes the following condition for the strategy profile β and belief system μ:

$$\mathbb{E}^{\beta^i_{t:T}\beta^{-i}_{t:T}}_{\mu_t} \left\{ \sum_{n=t}^{T} r^i_n(V, A_n)|h^i_t \right\} \geq \mathbb{E}^{\tilde{\beta}^i_{t:T}\beta^{-i}_{t:T}}_{\mu_t} \left\{ \sum_{n=t}^{T} r^i_n(V, A_n)|h^i_t \right\} \tag{2.5}$$

Sequential rationality ensures that at each information set h^i_t, each player's action is a best response to the strategy of others. This is formulated in equation (2.5), where β is the equilibrium strategy profile and $\tilde{\beta}^i$ is any other strategy of player i. The inequality indicates that player i gains more by playing β^i compared to $\tilde{\beta}^i$.

Notice that we have defined the belief μ^i_t to be a belief over the set of all histories at time t given the information set of player i at t. However, this is the most general belief that one could consider and depending on the specifics of the game, we can define other (simpler) types of beliefs that are sufficient for the players to act rationally. Note that for any types of beliefs that we consider, the update rule is the same as what was described here.

We note that PBE is not the only type of equilibrium that can be employed in this setting. Refinements of PBE, including trembling hand equilibrium and sequential equilibrium [77, 78] can also be considered. On the other hand, Bayes correlated equilibria [79], or their extensions to extensive-form games [80], may be a potential alternative; their complexity, however, can be much higher than the studied PBEs for games with long time horizons.

2.4 Structured PBE

The domain of the strategies $g_t^i(\cdot|h_t^i)$ is expanding in time. Finding such strategies is complicated with the complexity growing exponentially with the time horizon. For this reason, we consider summaries for $h_t^i \in \mathcal{H}_t^i$, i.e., $S(h_t^i)$, with time invariant domains [62]. Notice that the information set h_t^i is time variant and therefore, the summary $S(h_t^i)$ is also time variant. However, the domain of h_t^i, i.e., $\mathcal{H}_t^i$, is expanding by time while the domain of $S(h_t^i)$ is time invariant. We are interested in PBEs with strategies, $g_t^i(\cdot|h_t^i) = \psi_t^i(\cdot|S(h_t^i))$, that are functions of h_t^i only through the summaries $S(h_t^i)$. These PBEs are called structured PBEs [59]. Since the set of summaries does not grow in time, finding such structured PBEs is less complicated than a general PBE. According to [59], we can show that players can guarantee the same rewards by playing structured strategies compared to the general non-structured ones. In dynamic games with asymmetric information, summaries are usually the belief of players over the unknown variables of the game.

Define the private beliefs over the unknown state of the world V as

$$\xi_t^i(v) = \mathbb{P}^g(v|h_t^i) = \mathbb{P}^g(v|x_{1:t}^i, a_{1:t-1}). \tag{2.6}$$

We further define the conditional public belief over the private beliefs as follows

$$\pi_t(\xi_t|v) = \mathbb{P}^g(\xi_t|v, a_{1:t-1}). \tag{2.7}$$

Lemma 1 (Conditional Independence of Private Beliefs). *We have the following for the conditional public belief*

$$\pi_t(\xi_t|v) = \prod_{i\in\mathcal{N}} \pi_t^i(\xi_t^i|v), \tag{2.8}$$

where $\pi_t^i(\xi_t^i|v) = \mathbb{P}(\xi_t^i|v, a_{1:t-1})$*. Similarly, we have*

$$\mathbb{P}^g(x_{1:t}|v, a_{1:t-1}) = \prod_{i\in\mathcal{N}} \mathbb{P}^g(x_{1:t}^i|v, a_{1:t-1}). \tag{2.9}$$

Proof. *See Appendix A.1.*

Note that this conditional independence holds regardless of the strategy profiles g. Using this result, and with a slight abuse of notation[2], we can summarize the conditional public belief into the vector $\pi_t = [\pi_t^1, \ldots, \pi_t^N]$.

We are interested in strategies of the form $A_t^i \sim \psi_t^i(\cdot|\xi_t^i, \pi_t) = \gamma_t^i(\cdot|\xi_t^i)$, where $\gamma_t^i = \theta_t^i[\pi_t]$ and we will prove that such structured strategies form a PBE of the game. Note that with the above decomposition of the strategy ψ into partial strategies γ and the strategy θ, designing strategies ψ is equivalent to designing θ.

2.4.1 Belief Update

In this subsection, we present two lemmas regarding the beliefs and their update rules.

Lemma 2. *The private beliefs can be updated as $\xi_{t+1}^i = F^i[\xi_t^i, \pi_t^{-i}, \gamma_t^{-i}, a_t, x_{t+1}^i]$, where F^i is defined through*

$$\xi_{t+1}^i(v) = \frac{\displaystyle\int_{\xi_t^{-i}} \xi_t^i(v) \prod_{j\in -i} \pi_t^j(\xi_t^j|v)\gamma_t^j(a_t^j|\xi_t^j)Q_X^i(x_{t+1}^i|v, a_t)}{\displaystyle\int_{\xi_t^{-i},\tilde{v}} \xi^i(\tilde{v}) \prod_{j\in -i} \pi_t^j(\xi_t^j|\tilde{v})\gamma_t^j(a_t^j|\xi_t^j)Q_X^i(x_{t+1}^i|\tilde{v}, a_t)}, \tag{2.10}$$

for all v.

Proof. *See Appendix A.2.*

Note that this update depends on the strategy profile g only through the partial function γ_t^{-i}, i.e., it is independent of the strategy θ. We will also use the notation $\xi_{t+1} = F[\xi_t, \pi_t, \gamma_t, a_t, x_{t+1}]$ for the update function of the vector of private beliefs.

Lemma 3. *The conditional public beliefs can be updated as $\pi_{t+1}^i = F_\pi^i[\pi_t, \gamma_t, a_t]$, where F_π^i is defined through*

$$\pi_{t+1}^i(\xi_{t+1}^i|v) = \frac{\displaystyle\int_{\xi_t^i, x_{t+1}^i} \pi_t^i(\xi_t^i|v)\gamma_t^i(a_t^i|\xi_t^i)Q_X^i(x_{t+1}^i|v, a_t)\delta(\xi_{t+1}^i - F^i[\xi_t^i, \pi_t^{-i}, \gamma_t^{-i}, a_t, x_{t+1}^i])}{\displaystyle\int_{\tilde{\xi}_t^i} \pi_t^i(\tilde{\xi}_t^i|v)\gamma_t^i(a_t^i|\tilde{\xi}_t^i)}, \tag{2.11}$$

for all v and ξ_{t+1}^i.

[2]We will be using π_t to denote the joint conditional $\pi_t(\xi_t|v)$ as well as the vector of marginal conditionals $\pi_t = [\pi_t^1, \ldots, \pi_t^N]$. The distinction will be obvious from the context.

Proof. *See Appendix A.3.*

Similarly to the previous lemma, this update depends on the strategy profile g only through the partial function γ_t, i.e., it is independent of the strategy θ. We use the notation $\pi_{t+1} = F_\pi[\pi_t, \gamma_t, a_t]$ to denote the update function of the vector of conditional public beliefs.

2.4.2 Equilibrium Strategies

In this subsection, we will show that structured strategies of the form $\gamma_t^i(\cdot|\xi_t^i)$, where $\gamma_t^i = \theta_t^i[\pi_t]$ form structured PBE of the game. The following theorem formalizes this result and presents the fixed point equation characterizing the equilibrium strategies.

Theorem 1. *The strategy profile $\gamma_t^* = \theta_t[\pi_t]$ characterized by the following fixed point equation, forms a structured PBE of the game. For all $i \in \mathcal{N}$,*

$$\gamma_t^{*,i}(\cdot|\xi_t^i) \in \arg\max_{\gamma_t^i(\cdot|\xi_t^i)} \mathbb{E}[\widehat{r}_t^i(\pi_t, \xi_t^i, A_t^i) + J_{t+1}^i(F_\pi[\pi_t, \gamma_t^*, A_t], F^i[\xi_t^i, \pi_t, \gamma_t^{*,-i}, A_t, X_{t+1}^i]))|\pi_t, \xi_t^i], \tag{2.12a}$$

$$J_t^i(\pi_t, \xi_t^i) = \max_{\gamma_t^i(\cdot|\xi_t^i)} \mathbb{E}[\widehat{r}_t^i(\pi_t, \xi_t^i, A_t^i) + J_{t+1}^i(F_\pi[\pi_t, \gamma_t^*, A_t], F^i[\xi_t^i, \pi_t, \gamma_t^{*,-i}, A_t, X_{t+1}^i]))|\pi_t, \xi_t^i], \tag{2.12b}$$

where, $\widehat{r}_t^i(\pi_t, \xi_t^i, a_t^i) = \mathbb{E}\left[r_t^i(V, A_t)|\pi_t, \xi_t^i, a_t^i\right]$.

Proof. *See Appendix A.4.*

We remark here that in equation (2.12) the update rule of the public belief π_t is using the equilibrium strategies γ_t^* and therefore, for each time instance t, the collection of equations of the form (2.12a) for all $i \in \mathcal{N}$ constitutes a fixed point equation over the strategy profile γ_t^*. The reason for this is that in characterizing a PBE, one needs to fix the belief structure and then finds the equilibrium strategies corresponding to those beliefs. On the other hand, the beliefs have to be consistent with the equilibrium strategies. This creates a fixed point equation over $\gamma_t^{*,i}$. Furthermore, the above equation has to be solved simultaneously for all $i \in \mathcal{N}$, thus creating the fixed point equation over the strategy γ_t^*. Notice that equation (2.12) is a general formulation for finding structured PBE in dynamic games with the information structure considered in this work. All of such PBEs satisfy this equation and any solution of this equation, if a solution exists, is a structured PBE. An interesting question is the existence of a solution to (2.12). We first mention that existence results are very scarce in the literature for asymmetric information dynamic games. In [4, 59, 64, 81] existence is discussed under several simplifying assumptions. To this date the general question of existence in general dynamic games with asymmetric information is unresolved even for the independent-types case. The numerical results presented in section 2.6 provide positive evidence towards existence.

2.4.3 Discussion on belief hierarchy

In this section, we characterized the sufficient statistics of the histories of the considered dynamic game. As we mentioned in the Introduction, these summaries include private beliefs, ξ_t^i. One may wonder how we resolved the issue with the chain of private beliefs that was discussed in the Introduction. In other words, how did we resolve the issue of possibly requiring an infinite hierarchy of beliefs on beliefs. In the previous development, we actually proved that this chain stops at the second step. To see this, consider the introduction of private beliefs over others' private beliefs, i.e., $\mathbb{P}(\xi_t^{-i}|h_t^i)$. The results of Lemma 1 show that

$$\begin{aligned}
\mathbb{P}(\xi_t^{-i}|h_t^i) &= \int_{v,x_{1:t}^{-i}} \mathbb{P}(\xi_t^{-i}|v, h_t^i, x_{1:t}^{-i})\mathbb{P}(x_{1:t}^{-i}|v, h_t^i)\mathbb{P}(v|h_t^i) \\
&\overset{(a)}{=} \int_{v,x_{1:t}^{-i}} \mathbb{P}(\xi_t^{-i}|v, a_{1:t-1}, x_{1:t}^{-i})\mathbb{P}(x_{1:t}^{-i}|v, a_{1:t-1})\mathbb{P}(v|h_t^i) \\
&= \int_v \mathbb{P}(\xi_t^{-i}|v, a_{1:t-1})\mathbb{P}(v|h_t^i) \\
&= \int_v \pi_t(\xi_t^{-i}|v)\xi_t^i(v), \qquad (2.13)
\end{aligned}$$

where (a) is due to the definition of the private beliefs and (2.9). The above implies that these beliefs can be evaluated by the public information, π_t, and the first order private beliefs ξ_t^i. This is the exact reason why $\pi_t(\xi_t|v)$ was defined.

2.5 LQG Model

In this section, we study a specific instance of the model discussed so far which is the case where the unknown state of the world, V, is a Gaussian random variable, the private observation kernels are linear and Gaussian and the instantaneous reward is quadratic. Therefore we have an LQG model. The motivation for studying this model stems from the general development in the previous section. In particular we required that equilibrium strategies are generated based on private beliefs and public beliefs on beliefs. In the LQG setting these beliefs can be greatly simplified, thus enabling us to more succinctly characterize the equilibrium strategies discussed in the previous section.

In this model, we consider an unknown state of the world $V \sim \mathrm{N}(0, \mathbf{\Sigma})$ with size N_v. Each player has a private noisy observation X_t^i of V at every time step $t \in \mathcal{T}$

$$x_t^i = v + w_t^i, \tag{2.14}$$

where $W_t^i \sim \mathrm{N}(0, \mathbf{Q}^i)$ and all of the noise random vectors W_t^i are independent across i and t and also independent of V. The values of $\mathbf{\Sigma}$ and $\mathbf{Q}^i,\ \forall i \in \mathcal{N}$ are common knowledge between players. Note that in order to maintain the linearity of private observations, we have considered uncontrolled private observations unlike the general model in first part of the chapter. More discussion on this matter can be found in section 2.7. We have $a_t^i \in \mathcal{A}^i = \mathbb{R}^{N_a}$. The instantaneous reward[3] is given by

$$r_t^i(v, a_t) = \begin{bmatrix} v' & a_t' \end{bmatrix} \mathbf{R}_t^i \begin{bmatrix} v \\ a_t \end{bmatrix} = \mathrm{qd}(\mathbf{R}_t^i; \begin{bmatrix} v \\ a_t \end{bmatrix}), \tag{2.15}$$

where $\mathbf{R}_t^i$ is a symmetric negative definite matrix of appropriate dimensions.

2.5.1 Equilibrium Beliefs

In this setting, we will show that the private beliefs ξ_t^i are Gaussian and since any Gaussian belief can be expressed in terms of its mean and covariance matrix, we define the summaries such that they include the mean and covariance matrices of the beliefs of the players over V. The mean of each player's belief, i.e., her estimate of V, will be her private information. The covariance matrix, however, can be calculated publicly. We define the private estimate of players over V as follows. For all $i \in \mathcal{N},\ t \in \mathcal{T}$,

$$\widehat{v}_t^i = \mathbb{E}[V | h_t^i] = \mathbb{E}[V | x_{1:t}^i, a_{1:t-1}], \tag{2.16}$$

Since the private beliefs can be expressed in terms of their means and covariance matrices and since the covariance matrices are publicly calculated, the conditional public belief $\pi_t^i(\xi_t^i | v)$ is equivalent to a belief over the private estimates. Intuitively, each player, in addition to her own estimate of V, needs to interpret actions of others and predict their future actions. Hence, each player needs to have a belief over the estimates of other players on V. We will show that this latter belief is also Gaussian and therefore, one needs to keep track of only its mean and covariance. We define the

[3]Unlike more standard LQG setting we consider "rewards" instead of "costs" to maintain consistency with the general problem discussed earlier.

following quantity for all $i \in \mathcal{N},\ t \in \mathcal{T}$,

$$\tilde{v}_t^{i,j} = \mathbb{E}[\widehat{V}_t^j | h_t^i] = \mathbb{E}[\widehat{V}_t^j | x_{1:t}^i, a_{1:t-1}]. \tag{2.17}$$

The quantity $\widehat{v}_t^i$ is player i's best estimate of V given her observations up to time t. As mentioned before, this quantity is a private estimation for player i and is not measurable with respect to the sigma algebra generated by the observations of any other player j. Hence, player i should form an estimate over the private estimates of other players and this is the reason $\tilde{v}_t^{i,j}$ is defined. This in turn implies that players' strategies should also be a function of their estimates over others' estimates of V. Hence, the same argument as the one in the first part of the chapter about private beliefs holds and we need to define an estimate over estimates of players over other players' estimates of V. This argument continues as long as these estimates are private. Therefore, once again, we are faced with the problem of having to define a chain of private beliefs which are expressed as private estimates in this model. This chain stops whenever one of the estimates of players is public (or a public function of previous-step private estimates) and therefore, there is no need to form an estimate over it.

Indeed, we will show that $\tilde{v}_t^{i,-i}$ is a public linear function of $\widehat{v}_t^i$, hence, there is no need to include $\tilde{v}_t^{i,-i}$ in the private part of the summary $S(h_t^i)$ and therefore, no other player needs to form an estimate over it. The summary we use for h_t^i is defined as $S(h_t^i) = (\widehat{v}_t^i, P(h_t^i))$, where $P(h_t^i)$ is the public summary for h_t^i and it includes the covariance matrix of player i's belief over V and some other needed quantities that will be subsequently defined. We are interested in equilibria with strategies of the form $A_t^i \sim \psi_t^i(\cdot|\widehat{v}_t^i, P(h_t^i)) = \gamma_t^i(\cdot|\widehat{v}_t^i)$, where $\gamma_t^i = \theta_t^i[P(h_t^i)]$. In particular, we want to prove that pure linear strategies of the form $\gamma_t^i(a_t^i|\widehat{v}_t^i) = \delta(a_t^i - \mathbf{L}_t^i\widehat{v}_t^i - m_t^i)$, where $\mathbf{L}_t^i$ and m_t^i are matrices with appropriate dimensions and are functions of $P(h_t^i)$, form a PBE of the game.

In the next theorem, we show that when linear strategies are employed, the private beliefs are Gaussian.

Theorem 2. *Assuming pure linear strategies of the form $\gamma_t^i(a_t^i|\widehat{v}_t^i) = \delta(a_t^i - \mathbf{L}_t^i\widehat{v}_t^i - m_t^i)$, $\forall t \in \mathcal{T}$ and $\forall i \in \mathcal{N}$, the private belief ξ_t^i on V is Gaussian $N(\widehat{v}_t^i, \Sigma_t^i)$, where $\widehat{v}_t^i$ is the private estimate of player i of V and Σ_t^i is the corresponding covariance matrix, which can be evaluated publicly. Consequently, the public belief $\pi_t^i(\xi_t^i|v)$ can be reduced to a belief $\pi_t^i(\widehat{v}_t^i|v)$. Furthermore, $\pi_t^i(\widehat{v}_t^i|v)$ is Gaussian with mean $\mathbf{E}_t^i v + f_t^i$, where matrices $\mathbf{E}_t^i$, f_t^i can be evaluated publicly.*

Proof. *See Appendix A.5.*

In the following we summarize the parameters needed to update each of the quantities introduced in the proof of Theorem 2 and we introduce update functions for each one.

$$
\begin{aligned}
\widehat{v}^i_{t+1} &= F_{\widehat{v}}(\widehat{v}^i_t, \mathbf{\Sigma}^i_{t+1|t}, \mathbf{E}^{-i}_t, f^{-i}_t, \mathbf{L}^{-i}_t, m^{-i}_t, a^{-i}_t, x^i_{t+1}) && \text{(2.18a)}\\
\mathbf{\Sigma}^i_{t+1} &= F_{\mathbf{\Sigma}^i}(\mathbf{\Sigma}^i_{t+1|t}, \mathbf{L}^{-i}_t) && \text{(2.18b)}\\
\mathbf{\Sigma}_{t+2|t+1} &= F_{\mathbf{\Sigma}}(\mathbf{\Sigma}_{t+1|t}, \mathbf{E}_t, \mathbf{L}_t) && \text{(2.18c)}\\
\tilde{\mathbf{\Sigma}}_{t+2|t+1} &= F_{\tilde{\mathbf{\Sigma}}}(\tilde{\mathbf{\Sigma}}_{t+1|t}, \mathbf{\Sigma}_{t+1|t}, \mathbf{E}_t, \mathbf{L}_t) && \text{(2.18d)}\\
\mathbf{E}_{t+1} &= F_{\mathbf{E}}(\mathbf{E}_t, \mathbf{\Sigma}_{t+1|t}, \tilde{\mathbf{\Sigma}}_{t+1|t}, \mathbf{L}_t) && \text{(2.18e)}\\
f_{t+1} &= F_f(f_t, \mathbf{\Sigma}_{t+1|t}, \tilde{\mathbf{\Sigma}}_{t+1|t}, \mathbf{E}_t, \mathbf{L}_t, m_t, a_t), && \text{(2.18f)}
\end{aligned}
$$

where $F_{\widehat{v}}$ is defined in (A.18), $F_{\mathbf{\Sigma}^i}$ and $F_{\mathbf{\Sigma}}$ are defined in (A.20), $F_{\tilde{\mathbf{\Sigma}}}$ is defined in (A.28), and $F_{\mathbf{E}}$ and F_f are defined in (A.27). Equations (2.18a) and (2.18b) correspond to the private belief update and are similar in structure to the update function F^i of of ξ^i_t in Lemma 2 for the general case. The remaining update functions correspond to the public belief update F_π in Lemma 3 for the general case.

Note that according to the above equations, the quantities $\mathbf{\Sigma}_{t+1|t}$, $\tilde{\mathbf{\Sigma}}_{t+1|t}$, $\mathbf{E}_t$ are updated recursively using the strategy matrices $\mathbf{L}_t$. Hence, if one knows the strategies, one can calculate these quantities offline for the entire time horizon of the game. However, the quantity f_k is updated using the strategy matrices $\mathbf{L}_t$ and vectors m_k as well as the realized actions a_t and therefore, they cannot be evaluated offline.

We reiterate at this point that Theorem 2 implies that the estimate of player i over private estimates of players $-i$, i.e., $\tilde{v}^{i,-i}_t$, is a linear function of $\widehat{v}^i_t$,

$$
\begin{aligned}
\tilde{v}^{i,-i}_t &= \mathbb{E}[\widehat{V}^{-i}_t | h^i_t]\\
&= \mathbb{E}[\mathbb{E}[\widehat{V}^{-i}_t | V, A_{1:t-1}] | h^i_t]\\
&= \mathbb{E}[\mathbf{E}^{-i}_t V + f^{-i}_t | h^i_t]\\
&= \mathbf{E}^{-i}_t \widehat{v}^i_t + f^{-i}_t, && \text{(2.19)}
\end{aligned}
$$

with matrices $\mathbf{E}^{-i}_t$ and f^{-i}_t being public information. As a result, assuming linear strategies of the form $a^i_t = \mathbf{L}^i_t \widehat{v}^i_t + m^i_t$ at equilibrium, one can form the summary $S(h^i_t) = (\widehat{v}^{-i}_t, P(h^i_t))$ and base the selection of the matrices $\mathbf{L}^i_t$ and m^i_t on the public part of this summary, $P(h^i_t)$. In the next section we show that indeed linear strategies can form an equilibrium and provide a methodology to find the quantities $\mathbf{L}^i_t$ and m^i_t.

2.5.2 Linear Structured PBE

Theorem 2 implies that S_t^i is a jointly Gaussian random vector conditioned on player i's observation till time t, $\forall i \in \mathcal{N}, t \in \mathcal{T}$. This implies that the beliefs over V are jointly Gaussian and so players need only keep track of their belief's mean (estimation) and covariance matrices. Furthermore, this theorem implies that a player's belief over other players beliefs is also Gaussian and hence, players need to keep track of their estimation on other players' estimations, i.e., $\tilde{v}$. The important point of Theorem 2 is the statement that the estimation of players on others' estimations is a linear function of their own estimation and hence, in order to keep track of the estimation over other players' estimations, a player only needs to keep track of her own estimation over V. Therefore, $\widehat{v}_t^i$ is a sufficient statistic for player i's private observations till time t.

In terms of the public summary, we see four public quantities, $\mathbf{\Sigma}_{t+1|t}$, $\tilde{\mathbf{\Sigma}}_{t+1|t}$, $\mathbf{E}_t$ and f_t in (2.18). With some abuse of notation, we define $\mathbf{\Sigma}_t = [\mathbf{\Sigma}_{t+1|t}, \tilde{\mathbf{\Sigma}}_{t+1|t}]$. We will show that the tupple $(\mathbf{\Sigma}_t, \mathbf{E}_t, f_t)$ is the public summary of h_t^i, i.e., $P(h_t^i)$. Note that $\mathbf{E}_t$ and f_t are involved in the expression for the mean of the conditional public belief over $\widehat{v}_t$, hence, they correspond to the conditional public belief π_t in the first part of the chapter. The convariance matrices $\mathbf{\Sigma}_{t+1|t}$, $\tilde{\mathbf{\Sigma}}_{t+1|t}$ represent the covariance matrices of the private and conditional public beliefs. This implies that by having the tuple $(\widehat{v}_t^i, \mathbf{\Sigma}_t, \mathbf{E}_t, f_t)$, we have full characterization of the private and public belief and therefore, we have the summaries for the LQG game.

Therefore, we consider strategies of the form $\psi_t^i(\cdot|\widehat{v}_t^i, \mathbf{\Sigma}_t, \mathbf{E}_t, f_t) = \gamma_t^i(\cdot|\widehat{v}_t^i)$. In particular, we will now show that linear strategies of the form $\gamma_t^i(\cdot|\widehat{v}_t^i) = \delta(a_t^i - \mathbf{L}_t^i\widehat{v}_t^i - m_t^i)$, where $\mathbf{L}_t$ and m_t are derived from $(\mathbf{\Sigma}_t, \mathbf{E}_t, f_t)$, are PBE of the game.

Theorem 3. *The strategy profile $\psi_t^i(\cdot|\widehat{v}_t^i, \mathbf{\Sigma}_t, \mathbf{E}_t, f_t) = \gamma_t^i(\cdot|\widehat{v}_t^i)$ $\forall i \in \mathcal{N}$ where $\gamma_t^i(\cdot|\widehat{v}_t^i) = \delta(a_t^i - \mathbf{L}_t^i\widehat{v}_t^i - m_t^i)$, together with the corresponding Gaussian beliefs derived in Theorem 2, form a structured PBE of the game.*

The strategy matrices $\mathbf{L}_t$ and vectors m_t are constructed throughout the proof.

Proof. *See Appendix A.6.*

One important result from the proof of Theorem 3 is that the reward to go, $J_t^i(\widehat{v}_t^i, \mathbf{\Sigma}_t, \mathbf{E}_t, f_t)$ is quadratic with respect to $\widehat{v}_t^i$ and f_t, which are the only quantities in the summary that can not be evaluated offline, i.e., we have

$$J_t^i(\widehat{v}_t^i, \mathbf{\Sigma}_t, \mathbf{E}_t, f_t) = \mathrm{qd}(\mathbf{Z}_t^i; \begin{bmatrix} \widehat{v}_t^i \\ f_t \end{bmatrix}) + z_t^{i\prime} \begin{bmatrix} \widehat{v}_t^i \\ f_t \end{bmatrix} + o_t^i. \tag{2.20}$$

Therefore, if we have the quantities $\mathbf{Z}_t^i$, $z_t^{i\prime}$, and o_t^i we can evaluate the reward to go for every value of $\widehat{v}_t^i$ and f_t.

In the following, we propose a backward algorithm that evaluates the quantities $\mathbf{Z}_t^i$, z_t^i, and o_t^i as well as the strategy matrices $\mathbf{L}_t$, $\mathbf{M}_t$ and vectors $\overline{m}_t$ (we have $m_t^i = \mathbf{M}_t^i f_t + \overline{m}_t^i$, according to the proof of Theorem 3) as functions of $(\mathbf{\Sigma}_t, \mathbf{E}_t)$. Before stating the algorithm, we define the following functions.

$$\mathbf{L}_t = g_{\mathbf{L},t}(\mathbf{\Sigma}_t, \mathbf{E}_t) \tag{2.21a}$$

$$\mathbf{M}_t = g_{\mathbf{M},t}(\mathbf{\Sigma}_t, \mathbf{E}_t) \tag{2.21b}$$

$$\overline{m}_t = g_{\overline{m},t}(\mathbf{\Sigma}_t, \mathbf{E}_t) \tag{2.21c}$$

$$\mathbf{Z}_t = \psi_{\mathbf{Z},t}(\mathbf{\Sigma}_t, \mathbf{E}_t) \tag{2.21d}$$

$$z_t = \psi_{z,t}(\mathbf{\Sigma}_t, \mathbf{E}_t) \tag{2.21e}$$

$$o_t = \psi_{o,t}(\mathbf{\Sigma}_t, \mathbf{E}_t), \tag{2.21f}$$

where the first three functions are defined in equation (A.41) and the rest are defined in equation (A.43).

Backward Algorithm (Offline)

1. Set $t = T$. Set $\mathbf{Z}_{T+1} = \psi_{\mathbf{Z},T+1}(\mathbf{\Sigma}_{T+1}, \mathbf{E}_{T+1}) = \mathbf{0}$, $z_{T+1} = \psi_{z,T+1}(\mathbf{\Sigma}_{T+1}, \mathbf{E}_{T+1}) = \mathbf{0}$ and $o_{T+1} = \psi_{\mathbf{Z},T+1}(\mathbf{\Sigma}_{T+1}, \mathbf{E}_{T+1}) = \mathbf{0}$ for every $\mathbf{\Sigma}_{T+1}, \mathbf{E}_{T+1}$.

2. Calculate $\mathbf{L}_t = g_{\mathbf{L},t}(\mathbf{\Sigma}_t, \mathbf{E}_t)$, $\mathbf{M}_t = g_{\mathbf{M},t}(\mathbf{\Sigma}_t, \mathbf{E}_t)$, $\overline{m}_t = g_{\overline{m},t}(\mathbf{\Sigma}_t, \mathbf{E}_t)$, and $\mathbf{Z}_t = \psi_{\mathbf{Z},t}(\mathbf{\Sigma}_t, \mathbf{E}_t)$ for every $\mathbf{\Sigma}_t, \mathbf{E}_t$ and the corresponding $\psi_{\mathbf{Z},t+1}(\cdot,\cdot)$ according to equation (A.41) and (A.43).

3. Set $t = t - 1$.

4. If $t \geq 1$ Go to step 2. Else stop.

Using the functions defined above, one can run the following forward algorithm to find the strategy matrices $\mathbf{L}_t$, $\mathbf{M}_t$ and vectors $\overline{m}_t$ and the quantities $\mathbf{Z}_t^i$, $z_t^{i\prime}$, and o_t^i.

Forward Algorithm (Offline)

1. Set $t = 1$.

2. Initialize the value of $\mathbf{\Sigma}_1$ and $\mathbf{E}_1$ using Lemma 21.

3. Using $\mathbf{\Sigma}_t$ and $\mathbf{E}_t$, find $\mathbf{L}_t$, $\mathbf{M}_t$, $\overline{m}_t$ and the quantities $\mathbf{Z}_t^i$, $z_t^{i\prime}$, and o_t^i according to equation (2.21).

4. Using $\mathbf{\Sigma}_t$, $\mathbf{E}_t$ and $\mathbf{L}_t$, calculate $\mathbf{\Sigma}_{t+1}$ and $\mathbf{E}_{t+1}$ according to equation (2.18).

5. Set $t = t + 1$.

6. If $t \leq T$, Go to step 3. Else stop.

2.6 Example

In this section, we describe some numerical examples to show the equilibrium strategies discussed in this chapter. In these examples, we derive the equilibrium strategies by solving a fixed point equation for the entire time horizon using the following algorithm. Note that the superscript (k) in $A^{(k)}$ denotes the number of iterations performed. We define the convergence error as $\epsilon^{(k)} = \max(|\mathbf{L}_{1:T}^{(k+1)} - \mathbf{L}_{1:T}^{(k)}|, |\mathbf{M}_{1:T}^{(k+1)} - \mathbf{M}_{1:T}^{(k)}|, |\overline{m}_{1:T}^{(k+1)} - \overline{m}_{1:T}^{(k)}|)$.

Numerical Algorithm (Offline)

1. Set $k = 1$.

2. Initialize $\mathbf{L}_{1:T}^{(1)}$, $\mathbf{M}_{1:T}^{(1)}$, and $\overline{m}_{1:T}^{(1)}$ arbitrarily.

3. Using $\mathbf{L}_{1:T}^{(k)}$, evaluate $\mathbf{\Sigma}_{1:T}^{(k+1)}$, $\mathbf{E}_{1:T}^{(k+1)}$ according to equations (2.18) in a forward manner (using initial conditions $\mathbf{\Sigma}_1$ and $\mathbf{E}_1$ according to equations (A.20) and (A.28)).

4. Using $\mathbf{L}_{1:T}^{(k)}$, $\mathbf{M}_{1:T}^{(k)}$, $\overline{m}_{1:T}^{(k)}$, and $\mathbf{\Sigma}_{1:T}^{(k+1)}$, $\mathbf{E}_{1:T}^{(k+1)}$, evaluate $\mathbf{L}_{1:T}^{(k+1)}$, $\mathbf{M}_{1:T}^{(k+1)}$, and $\overline{m}_{1:T}^{(k+1)}$ according to the backward algorithm.
   $\mathbf{L}_t^{(k+1)} = g_{\mathbf{L},t}(\mathbf{\Sigma}_t^{(k+1)}, \mathbf{E}_t^{(k+1)}) = \text{bdp}_{\mathbf{L},t}(...)(\mathbf{\Sigma}_t^{(k+1)}, \mathbf{E}_t^{(k+1)})$

5. Evaluate $\epsilon^{(k)}$. If it is below the desired threshold, stop. Otherwise, go to step 4.

Note that in each step of the backward algorithm, one needs to solve a fixed point equation with respect to the strategy matrices and vectors to derive the functions defined in eq. (2.21) (see eq. (A.41) in Appendix A). However, in the numerical algorithm described above, we use the last iteration quantities for the right hand side of the equations and consequently, we do not need to solve any fixed point equations.

As a concrete example, we consider a setting where there is a project with an unknown attribute denoted by v. There are two agents working on this project exerting a costly effort a_t^i. The agents are rewarded based on the alignment of their effort with the project attribute, v, as well as based on their cooperation. At each time slot, the agents have private observations, x_t^i, of the project attribute. We consider two instances of the game where v is scalar in one and a two dimensional vector in the other, while the efforts are scalars in both.

2.6.1 Scalar State and Action

We model the considered scenario for scalar v and scalar actions a_t^i with the instantaneous rewards being $R_t^1(v, a_t) = a_t^1 v + \frac{1}{2}a_t^1 a_t^2 - (a_t^1)^2$ and $R_t^2(v, a_t) = a_t^2 v + \frac{1}{2}a_t^1 a_t^2 - (a_t^2)^2$. That is, we set $\mathbf{R}_t^1 = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & -1 & \frac{1}{4} \\ 0 & \frac{1}{4} & 0 \end{bmatrix}$ and $\mathbf{R}_t^2 = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & -1 \end{bmatrix}$. Note that the term $a_t^i v$ in the instantaneous rewards accounts for the alignment of a_t^i with v, and the term $a_t^1 a_t^2$ denotes the cooperation between the agents.

Case 1: If we assume that agents perfectly observe V, i.e., if we set $\mathbf{Q}^1 = 0$ and $\mathbf{Q}^2 = 0$, the following linear equilibrium strategy matrices and vectors are derived from the numerical analysis of this game for $T = 2$ and $\mathbf{\Sigma} = 1$

$$\begin{array}{ll} \mathbf{L}_1^1 = \frac{2}{3} & \mathbf{L}_1^2 = \frac{2}{3} \\ \mathbf{L}_2^1 = \frac{2}{3} & \mathbf{L}_2^2 = \frac{2}{3}. \end{array} \tag{2.22}$$

Furthermore, we have $\overline{m}_t^i = 0$ for $t = 1, 2$ and $i = 1, 2$. Note that since in this case, $f_t = 0$ for $t = 1, 2$, the strategy matrices $\mathbf{M}_t^i$ will not play any roles and are not presented here. These results imply that each agent will exert effort exactly equal to $\frac{2}{3}V$. As it turns out, these strategies are myopic, i.e., we also observe these strategies in the case $T = 1$. The reason for having myopic strategies is that the observations are perfect and hence, the actions have no effect in shaping the future beliefs.

Case 2: Consider agents with equally imperfect observations, $\mathbf{Q}^1 = \mathbf{Q}^2 = 1$. The following strategy matrices are derived

$$\begin{array}{ll} \mathbf{L}_1^1 = 0.6722 & \mathbf{L}_1^2 = 0.6722 \\ \mathbf{L}_2^1 = 0.5333 & \mathbf{L}_2^2 = 0.5333 \end{array} \tag{2.23a}$$

$$\begin{array}{ll} \mathbf{M}_1^1 = [0.0561 \quad 0.2620] & \mathbf{M}_1^2 = [0.2620 \quad 0.0561] \\ \mathbf{M}_2^1 = [0.0356 \quad 0.1422] & \mathbf{M}_2^2 = [0.1422 \quad 0.0356]\,, \end{array} \tag{2.23b}$$

together with $\overline{m}_t^i = 0$ for $t = 1, 2$ and $i = 1, 2$. Once more, it is observed that $\overline{m}_t^i = 0$ and as will be seen, the same is happening in all of the other cases studied as well. This could imply that it is sufficient to restrict attention to strategies with zero $\overline{m}_t^i$. We also observe that the value of the strategy matrices decrease with time.

Case 3: If one agent has better observations than the other, i.e., $\mathbf{Q}^1 = 1$, $\mathbf{Q}^2 = 2$, the strategy matrices are changed as follows.

$$\begin{array}{ll} \mathbf{L}_1^1 = 0.6700 & \mathbf{L}_1^2 = 0.6619 \\ \mathbf{L}_2^1 = 0.5224 & \mathbf{L}_2^2 = 0.5373 \end{array} \tag{2.24a}$$

$$\begin{array}{ll} \mathbf{M}_1^1 = [0.0520 \quad 0.2701] & \mathbf{M}_1^2 = [0.2738 \quad 0.0605] \\ \mathbf{M}_2^1 = [0.0348 \quad 0.1433] & \mathbf{M}_2^2 = [0.1393 \quad 0.0358] \end{array} \tag{2.24b}$$

and $\overline{m}_t^i = 0$ for $t = 1, 2$ and $i = 1, 2$. One can explain these results by paying attention to the interactions between the agents. At $t = 1$, agent one has a better estimation of V compared to agent two and therefore, she has higher $\mathbf{L}_1^1$. At $t = 2$, agent two has learned the estimation of agent one through her action at $t = 1$ and therefore, the two agents have almost equal estimations. But this time, agent two exerts slightly higher effort to compensate agent one's efforts at $t = 1$.

Case 4: The interaction between agents can also be seen in a scenario where one agent has perfect observations and the other one has partial observations, i.e., $\mathbf{Q}^1 = 0$, $\mathbf{Q}^2 = 2$. The strategy matrices are given as follows.

$$\begin{array}{ll} \mathbf{L}_1^1 = 0.7125 & \mathbf{L}_1^2 = 0.6781 \\ \mathbf{L}_2^1 = 0.5000 & \mathbf{L}_2^2 = 0.6250 \end{array} \tag{2.25a}$$

$$\begin{array}{ll} \mathbf{M}_1^1 = [0.0142 \quad 0.1808] & \mathbf{M}_1^2 = [0.1817 \quad 0.0452] \\ \mathbf{M}_2^1 = [0.0333 \quad 0.1667] & \mathbf{M}_2^2 = [0.1333 \quad 0.0417], \end{array} \tag{2.25b}$$

and $\overline{m}_t^i = 0$ for $t = 1, 2$ and $i = 1, 2$.

Case 5: Finally, consider a case where both agents have very noisy observations, that is $\mathbf{Q}^1$, $\mathbf{Q}^2$ are large numbers. In this case, $\widehat{v}_t^i = 0$ and $f_t = 0$. Therefore, the strategy matrices $\mathbf{L}_t^i$ and $\mathbf{M}_t^i$ do not play any roles and the actions will only follow $\overline{m}_t^i$. For this game we obtain $\overline{m}_t^i = 0$ for $t = 1, 2$ and $i = 1, 2$.

Case 6: We have also derived the strategy matrices of the game for larger values of T. In Figure 2.1, we can see the plot of the strategy matrices $\mathbf{L}_t^i$ with respect to time for the symmetric case of $\mathbf{Q}^1 = \mathbf{Q}^2 = 1$ and for $T = 10$. As before, we observe a trend where as time goes by, the values of the strategy matrices decrease. The intuition behind why such behavior is observed is that more public information is observed as time goes by. Therefore, the players estimation over others' estimations is mainly characterized by the public part of the state, f_t, rather than the private estimates. This indicates that the matrix $\mathbf{E}_t$ decreases with time and as it is observed in our numerical results in Figure 2.1, it converges to zero. One can also see that the strategies decrease as $\mathbf{E}_t$ decreases. Therefore, the strategy matrices $\mathbf{L}_t$ decrease as time passes and they converge to 0.5, which is the equilibrium of the game when $\mathbf{E}_t = \mathbf{0}$.

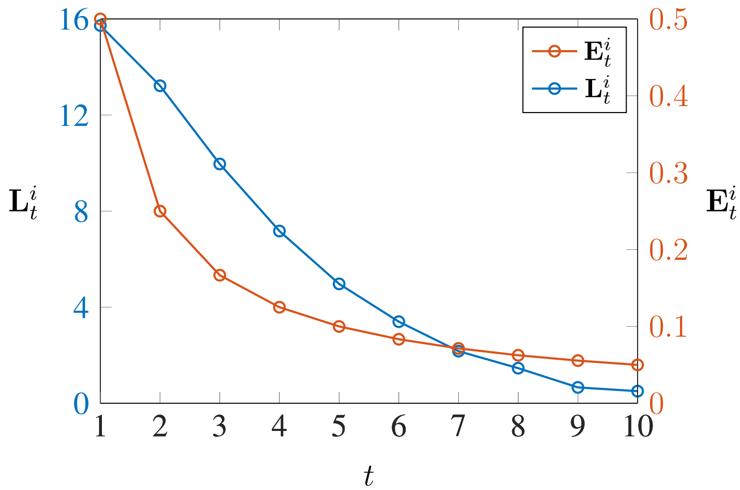


Figure 2.1: Strategy matrices $\mathbf{L}_t^i$ and quantities $\mathbf{E}_t^i$ for $T = 10$.

2.6.2 Game vs Centralized LQG

In this subsection, we have compared the total rewards per time obtained through the game by players for $\mathbf{Q}^1 = \mathbf{Q}^2 = 1$ with a scenario in which both actions are taken by a single decision maker and the sum of the two rewards are collected by her. We have done this comparison for different time horizons T and Figure 2.2 depicts the plot of the total rewards per time obtained, J^T, in the two considered scenarios. We notice that players are doing worse compared to the centralized decision maker. In order to comment on that we note that there are three possible scenarios one can consider in relation to the problem at hand: (i) the centralized problem, (ii) the decentralized team problem and (iii) the decentralized game. Problem (i) is considering a single (centralized) controller solving the optimal LQG problem with reward being the social utility (sum of rewards). Problem (ii) considers multiple decentralized controllers, all having the same goal of maximizing social utility but having the same information structure as in our model. This is a dynamic team problem and its solution is not at all clear (see Witsenhausen counter-example [68]). Finally, problem (iii) is the problem studied in this work. It should be clear that regarding achieving better social utility, (i) is better than (ii) and (ii) is better than (iii). The former is due to the decentralized nature of

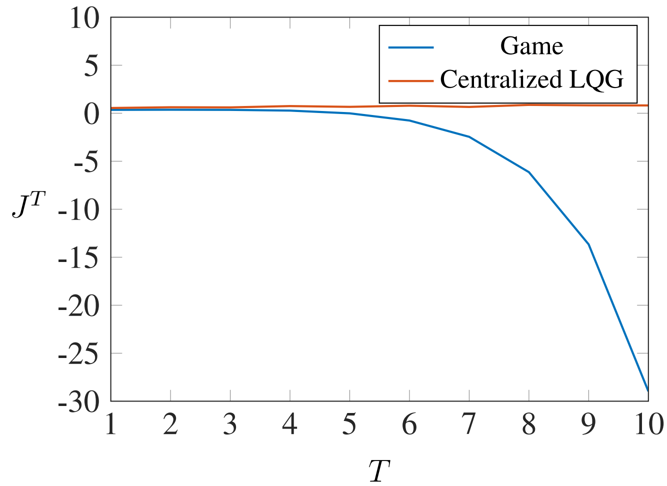


Figure 2.2: Total rewards per time obtained in game vs centralized LQG.

information in (ii), while the latter is due to what is called "Price of Anarchy" (PoA) in the literature, which is the strategic behavior of users in (iii) compared to the team problem (ii). Based on the above, the findings in Fig. 2.2 are not surprising and can be attributed to the PoA. Note, however, that through numerical analysis, we have found one of the possible equilibria of the game and the social reward might be better at other equilibria. We also notice that, the total reward in the centralized scenario is slightly increasing with the time horizon, while the total reward is decreasing with the time horizon in the game scenario. This is due to the fact that in the game scenario, the uncertainty in predicting the average reward-to-go increases drastically as time horizon increases. The centralized decision maker, however, benefits from time horizon increasing and her total reward per time converges to one. This is because as time goes by, the estimation over V becomes better and better and the reward converges to the one in the complete information case.

2.6.3 Two Dimensional State and Scalar Action

In this part, we consider a two dimensional attribute vector for the project, i.e., V is a two dimensional vector. Each agent tries to be aligned with one element of the attribute vector while maintaining the cooperation with the other agent. We can model this alignment and cooperation of agents with $R_t^1(v,a_t) = a_t^1 v(1) + a_t^1 a_t^2 - (a_t^1)^2$ and $R_t^2(v,a_t) = a_t^2 v(2) + a_t^1 a_t^2 - (a_t^2)^2$. That is, we set $\mathbf{R}_t^1 = \begin{bmatrix} 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & -1 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 \end{bmatrix}$ and $\mathbf{R}_t^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & -1 \end{bmatrix}$. We also set $\mathbf{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Case 1: The following linear equilibrium strategy matrices are derived for the full information case.

$$\begin{aligned} \mathbf{L}_1^1 &= \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \end{bmatrix} & \mathbf{L}_1^2 &= \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \end{bmatrix} \\ \mathbf{L}_2^1 &= \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \end{bmatrix} & \mathbf{L}_2^2 &= \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \end{bmatrix}, \end{aligned} \tag{2.26}$$

and $\overline{m}_t^i = 0$ for $t = 1, 2$ and $i = 1, 2$. Also, similar to the scalar case, M_t^i strategy matrices do not play any roles here since $f_t = 0$. We see that if V is perfectly observed, each agent will align her effort with a weighted average of $V(1)$ and $V(2)$ with the element corresponding to that agent having twice the weight. Also, similar to the scalar case, myopic strategies are played.

Case 2: Consider the partial information scenario with $\mathbf{Q}^1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\mathbf{Q}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. The following linear equilibrium strategy matrices are derived.

$$\begin{aligned} \mathbf{L}_1^1 &= [0.7224 \quad 0.2402] & \mathbf{L}_1^2 &= [0.2402 \quad 0.7224] \\ \mathbf{L}_2^1 &= [0.4858 \quad 0.0842] & \mathbf{L}_2^2 &= [0.0842 \quad 0.4858] \end{aligned} \tag{2.27a}$$

$$\mathbf{M}_1^1 = [0.2874 \quad 0.0780 \quad 0.1793 \quad 0.6054] \tag{2.27b}$$

$$\mathbf{M}_1^2 = [0.6054 \quad 0.1793 \quad 0.0780 \quad 0.2874] \tag{2.27c}$$

$$\mathbf{M}_2^1 = [0.1619 \quad 0.0281 \quad 0.0561 \quad 0.3239] \tag{2.27d}$$

$$\mathbf{M}_2^2 = [0.3239 \quad 0.0561 \quad 0.0281 \quad 0.1619], \tag{2.27e}$$

and $\overline{m}_t^i = 0$ for $t = 1, 2$ and $i = 1, 2$. Similar to the scalar scenario, we observe that the value of the strategy matrices decrease with time and again, $\overline{m}_t^i = 0$ for all of the cases.

Case 3: If each agent fully observes her corresponding element of the state and partially observes the other one, i.e., $\mathbf{Q}^1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ and $\mathbf{Q}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, we have the following linear equilibrium strategy matrices.

$$\begin{aligned} \mathbf{L}_1^1 &= [0.7198 \quad 0.4232] & \mathbf{L}_1^2 &= [0.4232 \quad 0.7198] \\ \mathbf{L}_2^1 &= [0.5071 \quad 0.2055] & \mathbf{L}_2^2 &= [0.2055 \quad 0.5071] \end{aligned} \tag{2.28a}$$

$$\mathbf{M}_1^1 = [0.3196 \quad 0.1506 \quad 0.3235 \quad 0.6293] \tag{2.28b}$$

$$\mathbf{M}_1^2 = [0.6293 \quad 0.3235 \quad 0.1506 \quad 0.3196] \tag{2.28c}$$

$$\mathbf{M}_2^1 = [0.1690 \quad 0.0685 \quad 0.1370 \quad 0.3380] \tag{2.28d}$$

$$\mathbf{M}_2^2 = [0.3380 \quad 0.1370 \quad 0.0685 \quad 0.1690] \tag{2.28e}$$

and $\overline{m}_t^i = 0$ for $t = 1, 2$ and $i = 1, 2$. An intuitive reason of why the second element and the first element of the strategy matrices $\mathbf{L}_t^1$ and $\mathbf{L}_t^2$, respectively, are larger than the previous case is that the second element and the first element of $\mathbf{E}_t^1$ and $\mathbf{E}_t^2$, respectively, have increased.

2.7 Model Extensions

In this section, we investigate alternative models that can be studied with the methodology introduced in this chapter and we explain how the results can be extended to such models.

As it is clear in equation (2.14), in the LQG model considered in this chapter, the private observations are not controlled by the actions, unlike the general model of the first part of the chapter. If we were to add control actions to equation (2.14), in order to maintain linearity, we would have added a term such as $B_t^i a_t$ and therefore, equation (2.14) would have looked like $x_t^i = v + w_t^i + B_t^i a_t$. Since the actions are publicly observed, the amount of information that player i extracts from V remains the same with or without the term $B_t^i a_t$. Hence, because the private observations serve only as measurements of V, adding control to equation (2.14) does not make any difference in the results.

Controlled private observations could make a difference in the LQG model if the private observations could affect the instantaneous rewards. That is, if the reward was $r_t^i(v, a_t, x_t^i) = \mathrm{qd}(\mathbf{R}_t^i; \begin{bmatrix} v \\ a_t \\ x_t^i \end{bmatrix})$. Note that the amount of information that x_t^i conveys about V is still the same as in the uncontrolled case. We can show that results similar to all of the ones in this chapter will hold for this model with controlled private observations and this type of instantaneous reward. Note that in this case, the strategies would be linear in both the private estimation and the latest private observation.

We can also extend our results of the first part of the chapter (the general model) to a model with the instantaneous reward being of the form of $r_t^i(v, a_t, x_t^i)$. In this case, x_t^i should be added to the summaries and the results will hold.

2.8 Conclusion

In this chapter, we studied a dynamic game with asymmetric information and dependent types and we characterized the structured perfect Bayesian equilibria of the game. We also studied a special case of our model that was Linear Quadratic Gaussian (LQG) non-zero-sum game and we characterized linear structured perfect Bayesian equilibria for the game. One of the important points that we made in this chapter was that due to the conditional independence of the private signals, the private belief chain stops at the second step and players beliefs over others' beliefs are public functions of their own beliefs. We further proved that these beliefs are Gaussian in the LQG case.

A future direction for this research could be investigating the models for which we have the same interesting features for the beliefs as we do in this chapter. That is, the models for which the private belief chain stops at two or any other given number of steps. Another important future direction is to investigate the existence conditions for the solution of fixed point equations presented in this chapter.

CHAPTER 3

Bayesian Learning with Non-myopic Agents

3.1 Introduction

When a new product/technology is deployed one cannot be certain about its quality in the early stages of the deployment. Many people together may form a more accurate prediction about its quality, but in a strategic environment players act selfishly and may not want to share their private information about the product/technology. Hence, other players' opinions (private information about the product quality) are revealed only indirectly through their actions, i.e., whether they bought the product (adopted the technology) or not. This means that from the perspective of a strategic player, waiting to see what other people have done may provide more certainty about the quality of the product. On the other hand, many products or trends which turn out to be beneficial are better to be adopted as early as possible, since their value can decay over time. This interaction can be formalized as a dynamic game with asymmetric information and a discounted reward. Players want to avoid buying a bad product, so they may postpone their decision to buy/adopt until more information is revealed, while at the same time they want to buy/adopt a good product as soon as possible. This scenario generalizes the classical problem of sequential Bayesian learning to a setting with forward-looking players and no predefined order of play.

Sequential learning has been extensively explored in the literature, with a special focus on a phenomenon known as an *informational cascade*. In two seminal papers [6, 7] the authors investigated the occurrence of fads in a social network, which was later generalized in [8]. Alternative learning models that have been studied in the literature include [82] where players only observe a random set of past actions, [83] where players observe the past actions through a noisy process, [84] where players observe only their immediate predecessor, and [85] where players are allowed to ask questions to a bounded subset of their predecessors.

The common assumption in all of these models is that players act only once in the game and there are informational externalities only, which allows for relatively easy computation of game equilibrium strategies. Some other works where all players act in each period but are myopic by design include [86, 87, 88, 89, 90, 91, 92]. In [93, 94, 95, 96], different models of Bayesian learning were studied where players do not observe the entire action history of the past players, but a "coarser" history. There are also works on non-Bayesian learning models where players do not update their beliefs in a Bayesian sense [97, 86, 98, 99, 100, 101], or do so only with some probability [102]. A survey of such models can be found in [103].

An informational cascade is a phenomenon where no player has an incentive to reveal her private information, hence learning stops in the system. This is an interesting case of herd behavior that happens even with fully rational players. While information cascades do not necessarily happen in all systems (see [104, 105]), they represent a universal phenomenon in sequential Bayesian learning where players act once in a sequence that is predefined before the game starts. In such systems, when the turn of a certain player arrives, she has no choice but to either buy the product if it seems profitable to her at the moment or forever forgo the opportunity. Hence, it is natural to ask whether cascades occur because this one-shot opportunity was forced upon the players. It is conceivable that if players had the freedom to choose to wait and gather more information about the product, a herd behavior, especially a wrong one, could be avoided. This question provides the motivation for studying information cascades in more complex environments. In [106], informational cascades were defined for a general dynamic scenario. However, no evidence for their occurrence was provided.

From a technical perspective, the sequential one-shot framework introduced in [6, 7] and followed in most of the subsequent literature, lends itself to relatively simple equilibrium analysis, since players do not have to account for how much their estimation on the value of the product is going to improve by waiting. This is simply because players are given a single opportunity to act, and cannot wait. In this case, players form a posterior belief on the value of the product based on their public and private signals. Consequently, the equilibrium consists of strategies that maximize each player's instantaneous reward based on this posterior belief.

In this chapter we consider a setting with a finite number of players with no predefined order of action. An exogenous process determines who enters the marketplace at each time epoch. Once a player is chosen, she is given the opportunity to buy the product (and leave the marketplace forever) or wait and have the opportunity to be called again at future times. In this setting strategic players take the future into account since they have multiple interactions with the environment. As a result, our players are typically non-myopic. This problem can be formulated as a dynamic game with asymmetric information.

In general, one appropriate solution concept for dynamic games with asymmetric information is the perfect Bayesian equilibrium (PBE) [76]. Finding a PBE is a crucial first step for establishing whether an informational cascade occurs. Finding a PBE in a general dynamic scenario with asymmetric information is an extremely challenging task. In [107, 59], the independence of players' types was exploited to introduce a sequential decomposition methodology to find PBE involving strategies with time-invariant domain. This sequential decomposition methodology was based on the common information approach in team problems [58] where the strategies are broken into two partial strategies and dynamic programming equations are used to generate the partial strategies to be applied to the private part of the history. The common information approach in games [107, 59, 108, 40, 41] is what we use in order to characterize PBE in this work.

The first contribution of this work is to characterize a class of PBE where strategies depend on the private observation, as well as the public history of previous actions summarized into a sufficient statistic, the size of which does not increase with time. As a result, equilibrium strategies have a time-invariant domain, and are characterized through the solution of a fixed-point equation (FPE). Furthermore, the domain of the value functions in the FPE we characterize is finite. The finite dimension of the FPE holds even though, for a system with N players, the belief by definition is a probability distribution over a set of size 2^{N+1} (all possible realizations of the quality of the good and players' private observation), and thus it is itself an infinite-dimensional object.

Although this sequential decomposition and the ensuing FPE reduce considerably the problem of finding a PBE, the FPE is still quite cumbersome since it has an exponential dimension in the number of players N (the dimensions of the domain of the value functions). Hence, solving the FPE to find PBE is infeasible for large-scale systems. The second contribution of this work is to show that by exploiting the structure of our model, we can further simplify the FPE such that the dimension of the domain of the value functions only grows quadratically with N.This simplification and the resulting summarizing variables have a very intuitive explanation that relates this model to the original sequential model of [6, 7] and highlights the fundamental differences between the two models. This quadratic-dimension FPE can be solved numerically in practice even for relatively large N. We present numerical results indicating that more collaborative equilibria emerge in this setting if players are sufficiently patient. In particular, players are willing to reveal their information even though they are quite certain that the value of the product is good and they would have bought it if they were acting myopically.

The third contribution of this work is to prove existence for the solution of the FPE and to characterize the structure of the solutions. Structural properties of the equilibrium strategies that apply to all of the solutions of the FPE are investigated. Specifically, the existence of a specific type of strategies, i.e., threshold policies, is proved.

The final contribution of this work is to study whether informational cascades can occur in this model. We study two settings. In the first setting, the discount factor, δ, is strictly below one. We show that in this case, the probability of a cascade approaches one as the number of players, N, approaches infinity. Moreover, the number of players who have revealed their information before the cascade occurs is small, which formalizes their inefficiency. The second setting involves a fixed number of players N with the discount factor approaching one. A surprising result emerges in this setting: when the product is bad, there exists a PBE where at least $\frac{N}{2}$ players reveal their information before the wrong cascade, when players buy the product, can occur. Since each revealing player is wrong with probability $p < \frac{1}{2}$, this implies that the probability for a wrong cascade vanishes with N. Furthermore, when the discount factor is exactly one and the product is bad, we show that there exists a PBE where a bad informational cascade does not happen at all.

The rest of this chapter is organized as follows. In section 3.2 we present the model and formulate the game of non-myopic players. In section 3.3 we characterize PBE through a FPE on appropriate beliefs. In section 3.4 we summarize the information contained in the aforementioned beliefs and provide characterization through FPEs with quadratic dimension in N. Existence results and further characterization of equilibrium strategies are presented in section 3.5. In section 3.6 we analyze informational cascades and we show that quite inefficient informational cascades happen with high probability (for large N) for discount factors strictly smaller than one. Furthermore, we show the surprising result that bad informational cascades can be avoided completely when the product is bad. Some numerical results are presented in section 3.7, while conclusions are drawn in section 3.8. Most of the proof of the theorems are relegated to the Appendices.

3.1.1 Notation

We use upper case letters for scalar and vector random variables. We use lower case letters for scalars and bold lower case letters for vectors. We denote the indicator function by $\mathbf{1}_a(b)$, such that $\mathbf{1}_a(b) = 1$ if $a = b$ and $\mathbf{1}_a(b) = 0$ otherwise. The space of distributions on a general set $\mathcal{A}$ is denoted as $\mathbb{P}(\mathcal{A})$.

3.2 Problem Formulation

Consider an infinite horizon dynamic game with N players in the set $\mathcal{N}$. Time is discrete and the current turn is denoted by t, starting from $t = 0$. At each turn, a player is chosen uniformly at random to act, independently between turns. Only a single player acts in each turn. The random index of the acting player at time t is denoted N_t, and its realization is n_t.

There is a product with a random state $V \in \mathcal{V} = \{-1, 1\}$ where $V = -1$ means that the product is bad and $V = 1$ means that the product is good[1]. We define $Q(v) = \mathbb{P}(V = v)$. In the following we assume for simplicity of exposition that $Q(1) = Q(-1) = 0.5$.

Each player has her own private information on the product. The private information of player n is the random variable $X^n \in \mathcal{X} \triangleq \{-1, 1\}$, with distribution

$$Q(x^n|v) = \mathbb{P}(X^n = x^n \,|\, V = v) = \begin{cases} 1-p & x^n = v \\ p & x^n \neq v \end{cases} \tag{3.1}$$

where $p \in (0, 1/2)$. Define the vector of private information as $\boldsymbol{X} = (X_1, ..., X_N)$. The private information is independent between players conditioned on the true value of V, so

$$\mathbb{P}(\boldsymbol{X} = (x^1, \dots, x^N)|V = v) = \prod_{n=1}^{N} Q(x^n|v). \tag{3.2}$$

Player n's action at turn t, denoted by a_t^n, is equal to 1 if player n buys the product at time t and 0 otherwise. Below, we restrict the action sets such that only player n_t can buy the product at time t, and she can do that only once.

Denote $\boldsymbol{a}_{0:t-1} = (\boldsymbol{a}_0, ..., \boldsymbol{a}_{t-1})$ and $n_{0:t} = (n_0, ..., n_t)$, where $\boldsymbol{a}_t = (a_t^n)_{n\in\mathcal{N}}$ is the action profile at time t. The total history of the game at time t is

$$\boldsymbol{h}_t = (v, \boldsymbol{x}, \boldsymbol{a}_{0:t-1}, n_{0:t}) \in \mathcal{H}_t. \tag{3.3}$$

We assume each player can observe all the previous actions taken by the other players, as well as their identities. Hence the common history at time t is

$$\boldsymbol{h}_t^c = (\boldsymbol{a}_{0:t-1}, n_{0:t}) \in \mathcal{H}_t^c. \tag{3.4}$$

The common history of actions provide her with additional information about the quality of the product. Together with her private information, they form the information set of player n at time t, denoted by

$$\boldsymbol{h}_t^n = (x^n, \boldsymbol{a}_{0:t-1}, n_{0:t}) \in \mathcal{H}_t^n. \tag{3.5}$$

We define $\boldsymbol{b}_t = (b_t^n)_{n\in\mathcal{N}}$ with b_t^n equal to 1 if and only if player n has already bought the product before time t. Clearly, $\boldsymbol{b}_t$ can be determined recursively through the publicly observed action profile history $\boldsymbol{a}_{0:t-1}$ and thus it is part of the common history of the players.

[1]In our model we assume the good has infinite many copies, or alternatively the good is a technology that can be adopted by all without scarcity constraints.

A player's pure strategy is a sequence of functions from the information sets of the game to the action space (i.e., a decision whether to buy or not). In this work, we consider pure strategies. Formally, player n's strategy is $\boldsymbol{s}^n = (s_t^n)_{t=0}^{\infty}$, with

$$s_t^n : \mathcal{H}_t^n \to \mathcal{A}^n\left(b_t^n, n_t\right) \tag{3.6}$$

where

$$\mathcal{A}^n\left(b_t^n, n_t\right) = \begin{cases} \{0,1\} & \text{if } b_t^n = 0\,, n_t = n \\ \{0\} & \text{else} \end{cases} \tag{3.7}$$

so that any player n can buy the product only once, and $a_t^n = 0$ for all t afterwards. In all the turns when player n does not act ($n_t \neq n$), she is restricted not to buy ("play zero").

Note that for player n, the unknown variables in $\boldsymbol{h}_t$ are X^{-n} and V. Hence, we define the private belief of player n on the history of the game as $\mu_t^n : \mathcal{H}_t^n \to \mathcal{P}(\mathcal{X}^{-n} \times \mathcal{V})$ and denote the sequence of private beliefs by $\boldsymbol{\mu}^n = (\mu_t^n)_{t\geq 0}$. Taking the expectation with respect to this belief and the strategies in (3.6), we define the expected reward-to-go of player n at time t as

$$R^n\left(\boldsymbol{s}_{t:\infty}, \mu_t^n, \boldsymbol{h}_t^n\right) = \mathbb{E}^{s,\mu_t^n}\left\{\sum_{t'=t}^{\infty} \delta^{t'-t} V A_{t'}^n \mid \boldsymbol{h}_t^n\right\}, \tag{3.8}$$

where $0 \leq \delta \leq 1$ is the discount factor. Note that at most a single term in the sum (3.8) can be non-zero, since $VA_{t'}^n = V$ only in the first time that player n buys the product, and 0 otherwise.

The strategies in (3.6) are functions of x^n, $\boldsymbol{a}_{0:t-1}$ and $n_{0:t}$. While $\boldsymbol{a}_{0:t-1}$ and $n_{0:t}$ are observed by all players, x^n is only known to player n. Throughout the chapter, it will be useful to decompose those strategies into their common and private components as follows.

Definition 1. *Player n at time t observes $\boldsymbol{h}_t^c$ and takes an action $a_t^n = \gamma_t^n\left(x^n\right)$, where $\gamma_t^n : \mathcal{X} \to \mathcal{A}^n\left(b_t^n, n_t\right)$ is the partial function from her private information to her action. These partial functions are generated through some policy*[2]

$$\psi_t^n : \mathcal{H}_t^c \to \{\mathcal{X} \to \mathcal{A}^n\} \qquad \forall n \in \mathcal{N} \tag{3.9}$$

which operates on $\boldsymbol{h}_t^c$ and returns a mapping from x^n to an action a_t^n, so $\gamma_t^n = \psi_t^n[\boldsymbol{h}_t^c]$ and $a_t^n = \psi_t^n[\boldsymbol{h}_t^c](x^n)$.

The above decomposition is a trivial consequence of the fact that any function $\mathcal{H}_t^c \times \mathcal{X} \to \mathcal{A}^n$ is equivalent to a function $\mathcal{H}_t^c \to \{\mathcal{X} \to \mathcal{A}^n\}$. In the first form, the strategy is a direct function of

[2]Throughout the chapter we use square brackets for mappings that produce functions.

both the public history $\boldsymbol{h}_t^c$ and the private signal x^n, so that $a_t^n = s_t^n(\boldsymbol{h}_t^c, x^n)$. In the second form, the strategy is decomposed into two steps: in the first step the public history produces a partial function $\gamma_t^n = \psi_t^n[\boldsymbol{h}_t^c]$, and in the second step this partial function is evaluated at the private signal to generate the final action $a_t^n = \gamma_t^n(x^n) = \psi_t^n[\boldsymbol{h}_t^c](x^n)$. Note that there are only four possible deterministic gamma functions γ_t^n: wait for any x^n (denoted by $\mathbf{0}$), buy for any x^n (denoted by $\mathbf{1}$), buy according to x^n (denoted by $\boldsymbol{I}$) and buy according to $-x^n$. The last one is clearly dominated by one of the other three so it is never considered. Hence, we are left with three possible partial strategies, namely, $\gamma_t^n \in \{\mathbf{0}, \mathbf{1}, \boldsymbol{I}\}$. Furthermore, since every non-acting player is essentially waiting (i.e., playing $\gamma_t^n = \mathbf{0}$ for $n \neq n_t$), in the following we will drop the superscript ^{n} and only refer to the acting player's partial function as $\gamma_t = \psi_t[\boldsymbol{h}_t^c]$.

We conclude this section by remarking that players' strategies and particularly their partial function γ_t are responsible for the revelation of the private information x^n to the rest of the community. Indeed, if a player plays according to $\gamma_t = \boldsymbol{I}$ then she reveals her private information x^n through her action a_t^n. Conversely, if she either plays according to $\gamma_t = \mathbf{0}$, or $\mathbf{1}$, her private information is not revealed. We note that "revealing" is a special case of "signaling", where the exact private information of a player can be inferred as opposed to only some Bayesian estimation of it [109].

3.3 Characterization of Structured Perfect Bayesian Equilibria

3.3.1 Perfect Bayesian Equilibrium

Our main goal is to study if an informational cascade occurs in the above setting. An informational cascade is defined as a state of the game where learning stops since actions no longer reveal new information. To do so, we first have to study the equilibrium strategies of this game. Since this is a dynamic game with asymmetric information, an appropriate solution concept is the PBE [76], defined as follows.

Definition 2. *A PBE with pure strategies is a pair $(\boldsymbol{s}^*, \boldsymbol{\mu}^*)$ of*

- *a strategy profile $\boldsymbol{s}^* = (\boldsymbol{s}^{*n})_{n\in\mathcal{N}}$,*
- *a belief profile sequence $\boldsymbol{\mu}^* = (\boldsymbol{\mu}^{*n})_{n\in\mathcal{N}}$,*

such that sequential rationality holds, i.e., for each $n \in \mathcal{N}$, $t \geq 0$ and $\boldsymbol{h}_t^n \in \mathcal{H}_t^n$, and each strategy $\boldsymbol{s}^n$:

$$R^n\left(s_{t:\infty}^{*n}, s_{t:\infty}^{*-n}, \mu_t^{*n}, \boldsymbol{h}_t^n\right) \geq R^n\left(s_{t:\infty}^{n}, s_{t:\infty}^{*-n}, \mu_t^{*n}, \boldsymbol{h}_t^n\right), \tag{3.10}$$

and the beliefs satisfy Bayesian updating whenever $\mathbb{P}^{s^*}(\boldsymbol{h}_t^n|\boldsymbol{h}_{t-1}^n) > 0$*, where we define*

$$\mathbb{P}^{s^*}(\boldsymbol{h}_t^n|\boldsymbol{h}_{t-1}^n) = \sum_{\boldsymbol{h}_t:\boldsymbol{h}_t^n}\sum_{\boldsymbol{h}_{t-1}} P^{s^*_{t-1}}(\boldsymbol{h}_t|\boldsymbol{h}_{t-1})\mu^{*n}_{t-1}(\boldsymbol{h}_{t-1}|\boldsymbol{h}_{t-1}^n), \tag{3.11}$$

and $P^{s^*_{t-1}}$ *is the kernel that describes the transition probability from* $\boldsymbol{h}_{t-1}$ *to* $\boldsymbol{h}_t$*. We define* $\boldsymbol{h}_t : \boldsymbol{h}_t^n$ *to be all* $\boldsymbol{h}_t$ *that ae consistent with* $\boldsymbol{h}_t^n$*.*

The Bayesian update of the beliefs is described in the following. For those $\boldsymbol{h}_t$ *that are consistent with* $\boldsymbol{h}_t^n$*, we write*

$$\mu_t^{*n}(\boldsymbol{h}_t|\boldsymbol{h}_t^n) = \mathbb{P}^{s^*}(\boldsymbol{h}_t|\boldsymbol{h}_t^n) = \frac{\sum_{\boldsymbol{h}_{t-1}} P^{s^*_{t-1}}(\boldsymbol{h}_t|\boldsymbol{h}_{t-1})\mu_t^{*n}(\boldsymbol{h}_{t-1}|\boldsymbol{h}_{t-1}^n)}{\sum_{\boldsymbol{h}_t:\boldsymbol{h}_t^n}\sum_{\boldsymbol{h}_{t-1}} P^{s^*_{t-1}}(\boldsymbol{h}_t|\boldsymbol{h}_{t-1})\mu_t^{*n}(\boldsymbol{h}_{t-1}|\boldsymbol{h}_{t-1}^n)}, \tag{3.12}$$

and if $\boldsymbol{h}_t$ *is not consistent with* $\boldsymbol{h}_t^n$*, we have* $\mu_t^{*n}(\boldsymbol{h}_t|\boldsymbol{h}_t^n) = 0$*.*

We remark that strategies and beliefs should be defined for all information sets, even those that occur with zero probability under equilibrium strategies (off-equilibrium paths). In our setting, there are both public and private off-equilibrium paths. The public off-equilibrium paths (i.e., paths where all players can confirm that there was a deviation from equilibrium) are those for which $a_{t-1}^{n_{t-1}} = 0$, but $s^{*n_{t-1}}(x^{n_{t-1}}, \boldsymbol{h}_{t-1}^c) = 1$, for all $x^{n_{t-1}}$ or similarly, $a_{t-1}^{n_{t-1}} = 1$, but $s^{*n_{t-1}}(x^{n_{t-1}}, \boldsymbol{h}_{t-1}^c) = 0$, for all $x^{n_{t-1}}$. In both of these situations, we have $\mathbb{P}^{s^*}(\boldsymbol{h}_t^n|\boldsymbol{h}_{t-1}^n) = 0$ and we pose no restriction on the belief updating. As will be shown in Lemma 1, in both of these cases, the beliefs are not updated for on-equilibrium actions, and so we choose to not update them even if the actions are not according to the equilibrium strategies. The beliefs at the continuation of the game from these points on, however, will be updated according to Bayes' rule if $\mathbb{P}^{s^*}(\boldsymbol{h}_t^n|\boldsymbol{h}_{t-1}^n) > 0$. The private off-equilibrium paths (i.e., paths where all players other than the acting player do not have a way to confirm if a deviation from equilibrium occurred) are when $s^{*n_{t-1}}(x^{n_{t-1}} = 1, \boldsymbol{h}_{t-1}^c) = 1$ and $s^{*n_{t-1}}(x^{n_{t-1}} = -1, \boldsymbol{h}_{t-1}^c) = 0$ (playing $\gamma_t^{n_t} = \boldsymbol{I}$) and the acting player played $a_{t-1}^{n_{t-1}} = 1$ with a private signal $x^{n_{t-1}} = -1$ or played $a_{t-1}^{n_{t-1}} = 0$ with a private signal $x^{n_{t-1}} = 1$, and she has not yet revealed her private information. In this situation, no player other than player n_{t-1} is aware of the deviation because both actions are possible. We impose the restriction on player n_{t-1}'s belief to not be updated at the time of her deviation, although other players update their beliefs about $x^{n_{t-1}}$ and consequently v. Intuitively, a player can not learn anything more by her own actions but she can induce different beliefs in others. One can refer to [76, 110] in order to justify this constraint on the off-equilibrium beliefs. Specifically, one of the conditions posed on off-equilibrium beliefs for PBE is referred to as "no signaling what you don't know" [110, p. 332]. This condition indicates that if one considers two different action profiles in which a specific player's action is the same, the belief about that player's type should be updated similarly for both action profiles. This implies that in our setting, the acting player should not change her belief about any other player's private signal because they are not playing. On the other hand, learning about v happens through players' private signals. If the belief about others' private signals does not change, the belief about v should not change either. So the acting player should not change her belief about neither v nor others' private signals when she is playing, no matter what she plays and whether she deviates or not.

Notice that in the above definition, we have defined the belief μ_t^n to be a belief over the set of all histories at time t given the information set of player n at t. However, this is the most general belief that one could consider and depending on the specifics of the game, we can define other (simpler) types of beliefs that are sufficient for the players to act rationally. Note that for any types of beliefs that we consider, the update rule is the same as what was described here.

In this work, we are interested in PBE that depend on the history of the game only through a summary in the form of the belief of the players about V and X. Hence, we formulate FPE for which the set of solutions is the set of these PBE, which are known as structured PBE [59]. Structured PBEs represent a more reasonable behavior since strategies that depend on sequentially updatable beliefs are more tractable than strategies that require tracking the whole history.

3.3.2 Characterization of Structured PBE

We now present a methodology for characterizing PBE where the strategy for the acting player n_t depends on the common history only through the common belief on the variables V, X (as well as the variable $\boldsymbol{B}_{t-1}$). In particular, we define the common belief $\pi_t \in \mathcal{P}\left(\mathcal{X}^N \times \mathcal{V}\right)$ where $\pi_t(\boldsymbol{x}, v) := \mathbb{P}^s\left(X=\boldsymbol{x}, V=v|\boldsymbol{a}_{0:t-1}, \boldsymbol{b}_{0:t-1}, n_{0:t}\right) = \mathbb{P}^{\psi}\left(X=\boldsymbol{x}, V=v|\boldsymbol{a}_{0:t-1}, \boldsymbol{b}_{0:t-1}, n_{0:t}, \gamma_{0:t-1}\right)$. For $t=0$, we set $\pi_0(\boldsymbol{x}, v) = Q(v)\prod_n Q(x^n|v)$. We first show that the belief π_t can be updated using only public information and that the update depends on ψ_t only through γ_t. Note that the dependence of the update equation on γ_t is the manifestation of "signaling" in our model. When the equilibrium strategy is $\gamma_t = \boldsymbol{I}$, acting player's action reveals her private information and changes the beliefs of other players about V and X.

Lemma 1. *There exists a function F such that the belief π_t can be updated as $\pi_{t+1} = F(\pi_t, \gamma_t, a_t^{n_t}, n_t)$. In particular, if $\gamma_t \neq \boldsymbol{I}$, the belief is not updated.*

Proof. *By simple application of Bayes' rule we have*

$$\pi_{t+1}\left(\boldsymbol{x}, v\right) = \mathbb{P}^s\left(\boldsymbol{x}, v|\boldsymbol{a}_{0:t}, \boldsymbol{b}_{0:t}, n_{0:t+1}\right) \tag{3.13a}$$

$$= \mathbb{P}^{\psi}\left(\boldsymbol{x}, v | \boldsymbol{a}_{0:t}, \boldsymbol{b}_{0:t}, n_{0:t+1}, \gamma_{0:t}\right) \tag{3.13b}$$

$$= \mathbb{P}^{\psi}\left(\boldsymbol{x}, v | \boldsymbol{a}_{0:t}, \boldsymbol{b}_{0:t-1}, n_{0:t}, \gamma_{0:t}\right) \tag{3.13c}$$

$$= \frac{\mathbb{P}^{\psi}\left(\boldsymbol{x}, v, \boldsymbol{a}_t | \boldsymbol{a}_{0:t-1}, \boldsymbol{b}_{0:t-1}, n_{0:t}, \gamma_{0:t}\right)}{\mathbb{P}^{\psi}\left(\boldsymbol{a}_t | \boldsymbol{a}_{0:t-1}, \boldsymbol{b}_{0:t-1}, n_{0:t}, \gamma_{0:t}\right)} \tag{3.13d}$$

$$= \frac{\mathbb{P}^{\psi}\left(\boldsymbol{a}_t | \boldsymbol{x}, v, \boldsymbol{a}_{0:t-1}, \boldsymbol{b}_{0:t-1}, n_{0:t}, \gamma_{0:t}\right) \mathbb{P}^{\psi}\left(\boldsymbol{x}, v | \boldsymbol{a}_{0:t-1}, \boldsymbol{b}_{0:t-1}, n_{0:t}, \gamma_{0:t}\right)}{\mathbb{P}\left(\boldsymbol{a}_t | \boldsymbol{a}_{0:t-1}, \boldsymbol{b}_{0:t-1}, n_{0:t}, \gamma_{0:t}\right)} \tag{3.13e}$$

$$= \frac{\boldsymbol{1}_{\gamma_t(x^{n_t})}\left(a_t^{n_t}\right) \pi_t\left(\boldsymbol{x}, v\right)}{\sum_{\boldsymbol{x'}, v'} \boldsymbol{1}_{\gamma_t(x'^{n_t})}\left(a_t^{n_t}\right) \pi_t\left(\boldsymbol{x'}, v'\right)}. \tag{3.13f}$$

Note that if γ_t is a constant function (i.e., $\gamma_t \neq \boldsymbol{I}$) the quantity $\boldsymbol{1}_{\gamma_t(x^{n_t})}(a_t^{n_t})$ cancels from numerator and denominator of the above expression, thus resulting in $\pi_{t+1} = \pi_t$. Furthermore, whenever the denominator is zero (off-equilibrium paths) we set $\pi_{t+1} = \pi_t$. Additionally, while $\pi_t(\boldsymbol{x}, v)$ depends on $\boldsymbol{a}_{0:t}$, $n_{0:t+1}$ and $\boldsymbol{b}_{0:t}$, the update function F only depends on $\pi_t, \gamma_t, a_t^{n_t}$ and n_t. By definition of the game, $a_t^m = 0$ for all $m \neq n_t$, and so factors of the form $\boldsymbol{1}_{\boldsymbol{0}}(a_t^m)$ are canceled from both numerator and denominator in the last equality.

The private beliefs of players on v on equilibrium paths are obtained by conditioning the public belief on V on players' private signal, X^n. More specifically, player n's private belief on equilibrium path is $\mu_t^n(v) = \pi_t(v|x^n) = \frac{\pi_t(x^n,v)}{\pi_t(x^n)}$, where $\pi_t(x^n, v)$ and $\pi_t(x^n)$ are marginal beliefs of $\pi_t(\boldsymbol{x}, v)$.

A player is only interested in the previous actions since they carry information about V. However, not every action reveals the private information of the acting player. For that to happen, the action that the player took must be determined by her private information. This motivates characterizing the beliefs using the following finite dimensional variables:

Definition 3. *Let $\tilde{x}_t^n \in \{0, -1, 1\}$ be the revealed information of player n up to time t, so $\tilde{x}_t^n = 0$ if the player has not yet revealed her private information, while $\tilde{x}_t^n = \pm 1$ if the player has already revealed her private signal and the value is as indicated. The quantity $\tilde{x}_t^n$ remains unchanged for non-acting players while it is recursively updated for the acting player as*

$$\widetilde{x}_t^n = f(\tilde{x}_{t-1}^n, \gamma_t, a_t^n) = \begin{cases} 2a_t^n - 1 & \gamma_t = \boldsymbol{I}, \tilde{x}_{t-1}^n = 0 \\ \widetilde{x}_{t-1}^n & o.w., \end{cases} \tag{3.14}$$

with the initial condition $\widetilde{x}_0^n = 0$. Note that $\widetilde{x}_t^n$ is a function of $\tilde{x}_{0:t-1}^n$, $\boldsymbol{a}_{0:t}$ and $n_{0:t}$, or equivalently of $\gamma_{0:t}$, $\boldsymbol{a}_{0:t}$ and $n_{0:t}$. We also define the function $\tilde{F}$ such that

$$\begin{aligned} \tilde{\boldsymbol{x}}_t = \left(\tilde{\boldsymbol{x}}_t^{-n_t}, \tilde{x}_t^{n_t}\right) &= \tilde{F}(\tilde{\boldsymbol{x}}_{t-1}, \gamma_t, a_t^{n_t}, n_t) \\ &\triangleq \left(\tilde{\boldsymbol{x}}_{t-1}^{-n_t}, f(\tilde{x}_{t-1}^{n_t}, \gamma_t, a_t^{n_t})\right) \end{aligned} \tag{3.15}$$

to summarize the recursive update of the entire vector $\tilde{\boldsymbol{x}}_t = (\tilde{x}_t^1, \ldots, \tilde{x}_t^N)$. *Note that only the acting player's component of this vector is updated. Furthermore,* $\tilde{x}_t^n$ *can be derived from the belief* π_t *since if* $\pi_t(x^n) = \boldsymbol{I}_k(x^n)$ *for* $k \in \{-1, 1\}$ *then* $\tilde{x}_t^n = k$ *and otherwise* $\tilde{x}_t^n = 0$.

Following the discussion after Definition 2 and by using $\tilde{\boldsymbol{x}}_{\boldsymbol{t}}$, we characterize the off-equilibrium private beliefs as follows:

$$\frac{\mu_t^n(v=1)}{\mu_t^n(v=-1)} = \frac{\pi_t(v=1)}{\pi_t(v=-1)} \left(\frac{1-p}{p}\right)^{-\tilde{x}_t^n + x^n} . \tag{3.16}$$

Intuitively, equation (3.16) says that if a player has not yet revealed her information ($\tilde{x}_t^n = 0$), then her private likelihood about V is the public likelihood amplified by the private factor $(\frac{1-p}{p})^{x^n}$. If however she has already revealed her information and she is on equilibrium $\tilde{x}_t^n = x^n$ then her private belief is the same as the public belief, which includes her private information since it was revealed. Finally, if she has already revealed her information and she is off-equilibrium $\tilde{x}_t^n = -x^n$ then her private likelihood has to correct for the erroneous public belief through the factor $(\frac{1-p}{p})^{-\tilde{x}_t^n}$ and then amplified by the true factor $(\frac{1-p}{p})^{x^n}$.

The following lemma shows that the common belief decomposes into a belief on v and a belief on x, and that each part can be updated recursively. Specifically, it proves that the private information variables $X^1, \ldots, X^N$ are conditionally independent given v , $\boldsymbol{h}_t^c$ and that the common belief can be expressed in terms of $\widetilde{\boldsymbol{x}}_t$ from Definition 3.

Lemma 2. *The public belief* $\pi_t(\boldsymbol{x}, v) = \mathbb{P}(X = \boldsymbol{x}, V = v | \boldsymbol{h}_t^c)$ *can be decomposed as follows*

$$\pi_t(\boldsymbol{x}, v) = \pi_t(v) \prod_{m=1}^{N} \pi_t(x^m | v) \tag{3.17}$$

where $\pi_t(v) \triangleq \mathbb{P}(V = v \,|\, \boldsymbol{h}_t^c)$ *and* $\pi_t(x^m|v) \triangleq \mathbb{P}(X^m = x^m \,|\, v, \boldsymbol{h}_t^c)$. *Furthermore,*

$$\pi_t(x^m|v) = \begin{cases} \boldsymbol{I}_{\widetilde{x}_t^m}(x^m), & \widetilde{x}_t^m \neq 0 \\ Q(x^m|v), & \widetilde{x}_t^m = 0 \end{cases} \tag{3.18}$$

and the belief on V *can be updated as*

$$\frac{\pi_{t+1}(1)}{\pi_{t+1}(-1)} = \frac{\pi_t(1)}{\pi_t(-1)} \times \begin{cases} q^{2a_t^{n_t}-1}, & \gamma_t = \boldsymbol{I} \text{ and } \tilde{x}_t^{n_t} = 0 \\ 1, & o.w., \end{cases} \tag{3.19}$$

with $q = \frac{1-p}{p}$*. Finally, the belief on* V *can be explicitly expressed as*

$$\frac{\pi_t(1)}{\pi_t(-1)} = q^{\sum_n \tilde{x}_t^n}. \tag{3.20}$$

Proof. *See Appendix B.1.*

We would like to characterize equilibrium strategies for the acting player $a_t^{n_t} = \psi_t[\boldsymbol{h}_t^c](x^{n_t})$ for which the ever-increasing common history $\boldsymbol{h}_t^c = (\boldsymbol{a}_{0:t-1}, n_{0:t})$ is summarized into the time-invariant quantities $(n_t, \pi_t, \boldsymbol{b}_t) \in \mathcal{N} \times \mathbb{P}(\mathcal{X}^N \times \mathcal{V}) \times \{0,1\}^N$, i.e., equilibrium strategies of the form $a_t^{n_t} = \theta[n_t, \pi_t, \boldsymbol{b}_t](x^{n_t})$. In other words, we seek equilibrium strategies where the partial functions are of the form $\gamma_t = \theta[n_t, \pi_t, \boldsymbol{b}_t]$. Thanks to Lemma 2 we know that the beliefs can be summarized using $\tilde{\boldsymbol{x}}$. Hence, with a slight abuse of notation, we can write $\gamma_t = \theta[n_t, \tilde{\boldsymbol{x}}_t, \boldsymbol{b}_t]$.

Using the above structural results for the beliefs, we can construct our finite-dimensional FPE.

Fixed-Point Equation 1 (Finite dimensional)**.** *For every* $n \in \mathcal{N}$*,* $\tilde{\boldsymbol{x}} \in \{-1,0,1\}^N$*,* $\boldsymbol{b} \in \{0,1\}^N$ *we evaluate* $\gamma^* = \theta\left[n, \tilde{\boldsymbol{x}}, \boldsymbol{b}\right]$ *as follows*

- *If* $b^n = 1$ *then* $\gamma^* = \boldsymbol{0}$*.*
- *If* $b^n = 0$ *then* γ^* *is the solution of the following system of equations,* $\forall x^n \in \mathcal{X}$

$$\gamma^*\left(x^n\right) = \arg\max \left\{ \underbrace{\frac{q^{\sum_m \tilde{x}^m - \tilde{x}^n + x^n} - 1}{q^{\sum_m \tilde{x}^m - \tilde{x}^n + x^n} + 1}}_{1=\text{“buy”}}, \underbrace{\frac{\delta}{N}\sum_{n'=1}^{N} V^n\left(x^n, n', \tilde{F}\left(\tilde{\boldsymbol{x}}, \gamma^*, 0, n\right), \boldsymbol{b}\right)}_{0=\text{“don't buy”}} \right\} \tag{3.21a}$$

where the value functions for all $m \in \mathcal{N}$ *satisfy*

$$V^m\left(x^m, n, \tilde{\boldsymbol{x}}, \boldsymbol{b}\right) =$$
$$\begin{cases} 0, & b^m = 1 \\ \frac{\delta}{N}\sum_{n'=1}^{N} V^m\left(x^m, n', \tilde{F}\left(\tilde{\boldsymbol{x}}, \gamma^*, 0, m\right), \boldsymbol{b}\right), & b^m = 0, n = m, \gamma^*\left(x^m\right) = 0 \\ \frac{q^{\sum_{m'} \tilde{x}^{m'} - \tilde{x}^m + x^m} - 1}{q^{\sum_{m'} \tilde{x}^{m'} - \tilde{x}^m + x^m} + 1}, & b^m = 0, n = m, \gamma^*\left(x^m\right) = 1 \\ \frac{\delta}{N}\sum_{n'=1}^{N} \mathbb{E}\left[V^m(x^m, n', \tilde{F}(\tilde{\boldsymbol{x}}, \gamma^*, \gamma^*(X^n), n), \boldsymbol{b}^{-n}B^n)\right], & b^m = 0, n \neq m, \end{cases} \tag{3.21b}$$

where expectation in (3.21b) *is w.r.t. the RVs* X^n *and* B^n *with*

$$\begin{aligned}&\mathbb{P}(X^n = x^n, B^n = b'^n | x^m, n, \tilde{\boldsymbol{x}}, \boldsymbol{b}) \\ &\quad = \mathbb{P}\left(B^n = b'^n | X^n = x^n, x^m, n, \tilde{\boldsymbol{x}}, \boldsymbol{b}\right) \mathbb{P}\left(X^n = x^n | x^m, n, \tilde{\boldsymbol{x}}, \boldsymbol{b}\right),\end{aligned} \tag{3.21c}$$

where

$$\mathbb{P}\left(B^n = 1 | X^n = x^n, x^m, n, \tilde{\boldsymbol{x}}, \boldsymbol{b}\right) = \begin{cases} 1 & \textit{, if } b^n = 1 \textit{ or } \gamma^*\left(x^n\right) = 1 \\ 0 & \textit{, else,} \end{cases} \tag{3.21d}$$

and

$$\mathbb{P}(X^n = x^n | x^m, n, \tilde{\boldsymbol{x}}, \boldsymbol{b}) = \begin{cases} \boldsymbol{1}_{\tilde{x}^n}(x^n) & \textit{, if } \tilde{x}^n \neq 0 \\ \frac{Q(x^n|-1) + Q(x^n|1) q^{\sum_{m'} \tilde{x}^{m'} - \tilde{x}^m + x^m}}{1 + q^{\sum_{m'} \tilde{x}^{m'} - \tilde{x}^m + x^m}} & \textit{, if } \tilde{x}^n = 0. \end{cases} \tag{3.21e}$$

■

Once the mapping $\theta[\cdot]$ has been found through the FPE 1, the PBE strategies and beliefs are generated through the following forward recursion.

- **Initialization**: Let $\tilde{\boldsymbol{x}}_0 = \boldsymbol{0} \in \mathbb{R}^N$.
- **For** $t = 0, 1, 2 \ldots$, $\forall n \in \mathcal{N}$, $\boldsymbol{h}_t^c \in \mathcal{H}_t^c$, $x^n \in \mathcal{X}$:

  1. Compute

  $$s_t^{*n}\left(\boldsymbol{h}_t^n\right) := \begin{cases} \theta\left[n_t, \tilde{\boldsymbol{x}}_t^*\left[\boldsymbol{h}_t^c\right], \boldsymbol{b}_{t-1}\right](x^n) & n = n_t \\ 0 & o.w. \end{cases} \tag{3.22a}$$

  2. Compute π_t^* according to Lemma 2.
  3. Generate the private beliefs μ_t^{*n} from π_t^* by as

  $$\mu_t^{*n}\left(\boldsymbol{x}^{-n}, v\right) = \pi_t^*(\boldsymbol{x}^{-n}|v)\mu_t^{*n}(v) \tag{3.22b}$$

  where

  $$\frac{\mu_t^{*n}(v=1)}{\mu_t^{*n}(v=-1)} = \frac{\pi_t^*(v=1)}{\pi_t^*(v=-1)}\left(\frac{1-p}{p}\right)^{-\tilde{x}^n + x^n}. \tag{3.22c}$$

4. Let $a_t^{n_t} = s_t^{*n}(\boldsymbol{h}_t^n)$. For every $n_{t+1} \in \mathcal{N}$, let $\boldsymbol{h}_{t+1}^c = (\boldsymbol{h}_t^c, a_t^{n_t}, n_{t+1})$ and compute:

$$\tilde{\boldsymbol{x}}_{t+1}^* \left[\boldsymbol{h}_{t+1}^c\right] := \tilde{F}\left(\tilde{\boldsymbol{x}}_t^* \left[\boldsymbol{h}_t^c\right], \theta\left[n_t, \tilde{\boldsymbol{x}}_t^* \left[\boldsymbol{h}_t^c\right], \boldsymbol{b}_{t-1}\right], a_t^{n_t}, n_t\right). \tag{3.22d}$$

The following theorem establishes that the above construction generates a PBE.

Theorem 1. *Whenever FPE 1 has a solution, the forward construction described in* (3.22) *generates a PBE.*

Proof. *See Appendix B.2.*

FPE 1, and in particular in (3.21a) is akin to a dynamic programming FPE in an infinite-horizon stopping-time problem. There is however a significant difference: although player n is deciding about her strategy which will lead to an action by maximizing the reward between buying and waiting, we use the equilibrium strategy γ^* in the update function of the belief π. The reason for this twist is shown in the proof of Theorem 1. This proof shows that player n faces an MDP only if every other player plays according to γ^*, and also, most crucially, if the update of π is according to γ^*. Hence, if these two requirements hold, the best response of player n will give us the PBE strategies, γ^*. Therefore, we have a FPE that contains γ^* in both the left- and right-hand side of the equation. In other words, γ_t^* is an equilibrium strategy only if it is the best response assuming that the belief update $\pi_{t+1} = F(\pi_t, \gamma_t^*, a_t^{n_t}, n_t)$(or equivalently $\tilde{\boldsymbol{x}}_{t+1} = \tilde{F}(\tilde{\boldsymbol{x}}_t, \gamma_{t+1}^*, a_{t+1}^{n_{t+1}}, n_{t+1})$) is evaluated using the equilibrium strategy.

We now provide intuition for the expressions in (3.21b). The first equation describes the case where a player has already bought the product so there is no additional expected reward. The second equation refers to the case where the acting player chooses to wait and so the future reward is averaged over all acting players at time $t+1$ with the beliefs being updated according to the equilibrium strategy γ^* and the action 0. The third equation refers to the case where the acting player chooses to buy the product and thus it receives the expected value estimated by her private belief. Finally, the last equation refers to non-acting players who evaluate their future rewards by taking expectation over all possible acting players at the next stage, as well as the private information of the currently acting player and whether she will buy the product or not.

The domain of the value functions $V^m(\cdot)$ in FPE 1 is finite, with size $2 \times N \times 3^N \times 2^N$. For practical systems with a large number of users N, the exponential dimension of FPE 1 renders the computation of the PBE infeasible. In the next section we show that using the structure of the problem, these equations can be simplified considerably, resulting in quadratic dimension in N. Then, the efficient computation of the PBE would allow characterizing informational cascades in large systems where the implication of a cascade can be dramatic.

3.4 Computing a PBE though a quadratic-dimensional FPE

In this section, we exploit the structure of the problem to simplify FPE 1. This simplification is done in two steps. The first step results in a FPE with value functions having domain that grows polynomially with N, and in particular as $\sim N^4$. However, we only present this result in Appendix B.3 for completeness. The second step results in an even more drastic simplification with strategies and value functions having domain that grows only quadratically with N. The key observation here is that the indexing of the players has no effect on the future reward a player estimates she would get by waiting. Since $\widetilde{\boldsymbol{x}}$ contains this information, it can be reduced to the following two quantities:

Definition 4. *Define the aggregated state information as*

$$y_t = \sum_{n=1}^{N} \widetilde{x}_t^n \in \mathcal{Y} = \{-N, \ldots, N\}. \tag{3.23}$$

Further, define the indicator that player n has revealed her private information as $r_t^n = |\tilde{x}_t^n|$. Using $z_t^n = \max\{r_t^n, b_t^n\}$, define the number of players who cannot reveal their private information after turn t by

$$w_t = \sum_{n=1}^{N} z_t^n \in \mathcal{W} = \{0, \ldots, N\}. \tag{3.24}$$

These are the players that have already revealed their private information or have already bought the product and cannot buy it again.

Since the value function and strategy of players with $b^n = 1$ are evidently 0 and $\gamma^* = \mathbf{0}$, respectively, we only argue for the players with $b^n = 0$ and drop b^n from the state variables. We define the functions $U_a : \mathcal{X} \times \{0,1\} \times \mathcal{Y} \times \mathcal{W} \to \mathbb{R}$ and $U_{na}^{\tilde{r}} : \mathcal{X} \times \{0,1\} \times \mathcal{Y} \times \mathcal{W} \to \mathbb{R}\, \forall \tilde{r} \in \{0,1\}$ as follows. $U_a(x, r, y, w)$ is the value function of the acting player n whose private information is $x^n = x$, she has revealed if $r = 1$ and the aforementioned state variables are $(y_t, w_t) = (y, w)$. Similarly, $U_{na}^{\tilde{r}}(x, z, y, w)$ is the value function of a non-acting player m, whose private information is $x^m = x$, she has revealed if $\tilde{r} = 1$ with an acting player n who can reveal her private information if $z = 0$, and y, w as before.

Finally, define the update functions G^r, G^z, G^y, G^w as follows

$$G^r(r, \gamma) = \begin{cases} 1, & r = 0 \text{ and } \gamma = \mathbf{I} \\ r, & \text{else} \end{cases} \tag{3.25a}$$

$$G^z(z,\gamma,a) = \begin{cases} 1, & z=0 \text{ and } (a=1 \text{ or } \gamma = \mathbf{I}) \\ z, & \text{else} \end{cases} \tag{3.25b}$$

$$G^y(z,y,\gamma,a) = \begin{cases} y+(2a-1), & z=0 \text{ and } \gamma = \mathbf{I} \\ y, & \text{else} \end{cases} \tag{3.25c}$$

$$G^w(z,w,\gamma,a) = w + G^z(z,\gamma,a) - z. \tag{3.25d}$$

We now can formulate the alternative FPE 2.

Fixed-Point Equation 2 (Quadratic dimension)**.** *For every $r \in \{0,1\}$, $y \in \mathcal{Y}$, $w \in \mathcal{W}$, we evaluate $\gamma^* = \phi[r,y,w]$ as follows*

- *γ^* is the solution of*

$$\gamma^*(x) = \arg\max \left\{ \underbrace{\frac{q^{y+r+x}-1}{q^{y+r+x}+1}}_{1=\textit{buy}}, \underbrace{A}_{0=\textit{don't buy}} \right\} \forall x \in \mathcal{X}, \tag{3.26a}$$

where

$$A = \frac{\delta}{N} U_a(x,r',y',w') + \frac{\delta}{N}(N-w-1+r)U_{na}^{r'}(x,0,y',w') + \frac{\delta}{N}(w-r)\,U_{na}^{r'}(x,1,y',w'). \tag{3.26b}$$

where the next state variables are

$$r' = G^r(r,\gamma^*) \tag{3.26c}$$

$$y' = G^y(r,y,\gamma^*,0) \tag{3.26d}$$

$$w' = G^w(r,w,\gamma^*,0). \tag{3.26e}$$

The value functions satisfy

$$U_a(x,r,y,w) = \begin{cases} A & \gamma^*(x)=0 \\ \frac{q^{y+r+x}-1}{q^{y+r+x}+1} & \gamma^*(x)=1 \end{cases} \tag{3.26f}$$

and for all $\tilde{r} \in \{0,1\}$

$$U_{na}^{\tilde{r}}(x,z,y,w) = \frac{\delta}{N}\mathbb{E}\left\{U_a\left(x,\tilde{r},\tilde{Y},\tilde{W}\right)\right\} + \frac{\delta}{N}\mathbb{E}\left\{U_{na}^{\tilde{r}}\left(x,\tilde{Z},\tilde{Y},\tilde{W}\right)\right\}$$

$$+ \frac{\delta}{N}(w - z - \tilde{r})\,\mathbb{E}\left\{U_{na}^{\tilde{r}}\left(x, 1, \tilde{Y}, \tilde{W}\right)\right\}$$
$$+ \frac{\delta}{N}(N - w - 2 + z + \tilde{r})\,\mathbb{E}\left\{U_{na}^{\tilde{r}}\left(x, 0, \tilde{Y}, \tilde{W}\right)\right\}, \tag{3.26g}$$

where the (random) next state variables from the point of view of a non-acting player are:

$$\tilde{Z} = G^{z}\left(z, \gamma^{*}, \gamma^{*}\left(X^{n}\right)\right) \tag{3.26h}$$
$$\tilde{Y} = G^{y}\left(z, y, \gamma^{*}, \gamma^{*}\left(X^{n}\right)\right) \tag{3.26i}$$
$$\tilde{W} = G^{w}\left(z, w, \gamma^{*}, \gamma^{*}\left(X^{n}\right)\right) \tag{3.26j}$$

and the expectation is w.r.t. the RV X^n, where

$$\mathbb{P}(X^{n} = x^{n}|\widetilde{r}, x, w, y) = \frac{Q(x^{n}|-1) + Q(x^{n}|1)q^{y+r+x}}{1 + q^{y+r+x}}.$$

Specifically, for $z = 1$ the above becomes

$$U_{na}^{\tilde{r}}(x, 1, y, w) = \frac{\delta}{N}U_{a}(x, \tilde{r}, y, w) + \frac{\delta}{N}(w - z - \tilde{r} + 1)\,U_{na}^{\tilde{r}}(x, 1, y, w)$$
$$+ \frac{\delta}{N}(N - w - 2 + z + \tilde{r})\,U_{na}^{\tilde{r}}(x, 0, y, w). \tag{3.26k}$$

■

The intuitive explanation for FPE 2 is as follows. Equation (3.26a) quantifies the decision between buying now or waiting, given the quality of information about V evaluated through y. Specifically, the reward-to-go for waiting in (3.26b) averages out the rewards obtained by whether the acting player will also be acting at the next epoch (first term), or whether she will be non-acting and the acting player can reveal her private information or not (the two terms with $z = 0, 1$). Similarly, a non-acting player updates her value function in (3.26g) by averaging out four possibilities for the next epoch: whether she will be the acting player (first term), whether she will be non-acting but the acting player will be the same as in the current epoch (second term), and whether she will be non-acting and the acting player will be some other than herself and the current acting player (last two terms). Specifically, if the current acting player has either bought the product or revealed her private information ($z = 1$) the second term in this equation is absorbed into the third one as shown in (3.26k).

The next Theorem shows that by finding a solution to FPE 2, we obtain a solution to FPE 1. Since equations (3.26) have quadratic dimension in N, this significantly reduces the complexity of solving FPE 1. Specifically, given the solution U^* of FPE 2 (together with ϕ) we construct the following strategies and value functions.

$$\gamma^* = \theta\left[n, \tilde{\boldsymbol{x}}, b^n, \boldsymbol{b}^{-n}\right] = \begin{cases} \phi\left[r^n, y, w\right], & b^n = 0 \\ \mathbf{0}, & b^n = 1 \end{cases} \tag{3.27}$$

$$\tilde{V}^m\left(\cdot, n, \tilde{\boldsymbol{x}}, b^m, \boldsymbol{b}^{-m}\right) = \begin{cases} U_a\left(\cdot, |\tilde{x}^n|, y, w\right), & b^m = 0, m = n \\ U_{na}^{|\tilde{x}^m|}\left(\cdot, \max\{|\tilde{x}^n|, b^n\}, y, w\right), & b^m = 0, m \neq n \\ 0, & b^m = 1, \end{cases} \tag{3.28}$$

where we note that y, w and r^n are all determined by $\tilde{\boldsymbol{x}}$ and n through (3.23) and (3.24). We will show that these value functions are solutions of the original FPE 1.

Theorem 2. *The value functions* $(\tilde{V}^m)_{m\in\mathcal{N}}$ *in* (3.28) *together with the strategy mapping* γ^* *in* (3.27) *satisfy FPE 1.*

Proof. *See Appendix B.4.*

3.5 Equilibrium Analysis

The convenient form of FPE 2 allows us to analyze properties of the PBE and even to verify intuitive PBE solutions. We first present an intermediate lemma which will be useful in proving subsequent results.

Lemma 3. *The following are true for all solutions of FPE 2.*

- *For $\delta = 1$, players with $x = 1$ are indifferent between buying and waiting for $y \geq -1$ and all w and r. Furthermore, players with $x = -1$ are indifferent between buying and waiting for $y + w \geq N$ and all r.*

- *For $\delta < 1$, players with $x = 1$ prefer buying over waiting for $y \geq 0$ and all w and r and for $y = -1$, $r = 1$, and are indifferent between buying and waiting for $y = -1$, $r = 0$. Also, players with $x = -1$ prefer buying over waiting for $y \geq 2$, $y + w \geq N$ and all r and also for $y = 1$, $w \geq N - 1$ and $r = 1$, and are indifferent for $y = 1$, $r = 0$ and $y = 0$, $r = 1$.*

- *For all $\delta \leq 1$, for $y < -1$, all w and $r = 0$ and for both values of x, players prefer to wait. Similarly, for $y < -2$, all w and $r = 1$ and both values of x, players prefer to wait. Finally,* *For $y=-2$, all w and $r=1$ players with $x=1$ are indifferent between buying and waiting and players with $x=-1$ prefer to wait.*

Proof. *See Appendix B.5.*

We comment at this point that the usefulness of the above lemma is the very fact that these statements are proved without explicitly solving FPE 2. Specifically, note that both the right hand side and the left hand side of equation (3.26a) depend on the solution γ^*. However, Lemma 3 claims that for all of the solutions γ^*, the aforementioned properties hold. Most of the remaining results of this section hinge on the above lemma.

The next theorem presents a solution of FPE 2 for all values of $\delta \leq 1$ including $\delta = 0$. For $\delta = 0$, our scenario coincides with the original myopic scenario from [7], up to the fact that players who do not buy get another opportunity to play. Therefore, we refer to the strategy profile in Theorem 3 as the myopic solution, even though it is a solution for all δ.

Theorem 3 (Existence)**.** *The following strategy profile is a solution of FPE 2 for all δ.*
For $r=0$, and all w,

$$\gamma^* = \phi[r,y,w] = \begin{cases} \mathbf{1}, & y \geq 2 \\ \mathbf{0}, & y \leq -2 \\ \boldsymbol{I}, & -1 \leq y \leq 1. \end{cases} \tag{3.29a}$$

For $r=1$, and all w,

$$\gamma^* = \phi[r,y,w] = \begin{cases} \mathbf{1}, & y \geq 2 \\ \mathbf{0}, & y \leq -2 \\ \boldsymbol{I}, & -1 \leq y \leq 0. \end{cases} \tag{3.29b}$$

Finally, for $r=1$, $y=1$, and all w, $\gamma^* = \phi[r,y,w]$ can be chosen appropriately, as a function of δ (it is either $\mathbf{1}$ or $\boldsymbol{I}$).

Proof. *See Appendix B.6.*

This strategy profile is depicted in Fig. 3.1 (in all such figures we present the case for $r=0$ and the case for $r=1$ with $\tilde{x}=-1$, since if a player has revealed and her private information is 1 it means that she has bought the product already). Notice that it mostly consists of strategies $\gamma^* = \mathbf{1}$ and $\gamma^* = \mathbf{0}$ which implies that players do not tend to reveal their private signal. Intuitively, if a player knows that others do not reveal their private signal, she does not gain from waiting for more information. Hence, revelation of private signals, which occur when $\gamma^* = \boldsymbol{I}$ is played, does not

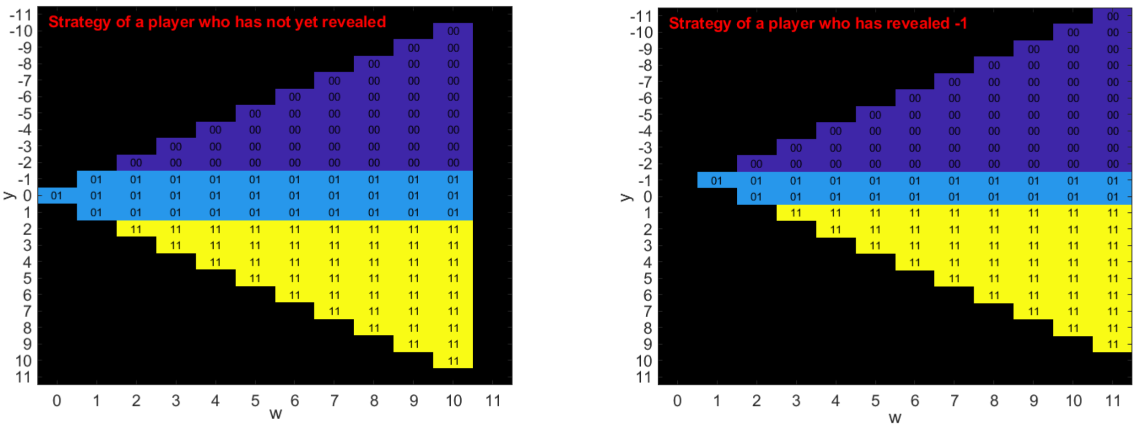


Figure 3.1: Equilibrium strategies for $N = 11$ and all $\delta \leq 1$, including $\delta = 0$. "00", "01", and "11" denote strategies $\mathbf{0}$, $\boldsymbol{I}$, and $\mathbf{1}$, respectively. The strategies for $r = 1$ and $y = 1$ are specifically for $\delta = 0$.

happen when both players with $x = 1$ and $x = -1$ have positive instantaneous reward. Therefore, for all values of δ, acting myopically is always an equilibrium.

Although the strategy profile of Theorem 3 is referred to as myopic, it captures the non-myopic aspect of the game too. For instance, at $y = 1$, a player with $r = 0$ and $x = -1$, does not buy the product because her value function is positive by not buying and therefore, she gains from waiting. But in the myopic case, her valuation is 0 for both buying and not buying so the player is indifferent between playing $a = 1$ and $a = 0$. This implies that if we change the apriori belief about V from $Q(v = 1) = 0.5$ to $Q(v = 1) = 0.5 + \epsilon$ for small enough ϵ, this strategy profile is an equilibrium for $\delta \neq 0$ but not for $\delta = 0$. This follows since a player with $x = -1$ at $y = 1$ strictly prefers to buy at the myopic setting, while she still prefers to wait at the non-myopic setting if ϵ is small enough.

We next investigate a solution for FPE 2 for both $\delta = 1$ and large enough $\delta < 1$.

Theorem 4. *The following strategy profile is a solution of FPE 2,*

- *For* $\delta = 1$,

$$\gamma^* = \phi[r, y, w] = \begin{cases} \mathbf{0}, & y \leq -2 \\ \boldsymbol{I}, & y \geq -1, w < N \\ \mathbf{1}, & y \geq 1, w = N, r = 1 \\ \boldsymbol{I}, & y \in \{0, -1\}, w = N, r = 1 \end{cases} \tag{3.30a}$$

- *For large enough $\delta < 1$ (which depends on N and other parameters of the game),*

$$\gamma^* = \phi[r, y, w] = \begin{cases} \mathbf{0}, & y \leq -2 \\ \boldsymbol{I}, & y \geq -1, y + w < N \\ \boldsymbol{I}, & y = 1, w = N - 1, r = 0 \\ \boldsymbol{I}, & y = 0, w = N, r = 1 \\ \mathbf{1}, & y \geq 2, y + w \geq N \\ \mathbf{1}, & y = 1, w \geq N - 1, r = 1 \end{cases} \tag{3.30b}$$

Proof. *See Appendix B.7*

The strategy profiles presented in Theorem 4 are depicted in Fig. 3.2 and 3.3 for $N = 11$ and $\delta = 1$ and large enough $\delta < 1$, respectively. Note that the strategy $\gamma^* = \boldsymbol{I}$ (denoted by 01) is extended throughout all the states with $y \geq -1$ and $w < N$ for $\delta = 1$.

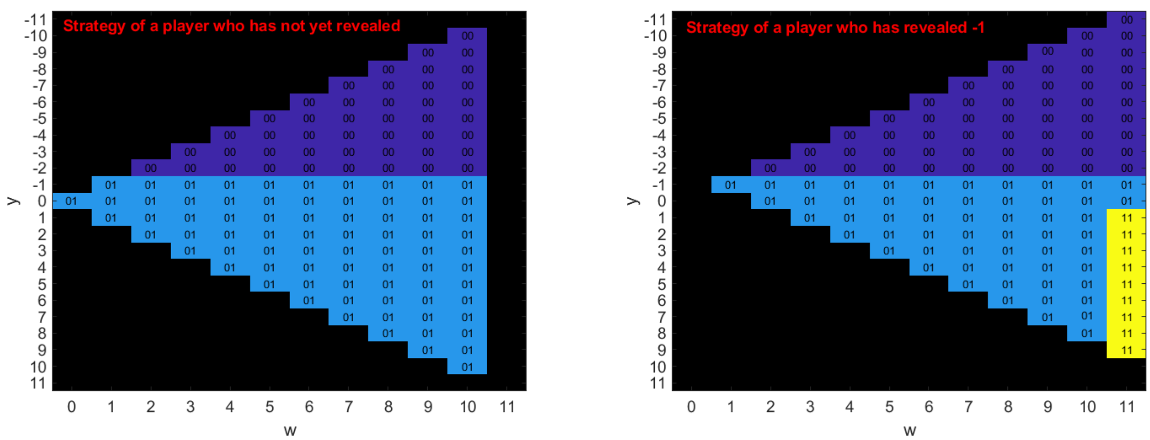


Figure 3.2: Equilibrium strategies for $N = 11$ and $\delta = 1$. "00", "01", and "11" denote strategies $\mathbf{0}$, $\boldsymbol{I}$, and $\mathbf{1}$, respectively.

The FPE 2 may exhibit more PBE than the PBE of Theorem 3. Nevertheless, all these potential PBE share similar structure, as the next theorem shows. We have also presented the existence results for the solutions that are threshold policies w.r.t. w and y in the next theorem.

Theorem 5. *The following properties hold for the solutions of FPE 2 for $b = 0$:*

- *All of the solutions of FPE 2 are either threshold policies (from 0 to 1) w.r.t. w or there exists a threshold policy w.r.t. w corresponding to a solution that is not of this type.*

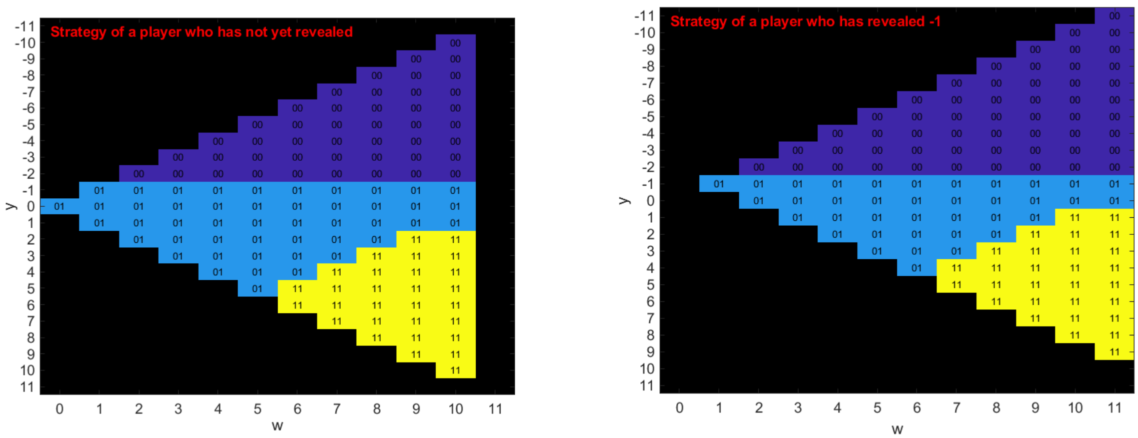


Figure 3.3: Equilibrium strategies for $N = 11$ and large enough $\delta < 1$. "00", "01", and "11" denote strategies $\mathbf{0}$, $\boldsymbol{I}$, and $\mathbf{1}$, respectively.

- *For $\delta < 1$, all of the solutions of FPE 2 that are threshold functions w.r.t. w, must be threshold functions w.r.t. y for $r = 0$, when all other parameters are fixed. This implies that if $\gamma^*(x) = \phi[0, y, w](x) = 1$, then $\gamma^*(x) = \phi[0, y', w'](x) = 1$ for $y' \geq y$ and $w' \geq w$. Further whenever the solution is threshold policy w.r.t. y for $r = 0$, the solutions can also be threshold policy w.r.t. y for $r = 1$.*

Further, for all of the solutions of FPE 2, we have the following properties:

- *They are threshold functions w.r.t. y for $x = 1$ and $r = 0$, and the threshold is either $y = -1$ or $y = 0$ for all w.*

- *They are such that $\gamma^* = \phi[r, y, w] = \mathbf{0}$ for $y \leq -3$ and all other parameters, and for $y = -2$ and $r = 0$. Also, $\gamma^* = \phi[0, y, w] \neq \mathbf{0}$ for $y \geq 0$.*

- *We have $\gamma^* = \phi[0, 0, w] = \boldsymbol{I}$.*

- *For $y \neq -1$, $\gamma^* = \phi[0, y, w] = \mathbf{0}$ for all w (constant w.r.t. w) or can only be either $\gamma^* = \phi[0, y, w] = \boldsymbol{I}$ or $\gamma^* = \phi[0, y, w] = \mathbf{1}$ for all w. It implies that by changing w and fixing other parameters, the equilibrium strategies either do not change and are always $\mathbf{0}$, or they can change between $\boldsymbol{I}$ and $\mathbf{1}$.*

- *For $y = -1$, both $\gamma^* = \phi[0, -1, w] = \boldsymbol{I}$ and $\gamma^* = \phi[0, -1, w] = \mathbf{0}$ are always solutions for all w.*

- *For $y = -2$, both $\gamma^* = \phi[1, -2, w] = \boldsymbol{I}$ and $\gamma^* = \phi[1, -2, w] = \mathbf{0}$ are always solutions for all w.*

Proof. *See Appendix B.8.*

The first two parts of this theorem imply that there exist solutions of FPE 2 that by increasing y or w, the equilibrium strategies change from $\mathbf{0}$'s to $\boldsymbol{I}$'s and then to $\mathbf{1}$'s. This is evident in all of the solutions that we have proposed in this chapter (Fig. 3.1, 3.2, 3.3). Other parts present more general statements about the solutions. For instance, as we can see in the proposed solutions, the equilibrium strategies are $\gamma^* = \phi[r, y, w] = \mathbf{0}$ for $y \leq -2$, which is because the instantaneous reward of players is non-positive. One can also verify other parts of the theorem by the solutions proposed in this chapter.

The boundary of $|y| = 2$ is of special importance for the equilibria. The reason is that $y = 2$ is the smallest y for which the instantaneous reward is positive for all players, regardless of their private information. Similarly, $y = -2$ is the largest y for which the instantaneous reward is negative for all players regardless of their private information. For the myopic scenario, these facts determine the equilibrium strategies at $y \geq 2$ and $y \leq -2$, and this is a possible PBE in our non-myopic scenario as well, as Theorem 3 shows. However, in our non-myopic scenario, more intricate behaviors are also possible at equilibrium. For $y \leq -2$, waiting, which gives zero reward, is always better than buying. Therefore this side of the boundary behaves like the myopic scenario in all PBE. Nevertheless, for $y \geq 2$, players may choose to wait and not buy the product even if their instantaneous reward is positive. Therefore, we may have signaling strategies for $y \geq 2$ and hence, we observe different equilibrium strategies for these values of y in Theorem 4.

3.6 Informational Cascades

Our results from the previous section allow us to evaluate PBE of the game by solving equations with a quadratic number (in N) of variables. This methodology provides us with the necessary tools to investigate whether informational cascades occur in settings with large number of players.

Definition 5. *An informational cascade is a sequence of turns in our game, starting from some $t_0 \geq 0$, such that $\gamma_t \neq \boldsymbol{I}$ for all $t \geq t_0$. We say that an informational cascade is bad if it leads to the wrong decision: users choose $\gamma = \mathbf{0}$ when $V = 1$ or $\gamma = \mathbf{1}$ when $V = 0$.*

While the sequence of events in a realization of the game is random, given a PBE, we can identify the histories of the game at which an informational cascade occurs. Using Theorem 2, we can characterize these histories using only w and y.

According to the definition above, an informational cascade can affect any number of players, from 1 to N. Obviously, informational cascades that affect more players are more significant. A natural question is then how much damage a bad informational cascade causes to the network. Our FPE 2 with its variables (y, w) gives an easy way to tackle this question. If the cascade occurred at state (y, w), then two things affect the damage done to the community: the probability that the cascade is bad, and the number of players that received the worst possible reward if the cascade is bad. Interestingly enough, both of these numbers are characterized by w.

The players that participate in a bad cascade with $\gamma = \mathbf{0}$ receive 0 reward, which is the worst possible. The best reward is 1 up to the discounting in the first turn they get to act. The number of players that made the right decision and bought the product before the cascade is bounded from above by w.

The players that participate in a bad cascade with $\gamma = \mathbf{1}$ receive a reward of -1 up to the discounting in the first turn they get to act. This is the worst reward possible, while the best reward is 0. The number of players that initially made the right decision not to buy the product is bounded from above by w. Indeed, any such player must have played $\gamma = \boldsymbol{I}$ since otherwise, she would have started a $\gamma = \mathbf{0}$ cascade instead.

We conclude that in any bad cascade, at least $N - w$ players receive the worst reward possible. Hence, both the probability for a bad cascade and the damage it causes decrease with w. Using w, one can bound the system performance using a metric of choice (e.g., social welfare, or some notion of fairness). In the next section, we numerically evaluate the probability for a bad cascade as a function of w.

A direct consequence of our model, which induces players to be forward-looking instead of acting myopically, is a multitude of equilibrium behaviors for the players. This rich spectrum of behaviors includes the myopic strategies that have been reported in the literature and that lead to informational cascades, but also–and more importantly–includes more cooperative strategies that induce players to reveal their information with the potential of alleviating or even eliminating informational cascades. The next two subsections explore these two extremes by proving conclusively the above claims.

3.6.1 The case of $\delta < 1$ and $N \to \infty$

In this part, we employ the results of the previous section to conclude that an informational cascade indeed happens with probability approaching 1 as the number of players approaches infinity even in a non-myopic scenario for a fixed $\delta < 1$.

Our methodology consists of defining a Markov chain and studying its properties. Specifically this Markov chain is not defined on absolute time t, but on the random times when a new revelation happens (i.e., when a player plays strategy $\gamma_t = \boldsymbol{I}$). Towards this goal we provide the following definition.

Definition 6. *Let $\phi[\cdot]$ be a solution to FPE 2. Define the random variables $(D_t)_{t\geq 0}$ with realization*

$$d_t = \begin{cases} 1, & \phi[r_t^{n_t}, y_t, w_t] = \boldsymbol{I} \text{ and } r_t^{n_t} = 0 \text{ and } b_t^{n_t} = 0 \\ 0, & \text{else}, \end{cases} \tag{3.31}$$

which indicates if the player who acts at turn t reveals her private information. Let Y_t be the random aggregated state information at time t (see (3.23)). Let T_i be the random time of the i-th revealing, so $T_0 = 0$ and $T_i = \min\{t > T_{i-1} | D_t = 1\}$ for $i \geq 1$. We also define the random process $(\overline{Y}_i)_{i\geq 0}$ with $\overline{Y}_i = Y_{T_i}$ when $T_i < \infty$ and $\overline{Y}_i = \overline{Y}_{i-1}$ otherwise.

The next lemma characterizes the reason why cascades still occur in a non-myopic scenario.

Lemma 4. *Let $\phi[\cdot]$ be a solution to FPE 2. The induced process $(\overline{Y}_i)_{i\geq 0}$ is a Markov chain where, for large enough N, there exist absorbing states y_R, y_L such that for all $y_L < y < y_R$, if $T_{i+1} < \infty$ then*

$$\mathbb{P}\left(\overline{Y}_{i+1} = y' | \overline{Y}_i = y\right) = \begin{cases} \frac{p+(1-p)q^y}{q^y+1} & y' = y+1 \\ \frac{1-p+pq^y}{q^y+1} & y' = y-1 \end{cases}. \tag{3.32}$$

Proof. *First we show the Markovianity of $(\overline{Y}_i)_{i\geq 0}$*

$$\begin{aligned} \mathbb{P}\left(\overline{Y}_{i+1} = y' | \overline{Y}_{0:i} = y_{0:i}\right) &= \mathbb{P}\left(Y_{T_i} + X^{N_{T_i}} = y' \,|\, Y_{T_{0:i}} = y_{0:i}\right) && \text{(3.33a)} \\ &= \mathbb{P}\left(X^{N_{T_i}} = y' - y_i | Y_{T_{0:i}} = y_{0:i}\right) && \text{(3.33b)} \\ &= \frac{Q(y' - y_i|0) + Q(y' - y_i|1)q^{y_i}}{q^{y_i} + 1} && \text{(3.33c)} \\ &= \mathbb{P}\left(\overline{Y}_{i+1} = y' | \overline{Y}_i = y_i\right). && \text{(3.33d)} \end{aligned}$$

Now we characterize the absorbing states. For $\delta < 1$ and $Y_{\max} = \left\lceil 1 + \log_q(\frac{1+\delta}{1-\delta}) \right\rceil < N$, we have

$$\frac{q^{Y_{\max}+r_t+x} - 1}{q^{Y_{\max}+r_t+x} + 1} > \delta > \delta U_a\left(x, r_{t+1}, y_{t+1}, w_{t+1}\right). \tag{3.34}$$

Hence, either $y_R = Y_{\max}$ is absorbing or there exists a $y_R < Y_{\max}$ that is absorbing. In $Y_t = y_R$, all players, regardless of x, prefer to buy. Similarly, for $Y_{\min} = -2$ we have

$$\frac{q^{-1} - 1}{q^{-1} + 1} = 2p - 1 < 0 < \delta U_a\left(x, r_{t+1}, y_{t+1}, w_{t+1}\right). \tag{3.35}$$

Therefore, either $y_L = Y_{\min} = -2$ or $y_L = -1$ is absorbing. In $Y_t = y_L$, all players, regardless of x, prefer to wait. Hence, in $Y_t = y_L$ or $Y_t = y_R$ no more revealings occur and Y_t (and $\overline{Y}_i$) remains constant for all $t' > t$ with probability 1.

The absorbing states of the Markov chain we defined above are informational cascades. As a result, an informational cascade will occur with probability approaching 1 as N increases as in the gambler's ruin problem. However, an informational cascade that occurs after (almost) all players have revealed their private information is of little concern. In such a case (almost) all available information about V has been revealed, so w is close to N and the cascade affects only a few players and also has small probability to be bad. Unfortunately, for a fixed $\delta < 1$, the following theorem shows that this is far from being the case, as an informational cascade occurs early on.

Theorem 6. *For $\delta < 1$, the probability that an informational cascade occurs in finite time approaches 1 as $N \to \infty$.*

Furthermore, let M_N be a sequence such that $\lim_{N\to\infty} \frac{M_N}{\sqrt{N}} = 0$ and $\lim_{N\to\infty} M_N = \infty$.

1. *The probability that less than M_N players have revealed their private information before the cascade occurred approaches 1 as $N \to \infty$.*

2. *If, in addition, the solution is such that $\phi[r, y, w] = \mathbf{1}$ implies $\phi[r, y, \widehat{w}] = \mathbf{1}$ for all $\widehat{w} > w$ (according to Theorem 5, we know such solutions exist), then the cascade happens in less than M_N turns with a probability that approaches 1 as $N \to \infty$.*

Proof. *See Appendix B.9.*

Theorem 6 implies that an informational cascade will occur at some finite time with probability approaching 1 as N increases. Secondly, the theorem implies that the cascade happens too early. This follows since when a cascade occurs, with high probability, less than M_N players have revealed their information for any increasing sequence that grows slower than $\sqrt{N}$ (e.g., $M_N = \log N$). Hence, a minuscule amount of the available information about V has been revealed before a cascade occurs (for large N). This is undesirable, since it means that the probability for a bad cascade can be significant, and that the cascade will affect almost all of the players.

3.6.2 The case of $\delta = 1$ or large enough $\delta < 1$ and finite N

In this subsection, we study informational cascades for a fixed N and for either $\delta = 1$ or large enough $\delta < 1$. We refer to these cases as infinitely patient and sufficiently patient players, respectively. As it will be shown, a very surprising result emerges in this setting. For $\delta = 1$ and

$V = -1$, there exists a PBE that completely avoids bad information cascades. For $V = -1$ and with large enough $\delta < 1$, there exists a PBE that has a vanishing probability (in N) for a bad information cascade, since it is guaranteed that at least half of the players will reveal their private information. The next two theorems formalize these results.

Theorem 7. *For $\delta = 1$, there exists a PBE in which there is no bad informational cascade for $V = -1$.*

Proof. *Consider the strategy profile of Theorem 4 for $\delta = 1$ (depicted in Fig. 3.2). There is no strategy $\gamma^* = \phi[r = 0, y, w] = \mathbf{1}$. This means that for $V = -1$, bad informational cascades never happen for this strategy profile.*

Although Theorem 7 states that bad informational cascades can be avoided for $V = -1$, they will always happen for $V = 1$ with positive probability due to the strategies $\gamma^* = \phi[r = 0, y, w] = \mathbf{0}$ that are played for $y \leq -2$ and all w.

Theorem 8. *For sufficiently large $\delta < 1$ (which depends on N) there exists a PBE for which bad informational cascades for $V = -1$ happen only when at least half of the players have revealed their private information. Consequently, the probability that a bad informational cascade for $V = -1$ happens is bounded from above by $e^{-\frac{N}{4}(1-2p)^2}$.*

Proof. *Assume that $\delta < 1$ is large enough such that the strategy profile of the second part of Theorem 4 (depicted in Fig. 3.3) is a PBE. This strategy profile consists of strategies $\gamma^* = \phi[0, y, w] = \mathbf{1}$ for $y \geq 2$ and $y + w \geq N$ (yellow cells in Fig. 3.3). This implies that for $V = -1$, a bad informational cascade happens only when $y \geq 2$ and $y + w \geq N$. This in turn means that a bad informational cascade happens when at least $w = \frac{N}{2}$. Since the initial value of y is 0 and the strategies played before reaching $y \geq 2$ and $y + w \geq N$, are all $\gamma^* = \phi[r, y, w] = \boldsymbol{I}$, then w is equal to the number of players who have revealed. Therefore, a bad cascade happens only when at least $\frac{N}{2}$ players have revealed their private information.*

Let $\mathcal{T}$ be the set of turns when players revealed their private information. Let $R = |\mathcal{T}|$. Let Y_∞ be the random value of y when an information cascade occurs, such that $Y_\infty = \infty$ if it does not occur. Let E be the error event, in which a bad information cascade happens. Then, using that

$p < \frac{1}{2}$ we have

$$
\begin{aligned}
\mathbb{P}(E) &\overset{(a)}{\leq} \mathbb{P}(Y_\infty = \sum_{t \in T} x^{n_t} \leq 0 \mid V = 1)\mathbb{P}(V = 1) + \mathbb{P}(Y_\infty = \sum_{t \in T} x^{n_t} \geq 0 \mid V = -1)\mathbb{P}(V = -1) \\
&= \frac{1}{2}\mathbb{P}(\sum_{t \in T} x^{n_t} - (1-2p)\,R \leq -(1-2p)\,R \mid V = 1) \\
&\quad + \frac{1}{2}\mathbb{P}(\sum_{t \in T} x^{n_t} + (1-2p)\,R \geq (1-2p)\,R \mid V = -1) \\
&\overset{(b)}{\leq} e^{-\frac{R}{2}(1-2p)^2} \overset{(c)}{\leq} e^{-\frac{N}{4}(1-2p)^2}
\end{aligned} \tag{3.36}
$$

where (a) follows since a bad information cascade can only occur if VY_∞ is non-positive, (b) is Hoeffding's inequality for bounded random variables, and (c) uses that $R \geq \frac{N}{2}$.

3.7 Numerical Results

In this section, we present numerical results for the solution of FPE 2. The results were obtained as follows. First an iterative algorithm was used to solve the FPE, much like the value iteration algorithm used in the solution of Markov Decision Processes. The iterative process was run until the value functions converged numerically. In order to verify without a doubt that this solution is an equilibrium, a second step was followed. At the second step, the equilibrium strategy obtained by this iterative process was fixed and a linear system of equations was formulated with unknowns being all value functions. This system was solved using infinite precision arithmetic (through rational number representation) and the exact value functions were obtained corresponding to this strategy profile. The final step involved checking if sequential rationality is satisfied for the obtained value functions, i.e., if all inequalities in (3.26) are satisfied.

In the following we present results for $N = 11$, $p = 0.1$ and three different cases: $\delta = 0$, $\delta = 0.999$, and $\delta = 1$. The first case ($\delta = 0$) is essentially the case of myopic players and the results in Fig. 3.4 confirm the ones in [7]. Regardless of the value of w, players who have not yet revealed their information, wait for $y \leq -2$, buy for $y \geq 2$ and reveal their information for $-1 \leq y \leq 1$. Note that for $y = 1$ a non-revealing player is indifferent between $\gamma = \boldsymbol{I}$ and $\gamma = \mathbf{1}$, and similarly for $y = -1$. We resolve the tie by assuming that the player always reveals. In addition, for $y = 0$ a player who has already revealed is indifferent between any action, and we resolve this ambiguity by assuming that she always reveals.

The second case ($\delta = 0.999$) studies more patient players and the results are depicted in Fig. 3.5. Not surprisingly, players are willing to wait more before committing to a buying decision. In

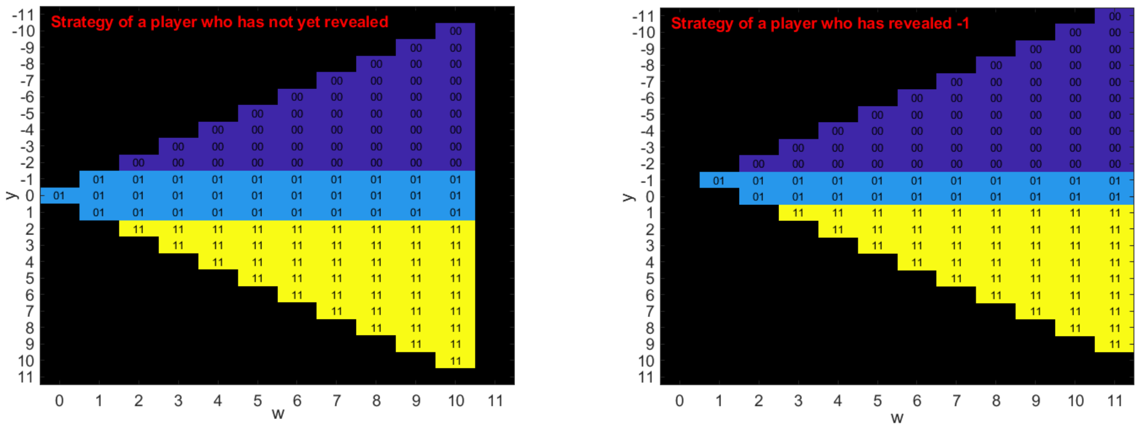


Figure 3.4: Equilibrium strategies for $N = 11$, $p = 0.1$, $\delta = 0$. "00", "01", and "11" denote strategies $\mathbf{0}$, $\boldsymbol{I}$, and $\mathbf{1}$, respectively.

fact, for values of $w = 2$ to $w = 5$ and with a believed product quality of $y = 2$ a player is not committing to buy (i.e., to play $\gamma = \mathbf{1}$) but the equilibrium strategy is to reveal her information ($\gamma = \boldsymbol{I}$). Similarly, with a believed product quality of $y = 2$ a player who has already revealed her private information $X^n = -1$ chooses to wait ($\gamma = \mathbf{0}$).

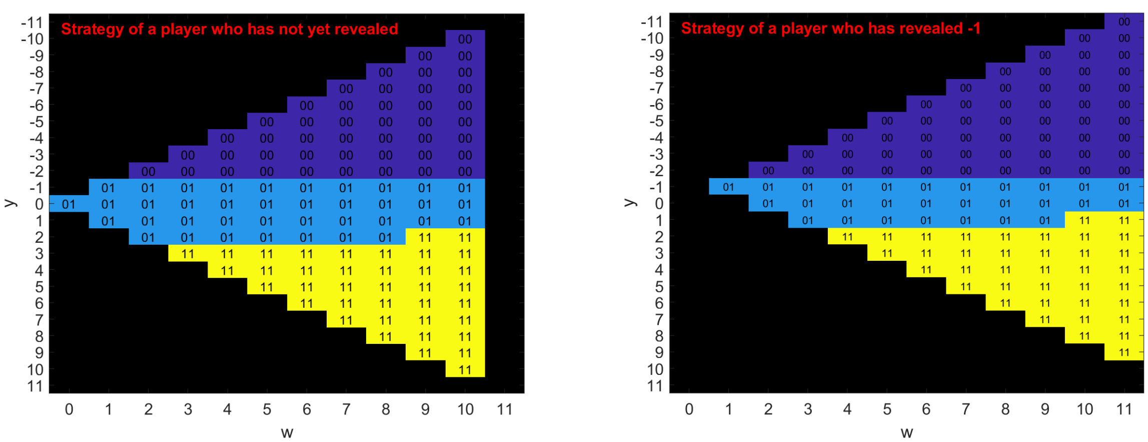


Figure 3.5: Equilibrium strategies for $N = 11$, $p = 0.1$, $\delta = 0.999$. "00", "01", and "11" denote strategies $\mathbf{0}$, $\boldsymbol{I}$, and $\mathbf{1}$, respectively.

The third case ($\delta = 1$) studies infinitely patient players and the results are depicted in Fig. 3.6. As intuition suggests, players are willing to wait more before committing to a buying decision. In fact, for $w = 5$ and with a believed product quality of $y = 5$ a player is not committing to buy (i.e., to play $\gamma = \mathbf{1}$) but the equilibrium strategy is to reveal her information ($\gamma = \boldsymbol{I}$). Similarly, for $w = 6$ and with a believed product quality of $y = 4$ a player who has already revealed her private information $X^n = -1$ chooses to wait ($\gamma = \mathbf{0}$). Clearly, as w increases and we are approaching the end of the game, players become more aggressive, as there is less information to be learnt by waiting, and at $w = N$ the equilibrium strategies for $\delta = 0$ and $\delta = 1$ coincide. Nevertheless, in the case of patient players a more cooperative equilibrium emerges (see strategies indicated in the red triangle in Fig. 3.6) where players are willing to help each other learn the unknown state V by revealing their private information.

We remark that these results are not inconsistent with Theorem 4 since the theorem claims existence of specific solutions to the FPE but not uniqueness. Indeed, although this is the case of $\delta = 1$ our numerical algorithm converges to the equilibrium described in (3.30b) and also depicted in Fig. 3.3.

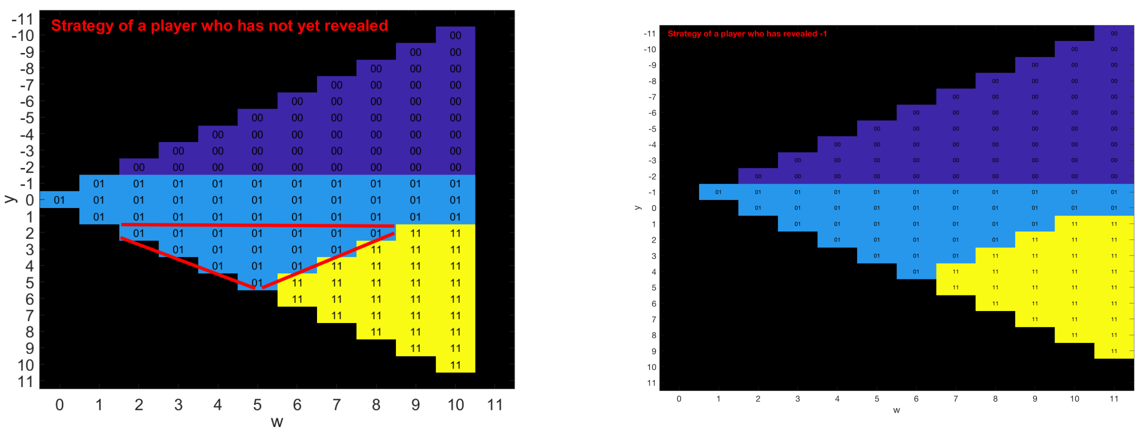


Figure 3.6: Equilibrium strategies for $N = 11$, $p = 0.1$, $\delta = 1$. “00”, “01”, and “11” denote strategies $\mathbf{0}$, $\boldsymbol{I}$, and $\mathbf{1}$, respectively.

The next set of figures shows the effect of the quality of information. In Fig. 3.7 the equilibrium for the case of $\delta = 0.999$ and $p = 0.4$ is depicted. This is a much noisier private observation compared to the one depicted in Fig. 3.5. As a result, equilibrium behavior is “softer”: players are willing to wait more and reveal their information, since a single observation is now of lower quality than before.

The last figure shows the probability of a bad cascade for the two different values of V, for different values of p and for a larger number of users $N = 21$. We further disaggregate this probability according to the value of W when a cascade occurs. We depict this information as cumulative bad cascade probabilities with $W \leq w$ for $w \in \{0, \ldots, N\}$ in Fig. 3.8. The figure

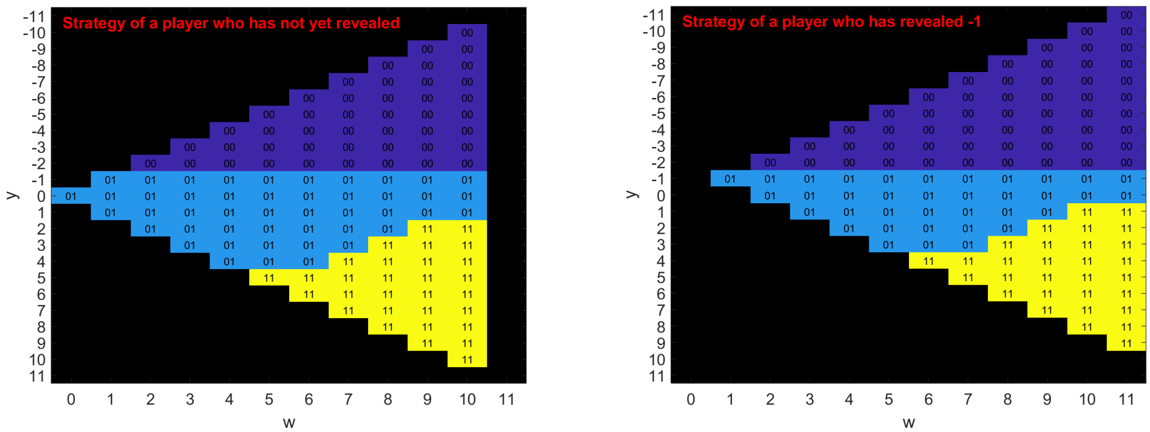


Figure 3.7: Equilibrium strategies for $N = 11$, $p = 0.4$, $\delta = 0.999$. “00”, “01”, and “11” denote strategies $\mathbf{0}$, $\boldsymbol{I}$, and $\mathbf{1}$, respectively.

shows that cascading behaviour is significantly asymmetric for the values $V = 1$ and $V = -1$ and it is more severe for $V = 1$, i.e., when the product is good and players opt to not buy it. This is due to the asymmetry of the sets of values of (y, w) for which the equilibrium is $\gamma = \mathbf{0}$ vs. that for which the equilibrium is $\gamma = \mathbf{1}$.

3.8 Conclusions

We studied a Bayesian learning scenario with non-myopic players. Our model generalizes the classic myopic and sequential one-shot scenario where informational cascades were first reported. In order to analyze information cascades in this scenario, an intricate analysis of the PBE of the dynamic game was needed. To that end, we first constructed FPEs that involve value functions defined on a finite domain. By further exploiting the structure of the model, we constructed FPEs with intuitive interpretations where the value functions has domain that is only quadratic in the number of players N. Building on the tractability of these equations, we investigated their solutions in two regimes. The first is for a fixed $\delta < 1$ and asymptotically large N. The second is for a fixed N and $\delta = 1$ or asymptotically approaching 1. For the first regime, we proved that an informational cascade eventually happens with high probability for large N. In these informational cascades, only a small portion of the information has been revealed, with high probability, making these cascades inefficient. For the second regime we proved that, surprisingly, infinitely patient players can completely avoid bad cascades when the product is bad. Furthermore, for sufficiently patient players when a bad cascade occurs (for a bad product) at least half of the players have already

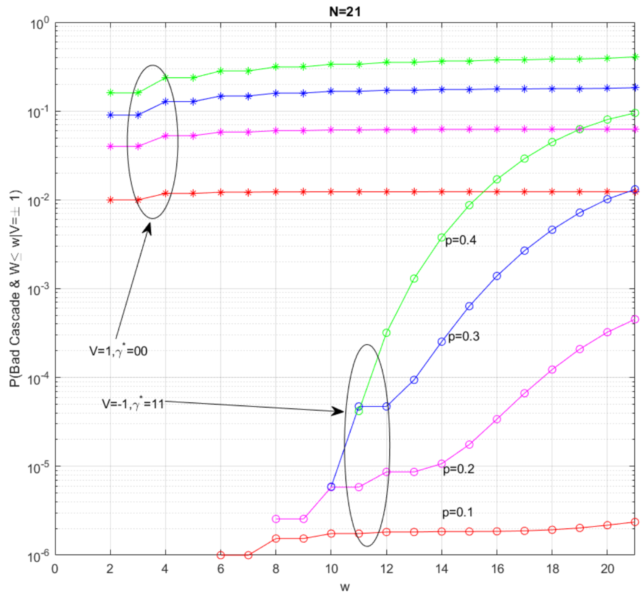


Figure 3.8: Bad cascade probability for $N = 21$, $\delta = 0.999999$, $p \in \{0.1, 0.2, 0.3, 0.4\}$.

revealed their private information, which guarantees an error probability that vanishes with N. Numerical solutions of the developed FPEs show that players exhibit a non-myopic behavior that is much more intricate than in the myopic case we generalized.

We were able to compress the fixed point equation based on the symmetry and structure of the problem. It could be interesting to apply the techniques introduced here to when the observation model $Q\left(x^n|v\right)$ is different between players. Extending FPE 1 is relatively straightforward. Extending FPE 2 is possible if there is a discrete set of available $Q\left(x^n|v\right)$. Then we can add a pair of w, y variables to count players that have the same $Q\left(x^n|v\right)$. As expected, the dimension of the FPE would then increase with the complexity of the scenario.

Part II

Design

CHAPTER 4

Distributed Mechanism Design for Network Resource Allocation Problems

4.1 Introduction

Allocation of scarce resources in networks has been a topic of intensive research in the last fifty years. This problem is often formulated as a network utility maximization (NUM) problem [16, Ch. 2] where the designer is seeking the optimal allocation of resources that maximizes the social welfare. The complexity of this problem, especially for large networks of heterogeneous and strategic agents with privacy constraints stems from the fact that agents may not be willing to share some of their private information related to their utilities. Hence, appropriate incentives (taxes/subsidies) have to be put in place to incentivize agents to reveal some part of their private information relevant to the welfare optimization problem. A useful mathematical framework for this setting is mechanism design (MD) [11, 12] that has been widely utilized in such areas of research as market allocations [13, 14, 15], rate and resource allocations [16, Sec. 2.7][17, 18, 19, 20, 21], spectrum sharing [22, 23, 24], data security [25], power allocation in wireless networks [26, 27], demand-side management in the power grid [28, 29, 30], etc.

In the standard MD framework (Hurwicz-Reiter [11]) agents are transmitting messages to a central authority. The central authority, upon receiving all these messages, determines allocation and taxes/subsidies for the agents of the network. Equivalently, all agents broadcast their messages to everyone else and then the allocation and taxes can be generated (and verified) by everyone in the network. Clearly, this arrangement results in a significant communication overhead due to message transmission of agents to the central authority or to each other, and this problem becomes more pronounced the larger the network. The motivation for this work is the more realistic scenario where such message transmission to a central authority (or equivalently, broadcasting of messages) cannot take place due to network communication constraints. To investigate this problem, we consider a setting in which agents are only allowed to transmit their messages to their neighboring agents. In this setting, neighborhoods are defined through an underlying message-exchange network. Consequently, each agent can determine her allocation and tax based on the messages she hears locally and therefore, there is no need for a central authority to evaluate these functions. This implies that, unlike standard mechanisms, the designed allocation and tax functions cannot have the whole message space as their domain; rather the allocation and tax function for each agent should only depend on the neighborhood messages. This additional restriction gives rise to a novel research direction that we call "Distributed Mechanism Design" (DMD).

A complementary view of DMD stems from the literature of distributed optimization (e.g., [31, 32, 33, 34, 35, 36, 37]). In distributed optimization agents do exchange local messages in order to solve a centralized allocation problem. It is assumed however, that agents are not strategic–in fact they are automata–and follow a predefined message exchange algorithm. DMD can be thought of as the generalization of distributed optimization to account for strategic agents, i.e., for settings where we can no longer assume that agents will follow a distributed message passing algorithm unless the designer puts in place appropriate incentives for them to do so.

In this work, we consider two network resource allocation problems with increasing degree of sophistication, formulated as NUM problems. Although the models presented are abstract, we present them through two concrete applications. In particular, we consider rate allocation for data transmission on a network which operates either with a unicast transmission protocol (UTP) or a multirate multicast transmission protocol (MMTP). For each of these protocols, a distributed mechanism is proposed for efficient rate allocation. The distributed mechanism proposed for MMTP is an extended and modified version of the distributed mechanism for UTP. For this reason, we present these two mechanisms side-by-side and highlight the main techniques used and the differences between them throughout the presentation.

The contributions of this work are as follows. Both of the proposed mechanisms are (a) distributed, (b) they fully implement the optimal allocation in Nash equilibria (NE) (i.e., there are no extraneous non-efficient equilibria in the induced game), and (c) their total message space dimension grows linearly with respect to the number of network agents. Furthermore, the mechanisms are (d) individually rational and (e) weak budget balanced at NE.

4.1.1 Relevant Literature

A non-distributed mechanism for efficient rate allocation with UTP has been proposed in [18, 20, 111] and with MMTP in [19, 21]. The models we consider in this work, closely follow the ones in [18, 20, 111, 19, 21] but the mechanisms differ in a fundamental way since our focus is designing distributed mechanisms. The current work builds on distributed mechanisms for Walrasian and Lindahl allocation in private and public goods, respectively, that were proposed in [112][113, Ch. 4]. We have utilized an idea similar to the radial allocation [15, 17, 21] to achieve feasibility at

NE. Unlike the mechanism in [112] with message dimensionality per user that grows linearly with the number of users in the whole network, in this work, the message dimensionality of each agent grows linearly only with respect to the size of her neighborhood.

There is a line of research in the computer science literature by the name distributed algorithmic mechanism design (DAMD) [114, 115, 116][117, Ch. 14]. We caution the reader not to confuse DAMD with DMD. The mechanisms in DAMD impose no restrictions on the domain of the allocation and tax functions. Indeed these functions can depend on the entire message from all users. The "distributed" aspect of DAMD pertains to the fact that an algorithm is designated to collect and disseminate all these messages to the users of the network so that they can all evaluate these complex functions. In DMD however, the allocation and tax functions are explicitly designed so that they only depend on messages from neighboring agents. Another difference is that in DAMD the message exchange network coincides with the network implied by the allocation problem (e.g., in [114] messages are exchanged between neighboring agents on the multicast tree) while in DMD, as will be evident in Section 4.2, the message exchange network can be arbitrary.

A related line of work in distributed optimization attempts to resolve "privacy" issues by means of dithering (i.e., adding noise to) the exchanged messages or the objective function as in [118, 119, 120, 121]. In these works, agents are given some privacy guarantees in that the distributed algorithm does not fully reveal their private information. However, the agents are still considered non-strategic automata, i.e., it is assumed that they follow the prescription of the algorithm even if there is an incentive to deviate.

We conclude the discussion on the relevant literature by pointing out that the games induced by DMD fall under the class of "graphical games" [122, 123]. This property may have some consequences on the complexity of off-line evaluation of the NE, which however, is not of importance in our work.

The rest of this chapter is structured as follows. In section 4.2, the model and problem formulation for both of the network transmission protocols are discussed. Section 4.3 presents both of the distributed mechanisms by characterizing the message-exchange network and defining messages, allocation and tax functions. A specific example is discussed in detail in section 4.4 in order to clarify the general definitions of the distributed algorithms. In section 4.5, the properties of the designed mechanisms are derived and the main results are presented. In section 4.6, for each of the two mechanisms, an alternative mechanism is presented by relaxing an assumption on the message-exchange network. We conclude and comment on the message dimensions in section 4.7. All of the proofs of intermediate lemmas can be found in Appendix C.

4.2 Model

The two abstract centralized optimization problems for which distributed mechanisms will be developed in the following sections are precisely defined in (4.1) and (4.3), respectively. In order to make the discussion more relevant to resource allocation in data transmission networks, we now present two concrete applications that will serve as prototypes in the subsequent discussion. In particular, we consider a data transmission network with two different transmission protocols, Unicast Transmission Protocol (UTP) and Multirate Multicast Transmission Protocol (MMTP) and for each one of them an optimization problem for efficient data rate allocation is presented. We are following closely the models developed in [16, Sec. 2.7][18, 19, 20, 21]. The network consists of multiple sources in the set $\mathcal{K} = \{1, ..., K\}$ and strategic receivers in the set $\mathcal{N} = \{1, ..., N\}$, which will be referred to as agents. Each agent has one designated source from which it receives content, and each source can send content to multiple agents possibly with different data rates. The vector of allocated rates to all agents is denoted by $x = (x_1, ..., x_N) \in \mathbb{R}_+^N$, where x_i is the data rate of agent i. Based on its allocated rate, x_i, each agent receives a private utility (satisfaction) $v_i(x_i)$. The following assumptions are imposed on the utility functions. For every $i \in \mathcal{N}$, $v_i(\cdot) \in \mathcal{V}_0$, where $\mathcal{V}_0$ is the set of strictly concave, monotonically increasing, twice differentiable $\mathbb{R}_+ \to \mathbb{R}$ functions with continuous second derivatives. The valuation function $v_i(.)$ is the private information of agent i. The network links are denoted by $\mathcal{L} = \{1, ..., L\}$, each of which has capacity $c^l > 0$. Agent i's data stream is transmitted via links $\mathcal{L}_i \subset \mathcal{L}$ with $|\mathcal{L}_i| = L_i$. The routing has been established in advance and it is considered fixed for the problem of interest in this work. For each link l, agents using it are denoted by $\mathcal{N}^l$ with $|\mathcal{N}^l| = N^l$. The designer's goal is to maximize the social welfare which is the summation of agents' valuation functions. This is done by determining the efficient x that is consistent with the capacity constraints of the network which arise from the specific transmission protocol utilized. In the following, we provide more concrete details for UTP and MMTP models.

4.2.1 Unicast Transmission Protocol (UTP)

In UTP, a separate data stream is established between each source-receiver pair, regardless of whether the same data content is transmitted multiple times over some links. An example of a network utilizing UTP is depicted in Fig. 4.1. We assume $N^l \geq 2$, that is at least two agents use any link l. This mild assumption is made so that there is competition between agents for using any of the links and it will help us avoid corner cases that distract from the main message of the paper.

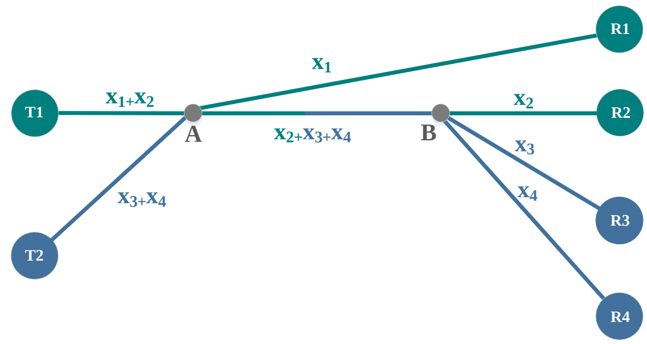


Figure 4.1: Network with Unicast Transmission Protocol (UTP). Both R1 and R2 use link T1-A, which, due to using UTP, is loaded with the sum data rate $x_1 + x_2$.

The centralized optimization problem that formulates the design goal for UTP is as follows

$$\max_{x} \sum_{i \in \mathcal{N}} v_i(x_i) \tag{4.1a}$$

$$\text{s.t.} \quad x_i \geq 0 \quad \forall i \in \mathcal{N} \tag{4.1b}$$

$$\text{and} \quad \sum_{j \in \mathcal{N}^l} x_j \leq c^l \quad \forall l \in \mathcal{L}. \tag{4.1c}$$

In order to characterize the solution of problem (4.1), we use dual variables $\lambda = \{\lambda^l, l \in \mathcal{L}\}$ and write the KKT conditions for this problem. Since the valuation functions are concave and all of the constraints of problem (4.1) are affine, problem (4.1) is a convex optimization problem and so KKT conditions are necessary and sufficient. These conditions at the optimal point (x^*, λ^*) are

(a) Primal Feasibility: x^* satisfies (4.1b) and (4.1c).

(b) Dual Feasibility: $\lambda^{l*} \geq 0 \quad \forall l \in \mathcal{L}$.

(c) Complimentary Slackness:

$$\lambda^{l*}(c^l - \sum_{j \in \mathcal{N}^l} x_j^*) = 0 \quad \forall l \in \mathcal{L}. \tag{4.2a}$$

(d) Stationarity:

$$v_i'(x_i^*) = \sum_{l \in \mathcal{L}_i} \lambda^{l*} \quad \text{if} \quad x_i^* > 0 \tag{4.2b}$$

$$v_i'(x_i^*) \leq \sum_{l \in \mathcal{L}_i} \lambda^{l*} \quad \text{if} \quad x_i^* = 0. \tag{4.2c}$$

4.2.2 Multirate Multicast Transmission Protocol (MMTP)

In MMTP model, agents are classified into K groups based on their data source and each group is denoted by $\mathcal{G}_k \subset \mathcal{N}$. Since agents in each group receive data from the same source (albeit with possibly different data rates), instead of transmitting a separate data stream to each agent, in each link only a single stream is transmitted for each group utilizing that link. Furthermore, the data rate of that stream is the maximum demanded rate among users in the group on that link. In other words, each source transmits the common data of agents by the best quality demanded in each link and each agent can regenerate her own data by down-sampling from the received data stream to get her desirable quality. This scenario is as if agents inside a group share the bandwidth with each other (public good) but they have competition for it with other groups (private good), a situation referred to as intergroup competition and intragroup sharing in [21]. An example of a network utilizing MMTP is illustrated in Fig. 4.2. We further define the following quantities. The group of agent i is

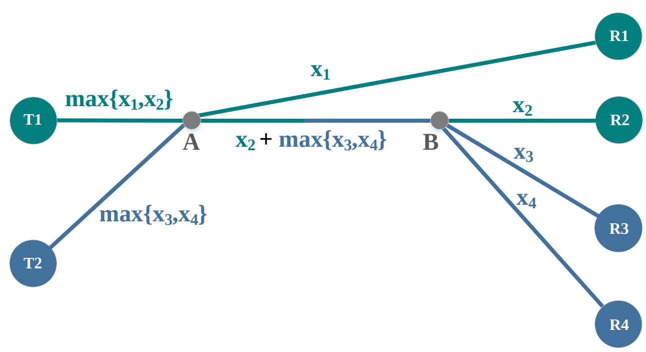


Figure 4.2: Network with Multirate Multicast Transmission Protocol (MMTP). Even though both R1 and R2 use link T1-A, due to using MMTP, it is only loaded with the maximum rate $\max\{x_1, x_2\}$.

denoted by $k(i)$. The set of users in group k using link l is denoted by $\mathcal{G}_k^l = \mathcal{N}^l \cap \mathcal{G}_k$. Further, $\mathcal{K}^l$ is the subset of groups that are using link l (groups that contain at least one agent using link l) with $|\mathcal{K}^l| = K^l$. For the same reason as in UTP, we assume $K^l \geq 2$, that is, at least two groups use each link l and this is for ensuring competition at each link.

The centralized optimization problem that formulates the design goal for MMTP is as follows

$$\max_x \sum_{i \in \mathcal{N}} v_i(x_i) \tag{4.3a}$$

$$\text{s.t.} \quad x_i \geq 0 \quad \forall i \in \mathcal{N} \tag{4.3b}$$

$$\text{and} \quad \sum_{k \in \mathcal{K}^l} \max_{i \in \mathcal{G}_k^l} \{x_i\} \leq c^l \quad \forall l \in \mathcal{L}. \tag{4.3c}$$

As in the case of UTP, we utilize KKT conditions to characterize the solution of problem (4.3). We first need to transform it into a convex optimization problem. Towards this goal, we introduce the variable b_k^l for each $l \in \mathcal{L}$ and $k \in \mathcal{K}^l$ that represents the maximum demanded rate of agents in group k that use link l, which we refer to as the "group rate". It is straightforward to show the equivalence of problem (4.3) with the one below

$$\max_x \sum_{i \in \mathcal{N}} v_i(x_i) \tag{4.4a}$$

$$\text{s.t.} \quad x_i \geq 0 \quad \forall i \in \mathcal{N} \tag{4.4b}$$

$$\text{and} \quad \sum_{k \in \mathcal{K}^l} b_k^l \leq c^l \quad \forall l \in \mathcal{L} \tag{4.4c}$$

$$\text{and} \quad x_i \leq b_k^l \quad \forall l \in \mathcal{L},\ k \in \mathcal{K}^l,\ i \in \mathcal{G}_k^l. \tag{4.4d}$$

Similar to problem (4.1), problem (4.4) is a convex optimization problem and hence, KKT conditions are necessary and sufficient for its solution. We use dual variables $\lambda = \{\lambda^l, l \in \mathcal{L}\}$, each of which corresponds to one of the constraints in (4.4c) and $\mu = \{\mu_i^l, \forall l \in \mathcal{L}, i \in \mathcal{N}^l\}$, each of which corresponds to one of the constraints in (4.4d). We can write the KKT conditions at the optimal point $(x^*, b^*, \lambda^*, \mu^*)$ as

(a) Primal Feasibility: x^* and b^* satisfy (4.4b) and (4.4c) and (4.4d).

(b) Dual Feasibility: $\lambda^{l*} \geq 0\ \forall l \in \mathcal{L}$ and $\mu_i^{l*} \geq 0\ \forall l \in \mathcal{L}, i \in \mathcal{N}^l$.

(c) Complimentary Slackness:

$$\lambda^{l*}(c^l - \sum_{k \in \mathcal{K}^l} b_k^{l*}) = 0 \quad \forall l \in \mathcal{L} \tag{4.5a}$$

$$\mu_i^{l*}(x_i^* - b_k^{l*}) = 0 \quad \forall l \in \mathcal{L}, k \in \mathcal{K}^l, i \in \mathcal{G}_k^l. \tag{4.5b}$$

(d) Stationarity:

$$v_i'(x_i^*) = \sum_{l \in \mathcal{L}_i} \mu_i^{l*} \quad \forall i \in \mathcal{N} \quad \text{if} \quad x_i^* > 0 \tag{4.5c}$$

$$v_i'(x_i^*) \leq \sum_{l \in \mathcal{L}_i} \mu_i^{l*} \quad \forall i \in \mathcal{N} \quad \text{if} \quad x_i^* = 0 \tag{4.5d}$$

$$\lambda^{l*} = \sum_{i \in \mathcal{G}_k^l} \mu_i^{l*} \quad \forall l \in \mathcal{L}, k \in \mathcal{K}^l. \tag{4.5e}$$

Note that from the above KKT conditions it is obvious that MMTP gives rise to the "free-riding" problem that is commonly encountered in public-goods problems [124, Sec. 11.C]: if an agent i on link l is not requesting the highest rate among the agents in the same group $k(i)$, then his contribution μ_i^{l*} is zero in the "price" λ^{l*} for this link and thus she can free-ride on the other agent(s) in the group requesting the highest rate.

The optimization problems (4.1) and (4.4) cannot be solved in a centralized manner because the valuation functions of agents are private information of the agents. In the next section, we present two distributed mechanisms that aim to reach the solution of optimization problems (4.1) and (4.4) in a decentralized fashion in the presence of strategic agents.

4.3 Distributed Mechanism

A mechanism consists of a message space $\mathcal{M}_i$ for each agent $i \in \mathcal{N}$ giving rise to a total message space $\mathcal{M} = \mathcal{M}_1 \times \ldots, \times \mathcal{M}_N$, and allocation and tax functions that are denoted by $\widehat{x}_i : \mathcal{M} \to \mathbb{R}_+^N$ and $\widehat{t}_i : \mathcal{M} \to \mathbb{R}^N$, respectively. Hence, a mechanism can be characterized completely by specifying the tuple $(\mathcal{M}, (\widehat{x}_i)_{i \in \mathcal{N}}, (\widehat{t}_i)_{i \in \mathcal{N}})$. The mechanism induces the game $\mathfrak{G} = (\mathcal{N}, \mathcal{M}, (\widehat{u}_i)_{i \in \mathcal{N}})$, where we consider a quasi-linear environment with $\widehat{u}_i(m) = v_i(\widehat{x}_i(m)) - \widehat{t}_i(m)$ for any $m \in \mathcal{M}$. In the following we will use superscripts $\mathfrak{u}$ and $\mathfrak{m}$ to specify quantities for UTP and MMTP, respectively. Thus, we will use notations $\mathcal{M}^{\mathfrak{u}}$, $\widehat{x}_i^{\mathfrak{u}}$, $\widehat{t}_i^{\mathfrak{u}}$, $m^{\mathfrak{u}}$, $\mathfrak{G}^{\mathfrak{u}}$, $\widehat{u}_i^{\mathfrak{u}}$ for UTP, and similarly, $\mathcal{M}^{\mathfrak{m}}$, $\widehat{x}_i^{\mathfrak{m}}$, $\widehat{t}_i^{\mathfrak{m}}$, $m^{\mathfrak{m}}$, $\mathfrak{G}^{\mathfrak{m}}$, $\widehat{u}_i^{\mathfrak{m}}$ for MMTP. In the following we formally describe "distributed" mechanisms, i.e., mechanisms for which the allocation and tax functions depend only on neighboring agents' messages as opposed to the entire message space $\mathcal{M}$.

4.3.1 Message-Exchange Network

We describe the local exchange of messages through a "message-exchange network", which is modeled as an undirected acyclic (tree) connected graph $\mathcal{GR} = (\mathcal{N}, \mathcal{E})$, in which agents are denoted by nodes and an edge between two agents indicates that these two agents receive each others' messages. For all $i \in \mathcal{N}$, $\mathcal{N}(i)$ is the set of neighbors of agent i in $\mathcal{GR}$ and $|\mathcal{N}(i)| = N(i)$. Further, $n(i,j)$ is agent i's neighbor which is on the shortest path from i to j. Also, $\mathcal{N}^l(i) = \mathcal{N}(i) \cap \mathcal{N}^l$ and $|\mathcal{N}^l(i)| = N^l(i)$. For each agent $i \in \mathcal{N}$, the function $\phi(i)$ arbitrarily chooses one of agent i's neighbors. We define the set $\mathcal{I}_i = \{h \in \mathcal{N}(i) : \phi(h) = i\}$. The role of this function will become evident in the description of the allocation and tax functions.

Notice that the "message-exchange network" is not to be confused with the "data transmission network" related to UTP and MMTP and modeled through the centralized problems in (4.1) or (4.3). The former enables the decentralized solution of those problems by means of exchanging messages between neighboring nodes, while the latter describes the relation between agents dictated by the common links utilized by their data flows. These two networks are illustrated in Fig. 4.3.

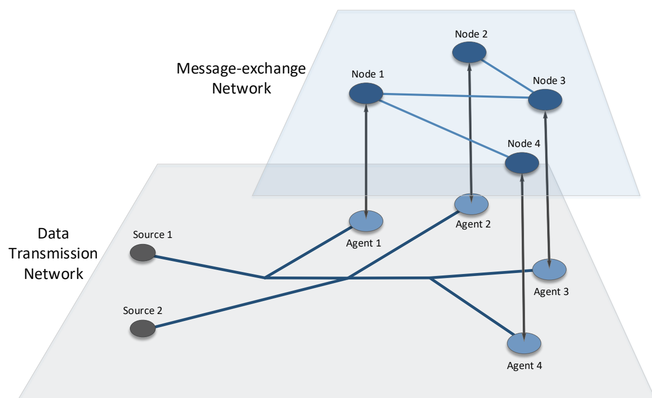


Figure 4.3: Message-exchange network vs. Data transmission network

In the following, we state two assumptions for the message-exchange network where Assumption 1 holds for both UTP and MMTP mechanisms and Assumption 2 only holds for MMTP mechanism. These assumptions simplify the exposition of the mechanisms. We will further relax Assumption 1 for both mechanisms in section 4.6 and two alternative mechanisms will be proposed.

Assumption 1. *(UTP/MMTP) For each link $l \in \mathcal{L}$, the sub-graph consisting of agents $i \in \mathcal{N}^l$ is a connected graph.*

This assumption essentially requires that a connected path exists for message passing between agents using the same link, eventually enabling them to form a consensus about some of the exchanged messages, e.g., the price for using each link.

Assumption 2. *(MMTP) For each link $l \in \mathcal{L}$ and group $k \in \mathcal{K}^l$, there is at least one node $i \in \mathcal{G}_k^l$ that is connected to all other nodes $j \in \mathcal{G}_k^l$ in a single hop. This node will be referred to as the "group leader" of group k in link l and will be denoted by $c(k, l)$.*

For each agent $i \in \mathcal{N}$, the set $\mathcal{C}_i$ is defined as the set of links l for which $c(k(i), l) = i$, i.e., this set contains all links for which agent i is a group leader.

The reason for this assumption is that in MMTP due to the free-riding problem, we require that there exists a user in each group $\mathcal{G}_k^l$ who has access to some necessary information about her group (e.g., group demand, group price) and can announce this information to the rest of the agents in $\mathcal{G}_k^l$.

4.3.2 Message Components

4.3.2.1 UTP

Agent i announces the message $m_i^{\mathrm{u}} = (y_i, n_i, q_i, p_i)$. The first message $y_i \in \mathbb{R}_+$ is a proxy for her demanded rate. The second message, $n_i = (n_{i,j}^l, j \in \mathcal{N}(i), l \in \mathcal{L}) \in \mathbb{R}_+^{L \times N(i)}$ consists of components $n_{i,j}^l$, which are referred to as "summary" messages, each of which is a proxy for the sum of demands of the agents $h \in \mathcal{N}^l$ with $n(i,h) = j$. In other words, $n_{i,j}^l$ is the sum of demanded rates for all users that are connected to i through her neighbor j. These messages are required by agent i in order to assess the total demand on each link she is using. The third message, $q_i = (q_{i,j}, j \in \mathcal{I}_i) \in \mathbb{R}_+^{|\mathcal{I}_i|}$ is a vector of components $q_{i,j}$, each of which is a proxy for the demand of neighboring agent $j \in \mathcal{I}_i$. The purpose of these messages will become evident in the allocation function (4.7) and can intuitively be explained as follows: in order to allocate rate to agent i in such a way that the capacity constraint at each link is satisfied, her demanded rate needs to be scaled by the sum of rates in that link. However, in evaluating the sum of rates, the rate of agent i should not be controlled by her; instead an arbitrarily chosen neighboring agent j quotes her rate through the message $q_{j,i}$. Clearly, we want the message $q_{j,i}$ to agree with y_i at NE. Finally, the message $p_i = (p_i^l, l \in \mathcal{L}_i) \in \mathbb{R}_+^{L_i}$ is the price (per unit of rate) that agent i suggests for using each link l. This message is essentially a proxy for the dual variable λ^{l*} that appears in the KKT conditions (4.2).

The message components for UTP are summarized in Table 4.1.

Table 4.1: Message components of agent $i \in \mathcal{N}$ in UTP mechanism

Message Component	Definition	Functionality
y_i	-	Demand for the data rate
$n_{i,j}^l$	$j \in \mathcal{N}(i)$ $l \in \mathcal{L}$	Summary for demands of agents on link l, connected to i via j
$q_{i,j}$	$j \in \mathcal{I}_i$	Proxy for y_j
p_i^l	$l \in \mathcal{L}_i$	Suggested price for using link l

4.3.2.2 MMTP

Agent i announces the message $m_i^{\mathrm{m}} = (y_i, n_i, q_i, p_i, s_i, w_i, z_i, a_i)$. The reason for the larger message compared to UTP stems form the fact that (a) all agents within a group $\mathcal{G}_k^l$ need access to the maximum demanded rate in that group (this is required due to the free-riding problem that is inherent in the MMTP scenario) and (b) this information needs to be disseminated to all agents in the network while satisfying the communication constraints. In the following we give intuitive explanations for the meaning of each of the eight message components. The first message $y_i \in \mathbb{R}_+$ is a proxy for the agent i's demanded rate. The second message, $n_i = (n_{i,j}^l, j \in \mathcal{N}(i), l \in \mathcal{L}) \in \mathbb{R}_+^{L \times N(i)}$ consists of components $n_{i,j}^l$ where, similar to UTP, are referred to as "summary" messages, with the only difference being that each of them is a proxy for the sum of group demands (not individual demands) of the agents $h \in \mathcal{N}^l$ with $n(i,h) = j$. The third message, $q_i = (q_{i,j}, j \in \mathcal{I}_i) \in \mathbb{R}_+^{|\mathcal{I}_i|}$ consists of elements $q_{i,j}$, each of which is a proxy for agent j's demand, and its role is similar to UTP. The fourth message consists of two components $p_i = (p_i^1, p_i^2)$. The first component is defined as $p_i^1 = (p_i^{1,l}, l \in \mathcal{L}_i) \in \mathbb{R}_+^{L_i}$, where each variable $p_i^{1,l}$ is the price that agent i is willing to pay for using link l and is similar to the message p_i^l in UTP. This is essentially a proxy for the dual variable μ_i^{l*} that appears in the KKT conditions (4.5). The second component is defined as $p_i^2 = (p_{i,j}^{2,l}, j \in \mathcal{I}_i, l \in \mathcal{L}_j) \in \mathbb{R}_+^{(\sum_{j \in \mathcal{I}_i} L_j)}$, where each variable $p_{i,j}^{2,l}$ is the price that agent i thinks agent j should pay for using link $l \in \mathcal{L}_j$. The reason why user i quotes a price relevant to user j is the same as in the case of the q messages explained above in the UTP scenario. We now define the new messages that are present in MMTP and give intuitive explanations for their role. The fifth message is defined as $s_i = (s_i^l, l \in \mathcal{L}_i) \in \mathbb{R}_+^{L_i}$, where, each of the variables s_i^l is capturing whether the specific agent belongs in the group of agents that demand the maximum rate within the group $\mathcal{G}_{k(i)}^l$. We call these messages as proxies of the "group demand". Specifically, at NE, the message s_i^l becomes zero if agent i is not in the max group in link l, and otherwise, it will be the maximum demanded rate of group $\mathcal{G}_{k(i)}^l$ divided by the number of users in the max group in link l. The sixth message $w_i = (w_i^l, l \in \mathcal{L}_i) \in \mathbb{R}_+^{L_i}$ consists of components w_i^l, each of which is a proxy for the price that group $k(i)$ is willing to pay for link l and will be referred to as group price of group $k(i)$ for link l. These messages have to converge at NE to the dual variables $\lambda_l{}^*$ in the KKT conditions (4.5) for all users $i \in \mathcal{N}^l$. The seventh message is defined as $z_i = (z_i^1, z_i^2)$. The first component $z_i^1 = (z_i^{1,l},\ l \in \mathcal{C}_i) \in \mathbb{R}_+^{|\mathcal{C}_i|}$ consists of elements $z_i^{1,l}$, each of which is a proxy for the maximum value of demands of agents in $\mathcal{G}_{k(i)}^l$, i.e., the total group demand of agents in the group $\mathcal{G}_{k(i)}^l$. Further, $z_i^2 = (z_i^{2,l},\ l \in \mathcal{C}_i) \in \mathbb{R}_+^{|\mathcal{C}_i|}$ consists of elements $z_i^{2,l}$, each of which is a proxy for the number of agents that have maximum demand in $\mathcal{G}_{k(i)}^l$. These messages are quoted by user i for every link for which i is the group leader of $\mathcal{G}_{k(i)}^l$. Finally, the eighth message, $a_i = (a_i^1, a_i^2)$, consists of two components. The first component is defined as $a_i^1 = (a_i^{1,l}, l \in \mathcal{L}_i) \in \mathbb{R}_{++}^{L_i}$. Its role is quite technical and will become evident in the proof of efficiency of the NE of this mechanism. The second component, $a_i^2 = (a_{i,j}^{2,l}, j \in \mathcal{I}_i, l \in \mathcal{L}_j) \in \mathbb{R}_{++}^{(\sum_{j \in \mathcal{I}_i} L_j)}$, consists of the elements $a_{i,j}^{2,l}$, each of which is a proxy for the message $a_j^{1,l}$. The reason for user i quoting such message relevant to user j is the same as in the case of q messages explained earlier in the UTP scenario.

The message components for MMTP are summarized in Table 4.2.

Table 4.2: Message components of agent $i \in \mathcal{N}$ in MMTP mechanism

Message Component	Definition	Functionality
y_i	-	Demand for the data rate
$n_{i,j}^l$	$j \in \mathcal{N}(i)$ $l \in \mathcal{L}$	Summary for group demands of agents on link l and connected to i via j
$q_{i,j}$	$j \in \mathcal{I}_i$	Proxy for y_j
$p_i^{1,l}$	$l \in \mathcal{L}_i$	Suggested price for using link l
$p_{i,j}^{2,l}$	$j \in \mathcal{I}_i$ $l \in \mathcal{L}_j$	Proxy for $p_j^{1,l}$
s_i^l	$l \in \mathcal{L}_i$	Proxy for group demand on link l and group $k(i)$
w_i^l	$l \in \mathcal{L}_i$	Proxy for group price of group $k(i)$ for link l
$z_i^{1,l}$	$l \in \mathcal{C}_i$	Proxy for the total group demand on link l and group $k(i)$
$z_i^{2,l}$	$l \in \mathcal{C}_i$	Proxy for the number of agents with max demand on link l and group $k(i)$
$a_i^{1,l}$	$l \in \mathcal{L}_i$	Technical point in the proof
$a_{i,j}^{2,l}$	$j \in \mathcal{I}_i$ $l \in \mathcal{L}_j$	Proxy for $a_j^{1,l}$

4.3.3 Allocation Functions

Let us first define some auxiliary variables. For the UTP scenario, for each agent $i \in \mathcal{N}$ and for every $l \in \mathcal{L}$, y_i^l is defined as $y_i^l = \mathbf{1}_{\mathcal{L}_i}(l) y_i$, where $\mathbf{1}_{\mathcal{A}}$ is the indicator function of the set $\mathcal{A}$. Similarly, for the MMTP scenario, we define y_i^l as $y_i^l = \mathbf{1}_{\mathcal{L}_i}(l) s_i^l$. Further, for each agent $i \in \mathcal{N}$ and $l \in \mathcal{L}_i$, the auxiliary quantities $\overline{z}_i^{1,l}$ and $\overline{z}_i^{2,l}$ are defined as

$$\overline{z}_i^{1,l} = \begin{cases} \max\{q_{\phi(i),i}, \max_{j \in \mathcal{G}^l_{k(i)}, j \neq i}\{y_j\}\} & \text{if } l \in \mathcal{C}_i \\ z^{1,l}_{c(k(i),l)} & \text{if } l \notin \mathcal{C}_i \end{cases} \tag{4.6a}$$

$$\overline{z}_i^{2,l} = \begin{cases} \mathbf{1}_{\{q_{\phi(i),i}\}}(\overline{z}_i^{1,l}) + \sum_{j \in \mathcal{G}^l_{k(i)}, j \neq i} \mathbf{1}_{\{y_j\}}(\overline{z}_i^{1,l}) & \text{if } l \in \mathcal{C}_i \\ z^{2,l}_{c(k(i),l)} & \text{if } l \notin \mathcal{C}_i. \end{cases} \tag{4.6b}$$

The meaning of these quantities is as follows. The quantity $\overline{z}_i^{1,l}$ encodes the maximum demanded rate in the group $\mathcal{G}^l_{k(i)}$. If user i is not the leader of that group, then the leader $c(k(i), l)$ quotes this message through $z^{1,l}_{c(k(i),l)}$. On the other hand, if i is the leader of the group then she has to compute the maximum demand from all other members of the group including her own demand which is now quoted by a proxy through $q_{\phi(i),i}$. Similarly, the quantity $\overline{z}_i^{2,l}$ encodes the number of agents with maximum demand among all of the agents in the group $\mathcal{G}^l_{k(i)}$.

We utilize an idea similar to the radial allocation [21] to have feasible allocation at NE. With this goal in mind, for message vectors $m_i^{\mathtt{u}} = (y_i, n_i, q_i, p_i)$ and $m_i^{\mathtt{m}} = (y_i, n_i, q_i, p_i, s_i, w_i, z_i, a_i)$, the allocation functions for the two mechanisms are defined as appropriately scaled versions of the demanded rates as follows

$$\widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}}) = r_i^{\mathtt{u}}\, y_i \tag{4.7a}$$

$$\widehat{x}_i^{\mathtt{m}}(m^{\mathtt{m}}) = r_i^{\mathtt{m}}\, y_i, \tag{4.7b}$$

where $r_i^{\mathtt{u}}$ and $r_i^{\mathtt{m}}$ are agent i's radial allocation factor in the two protocols and they are defined as

$$r_i^{\mathtt{u}} = \min_{l \in \mathcal{L}} \frac{c^l}{f_i^{\mathtt{u},l}} \tag{4.7c}$$

$$r_i^{\mathtt{m}} = \min_{l \in \mathcal{L}} \frac{c^l}{f_i^{\mathtt{m},l}}, \tag{4.7d}$$

where for $l \in \mathcal{L}_i$, $f_i^{\mathtt{u},l}$ and $f_i^{\mathtt{m},l}$ are defined as

$$f_i^{\mathtt{u},l} = q_{\phi(i),i} + \sum_{j \in \mathcal{N}(i)} (y_j^l + \sum_{h \in \mathcal{N}(j), h \neq i} n^l_{j,h}) \tag{4.7e}$$

$$f_i^{\mathtt{m},l} = \frac{q_{\phi(i),i} \mathbf{1}_{\{q_{\phi(i),i}\}}(\overline{z}_i^{1,l})}{\overline{z}_i^{2,l}} + \sum_{j \in \mathcal{N}(i)} (y_j^l + \sum_{h \in \mathcal{N}(j), h \neq i} n^l_{j,h}), \tag{4.7f}$$

and for $l \notin \mathcal{L}_i$,

$$f_i^{\mathrm{u},l} = f_i^{\mathrm{m},l} = \sum_{j \in \mathcal{N}(i)} (y_j^l + \sum_{h \in \mathcal{N}(j), h \neq i} n_{j,h}^l). \tag{4.7g}$$

The role of the messages n and q should now be clear from the above description. The quantity $f_i^{\mathrm{u},l}$ is the total demanded rate on link l by all agents (from agent i's viewpoint). However, since agent i does not have access to other agents' demands outside her neighborhood, utilizes the summary messages n_j, $j \in \mathcal{N}(i)$ for this purpose. In addition, her own demand is quoted by a pre-specified neighboring agent $\phi(i)$ through $q_{\phi(i),i}$. This is done so that the quantities $f_i^{\mathrm{u},l}$ and $f_i^{\mathrm{m},l}$ do not depend on agent i's messages and the only part that agent i can control in her allocation is y_i in (4.7a). This choice will greatly simplify our proofs of efficiency of the mechanisms.

The additional complication in (4.7f) is due to the fact that in MMTP, if there are more than one agents who quote the maximum rate in a group, they should only load the corresponding link by that maximum rate and not the sum of the maximum rates, thus the normalization by $\overline{z}_i^{2,l}$. This is exactly the reason for the introduction of the z messages in MMTP mechanism.

4.3.4 Tax Functions

4.3.4.1 UTP

The tax function is $\widehat{t}_i^{\mathrm{u}}(m^{\mathrm{u}}) = \sum_{l \in \mathcal{L}} \widehat{t}_i^{\mathrm{u},l}(m^{\mathrm{u}})$ and for each component $\widehat{t}_i^{\mathrm{u},l}(m^{\mathrm{u}})$ we have different cases.

For $l \in \mathcal{L}_i$ we have

$$\begin{aligned} \widehat{t}_i^{\mathrm{u},l}(m^{\mathrm{u}}) =& \overline{p}_{-i}^l \widehat{x}_i^{\mathrm{u}}(m^{\mathrm{u}}) + \sum_{j \in \mathcal{N}(i)} (n_{i,j}^l - y_j^l - \sum_{h \in \mathcal{N}(j), h \neq i} n_{j,h}^l)^2 \\ &+ \sum_{j \in \mathcal{I}_i} (q_{i,j} - y_j)^2 + (p_i^l - \overline{p}_{-i}^l)^2 + (p_i^l - \overline{p}_{-i}^l)\overline{p}_{-i}^l (c^l - r_i^{\mathrm{u}} f_i^{\mathrm{u},l})^2, \end{aligned} \tag{4.8a}$$

where $\overline{p}_{-i}^l$ is the average price for link l quoted by the neighbors of i and is defined as

$$\overline{p}_{-i}^l = \frac{1}{N^l(i)} \sum_{j \in \mathcal{N}^l(i)} p_j^l. \tag{4.8b}$$

For $l \notin \mathcal{L}_i$ we have

$$\widehat{t}_i^{\mathrm{u},l}(m^{\mathrm{u}}) = \sum_{j \in \mathcal{N}(i)} (n_{i,j}^l - y_j^l - \sum_{h \in \mathcal{N}(j), h \neq i} n_{j,h}^l)^2. \tag{4.8c}$$

Intuitively, the tax functions provide some penalties to incentivize agents for quoting messages in a desirable way. With this goal in mind, taxes contain three types of terms. The first type is a rate times price component (first term in (4.8a)). Since agent i controls her allocation through y_i we do not allow her to control her price as well and so the price that she pays is dictated by her neighbors through $\overline{p}^l_{-i}$. The second type consists of quadratic terms that at NE will become zero and thus can be thought of as incentivizing agents to come to a consensus (second to fourth terms in (4.8a)). This enables the mechanism to provide proxies for the missing information of agents at NE, in addition to the requirements of having efficient allocation at NE. The third type relates to the complimentary slackness conditions in (4.2) (fifth term in (4.8a)). The reason of defining two different tax functions is that different incentives are required for agents that utilize a link versus ones that do not. For instance, each agent i has to pay a tax even for links $l \notin \mathcal{L}_i$, which is required for consensus about the "summary messages". The intuition about each tax term will become more evident from the results in the lemmas of section 4.5.

4.3.4.2 MMTP

The tax function is defined as $\widehat{t}^{\mathfrak{m}}_i(m^{\mathfrak{m}}) = \widehat{t}^{\mathfrak{m},\mathfrak{c}}_i(m^{\mathfrak{m}}) + \sum_{l\in\mathcal{L}} \widehat{t}^{\mathfrak{m},l}_i(m^{\mathfrak{m}})$, where the first term is defined as

$$\widehat{t}^{\mathfrak{m},\mathfrak{c}}_i(m^{\mathfrak{m}}) = \sum_{j\in\mathcal{I}_i}\sum_{l\in\mathcal{L}_j}((p^{2,l}_{i,j} - p^{1,l}_j)^2 + (a^{2,l}_{i,j} - a^{1,l}_j)^2) + \sum_{j\in\mathcal{I}_i}(q_{i,j} - y_j)^2, \tag{4.9a}$$

and for each component $\widehat{t}^{\mathfrak{m},l}_i(m^{\mathfrak{m}})$, we consider different cases. For $l \in \mathcal{L}_i, l \notin \mathcal{C}_i$, we have

$$\begin{aligned}\widehat{t}^{\mathfrak{m},l}_i(m^{\mathfrak{m}}) =& p^{2,l}_{\phi(i),i}\widehat{x}^{\mathfrak{m}}_i(m^{\mathfrak{m}}) + \sum_{j\in\mathcal{N}(i)}(n^l_{i,j} - y^l_j - \sum_{h\in\mathcal{N}(j),h\neq i} n^l_{j,h})^2 \\ &+ (s^l_i - \frac{q_{\phi(i),i}\mathbf{1}_{\{q_{\phi(i),i}\}}(\overline{z}^{1,l}_i)}{\overline{z}^{2,l}_i})^2 + \overline{w}^l_{-i}(\widehat{w}^l_i - \overline{w}^l_{-i})(c^l - r^{\mathfrak{m}}_i f^{\mathfrak{m},l}_i)^2 \\ &+ (\widehat{w}^l_i - \overline{w}^l_{-i})^2 + p^{2,l}_{\phi(i),i}(p^{1,l}_i - p^{2,l}_{\phi(i),i})(\overline{z}^{1,l}_i - q_{\phi(i),i})^2 + (w^l_i - w^l_{c(k(i),l)})^2.\end{aligned} \tag{4.9b}$$

For $l \in \mathcal{L}_i, l \in \mathcal{C}_i$, we have

$$\begin{aligned}\widehat{t}^{\mathfrak{m},l}_i(m^{\mathfrak{m}}) =& p^{2,l}_{\phi(i),i}\widehat{x}^{\mathfrak{m}}_i(m^{\mathfrak{m}}) + \sum_{j\in\mathcal{N}(i)}(n^l_{i,j} - y^l_j - \sum_{h\in\mathcal{N}(j),h\neq i} n^l_{j,h})^2 \\ &+ (s^l_i - \frac{q_{\phi(i),i}\mathbf{1}_{\{q_{\phi(i),i}\}}(\overline{z}^{1,l}_i)}{\overline{z}^{2,l}_i})^2 + (z^{1,l}_i - \overline{z}^{1,l}_i)^2 + (z^{2,l}_i - \overline{z}^{2,l}_i)^2\end{aligned}$$

$$+ \overline{w}^l_{-i}(\widehat{w}^l_i - \overline{w}^l_{-i})(c^l - r^{\mathtt{m}}_i f^{\mathtt{m},l}_i)^2 + (\widehat{w}^l_i - \overline{w}^l_{-i})^2$$
$$+ p^{2,l}_{\phi(i),i}(p^{1,l}_i - p^{2,l}_{\phi(i),i})(\overline{z}^{1,l}_i - q_{\phi(i),i})^2 + (w^l_i - p^{2,l}_{\phi(i),i} - \sum_{j\in\mathcal{G}^l_{k(i)},j\neq i} p^{1,l}_j)^2, \qquad (4.9c)$$

where for each link l and agent $i \in \mathcal{N}^l$, $\widehat{w}^l_i$ is defined as

$$\widehat{w}^l_i = \begin{cases} \sum_{j\in\mathcal{G}^l_{k(i)}} p^{1,l}_j + (a^{1,l}_i - a^{2,l}_{\phi(i),i}) & \text{if } l \in \mathcal{C}_i \\ w^l_{c(k(i),l)} - p^{2,l}_{\phi(i),i} + p^{1,l}_i + (a^{1,l}_i - a^{2,l}_{\phi(i),i}) & \text{if } l \notin \mathcal{C}_i, \end{cases} \qquad (4.9d)$$

and, $\overline{w}^l_{-i}$ is defined as

$$\overline{w}^l_{-i} = \frac{1}{N^l(i)} \sum_{j\in\mathcal{N}^l(i)} w^l_j. \qquad (4.9e)$$

Finally for $l \notin \mathcal{L}_i$, the tax term is defined as

$$\widehat{t}^{\mathtt{m},l}_i(m^{\mathtt{m}}) = \sum_{j\in\mathcal{N}(i)} (n^l_{i,j} - y^l_j - \sum_{h\in\mathcal{N}(j),h\neq i} n^l_{j,h})^2. \qquad (4.9f)$$

The intuitive explanation of the terms appearing in the tax function is very similar to the one given above for the UTP scenario. The additional complication stems from the fact that we need to keep track of two types of prices, p and w, corresponding to dual variables μ and λ, respectively.

4.4 A concrete example with UTP

In this section, we provide a simple but not trivial example of UTP model and the corresponding mechanism for that. Assume that we have three agents $\mathcal{N} = \{1, 2, 3\}$ using a single link (link 1) with capacity $c^1 = 1$. The valuation function of agent $i \in \mathcal{N}$ is given by $v_i(x_i) = i\ln(x_i)$. The optimization problem (4.1) will be of the form

$$\max_x \quad \ln(x_1) + 2\ln(x_2) + 3\ln(x_3) \qquad (4.10a)$$
$$\text{s.t.} \quad x_i \geq 0 \quad \forall i \in \mathcal{N} \qquad (4.10b)$$
$$\text{and} \quad x_1 + x_2 + x_3 \leq 1. \qquad (4.10c)$$

By writing the KKT conditions for this problem, one can easily calculate the solution to this optimization problem to be $x^* = (\frac{1}{6}, \frac{1}{3}, \frac{1}{2})$ and the optimal dual variable is $\lambda = 6$. We will show that the mechanism for UTP, has Nash equilibria, all of which result in this efficient allocation x^*.

We consider the message network of Figure 4.4 between the agents. Note that this message network satisfies Assumption 1 and it is a tree graph. We know that $\phi(2) = 1$ and $\phi(3) = 1$ (they

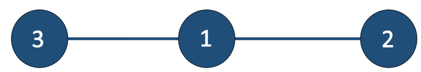


Figure 4.4: Message-exchange network

only have one option). Assume that $\phi(1) = 2$. The above means that agent 1 uses agent 2 for quoting a proxy of his demand, while agent 2 uses agent 1 and agent 3 uses agent 1 for the same. The message components of agents are $m_1^{\mathrm{u}} = (y_1, p_1^1, n_{1,2}^1, n_{1,3}^1, q_{1,2}, q_{1,3})$, $m_2^{\mathrm{u}} = (y_2, p_2^1, n_{2,1}^1, q_{2,1})$ and $m_3^{\mathrm{u}} = (y_3, p_3^1, n_{3,1}^1)$. In this simple, single-link setting, superscripts ^{1} are redundant but we maintain them for notational uniformity with the general description. In this network, agent 1 can listen to all messages, while agent 2 cannot listen to m_3, and similarly, agent 3 cannot listen to m_2. The allocation functions are as follows

$$\widehat{x}_1^{\mathrm{u}}(m^{\mathrm{u}}) = r_1^{\mathrm{u}} y_1 \tag{4.11a}$$
$$\widehat{x}_2^{\mathrm{u}}(m^{\mathrm{u}}) = r_2^{\mathrm{u}} y_2 \tag{4.11b}$$
$$\widehat{x}_3^{\mathrm{u}}(m^{\mathrm{u}}) = r_3^{\mathrm{u}} y_3, \tag{4.11c}$$

where

$$r_1^{\mathrm{u}} = \frac{1}{q_{2,1} + y_2 + y_3} \tag{4.12a}$$
$$r_2^{\mathrm{u}} = \frac{1}{y_1 + q_{1,2} + n_{1,3}^1} \tag{4.12b}$$
$$r_3^{\mathrm{u}} = \frac{1}{y_1 + n_{1,2}^1 + q_{1,3}}. \tag{4.12c}$$

Observe the roles of the q and the n message components. All agents would have to scale their messages by the total demand $y_1 + y_2 + y_3$. We do not want agent 1 to control her scaling factor and thus we ask agent 2 to quote a proxy $q_{2,1}$ for her demand y_1. A similar argument for agents 2 and 3 justifies the presence of the messages $q_{1,2}$ and $q_{1,3}$. In addition, agent 2 does not have access to the demand quoted by agent 3 and that's why she is using the summary message $n_{1,3}^1$ quoted by agent 1 for this purpose. Similarly for agent 3. Finally, note that summary messages $n_{2,1}^1$ and $n_{3,1}^1$ are redundant and are not used in this example. Since agents 2 and 3 are at the leafs of the tree they do not need to pass any information downstream, so these messages are not used in the mechanism.

The tax functions can be written as follows

$$\begin{aligned} t_1(m^{\mathrm{u}}) =& \overline{p}^1_{-1}\widehat{x}^{\mathrm{u}}_1 + (n^1_{1,2} - y_2)^2 + (n^1_{1,3} - y_3)^2 + (q_{1,2} - y_2)^2 + (q_{1,3} - y_3)^2 \\ &+ (p^1_1 - \overline{p}^1_{-1})^2 + (p^1_1 - \overline{p}^1_{-1})\overline{p}^1_{-1}(1 - r^{\mathrm{u}}_1(q_{2,1} + y_2 + y_3))^2 \end{aligned} \tag{4.13a}$$

$$\begin{aligned} t_2(m^{\mathrm{u}}) =& \overline{p}^1_{-2}\widehat{x}^{\mathrm{u}}_2 + (n^1_{2,1} - y_1 - n^1_{1,3})^2 + (q_{2,1} - y_1)^2 \\ &+ (p^1_2 - \overline{p}^1_{-2})^2 + (p^1_2 - \overline{p}^1_{-2})\overline{p}^1_{-2}(1 - r^{\mathrm{u}}_2(y_1 + q_{1,2} + n^1_{1,3}))^2 \end{aligned} \tag{4.13b}$$

$$\begin{aligned} t_3(m^{\mathrm{u}}) =& \overline{p}^1_{-3}\widehat{x}^{\mathrm{u}}_3 + (n^1_{3,1} - y_1 - n^1_{1,2})^2 \\ &+ (p^1_3 - \overline{p}^1_{-3})^2 + (p^1_3 - \overline{p}^1_{-3})\overline{p}^1_{-3}(1 - r^{\mathrm{u}}_3(y_1 + n^1_{1,2} + q_{1,3}))^2, \end{aligned} \tag{4.13c}$$

where

$$\overline{p}^1_{-1} = \frac{p^1_2 + p^1_3}{2} \quad \overline{p}^1_{-2} = p^1_1 \quad \overline{p}^1_{-3} = p^1_1. \tag{4.14}$$

Since the n and q messages only appear once in the tax function of each agent, each agent has the ability to minimize the tax terms by zeroing out the corresponding quadratic terms (four terms for agent 1, two terms for agent 2 and one term for agent 3). So, at NE, we have

$$\begin{aligned} &n^1_{1,2} = y_2 \quad n^1_{1,3} = y_3 \quad n^1_{2,1} = y_1 + n^1_{1,3} \quad n^1_{3,1} = y_1 + n^1_{1,2} \\ &q_{1,2} = y_2 \quad q_{1,3} = y_3 \quad q_{2,1} = y_1. \end{aligned} \tag{4.15}$$

This means that at NE, we have

$$r^{\mathrm{u}}_1 = r^{\mathrm{u}}_2 = r^{\mathrm{u}}_3 = \frac{1}{y_1 + y_2 + y_3}, \tag{4.16}$$

and therefore,

$$\widehat{x}^{\mathrm{u}}_i(m^{\mathrm{u}}) = \frac{y_i}{y_1 + y_2 + y_3}, \qquad i \in \mathcal{N}. \tag{4.17}$$

The above further implies that at NE, $\widehat{x}^{\mathrm{u}}_1(m^{\mathrm{u}}) + \widehat{x}^{\mathrm{u}}_2(m^{\mathrm{u}}) + \widehat{x}^{\mathrm{u}}_3(m^{\mathrm{u}}) = 1$ and the link is fully loaded, and as a result, the last term in all three tax functions will be zero. Consequently, each agent has now the ability to minimize the tax terms by zeroing out the remaining quadratic term that depends on the quoted price. We can conclude that at NE, $p^1_1 = \overline{p}^1_{-1},\ p^1_2 = \overline{p}^1_{-2} = p^1_1,\ p^1_3 = \overline{p}^1_{-3} = p^1_1$, which implies that all price messages are equal (to some yet unspecified price p) at NE, i.e., $p^1_1 = p^1_2 = p^1_3 = p$.

Furthermore, by deriving the best response for the messages y_1, y_2 and y_3 we have the following equation for $i \in \mathcal{N}$,

$$\frac{du_i(m^{\mathfrak{u}})}{dy_i} = 0 \quad \text{if } y_i > 0 \tag{4.18a}$$

$$\frac{du_i(m^{\mathfrak{u}})}{dy_i} < 0 \quad \text{if } y_i = 0, \tag{4.18b}$$

which implies that

$$v_i'(\widehat{x}_i^{\mathfrak{u}}) = p \quad \text{if } y_i > 0 \tag{4.19a}$$

$$v_i'(\widehat{x}_i^{\mathfrak{u}}) < p \quad \text{if } y_i = 0. \tag{4.19b}$$

By solving these equations for $\widehat{x}_i^{\mathfrak{u}}$ and p, we have $p = 6$ (it equals to λ in the optimization problem (4.10)) and $\widehat{x}_i^{\mathfrak{u}} = \frac{i}{6}$ which means that $\widehat{x}_1^{\mathfrak{u}} = \frac{1}{6}$, $\widehat{x}_2^{\mathfrak{u}} = \frac{2}{6}$ and $\widehat{x}_3^{\mathfrak{u}} = \frac{3}{6}$. Hence, the allocation at NE is efficient. The equilibrium y messages can be derived from the following equation

$$\frac{y_i}{y_1 + y_2 + y_3} = \frac{i}{6}, \tag{4.20}$$

which implies that $y_i = k\frac{i}{6}$ for any constant number $k > 0$. This shows that there are infinitely many NE, but all of them have the same and efficient allocation.

4.5 Mechanism Properties

Fact 1. *The mechanisms* $(\mathcal{M}^{\mathfrak{u}}, \widehat{x}^{\mathfrak{u}}, \widehat{t}^{\mathfrak{u}})$ *and* $(\mathcal{M}^{\mathfrak{m}}, \widehat{x}^{\mathfrak{m}}, \widehat{t}^{\mathfrak{m}})$ *are distributed.*

This can be obviously derived from the definition of allocation and tax functions. Clearly, they depend only on each agent's own messages and the messages of her neighboring agents.

Theorem 4. *(Full Implementation, Individual Rationality and Weak Budget Balance - UTP) The game* $\mathfrak{G}^{\mathfrak{u}}$ *has infinitely many Nash equilibria. At every Nash equilibrium* $m^{\mathfrak{u}} \in \mathcal{NE}^{\mathfrak{u}}$ *of the game* $\mathfrak{G}^{\mathfrak{u}}$*, the allocation vector* $\widehat{x}^{\mathfrak{u}}(m^{\mathfrak{u}})$ *is efficient, i.e., it is equal to the solution,* x^**, of problem* (4.1). *In addition, for each agent, individual rationality is satisfied at all NE. Further, the game* $\mathfrak{G}^{\mathfrak{u}}$ *is weak budget balanced at all NE, i.e.,* $\sum_{i \in \mathcal{N}} \widehat{t}_i^{\mathfrak{u}}(m^{\mathfrak{u}}) \geq 0$.

Theorem 5. *(Full Implementation, Individual Rationality and Weak Budget Balance - MMTP) The game* $\mathfrak{G}^{\mathfrak{m}}$ *has infinitely many Nash equilibria. At every Nash equilibrium* $m^{\mathfrak{m}} \in \mathcal{NE}^{\mathfrak{m}}$ *of the game* $\mathfrak{G}^{\mathfrak{m}}$*, the allocation vector* $\widehat{x}^{\mathfrak{m}}(m^{\mathfrak{m}})$ *is efficient, i.e., it is equal to the solution,* x^**, of problem* (4.3). *In addition, for each agent, individual rationality is satisfied at all NE. Further, the game* $\mathfrak{G}^{\mathfrak{m}}$ *is weak budget balanced at all NE, i.e.,* $\sum_{i \in \mathcal{N}} \widehat{t}_i^{\mathfrak{m}}(m^{\mathfrak{m}}) \geq 0$.

Regarding the multiplicity of Nash equilibria in the induced games, we point out that there are two reasons for this behavior. The first reason of not having a unique Nash equilibrium is that the dual variables are not generally unique and therefore, for each dual variable solution (price messages in the mechanism), we can construct a different Nash equilibrium. The second reason is that the demand vector in each of the Nash equilibria is a scaled version of the efficient allocation and every Nash equilibrium corresponds to a scaled version of the allocation. However, the resulting allocation in all of these Nash equilibria is efficient as stated in the theorems. Since for both problems (4.1) and (4.3) the solution is unique, then, according to Theorems 4 and 5, for all $m^{\mathfrak{u}} \in \mathcal{NE}^{\mathfrak{u}}$, the allocation vector $\widehat{x}^{\mathfrak{u}}(m^{\mathfrak{u}})$ is unique, and for all $m^{\mathfrak{m}} \in \mathcal{NE}^{\mathfrak{m}}$, the allocation vector $\widehat{x}^{\mathfrak{m}}(m^{\mathfrak{m}})$ is unique.

Before proving Theorems 4 and 5, some lemmas are presented that are necessary for their proof. The basic idea behind the proof is to show, through a series of lemmas, certain necessary conditions that all NE of the induced games $\mathfrak{G}^{\mathfrak{u}}$ and $\mathfrak{G}^{\mathfrak{m}}$ should satisfy. These necessary conditions essentially lead to showing that the allocations and prices at NE are satisfying the KKT conditions for problems (4.1) and (4.4), respectively. The proof is concluded by showing that indeed there exists such an equilibrium by constructing it.

Lemma 4. *(Concavity) The function $\widehat{u}_i^{\mathfrak{u}}(m_i^{\mathfrak{u}}, m_{-i}^{\mathfrak{u}})$ is strictly concave w.r.t. $m_i^{\mathfrak{u}}$. Similarly, the function $\widehat{u}_i^{\mathfrak{m}}(m_i^{\mathfrak{m}}, m_{-i}^{\mathfrak{m}})$ is strictly concave w.r.t. $m_i^{\mathfrak{m}}$.*

Proof. *See Appendix C.1.*

The strict concavity of $\widehat{u}_i^{\mathfrak{u}}(m_i^{\mathfrak{u}}, m_{-i}^{\mathfrak{u}})$ and $\widehat{u}_i^{\mathfrak{m}}(m_i^{\mathfrak{m}}, m_{-i}^{\mathfrak{m}})$ w.r.t. $m_i^{\mathfrak{u}}$ and $m_i^{\mathfrak{m}}$, respectively, helps us calculate the best response functions for player i in each of the games $\mathfrak{G}^{\mathfrak{u}}$ and $\mathfrak{G}^{\mathfrak{m}}$ by setting the gradient of $\widehat{u}_i^{\mathfrak{u}}(m_i^{\mathfrak{u}}, m_{-i}^{\mathfrak{u}})$ w.r.t. $m_i^{\mathfrak{u}}$ and $\widehat{u}_i^{\mathfrak{m}}(m_i^{\mathfrak{m}}, m_{-i}^{\mathfrak{m}})$ w.r.t. $m_i^{\mathfrak{m}}$ to be equal to zero, respectively. Whenever an element of the gradient cannot be set to zero, it is either always positive or always negative. If any of the elements is always positive, then as message spaces are unbounded from above, there is no best response. Otherwise, if any of the elements of the gradient vector is always negative, the best response would be zero for that element.

The next two lemmas are related with the quadratic terms in the tax functions of UTP and MMTP mechanisms. As mentioned earlier, at NE, agents force these quadratic terms to zero thus achieving consensus. For instance, each message component $q_{i,j}$ can be used as a proxy for y_j by agent j and yet is not controlled by agent j. Furthermore, these lemmas explain how summary messages n are designed to sum up the demands (UTP) or group demands (MMTP) of all agents using link l at NE.

Lemma 5. *At any $m \in \mathcal{NE}^{\mathfrak{u}}$ we have*

$$q_{i,j} = y_j, \quad \forall j \in \mathcal{I}_i, \tag{4.21a}$$

and

$$n^l_{i,j} = y^l_j + \sum_{h \in \mathcal{N}(j), h \neq i} n^l_{j,h}. \tag{4.21b}$$

This implies that at any NE,

$$n^l_{i,j} = \sum_{h \in \mathcal{N}, n(i,h)=j} y^l_h. \tag{4.21c}$$

Proof. *See Appendix C.2.*

Regarding (4.21c), note that since the message graph is a tree, each node is connected to any other node only by one path, and therefore, the demand of each agent is counted once. This ensures no double counting of demands.

Lemma 6. *At any $m^{\mathfrak{m}} \in \mathcal{NE}^{\mathfrak{m}}$, the following equations hold for any $i \in \mathcal{N}$,*

$$q_{i,j} = y_j, \ \forall j \in \mathcal{I}_i \tag{4.22a}$$

$$s^l_i = \frac{q_{\phi(i),i} \boldsymbol{1}_{\{q_{\phi(i),i}\}}(\overline{z}^{1,l}_i)}{\overline{z}^{2,l}_i}, \ \forall l \in \mathcal{L}_i \tag{4.22b}$$

$$p^{2,l}_{i,j} = p^{1,l}_j, \ \forall j \in \mathcal{I}_i, l \in \mathcal{L}_j \tag{4.22c}$$

$$w^l_i = \begin{cases} w^l_{c(k(i),l)} & \textit{if} \quad l \in \mathcal{L}_i, l \notin \mathcal{C}_i \\ p^{2,l}_{\phi(i),i} + \sum_{j \in \mathcal{G}^l_{k(i)}, j \neq i} p^{1,l}_j & \textit{if} \quad l \in \mathcal{L}_i, l \in \mathcal{C}_i \end{cases} \tag{4.22d}$$

$$n^l_{i,j} = y^l_j + \sum_{h \in \mathcal{N}(j), h \neq i} n^l_{j,h}, \ \forall l \in \mathcal{L}, j \in \mathcal{N}(i) \tag{4.22e}$$

$$z^{1,l}_i = \overline{z}^{1,l}_i = \max\{q_{\phi(i),i}, \max_{j \in \mathcal{G}^l_{k(i)}, j \neq i} y_j\}, \ \forall l \in \mathcal{L}_i, l \in \mathcal{C}_i \tag{4.22f}$$

$$z^{2,l}_i = \overline{z}^{2,l}_i = \boldsymbol{1}_{\{q_{\phi(i),i}\}}(\overline{z}^{1,l}_i) + \sum_{j \in \mathcal{G}^l_{k(i)}, j \neq i} \boldsymbol{1}_{\{y_j\}}(\overline{z}^{1,l}_i), \ \forall l \in \mathcal{L}_i, l \in \mathcal{C}_i \tag{4.22g}$$

$$a^{2,l}_{i,j} = a^{1,l}_j, \ \forall j \in \mathcal{I}_i, l \in \mathcal{L}_j. \tag{4.22h}$$

Proof. *See Appendix C.3.*

In the next lemma it is shown how radial allocation (whereby the actual allocation is a scaled version of the requested allocation by all agents) ensures feasibility at NE.

Lemma 7. *(Primal Feasibility) At any $m^{\mathfrak{u}} \in \mathcal{NE}^{\mathfrak{u}}$, of the game $\mathfrak{G}^{\mathfrak{u}}$, the allocation vector $\widehat{x}^{\mathfrak{u}}(m^{\mathfrak{u}})$ is feasible for problem* (4.1). *Similarly, at any $m^{\mathfrak{m}} \in \mathcal{NE}^{\mathfrak{m}}$, of the game $\mathfrak{G}^{\mathfrak{m}}$, the allocation vector $\widehat{x}^{\mathfrak{m}}(m^{\mathfrak{m}})$ is feasible for problem* (4.3).

Proof. *See Appendix C.4.*

The next two lemmas show how different agents form a consensus on the price variables for each link l. For example it is shown that all quoted prices p_i^l end up being equal to a price p^l at NE for the UTP scenario. This price will play the role of the dual variable λ^l in the KKT conditions (4.2). Similarly, for the MMTP scenario, it is shown how different groups form a consensus on the group prices $\widehat{w}_i^l$ for each link l which becomes equal to w^l at NE. Furthermore, in both lemmas, equilibrium expressions are derived that resemble the complimentary slackness terms of the KKT conditions in (4.2) and (4.5).

Lemma 8. *At any $m^{\mathrm{u}} \in \mathcal{NE}^{\mathrm{u}}$, of the game $\mathfrak{G}^{\mathrm{u}}$,*

$$p_i^l = p^l, \forall i \in \mathcal{N}, l \in \mathcal{L}_i. \tag{4.23a}$$

Also,

$$p^l(c^l - \sum_{i \in \mathcal{N}^l} \widehat{x}_i^{\mathrm{u}}) = 0 \quad \forall l \in \mathcal{L}. \tag{4.23b}$$

Proof. *See Appendix C.5.*

Lemma 9. *At any $m^{\mathrm{m}} \in \mathcal{NE}^{\mathrm{m}}$, of the game $\mathfrak{G}^{\mathrm{m}}$, the following constraints hold for all $i \in \mathcal{N}$ and $l \in \mathcal{L}_i$,*

$$\widehat{w}_i^l = w^l \tag{4.24a}$$

$$w^l(c^l - r_i^{\mathrm{m}} \sum_{i \in \mathcal{N}} y_i^l) = 0 \tag{4.24b}$$

$$p_i^{1,l}(y_i - \overline{z}_i^{1,l}) = 0. \tag{4.24c}$$

Proof. *See Appendix C.6.*

The next two lemmas conclude the necessary conditions by showing that NE implies the stationary conditions in (4.2) and (4.5).

Lemma 10. *(Stationarity - UTP) At any $m^{\mathrm{u}} \in \mathcal{NE}^{\mathrm{u}}$, of the game $\mathfrak{G}^{\mathrm{u}}$, the following constraints are satisfied,*

$$v_i'(\widehat{x}_i^{\mathrm{u}}(m^{\mathrm{u}})) = \sum_{l \in \mathcal{L}_i} p^l \quad \textit{if} \quad \widehat{x}_i^{\mathrm{u}}(m^{\mathrm{u}}) > 0 \tag{4.25a}$$

$$v_i'(\widehat{x}_i^{\mathrm{u}}(m^{\mathrm{u}})) \leq \sum_{l \in \mathcal{L}_i} p^l \quad \textit{if} \quad \widehat{x}_i^{\mathrm{u}}(m^{\mathrm{u}}) = 0. \tag{4.25b}$$

Proof. *See Appendix C.7.*

Lemma 11. *(Stationarity - MMTP) The following constraints hold at any $m^{\mathfrak{m}} \in \mathcal{NE}^{\mathfrak{m}}$, of the game $\mathfrak{G}^{\mathfrak{m}}$,*

$$v_i'(\widehat{x}_i^{\mathfrak{m}}(m^{\mathfrak{m}})) = \sum_{l \in \mathcal{L}_i} p_i^{1,l} \quad \textit{if} \quad \widehat{x}_i^{\mathfrak{m}}(m^{\mathfrak{m}}) > 0 \tag{4.26a}$$

$$v_i'(\widehat{x}_i^{\mathfrak{m}}(m^{\mathfrak{m}})) \leq \sum_{l \in \mathcal{L}_i} p_i^{1,l} \quad \textit{if} \quad \widehat{x}_i^{\mathfrak{m}}(m^{\mathfrak{m}}) = 0. \tag{4.26b}$$

Proof. *See Appendix C.8.*

As mentioned earlier, the overall strategy for proving our main result is to show that any NE satisfies the KKT conditions of the original problem and then showing that such an equilibrium indeed exists. This last step is shown in the following lemma.

Lemma 12. *(Existence of NE) There exist infinitely many Nash equilibria $m^{\mathfrak{u}} \in \mathcal{NE}^{\mathfrak{u}}$, for the game $\mathfrak{G}^{\mathfrak{u}}$. Also, there exist infinitely many Nash equilibria $m^{\mathfrak{m}} \in \mathcal{NE}^{\mathfrak{m}}$, for the game $\mathfrak{G}^{\mathfrak{m}}$.*

Proof. *See Appendix C.9.*

The above series of lemmas is sufficient to prove the full implementation result for the two mechanisms for UTP and MMTP. Individual rationality and weak budget balance are shown separately in the following lemma.

Lemma 13. *(Individual Rationality and Weak Budget Balance) At any NE of the games $\mathfrak{G}^{\mathfrak{u}}$ and $\mathfrak{G}^{\mathfrak{m}}$, individual rationality is satisfied*

$$v_i(\widehat{x}_i^{\mathfrak{u}}(m^{\mathfrak{u}})) - \widehat{t}_i^{\mathfrak{u}}(m^{\mathfrak{u}}) \geq v_i(0),\ \forall i \in \mathcal{N} \tag{4.27a}$$

$$v_i(\widehat{x}_i^{\mathfrak{m}}(m^{\mathfrak{m}})) - \widehat{t}_i^{\mathfrak{m}}(m^{\mathfrak{m}}) \geq v_i(0),\ \forall i \in \mathcal{N}. \tag{4.27b}$$

Furthermore, both of the mechanisms are weak budget balanced

$$\sum_{i \in \mathcal{N}} \widehat{t}_i^{\mathfrak{u}}(m^{\mathfrak{u}}) \geq 0 \tag{4.28a}$$

$$\sum_{i \in \mathcal{N}} \widehat{t}_i^{\mathfrak{m}}(m^{\mathfrak{m}}) \geq 0. \tag{4.28b}$$

Proof. *See Appendix C.10.*

We are now ready to state the proofs of Theorems 4 and 5.

Proof of Theorem 4. In the proof of Lemma 12, we showed that the message associated to the solution of problem (4.1), (x^*, λ^*), is a NE of the game $\mathfrak{G}^{\mathfrak{u}}$. Now, we want to prove that all of the NE of the game $\mathfrak{G}^{\mathfrak{u}}$ generate allocation and prices that are efficient for problem (4.1). Consider any $m^{\mathfrak{u}} \in \mathcal{NE}^{\mathfrak{u}}$, if $\widehat{x}^{\mathfrak{u}}(m^{\mathfrak{u}})$ is used as the primal variables vector and $p = \{p^1, ..., p^L\}$ is used as the dual variables vector, all of the KKT conditions (4.2) are satisfied due to Lemmas 5, 7, 8, and 10. Therefore, $\widehat{x}^{\mathfrak{u}}(m^{\mathfrak{u}}) = x^*$ for any $m^{\mathfrak{u}} \in \mathcal{NE}^{\mathfrak{u}}$ and so, full implementation is obtained. Furthermore, Lemma 13 proves individual rationality and weak budget balance at any NE of the game $\mathfrak{G}^{\mathfrak{u}}$. $\square$

The proof of Theorem 5 is very similar to the proof of Theorem 4 and it is stated below for completeness.

Proof of Theorem 5. Let $(x^*, b^*, \lambda^*, \mu^*)$ be the solution of problem (4.4) and consider any $m^{\mathfrak{m}} \in \mathcal{NE}^{\mathfrak{m}}$. Due to Lemmas 6, 7, 9 and 11, the allocation vector, $\widehat{x}^{\mathfrak{m}}(m^{\mathfrak{m}})$ as x^*, $r_i^{\mathfrak{m}}\overline{z}_i^{1,l}$ as $b_{k(i)}^{l*}$ (any $r_j^{\mathfrak{m}1}\overline{z}_j^l$, $j \in \mathcal{G}_{k(i)}^l$ could work too) and the variables $p_i^{1,l}$ and w_i^l (or any w_j^l for $j \in \mathcal{N}^l$) as μ_i^{l*} and λ^{l*}, respectively, satisfy the KKT conditions (4.5). Therefore, $\widehat{x}^{\mathfrak{m}}(m^{\mathfrak{m}}) = x^*$ for any $m^{\mathfrak{m}} \in \mathcal{NE}^{\mathfrak{m}}$ and hence, the allocation at all NE is unique and efficient. Also, due to Lemma 12, we know at least one NE exists and therefore, the mechanism fully implements problem (4.3) at its Nash equilibria. Furthermore, Lemma 13 proves individual rationality and weak budget balance. $\square$

4.6 Relaxing Assumption 1 on Message-Exchange Network

The primary reason of imposing Assumption 1 is that there should be a consensus on the prices of different agents (UTP) or groups (MMTP) using link l at NE and this is not implementable by the proposed mechanism if the sub-graph of the agents using link l is not connected.

On the other hand, a message-exchange network that satisfies the required properties may be hard or even impossible to construct. Consider the special case of having only one link (constraint) in the UTP/MMTP optimization problem. Then, the message-exchange network should be the tree that contains all of the agents of the network. But in general there are more than one links in the UTP/MMTP optimization problem and the message-exchange network should consist of multiple connected sub-graphs (each corresponding to one constraint) and should still be a tree. Constructing such message-exchange network is hard and may be impossible. In this section, we propose an alternative extended mechanism so that there is no need for imposing Assumption 1 on the message-exchange network.

In the alternative mechanism, we extend the agents that quote message p_i^l, in UTP, and w_i^l, in MMTP, from the agents using link l to a bigger group of agents as follows. For every link l, consider a connected sub-graph $\mathcal{GR}^l = (\widehat{\mathcal{N}}^l, \mathcal{E}^l)$ consisting of all agents $i \in \mathcal{N}^l$ in addition to the minimum number of other agents that do not use link l and are required to make the sub-graph connected. This connected sub-graph is called link l's sub-graph and we know that it exists due to the connectivity of the message graph. For each agent i, the set of links $l \notin \mathcal{L}_i$ which $i \in \widehat{\mathcal{N}}^l$ are denoted by $\widehat{\mathcal{L}}_i$ with $|\widehat{\mathcal{L}}_i| = \widehat{L}_i$. The definition of $\mathcal{N}^l(i)$ is modified as

$$\mathcal{N}^l(i) = \{j: \ j \in \mathcal{N}(i) \cap \widehat{\mathcal{N}}^l\}, \ \forall i \in \mathcal{N}, l \in \mathcal{L}_i \cup \widehat{\mathcal{L}}_i. \tag{4.29}$$

In the game $\mathfrak{G}^{\mathfrak{u}}$, the extended definition of message p_i is $p_i = (p_i^l, l \in \mathcal{L}_i \cup \widehat{\mathcal{L}}_i)$, while in the game $\mathfrak{G}^{\mathfrak{m}}$, the extended definition of message w_i is $w_i = (w_i^l, l \in \mathcal{L}_i \cup \widehat{\mathcal{L}}_i)$. The tax functions are also modified for $l \in \widehat{\mathcal{L}}_i$ according to

$$\widehat{t}_i^{\mathfrak{u},l}(m^{\mathfrak{u}}) = \sum_{j \in \mathcal{N}(i)} (n_{i,j}^l - y_j^l - \sum_{h \in \mathcal{N}(j), h \neq i} n_{j,h}^l)^2 + (p_i^l - \overline{p}_{-i}^l)^2 + (p_i^l - \overline{p}_{-i}^l)\overline{p}_{-i}^l(c^l - r_i^{\mathfrak{u}} f_i^{\mathfrak{u},l})^2 \tag{4.30a}$$

$$\widehat{t}_i^{\mathfrak{m},l}(m^{\mathfrak{m}}) = \sum_{j \in \mathcal{N}(i)} (n_{i,j}^l - y_j^l - \sum_{h \in \mathcal{N}(j), h \neq i} n_{j,h}^l)^2 + (w_i^l - \overline{w}_{-i}^l)^2 + \overline{w}_{-i}^l(w_i^l - \overline{w}_{-i}^l)(c^l - r_i^{\mathfrak{m}} f_i^{\mathfrak{m},l})^2. \tag{4.30b}$$

Intuitively, since the sub-graph of agents using each link l may not be connected, we need other agents $i \notin \mathcal{N}^l$ to quote p_i^l messages in the game $\mathfrak{G}^{\mathfrak{u}}$ and w_i^l messages in the game $\mathfrak{G}^{\mathfrak{m}}$. This helps the agents $j \in \mathcal{N}^l$ in forming a consensus on the prices or group prices of using link l. This is why two terms have been added to the tax functions above that impose required conditions for the messages p_i^l and w_i^l in the two games. In both games, the tax function does not change for $l \in \mathcal{L}_i$. For $l \notin \mathcal{L}_i \cup \widehat{\mathcal{L}}_i$, the tax function is the same as the $l \notin \mathcal{L}_i$ case for the original mechanism. It is straightforward to prove almost the same results (with some minor changes to cover the new messages) for these mechanisms. Therefore, these mechanisms also fully Nash implement problem (4.1) and (4.3), respectively, and are individually rational and weak budget balanced at NE.

4.7 Discussion and Conclusion

We proposed two distributed mechanisms for the networks with UTP and MMTP. As mentioned before, the mechanisms are applicable to a number of other optimization problems with linear/max constraints. The proposed mechanisms were proved to fully Nash implement the solution of problems (4.1) (UTP) and (4.3) (MMTP). The main feature of this work is that message transmission is done locally via an underlying message-exchange network, in contrast to the standard mechanism design framework that allows message transmission throughout the whole network.

The dimensionality of agent i's message in the mechanism for UTP is $M_i = 1+N(i)L+|\mathcal{I}_i|+L_i$. Since for each agent i, the function $\phi(i)$ chooses one agent $j \in \mathcal{N}(i)$, the average size of the set $\mathcal{I}_i$ is 1. Hence, the average dimensionality of an agent's message is $\mathbb{E}_{i\in\mathcal{N}}[M_i] = 2 + \mathbb{E}_{i\in\mathcal{N}}[N(i)]L + \mathbb{E}_{i\in\mathcal{N}}[L_i]$ and by denoting $\mathbb{E}_{i\in\mathcal{N}}[N(i)]$ and $\mathbb{E}_{i\in\mathcal{N}}[L_i]$ by $\overline{N}$ and $\overline{L}$ respectively, the average message dimensionality of the whole network is

$$\mathbb{E}[\sum_{i\in\mathcal{N}} M_i] = N(2 + L\overline{N} + \overline{L}).$$

Clearly, the dimensionality of message space grows linearly with N.

The dimensionality of agent i's message in the mechanism for MMTP is $M_i = 1 + 4L_i + N(i)L + |\mathcal{I}_i| + 2\sum_{j\in\mathcal{I}_i} L_j + 2|\mathcal{C}_i|$. Similar to UTP, the average size of the set $\mathcal{I}_i$ is 1. Further, the average value of $\sum_{j\in\mathcal{I}_i} L_j$ is $\frac{\sum_{i\in\mathcal{N}} L_i}{N}$. Also, we know that for each link l and group $k \in \mathcal{K}^l$, there is one agent denoted by $c(k,l)$ and hence, the average size of $|\mathcal{C}_i|$ is $\frac{\sum_{l\in\mathcal{L}} K^l}{N}$. Consequently, the average size of the whole network's message is

$$\mathbb{E}[\sum_{i\in\mathcal{N}} M_i] = N(2 + 4\overline{L} + \overline{N}L + 2\frac{\sum_{i\in\mathcal{N}} L_i}{N} + 2\frac{\sum_{l\in\mathcal{L}} K^l}{N}),$$

which, similar to UTP, grows linearly with the number of agents in the network, N.

For the alternative mechanism presented in section 4.6, the term $N\mathbb{E}_{i\in\mathcal{N}}(\widehat{L}_i)$ should be added to the average of the message dimensionality of the whole network. This is due to the extra messages that agents have to quote in the alternative mechanisms to preserve the connectivity of the message passing.

In terms of message dimensionality, the mechanisms proposed in this work are more efficient than the distributed mechanism proposed in [112] which has a message dimensionality that grows with N^2. This gain in dimensionality may be a consequence of the fact that the proposed mechanism in [112] has additional learning guarantees that our proposed mechanism does not possess.

We would like to emphasize that although the proposed mechanisms are proven to have efficient Nash equilibria, in the current work we do not propose any message exchange algorithm that is guaranteed to converge to these equilibria. In general, it is not even clear if such algorithm exists. One possible future research direction is redesigning these mechanisms to possess other features such as learning guarantees and convergence to NE. Such features would enable the mechanisms to converge to their NE in a dynamic learning process over a large set of possible dynamics followed by the agents [125, 112]. In addition, the study of the tradeoff between message dimensionality and convergence guarantees is an interesting open problem.

CHAPTER 5

Joint Information and Mechanism Design for Queues with Heterogeneous Users

5.1 Introduction

Decentralized information is an important and inevitable aspect of today's systems. Each agent in a system can own some information that others might be interested to know. On the other hand, agents usually act strategically and might not be willing to share their information with others. Therefore, incentives have to be put in place to motivate agents to reveal some parts of their information. There are two main approaches, mechanism design and information design, where the sharing of information between agents and their incentives of doing so is studied.

In mechanism design [126, 11, 20, 127, 51, 128, 129, 22, 52, 130, 131, 132, 133], there are a number of agents with some private information. There is a designer that designs messages together with allocation and tax/subsidy functions. Agents, as "senders" of information, send their messages, which could convey true or false information, to the central authority, acting as the "receiver" of information. Upon reception of these messages, the central authority will then determine their allocations and taxes/subsidies. The incentives for the agents to send truthful messages are created through allocation and tax functions. Note that the central authority commits to the allocation and tax function and can not change these functions after hearing the agents' messages.

In information design [38, 134, 135, 136, 137, 54], there is usually one "sender" who owns a piece of information. The sender shares some part of his information with a number of agents as "receivers" by sending messages to them. The messages are created according to a policy that is to be designed by the sender. The agents will then interpret the messages using the policy based on which the messages are generated and then they take some actions. The sender has to optimally choose his policy to steer agents' actions to a desired direction. Note that similar to the mechanism design framework, the sender commits to the policy he is using to create messages. The difference is that, the commitment in information design is from the sender while in mechanism design it is from the receiver. Information design problems with one sender and one receiver are usually referred to as "Bayesian persuasion", which was introduced in [38], where the authors present a geometric form of interpreting information design and show when it is profitable for the designer to not share some part of the information. However, when there are multiple receivers, the information design problem becomes more complex and notions of equilibria must be introduced to analyze the game. As it is shown in [138], an information design model with multiple receivers is in fact a game with incomplete information and the set of outcomes of the information design problem corresponds to the set of Bayes-correlated equilibria (BCE). In the definition of BCE in [138], the information designer follows a direct strategy where he directly recommends actions to the players. The strategy of the information designer has to satisfy an obedience condition, that is, each player has to be willing to follow the recommendation.

In this work, we combine the two approaches and study joint information and mechanism design for a queuing system. In our model, there is a queue with an unobservable backlog by the incoming users who have payoff relevant private types. There is a planner who observes the queue backlog and designs a mechanism and an admission policy for the users. A user, upon arrival, decides to either be admitted to the queue by the planner or join the queue or not on her own. If she decides to be admitted by the planner, she has to send a message, that is supposed to be her type, to the planner. The planner then creates an admission signal for that specific type of user. Information about the queue backlog is conveyed through the admission signal and the users have to pay a tax in exchange for the information they receive. The information sent through the admission signal and the tax function incentivize the users to reveal their true type. In this setting, the planner is both an information designer (designs the admission policy) and a mechanism designer (designs the tax function) and he has commitments to both his admission policy and the tax function. Note that the planner is a sender of the information in the information design aspect and a receiver in the mechanism design aspect of our model. We formulate an optimization problem that characterizes the solution of the joint design problem. We characterize the tax function that satisfies dominant strategy incentive compatibility and provide structural results, which are supported by numerical analysis, for the optimal admission policy.

The problem studied in this chapter can model many real world applications such as admission control for a customer service call center where the customers can decide to wait in line for a representative or leave the queue on their own, or they can pay a price to be admitted by a planner (and possibly avoid long wait times). Another example could be ride sharing apps such as Uber where the customers can wait (possibly a long time) for a ride or they can choose to pay a price and get in the line or be kicked out by the app to avoid long waits. In both of these two examples, the customers are unaware of the system backlog.

There are some works on information design where, similar to our model, the receivers have private information, e.g., private types. There are two approaches to these problems: without elicitation and with elicitation of the private information. In the case of information design without elicitation [139], the information designer has to send a list of suggestions for each possible type of the receivers. This setting is referred to as public persuasion by Kolotilin et al. in [139]. In the case of information design with elicitation [139, 39, 140], receivers report their types and instead of the obedience constraint, the decision rule of the information designer should satisfy an incentive constraint that makes sure each type of the receiver prefers her own recommendation over other recommendations that she can possibly hear if she reports her type untruthfully. Kolotilin et al. [139] refer to this setting as private persuasion. In [39], authors utilize monetary transfers, i.e., taxes/subsidies, to elicit the private types, as opposed to the model in [139] where elicitation is done without taxes. In [141, 142], the persuasion is done not only through information design, but also by using monetary transfers. However, the receiver does not have any private information. In [143], there is also some type of joint mechanism and information design but the information disclosure is public and not a function of the users' reported types. In addition, there are no monetary transfers. In [144], authors have studied the effect of a third-party data provider on simple mechanisms and in this sense, they have considered a joint information and mechanism design problem. They show that simple mechanisms fail to approximate the optimal revenue in the presence of a third-party signal.

Our formulation of joint information and mechanism design is similar in spirit to the one discussed in [39], where there are multiple players with private prior beliefs about a state of the world. The information designer offers a menu of experiments (that convey information) that players can choose from and they have to pay a tax in return. The information designer maximizes his revenue over the set of experiments and taxes subject to incentive compatibility and individual rationality constraints. Our setting can be considered a special case of the general framework discussed in [39]. The specifics of our model, such as users' utilities being linear in their private types, enable us to evaluate an explicit tax function and formulate a linear optimization problem for the planner.

The queuing system presented in this chapter builds on the model by [145] with the main difference being that in our model the users have private types where in [145], the incoming traffic is uniform (there are some discussions on the case of different user types but these types are assumed to be known to the information designer).

Information design problems can study dynamic or static systems. In static information design, there is no dynamic state in the system and there is no time involved. Therefore, the problem that the information designer faces is a static optimization problem [38, 138, 146, 147, 148, 149]. Dynamic information design problems deal with dynamic settings that usually involve time [150, 135, 151, 145]. Therefore, the strategy of the designer can be dynamic and the information can be disclosed sequentially. Dynamic programming techniques can therefore be used to characterize the optimal strategy. Although our model can be considered a dynamic information design problem, the strategy of the information designer is not time dependent. However, the strategy of the information designer affects the evolution of the state variable.

The rest of this chapter is structured as follows. In section 5.2, we discuss the model. In section 5.3 we characterize the users' strategies. The mechanism objectives are discussed in section 5.4. The tax functions are presented in section 5.5. We formulate the planner's optimization problem in section 5.6. We study the mechanism with individual rationality in section 5.7 together with the extreme cases of full information and no information mechanisms. We present numerical analysis in section 5.8 and we conclude in section 5.9.

5.2 Model

We consider a service provider with service rate 1. There is a queue with infinite capacity and users arrive at the queue according to a Poisson arrival process with rate $\lambda > 1$. We denote the number of users in the queue by x and we have $X \sim \mu(\cdot)$, where $\mu(\cdot)$ is the stationary distribution of the queue backlog. The users have payoff relevant private types $i \in \mathcal{I} = \{1, \cdots, N\}$ and the a-priori type distribution is known $I \sim P_I(\cdot)$. The queue backlog is unobservable by users. There is a planner who observes the queue backlog and sets up an admission control mechanism $\mathbf{M} \stackrel{\text{def}}{=} (\sigma, t)$ in which users can participate or not. The mechanism consists of an admission policy σ and a tax function t. Upon arrival, a user makes a decision of being admitted by the planner (join the mechanism) or decide to join the queue or not on her own. We define the following sequence of actions that users take at the time of arrival. Note that we refer to a user by she and to the planner by he.

- First, the user with type i who has arrived at the queue decides to participate in the admission mechanism or not by choosing the probability of participating in the mechanism, γ_i. Her action is denoted by d where $P(D = 1) = \gamma_i$. We denote $\gamma = (\gamma_1, \cdots, \gamma_N)$.

- If the user does not participate in the mechanism, she will make a decision to join the queue or not by deciding on the probability of joining the queue, α_i. Her decision is denoted by e where $P(E = 1) = \alpha_i$. We denote $\alpha = (\alpha_1, \cdots, \alpha_N)$. If a user does not participate in the mechanism, we say she has chosen the outside option.

- If the user participates in the mechanism, she has to send a message $m = f(i),\ m \in \mathcal{I}$ to the planner. The mechanism is a direct mechanism and the message m sent by a user with type i

is supposed to be her type. The planner determines a tax $t(m)$ that is to be paid by the user. He then generates a (randomized) admission signal s where $S \sim \sigma(\cdot|x, m)$ and announces it to the user, where σ is the admission policy. The user will then either join the queue if she is admitted or leave if she is not admitted.

Note that we assume all users with the same type choose the same strategy even if there is a tie. Based on the steps described above, we can denote the information set of a user by $h \in \mathcal{H}$, where the set of information sets is defined as $\mathcal{H} = \{(i, d = 0, e = 0)_{(i \in \mathcal{I})},\ (i, d = 0, e = 1)_{(i \in \mathcal{I})},\ (i, d = 1, m, s = 0)_{(i, m \in \mathcal{I})},\ (i, d = 1, m, s = 1)_{(i, m \in \mathcal{I})}\}$. The utility of a user for each of these information sets is denoted by $u(h)$, and is described in the following.

The users have to pay a price p for the service they receive if they opt out of the mechanism (choose the outside option) but enter the queue. As mentioned before, the users pay a tax $t(m)$ if they participate in the mechanism and they do not pay any other price for receiving service if they are admitted by the planner (the price of the service is included in the tax function). The user with type i receives a reward of $iv(x)$ by joining the queue of backlog x, where $v(\cdot)$ is a decreasing function, which can be negative for large enough x. Therefore, for $h = (i, d = 0, e = 1)$, the user receives the expected utility of $u(i, d = 0, e = 1) = i\overline{v} - p$ where $\overline{v} = \mathbb{E}[v(X)] = \sum_{x=0}^{\infty} v(x)\mu(x)$. For $h = (i, d = 0, e = 0)$, she leaves the queue and receives $u(i, d = 0, e = 0) = 0$. For $h = (i, d = 1, m, s = 1)$, the user receives the expected utility of $u(i, d = 1, m, s = 1) = i\mathbb{E}[v(X)|h] - t(m)$. Finally, for $h = (i, d = 1, m, s = 0)$, the user receives the utility $u(i, d = 1, m, s = 0) = -t(m)$. Figure 5.1 depicts the extended form of the game, where the black circles indicate decision points.

One can express the joint probability distribution of the random variables described in this model as follows.

$$\mathbb{P}(x, i, d, e, m, s) = \mu(x)P_I(i)\mathbb{P}(d, e, m|i)\sigma(s|x, m), \tag{5.1}$$

where the stationary distribution of X, $\mu(\cdot)$, depends on σ. Note that $\mathbb{P}(d, e, m|i)$ is determined by the strategy of the user with type i and we have

$$\mathbb{P}(m|i) = \mathbf{1}_{f(i)}(m) \tag{5.2}$$

$$\mathbb{P}(d = 1|i) = \gamma_i \tag{5.3}$$

$$\mathbb{P}(e = 1|i) = \alpha_i, \tag{5.4}$$

where $\mathbf{1}_a(b) = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{o.w.} \end{cases}$. In order to characterize the stationary distribution of X, we need to know the effective arrival rate to the queue. The arrival to the queue consists of both users joining

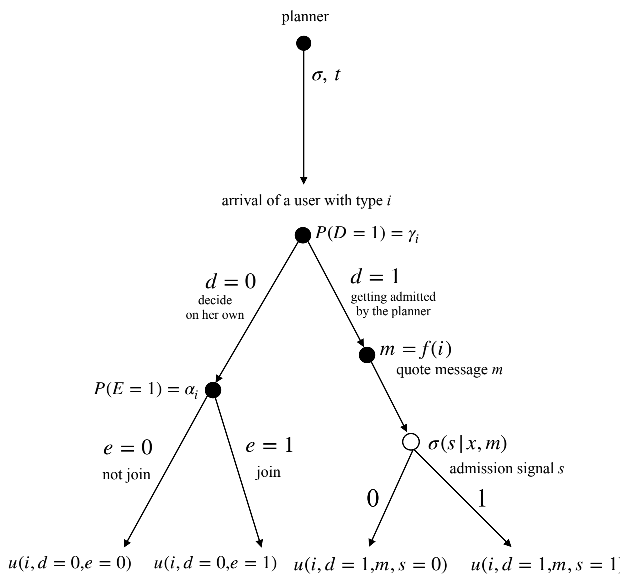


Figure 5.1: Extended form of the game.

the queue through the admission mechanism and users joining the queue on their own in the outside option. Therefore, in order to know the arrival rate, we need the admission policy, γ_i and α_i for all i. The stationary distribution of X, $\mu(\cdot)$, can be found using the following lemma.

Lemma 14. *The stationary distribution of X, i.e., $\mu(\cdot)$, is given by the following equation.*

$$\mu(x+1) = \lambda\mu(x) \sum_{i=1}^{N} P_I(i)(\gamma_i \sigma(1|x, f(i)) + (1-\gamma_i)\alpha_i)$$

$$\sum_{x=0}^{\infty} \mu(x) = 1.$$

Proof. *See Appendix D.1.*

To calculate the average utility $u(i, d = 1, m, s = 1)$, we find $\mathbb{E}[v(X)|h]$ using the joint probability distribution of the random variables as follows.

$$\mathbb{E}[v(X)|h] = \frac{1}{\mathbb{P}(S = 1|m)} \sum_{x=0}^{\infty} v(x)\mu(x)\sigma(1|x, m), \tag{5.6}$$

where $\mathbb{P}(S=1|m)=\sum_{x=0}^{\infty}\mu(x)\sigma(1|x,m)$.

The average utility that a user with type i predicts to receive if she participates in the mechanism (before making any decision or hearing any signal) is $\mathbb{E}[u(i,d=1,m,S)]$ and we have

$$\begin{aligned}\mathbb{E}[u(i,d=1,m,S)] &= \mathbb{P}(S=1|m)u(i,d=1,m,s=1)+\mathbb{P}(S=0|m)u(i,d=0,m,s=0)\\ &= i\sum_{x=0}^{\infty}v(x)\mu(x)\sigma(1|x,m)-t(m), \end{aligned} \tag{5.7}$$

We define

$$q(m)=\sum_{x=0}^{\infty}v(x)\mu(x)\sigma(1|x,m), \quad \forall m\in\mathcal{I}, \tag{5.8}$$

to be the "allocation" to a user that reports m as her type. Therefore, the average utility of a user with type i that reports m as her type is

$$\mathbb{E}[u(i,d=1,m,S)]=iq(m)-t(m). \tag{5.9}$$

This formulation enables us to solve a mechanism design problem with linear utilities with respect to allocations.

The average utility that the user predicts to receive from the outside option is the maximum of 0 and $i\overline{v}-p$ (depending on the action e, one of these utilities are received), and we denote it by $(i\overline{v}-p)^+$, where we define the plus operator as $(a)^+=\max(a,0)$.

5.3 Users' Strategies

In this section, we characterize the users' best response strategies to the planner's decision on σ and t. We first summarize the actions that are taken in the game. We have actions/decisions d, e, m, that are taken by the users and for a user with type i, $\mathbb{P}(D=1)=\gamma_i$, $\mathbb{P}(E=1)=\alpha_i$ and $m=f(i)$. We also have the admission policy, σ, that generates the admission signal, s, according to $S\sim\sigma(\cdot|x,m)$ and tax function t that are to be designed by the planner. The following lemma characterizes α_i, function f, and γ_i.

Lemma 15. *For a given σ and t, we have the following for α_i.*

- $\alpha_i=1$ *If* $i\overline{v}-p>0$
- $\alpha_i=0$ *If* $i\overline{v}-p<0$

- $\alpha_i \in [0,1]$ *If* $i\overline{v} - p = 0$.

Also, function f *is given by the following equation.*

$$f(i) = \arg\max_{m}\ iq(m) - t(m).$$

Finally, for γ_i *we have*

- $\gamma_i = 1$ *If* $iq(m) - t(m) > (i\overline{v} - p)^+$
- $\gamma_i = 0$ *If* $iq(m) - t(m) < (i\overline{v} - p)^+$
- $\gamma_i \in [0,1]$ *If* $iq(m) - t(m) = (i\overline{v} - p)^+$,

where $m = f(i)$.

Proof. *See Appendix D.2.*

5.4 Mechanism Objectives

The mechanism $\mathbf{M} \stackrel{\text{def}}{=} (\sigma, t)$ is designed by the planner to have the following properties:

- DSIC: The mechanism should be dominant strategy incentive compatible (DSIC). That is, all users should act truthfully in reporting their types, no matter what other users do, i.e., $m = f(i) = i$ for all $i \in \mathcal{I}$. For the mechanism to be DSIC we should have the following.

$$f(i) = \arg\max_{m}\ iq(m) - t(m) = i,\ \forall i \tag{5.10}$$

- The mechanism should maximize the planner's expected revenue. That is, the planner solves the following optimization problem.

$$\sigma^*, t^* \in \arg\max_{\sigma,t} \lambda\mathbb{E}[t(M)D], \tag{5.11}$$

  where $M = f(I)$. Note that the tax is only paid by the users who participate in the mechanism, i.e., $d = 1$.

Note that in DSIC constraint, the expected utility $\mathbb{E}(u(i, d = 1, m, S))$ depends on the strategy and actions of other players only through the stationary distribution of X. The DSIC constraint ensures that for any stationary distribution of X, acting truthful is the best strategy. Therefore, for any strategy of other players, truthful strategy is the best response and therefore, we have dominant strategy incentive compatibility.

Notice that we do not enforce individual rationality constraint for the mechanism as it might be more beneficial for the planner to not have all types of the users participate in the mechanism. The intuitive reason is that it might be too costly for the planner to incentivize all types to participate in the mechanism and so some types of the users might choose the outside option, i.e., choose on their own to join the queue or not.

Given the DSIC constraint, the planner's expected revenue can be simplified as follows.

$$\lambda \mathbb{E}[t(I)D] = \lambda \sum_{i=1}^{N} \gamma_i P_I(i) t(i). \tag{5.12}$$

5.5 Tax Function

In this section we introduce the tax function for the mechanism and prove DSIC. We also prove that this type of tax function maximizes the planner's revenue.

We consider the following tax function.

$$t(m) = t_0 + mq(m) - \sum_{j=1}^{m-1} q(j), \quad \forall m \in \mathcal{I}. \tag{5.13}$$

We refer to t_0 by the tax offset.

We will see in the next theorem that this types of tax function ensures dominant strategy incentive compatibility. Notice that there are two degrees of freedom in the tax function, t_0 and $q(\cdot)$ that is determined by σ and the planner will choose them to maximize his revenue.

Theorem 6. *The mechanism is dominant strategy incentive compatible if*

- $q(m)$ *is weakly increasing w.r.t.* m.
- *The tax function is given by equation* (5.13).

Furthermore, the given tax function maximizes the planner's revenue among all other DSIC tax functions.

Proof. *See Appendix D.3.*

5.6 Planner's Optimization Problem

Given the tax function described in the previous section and the fact that DSIC holds for a mechanism with such tax function, the objective of the planner is to maximize the following quantity.

$$\lambda \mathbb{E}[t(I)D] = \lambda \sum_{i=1}^{N} \gamma_i P_I(i)(t_0 + iq(i) - \sum_{j=1}^{i-1} q(j)) \tag{5.14a}$$

As we mentioned in the previous section, the tax function t is determined by σ and t_0 and therefore, the planner maximizes his revenue w.r.t. σ and t_0 instead of σ and t as we had in equation (5.11). Also, notice that we assume all users with the same type make the same actions even if there is a tie, and we assume that if there is a tie, the users will choose the action that benefits the planner the most. Therefore, one can say that the planner is also optimizing over α_i and γ_i subject to the constraints in Lemma 15 (function f is already determined as $f(i) = i$ according to DSIC condition). Hence, we can formulate the planner's optimization problem as follows.

$$\max_{\sigma, t_0, \alpha, \gamma} \quad \lambda \sum_{i=1}^{N} \gamma_i P_I(i)(t_0 + iq(i) - \sum_{j=1}^{i-1} q(j)) \tag{5.15a}$$

$$s.t. \quad q(i) \leq q(i+1), \quad \forall i \in \mathcal{I}, i \neq N \tag{5.15b}$$

$$\mu(x+1) = \lambda \mu(x) \sum_{i=1}^{N} P_I(i)(\gamma_i \sigma(1|x,i) + (1-\gamma_i)\alpha_i) \forall x \geq 0 \tag{5.15c}$$

$$\sum_{x=0}^{\infty} \mu(x) = 1. \tag{5.15d}$$

$$\alpha \text{ and } \gamma \text{ satisfy Lemma 15 for } \sigma \text{ and } t_0. \tag{5.15e}$$

In the following we give some illustration on how t_0 affects which types of players will participate in the mechanism and which ones do not. The utility of a user with type i that participates in the mechanism is as follows.

$$\sum_{j=1}^{i-1} q(j) - t_0. \tag{5.16}$$

The utility of the outside option for a user with type i is $(i\overline{v} - p)^+$. Figure 5.2 demonstrates an example of the outside option utility and the utility of the admission mechanism with respect to the type of the users for two different value of t_0, 0 and $-a$ for some $a > 0$. The types for which we have the admission mechanism utility below the outside option utility, do not participate in the mechanism and $\gamma_i = 0$. As the planner decreases t_0 from 0 to $-a$, more types of players participate in the mechanism. In the example of Figure 5.2, with $t_0 = 0$, users of type 2, 3 and 4 do not participate in the mechanism but with $t_0 = -a$, only type 3 does not participate.

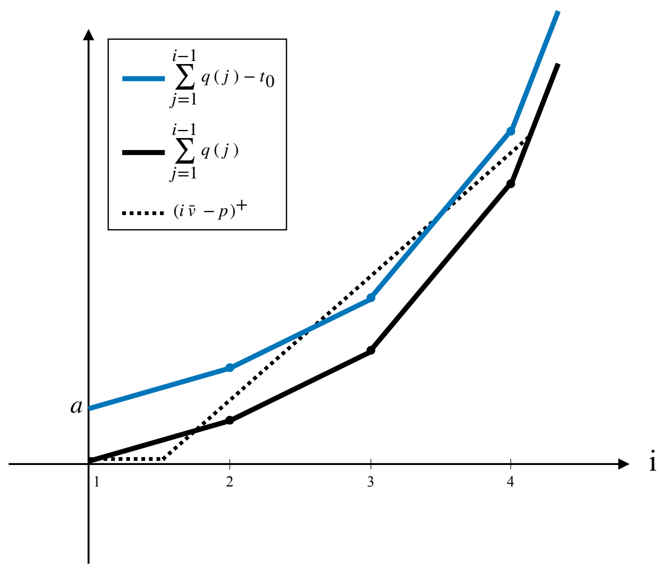


Figure 5.2: An example of how t_0 affects which types of users participate in the mechanism

The optimization problem (5.15) is not linear or convex and therefore, through numerical analysis, we first characterize the possible best values for α and γ and then we focus our attention to the optimization problem for the specific optimal values of α and γ. In Table 5.1, we have listed the optimal values of α and γ for the special case of $N = 2$ for different values of p. We have done the analysis for $\lambda = 1.2$, $v(x) = \frac{1}{50^2}(50^2 - x^2)$, and $P_I(1) = P_I(2) = \frac{1}{2}$. Note that since $v(x) \leq 1$, we consider different values of $p \in [0, 1]$.

The analysis in Table 5.1 shows that for most reasonable values of p (the values for which the outside option is indeed a reasonable option), the optimal value of the parameters γ_1 and γ_2 is 1. It indicates that it is most beneficial for the planner to have both types of the users participate in the mechanism. Notice that the parameter α_1 and α_2 are the decision of users in off-equilibrium decision points and although they are chosen so that the players would be best responding at those points, they do not play any role in the on-equilibrium system dynamics. Although this analysis is

p	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
α_1	0	0	0	0	0	0	0	0	0	0	0
α_2	1	1	1	1	1	1	1	1	1	1	1
γ_1	1	1	1	1	1	1	1	0.7	0.7	0.7	0.75
γ_2	1	1	1	1	1	1	1	1	1	1	1

Table 5.1: The optimal values of γ and α for a range of different values of p. The analysis is done for $N = 2$, $\lambda = 1.2$ and $v(x) = \frac{1}{50^2}(50^2 - x^2)$.

only done for the special $N = 2$ case, to simplify the planner's optimization problem, in the rest of the chapter, we focus our attention to the case of $\gamma_i = 1$ for all $i \in \mathcal{I}$, i.e., all types of users participate in the mechanism. Notice that if $\gamma_i = 1$, the value of α_i does not matter for the planner. The condition that enforces the mechanism to have all types of users participating in the mechanism is called individual rationality.

5.7 Mechanism with Individual Rationality

In this section, we reformulate the planner's optimization problem by enforcing the individual rationality constraint. That is, we assume the planner designs the mechanism such that users choose $\gamma_i = 1$ for all $i \in \mathcal{I}$. We further notice that even for a fixed value of $\gamma_i = 1$ for all $i \in \mathcal{I}$, the planner's optimization problem is not linear with respect to σ because of the stationary constraints of (5.15c). Therefore, we restate the problem in terms of the joint probability distribution on (S, I, X), denoted by $\gamma(S, I, X)$. Note that $q(i) = \sum_{x=0}^{\infty} v(x)\mu(x)\sigma(1|x, i) = \sum_{x=0}^{\infty} v(x)\frac{\gamma(s=1,i,x)}{P_I(i)}$. Therefore, we have the following linear optimization problem for the planner.

$$\max_{\gamma, t_0} \;\; \lambda t_0 + \lambda \sum_{i=1}^{N} \sum_{x=0}^{\infty} v(x)(i\gamma(s = 1, i, x) - \sum_{j=0}^{i-1} \frac{P_I(i)}{P_I(j)} \gamma(s = 1, j, x)) \tag{5.17a}$$

$$s.t. \;\; N \sum_{j=1}^{i-1} \sum_{x=0}^{\infty} v(x)\gamma(s = 1, j, x) - t_0 \geq i \sum_{x=0}^{\infty} v(x) \sum_{s,j} \gamma(s, j, x) - p, \; \forall i \in \mathcal{I} \tag{5.17b}$$

$$N \sum_{j=1}^{i-1} \sum_{x=0}^{\infty} v(x)\gamma(s = 1, j, x) - t_0 \geq 0, \; \forall i \in \mathcal{I} \tag{5.17c}$$

$$\sum_{x=0}^{\infty} v(x)\gamma(s = 1, i, x) \leq \sum_{x=0}^{\infty} v(x)\gamma(s = 1, i + 1, x), \forall i \in \mathcal{I}, i \neq N \tag{5.17d}$$

$$\sum_{s,i} \gamma(s, i, x + 1) = \lambda \sum_{i} \gamma(1, i, x), \quad \forall x \geq 0 \tag{5.17e}$$

$$\sum_{s} \gamma(s,i,x) = \frac{1}{N} \sum_{s,j} \gamma(s,j,x), \ \forall i \in \mathcal{I}, \ x \geq 0 \tag{5.17f}$$

$$\sum_{s,i,x} \gamma(s,i,x) = 1 \tag{5.17g}$$

$$\gamma(s,i,x) \geq 0, \quad \forall s \in \{0,1\}, \ i \in \mathcal{I}, \ x \geq 0. \tag{5.17h}$$

Note that constraints (5.17b) and (5.17c) correspond to linearized constraints of the individual rationality condition, while constraint (5.17f) is to ensure $\mathbb{P}(x,i) = \mu(x)P_I(i)$ according to (5.1).

Using the linear formulation of the planner's optimization problem, we will provide some structural properties for the optimal policy in the next section. Also, in the numerical analysis section, we will use this linear formulation to numerically solve the optimization problem using Matlab.

5.7.1 Structural Properties

In this section, we discuss some properties and behaviors of the optimal admission policy of the mechanism with individual rationality, i.e., the solution of the optimization problem (5.17). We first define the dominance of type i over type j in an admission policy σ as follows.

Definition 7 (Dominance)**.** *For a given admission policy σ, type i dominates type j if the following hold.*

- *If $v(x) > 0$ and $\sigma(s=1|j,x) > 0$, then $\sigma(s=1|i,x) = 1$.*
- *If $v(x) < 0$ and $\sigma(s=1|i,x) > 0$, then $\sigma(s=1|j,x) = 1$.*

Note that the dominance condition implies that the planner favors type i over type j in his admission policy.

In the next theorem, we present some results for the special case of two types of users and then generalize the results to multiple types.

Theorem 7. *Suppose t_0 and γ^* (or equivalently σ^*) are the solution of* (5.17) *for $N = 2$. Then, one of the following holds.*

- *Case 1: Type 2 dominates type 1.*
- *Case 2: There is a threshold $\tilde{x}$ such that for $x \geq \tilde{x}$ we have $\sigma^*(1|i,x) = 0$ for all $i \in \mathcal{I}$. Furthermore, for $x < \tilde{x}$, $\sigma^*(1|i,x) = 1$ for all $i \in \mathcal{I}$ except for some points in $\tilde{\mathcal{X}}$ =*

$\{x_1, x_2, \ldots\},\ x_k < \tilde{x}$ *for which we can have* $\sigma^*(1|1, x_k) < 1$, *or* $\sigma^*(1|2, x_k) < 1$, *where all* $x_k \in \tilde{\mathcal{X}}$ *satisfy the following condition. There exists* $\epsilon_1 > 0$ *and* ψ *such that*

$$(2\sum_{x=0}^{x_k} \lambda^x v(x))\epsilon_1 + (\sum_{x=0}^{x_k} \lambda^x)\psi = \sum_{x=0}^{x_k-1} \lambda^x v(x), \forall x_k \in \tilde{\mathcal{X}}. \tag{5.18}$$

Proof. *See Appendix D.4.*

The intuitive explanation of the above theorem is as follows. If we define x_0 as $v(x) > 0$ for $x < x_0$ and $v(x) < 0$ for $x > x_0$, then case 1 of the theorem implies a threshold behavior for the admission policy with the threshold being x_0. That is, for x below the threshold, the admission policy favors type 2 of the users and only allows type 1 to enter the queue if type 2 is admitted with probability 1. Similarly, for x above the threshold, type 1 is favored for entering the queue and type 2 is allowed to enter the queue if type 1 is allowed in with probability 1. Notice that users are not interested to enter the queue for $x > x_0$ because $v(x) < 0$. Therefore, type 2 is favored by the admission policy in both $x < x_0$ and $x > x_0$.

Case 2 of the theorem implies that the planner is sending the same signal (except for $x \in \tilde{\mathcal{X}}$) for both types of the users. Therefore, revenue due to the discrimination between the two user types can only be gained for states $x \in \tilde{\mathcal{X}}$. One question that arises is what is the size $|\tilde{\mathcal{X}}|$ of this set. Equation (5.18) indicates that for a given size $|\tilde{\mathcal{X}}|$, there are $|\tilde{\mathcal{X}}|$ equations to be satisfied and only two unknowns (ϵ_1 and ψ). As a result, it is highly unlikely that for a general utility function $v(\cdot)$, the size of the set is larger than 2. Evaluating the quantities x_1 and x_2 can be done systematically by first evaluating $\tilde{x}$ and then searching over all $O(\tilde{x}^2)$ cases for the values of x_1 and x_2 by checking if (5.18) is satisfied for some $\epsilon_1 > 0$ and ψ.

In the next theorem we extend the structural results to the general case of N types of users.

Theorem 8. *Suppose* t_0 *and* $\gamma^*(\cdot,\cdot,\cdot)$ *(or equivalently* $\sigma^*(\cdot|\cdot,\cdot)$*) are the solution of* (5.17)*. Then, for each* i_1 *and* i_2 *where* $i_2 > i_1$*, at least one of the following holds.*

- *Type* i_2 *dominates type* i_1.
- *If* $i_2 - i_1 > \frac{N}{2}$, *then* $q(i_2) = q(i_2 + 1)$, *or* $q(i_1) = q(i_1 - 1)$.
- *If* $i_2 - i_1 \le \frac{N}{2}$, *then* $q(i_2) = q(i_2 + 1)$, *or* $q(i_1) = q(i_1 - 1)$, *or there exists an* $\bar{i}$, *where* $i_1 < \bar{i} \le i_2$, *for which the mechanism utility matches that of the outside option, i.e.,* $\bar{i}q(\bar{i}) - t(\bar{i}) = (\bar{i}\bar{v} - p)^+$.

Proof. *See Appendix D.5.*

The dominance condition in case 1 of Theorem 8 is similar to the first case of Theorem 7, which indicated a threshold behavior for the optimal admission policy. Theorem 8 also states that for $i_2 > i_1$, the admission policy favors type i_2 over type i_1 unless we have two conditions (case two and three of the theorem) where the planner can not favor type i_2 over type i_1. The intuitive reason of why in case two and three the planner can not favor type i_2 over type i_1 is that he either cannot increase allocation of type i_2 since $q(i_2) = q(i_2 + 1)$ or he cannot decrease allocation of type i_1 that can happen due to two reasons. First, we might have $q(i_1) = q(i_1 - 1)$. Second, if the utility of some type, $\bar{i}$, between i_1 and i_2 matches her outside option utility, by decreasing $q(i_1)$, the utility of type $\bar{i}$ will become less than the outside option and it violates the individual rationality constraint.

5.7.2 No information and Full information extremes

In this section, we evaluate two extreme policies of "no information" and "full information" for the mechanism with individual rationality. The "no information" policy refers to a policy that conveys no information about x to the users through the admission signal. This happens when $\sigma(1|x, i) = \sigma(1|x', i)$ for all $x,\ x'$ and all i. Therefore, we can denote the admission policy by $\sigma(\cdot|i)$. One can evaluate the stationary distribution and the tax function described in the previous sections for the no information policy as follows.

$$\mu(x + 1) = \lambda\mu(x) \sum_{i=0}^{n-1} P_I(i)\sigma(1|i) = \lambda\mathbb{P}(S = 1)\mu(x) \tag{5.19a}$$

$$\sum_{x=0}^{\infty} \mu(x) = 1. \tag{5.19b}$$

Therefore, if we denote $\lambda^\sigma = \lambda\mathbb{P}(S = 1)$, the stationary distribution is given by the following equation.

$$\mu(x) = (\lambda^\sigma)^x(1 - \lambda^\sigma). \tag{5.19c}$$

The allocation in this case is given by $q(m) = \sigma(1|m) \sum_{x=0}^{\infty} v(x)\mu(x) = \sigma(1|m)\overline{v}$.

In the next lemma, we present an upper bound on the revenue of the planner in the no information scenario.

Theorem 9. *The revenue of the planner in the no information case is bounded from above by λp.*

Proof. *See Appendix D.6.*

Next, we consider the full information case. First, we notice that the planner can not convey full information about x only through the admission signal, as it is binary and the queue backlog, x, is not. Therefore, in order to analyze the full information scenario, we assume that the planner uses richer signals $s \in \{0, 1, \ldots\}$ and therefore, $\sigma(s|x, i) = \mathbf{1}(s = x)$. Note that the signal s in this scenario is not an admission signal any more. Therefore, the users will make a decision of joining or leaving the queue after hearing the signal. We denote this decision by $e = k(i, x)$, where $e = 1$ means the user enters the queue and $e = 0$ means she leaves.

Since the planner is not discriminating between the types in providing information for them, he can not charge them differently through taxes. Otherwise, every one pretends to be the type that pays the least. Therefore, we have $t(m) = t$. Hence, the message quoting is basically useless in this case. One can easily derive the equation for e to be $e = k(i, x) = \mathbf{1}(v(x) \geq 0)$. We do not include t in this equation because the user has already paid t when she is deciding about e. Therefore, the user with type i receives $iv(x)^+ - t$ through the mechanism.

Theorem 10. *The revenue of the planner in the full information case is less than or equal to* $-\lambda v(x_{neg}) \frac{\lambda^{x_{neg}}}{\sum_{x=0}^{x_{neg}} \lambda^x} + \lambda p$*, where* x_{neg} *is the smallest* x *for which we have* $v(x) < 0$.

Proof. *See Appendix D.7.*

5.8 Numerical Analysis

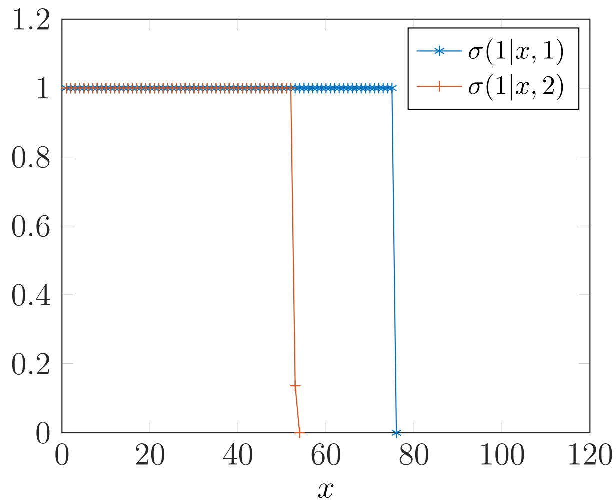
In this section, we present some numerical analysis of the model discussed in this chapter. We have numerically solved the linear optimization problem (5.17) using Matlab. In our analysis, we have set a maximum capacity for the queue such that it does not affect the stationary distribution of the queue backlog.

We consider $N = 2$, $\lambda = 1.2$, and $v(x) = \frac{1}{50^2}(50^2 - x^2)$ as the value function of the users. Fig. 5.3a represents the plot of the admission policy of the planner for the two types of users with respect to the queue backlog for $p = 0$ and the stationary distribution of the queue backlog is plotted is Fig. 5.3b. The revenue of the planner in this case is 0.0786. The admission policy in Fig. 5.3a confirms the results of Theorem 7 and we can see that type 2 dominates type 1.

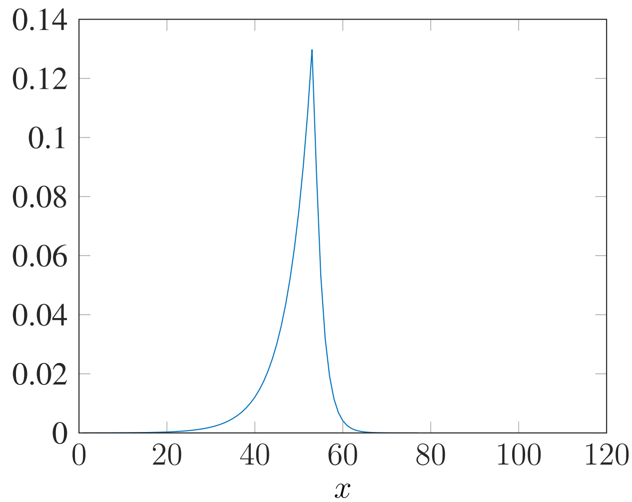
In Fig. 5.4, the plots are represented for $p = 0.2$ and the revenue of the planner is 0.2693. Similar to the $p = 0$ case, the results are consistent with case 1 of Theorem 7.

Fig. 5.5 shows the optimal admission policy for $N = 3$ and we see that for this admission policy, type 3 dominates type 2 and type 2 dominates type 1, which confirms the results of Theorem 8.

In order to have an evaluation of how good the planner is doing in terms of gaining revenue, we can calculate the revenue of the queue from the outside option, i.e., if there was no planner

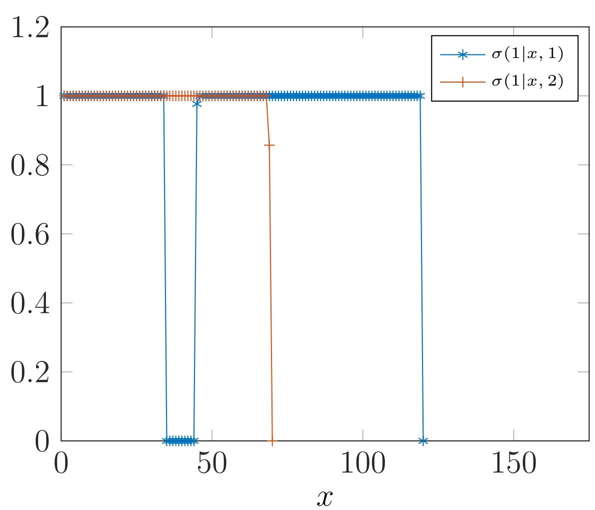


(a) Admission policy

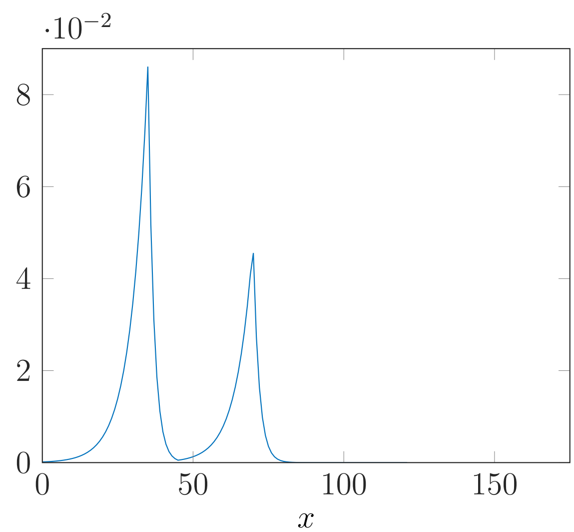


(b) Stationary distribution of the queue backlog

Figure 5.3: Numerical Results for $N = 2$ and $p = 0$

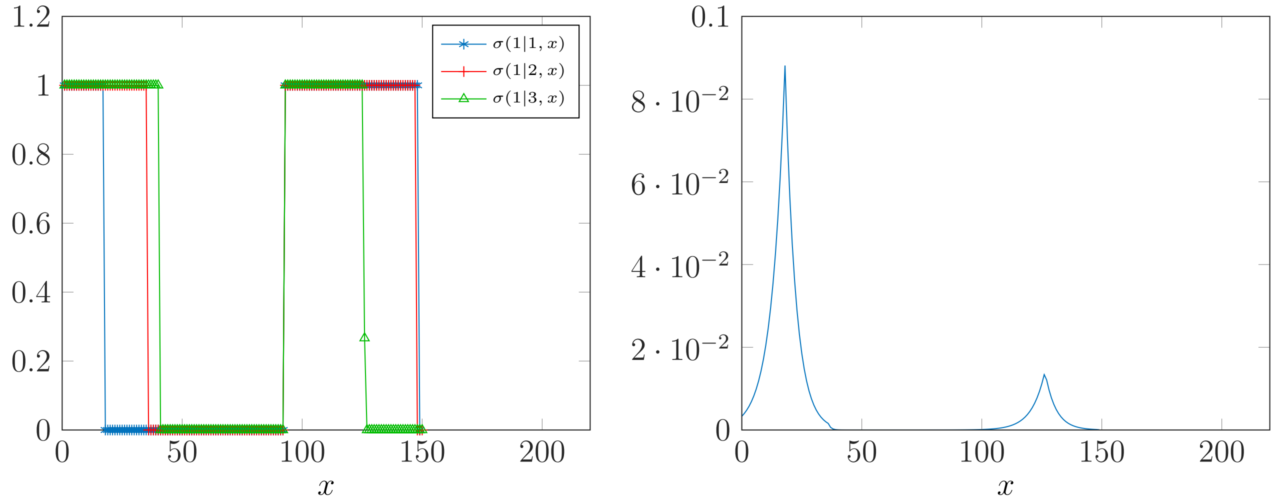


(a) Admission policy



(b) Stationary distribution of the queue backlog

Figure 5.4: Numerical Results for $N = 2$ and $p = 0.2$

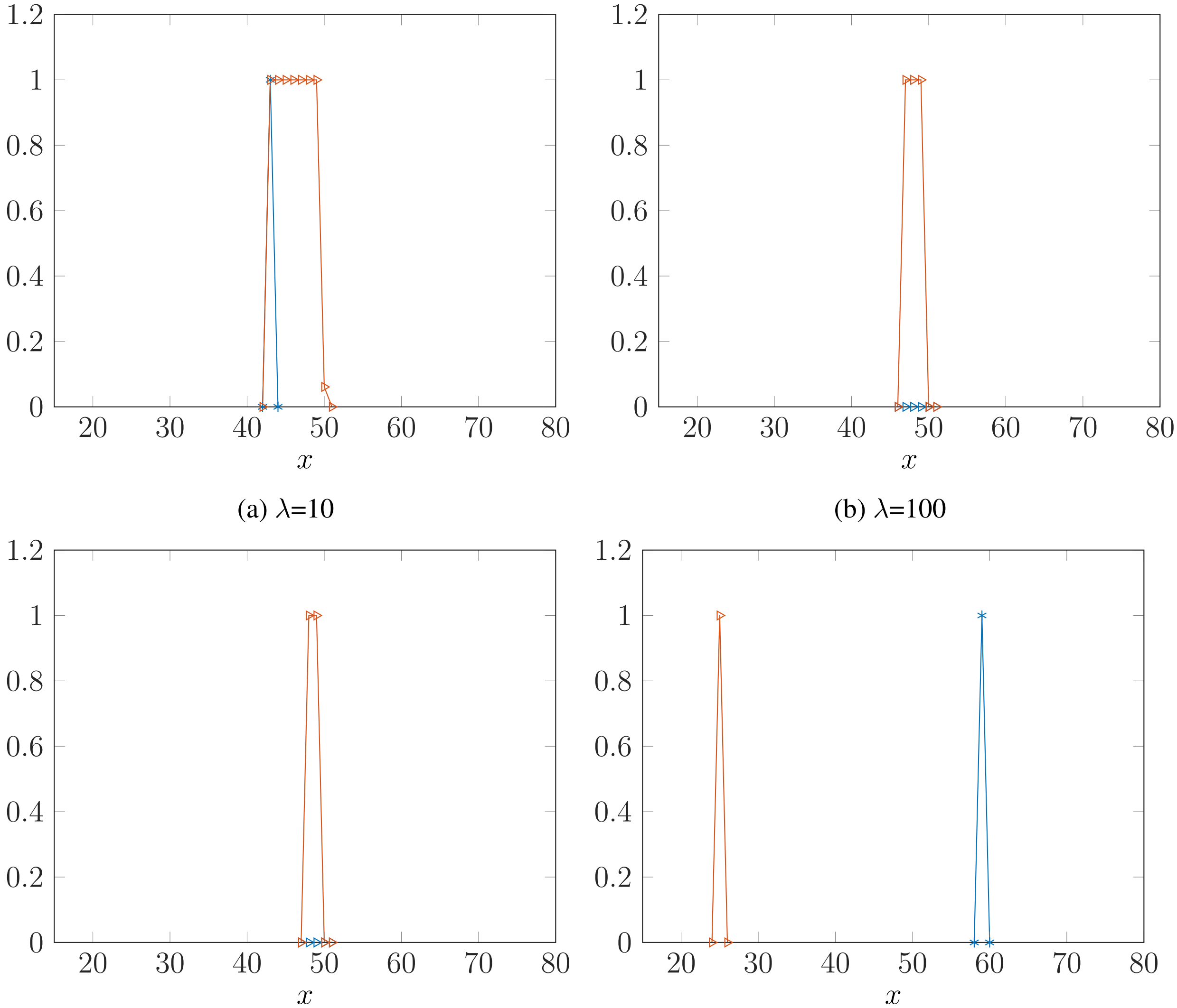


(a) Admission policy (b) Stationary distribution of the queue backlog

Figure 5.5: Numerical Results for $N = 3$ and $p = 0$

and the incoming traffic would choose to join the queue without any information. In the outside option for $N = 2$ and the given $v(\cdot)$, if we assume deterministic decisions for the users, only one type of the users will join the queue and the rate would become $\frac{\lambda}{2}$. Otherwise, the queue becomes unstable. Therefore, the revenue is $\frac{\lambda p}{2}$. Hence, for $p = 0$, the revenue of the outside option is 0 and for $p = 0.2$, the revenue of the outside option is 0.12 and clearly, the planner is doing better than the outside option. We also note that the revenue of the planner in both cases is greater than λp ($\lambda p = 0$ for $p = 0$ and $\lambda p = 0.24$ for $p = 0.2$), which is the upper bound for the revenue of the planner in the no information case. Similarly, the revenue of the planner in both cases is more than the upper bound on the revenue of the full information case, $-\lambda v(x_{neg}) \frac{\lambda^{x_{neg}}}{\sum_{x=0}^{x_{neg}} \lambda^x} + \lambda p$ (0.0081 for $p = 0$ and 0.2481 for $p = 0.2$).

We have also investigated the behavior of the admission policy when the arrival rate of the users increases. We have increased the value of λ to 50000 in four steps of $\lambda \in \{10, 100, 300, 50000\}$. The admission policies for these four values of λ are represented in Fig. 5.6. We observe that by increasing the arrival rate, at first, the interval over which the users are admitted to the queue decreases in length (comparing $\lambda = 1.2$ and $\lambda = 10$). Then, by further increasing the arrival rate, the users of type 1 are never admitted to the queue and the users of type 2 are admitted to the queue over an interval that is decreasing in length as λ increases. We also observe that for an extremely high value of $\lambda = 50000$, the admission policy consists of two single admission points, one for type 1 and one for type 2 of users.



(a) λ=10

(b) λ=100

(c) λ=300

(d) λ=50000

Figure 5.6: The admission policy for different values of the arrival rate and $p = 0$

5.9 Conclusion

In this chapter, we studied a joint information and mechanism design problem for a queuing system with heterogeneous users and an unobservable queue backlog. We investigated how the planner can design tax function and provide different information for different types of the users in order to gain the most revenue. We designed the tax function to ensure dominant strategy incentive compatibility. Through numerical analysis we observed that for most reasonable model parameters, the planner prefers all types of users to participate in his mechanism and we characterized some structural results for the optimal admission policy of the mechanism with individual rationality. Our structural results were also supported by numerical analysis.

CHAPTER 6

Information Design for a Non-atomic Service Scheduling Game

6.1 Introduction

Information asymmetry is inevitable in today's ever-growing systems and networks. Each agent in these systems faces decision makings in the presence of uncertainty toward some states of the world or other agent's information [41, 40]. Having access to as much information as possible enables these agents to make more profitable decisions. Information design [38, 134] studies how sharing information strategically with the agents can steer their actions towards a desirable direction. In the information design framework, there is a sender that possesses some private knowledge about the state of the world. There are possibly multiple receivers of the information. The information that the sender shares with the receivers is shaped in a way to align their objectives with that of the sender as much as possible.

There are different types of information design problems depending on whether there are multiple receivers or a single one, whether the system is dynamic or not, whether the receivers have private information or not, etc. The information design problems with a single receiver are referred to as "Bayesian persuasion" as first introduced in [38], where a geometric method of analyzing the information design problem is proposed. Information design problems with more than one receiver are usually more complex since the solution must induce an equilibrium between the receivers. It is shown in [138] that the set of outcomes in an information design problem with multiple receivers is indeed the set of Bayes-correlated equilibria, BCE, for the receivers. According to the definition of BCE in [138], the information shared with the receivers contains suggestions of what actions they should take. Therefore, an obedience condition has to be imposed on the strategy of the sender to make sure the receivers will follow the suggestions once they hear them. The obedience condition is the same as the conditions that are imposed in the definition of correlated equilibria.

Information design problems study dynamic or static systems. In static information design, the problem that the information designer faces is a static optimization problem [38, 138, 146, 147, 148, 149]. Dynamic information design problems [150, 135, 151, 145, 74, 53] deal with dynamic settings and therefore, the information can be disclosed sequentially. Dynamic programming techniques can therefore be used to characterize the optimal strategy.

An example of dynamic information design can be found in [150], where the receiver is awaiting the occurrence of a random event, e.g., the arrival of an email, so that she can check her email. The receiver is informed of the arrival of the email by a beep that is sent by the sender. The sender wants the receiver not to check her email for as long as possible. Therefore, the sender has to solve a dynamic information design problem to decide whether or not or when to send a beep and reveal the arrival of an email. The problem is solved in continuous time and then a discrete time generalization is presented.

In this chapter, we study an information design problem where there are not only multiple receivers, but they are non-atomic. That is, they form a continuum of population with unit total mass. A service scheduling problem is studied where the service start time is unknown to the agents who want to make decisions of when to join the queue in order to avoid long waits in the queue or not to arrive earlier than the service has started. The service starting time and agents' decisions are in continuous time. There is a planner that knows when the service starts and makes suggestions to the agents about when to join the queue. The suggestion profile has to satisfy the obedience condition. That is, an agent that has received the suggestion of joining at time t must be willing to obey that suggestion. Our model can be considered a dynamic information design problem because the planner makes suggestions for the whole dynamic arrival process of agents. However, each agent only receives one signal from the planner.

Our model of a continuum of agent population arriving at a queue and their cost function closely follows that of [152, 153]. The existence and uniqueness of the equilibrium arrival process is proved in [152] and [153], respectively. In these works, the agents have a preference of when to depart the queue while in our model, this preferred time coincides with the time the service starts and is also the same for all of the agents, although they do not know when that time is. This is where the information design aspect of our model plays its role. Information design for non-atomic agents has also been studied in [154], where a routing game has been considered in which the unknown states of the world affect the latency of the links. The problem has been shown to be a generalized problem of moments and a hierarchy of polynomial optimization is proposed to approximate the solution.

The contributions of this work are as follows. We formulate an information design problem for a service scheduling game consisting of non-atomic agents. We characterize the equilibrium in full information and no information extremes. We show some results on when the planner can do no better than revealing the full information to the agents. We impose some assumptions on our model that will allow us to express the information design problem as a generalized problem of moments (GPM) [55]. We use the computation tools for these problems such as Gloptipoly [56] to numerically solve the information design problem.

The rest of the chapter is structured as follows. Problem formulation is discussed in section 6.2. In section 6.3, the obedience condition is defined and simplified. We study the two extreme cases of full information and no information equilibria in section 6.4. A structural result is stated in section 6.5 for a specific type of arrival processes. We formulate the problem as a GPM in section 6.6 and we present numerical analysis in section 6.7. We conclude in section 6.8. The proofs of the lemmas and theorems can be found in Appendix E.

6.2 Problem Formulation

A service provider starts its service at a fixed rate $\mu \in (0,1)$ starting at some time $\tau \geq 0$ with probability distribution of $f_\tau(\cdot)$. A continuum of agent population of unit total mass needs this service. The action of an agent is the time t to join the service queue. The collection of actions of all the agents can be represented as a probability measure, m, on $\mathbb{R}_{\geq 0}$. Let the set of such measures be denoted as M. We usually refer to the measure m as the arrival process. Note that the support of m can include negative numbers. That is, the arrival times of agents can be a negative number which is due to the fact that the time origin is considered to be when the service can possibly start. For a given $m \in M$ and τ, we denote the queue length at time t by $q_{\tau,m}(t)$, which is given as follows.

$$q_{\tau,m}(t) = \int_{s=-\infty}^{t} m(s)ds - \mu(t-\tau)^+,$$

where $(a)^+ = \max(a, 0)$.

The cost of an agent with action $t \in \mathrm{supp}(m)$ and for a given τ and m is denoted by $c_{\tau,m}(t)$ and is the weighted sum of (i) time to wait in the queue until receiving service; and (ii) the difference between the time of service and realization of τ. Note that (i) includes the time to wait for the service to start in case $t < \tau$. Part (ii) is considered to capture the possibility of service quality deterioration by time. For example, the service quality degrades by time in a food distribution center since the food quality degrades. Part (ii) also captures the fact that agents might be impatient and want to get serviced as soon as possible. Therefore, we have the following cost function.

$$c_{\tau,m}(t) = c_1\left(\frac{q_{\tau,m}(t)}{\mu} + (\tau - t)^+\right) + c_2\left(t + \frac{q_{\tau,m}(t)}{\mu} + (\tau - t)^+ - \tau\right)$$

$$
\begin{aligned}
&\equiv \frac{q_{\tau,m}(t)}{\mu} + c(\tau - t)^+ + (1-c)(t-\tau)^+ \\
&= \frac{q_{\tau,m}(t)}{\mu} + (t-\tau)^+ - c(t-\tau),
\end{aligned} \tag{6.1}
$$

where $c = \frac{c_1}{c_1+c_2} \leq 1$, and c_1 and c_2 are the weights of the two parts of the cost function. Without loss of generality, we can assume c_1 and c_2 are between 0 and 1.

The social cost associated with an arrival process m and a τ is denoted by $s(\tau, m)$ and is defined as the sum of costs of all agents, i.e., $s(\tau, m) := \int_t m(t)c_{\tau,m}(t)\mathrm{d}t$.

The service rate μ and the probability distribution of τ, $f_\tau(\cdot)$, are common knowledge. The agents do not know the exact realization of τ, but there is a planner who does. The planner desires to utilize this information asymmetry to minimize expected social cost over all *obedient* direct signaling strategies. A direct signaling strategy is a map $\pi : \mathbb{R}_{\geq 0} \to \triangle M$, where $\triangle M$ is the set of probability distributions over M. That is, for a realization τ, the planner privately recommends actions to the agents consistent with a measure $m \in M$ sampled from $\pi(.|\tau)$. The obedience condition is defined in the next section. The objective of the planner is to minimize the average value of the social costs, $\overline{s}(\pi)$, which is given below.

$$
\overline{s}(\pi) := \int_{\tau,m} \int_t m(t)c_{\tau,m}(t)f_\tau(\tau)\pi(m|\tau)\mathrm{d}m\mathrm{d}\tau\mathrm{d}t. \tag{6.2a}
$$

Throughout this chapter, we impose different assumptions on the set of arrival processes M, to which the designer restricts his attention. In each section, it will be stated which assumption has been considered. Below is the list of these assumptions.

Assumptions:

(a) $m(t) \leq \mu, \forall t$.

(b) For all m with $\pi(m|\tau) > 0$, $m(t) = 0$ for $t \notin [\underline{t}_\tau, \overline{t}_\tau]$ and some $\underline{t}_\tau$ and $\overline{t}_\tau$ that are increasing with respect to τ.

(c) For all m with $\pi(m|\tau) > 0$, if $q_{\tau,m}(t) = 0$ and $m(s) > 0$ for some $s > t$ and $s < t$, then $m(t) = \mu$.

(d) m is piecewise continuous.

Note that assumption (c) is to make sure that the server works at its full capacity as long as there is yet agents to arrive. As we will see in section 6.4, the full information equilibrium arrival process satisfies all of the above assumptions. Further, the no information equilibrium arrival process that satisfies (d), also satisfies assumptions (a) and (b).

6.3 Obedience Condition

The agent that has received suggestion t, will form her posterior belief on τ and m which can be used to calculate the average cost of taking action s (arriving at time s). We denote this average cost by $\overline{c}_{t,\pi}(s)$. The posterior belief of an agent that has received the suggestion t is given below.

$$\beta(\tau, m|t, \pi) = \frac{f_\tau(\tau)\pi(m|\tau)m(t)}{\int_{\tau,m} f_\tau(\tau)\pi(m|\tau)m(t)}. \tag{6.3}$$

In order to calculate $\overline{c}_{t,\pi}(s)$, we define a quantity $\tilde{\tau}_m(t)$ as follows. For a given arrival process m and each $t \geq 0$, we define $\tilde{\tau}_m(t) \leq t$ as follows.

$$\begin{aligned}
&\forall \tau \leq \tilde{\tau}_m(t), \quad q_{\tau,m}(t) = 0 \\
&\forall \tau > \tilde{\tau}_m(t), \quad q_{\tau,m}(t) > 0.
\end{aligned} \tag{6.4}$$

Note that there might exist a t for which we have $q_{\tau,m}(t) > 0$, for all τ. In this case, we define $\tilde{\tau}_m(t) = 0$. Also, for $t < 0$, we define $\tilde{\tau}_m(t) = 0$.

Throughout the chapter, except for section 6.4, we assume m follows assumption (a). As we will see in section 6.4, both full information and no information equilibria satisfy this assumption.

The average cost $\overline{c}_{t,\pi}(s)$ is given in the following lemma.

Lemma 16. *$\overline{c}_{t,\pi}(s)$ which is the average value of the cost for an agent that has received suggestion t through the signaling strategy π is given as follows.*

$$\begin{aligned}
\overline{c}_{t,\pi}(s) &= \mathbb{E}\{c_{\tau,m}(s)|t,\pi\} \\
&= \frac{1}{\mu\overline{m}(t)} \int_{m,\tau>\tilde{\tau}_m(s)} f_\tau(\tau)\pi(m|\tau)m(t)\Big(\int_{l=-\infty}^{s} m(l)\mathrm{d}l - \mu c s\Big)\mathrm{d}\tau\mathrm{d}m \\
&\quad + \frac{1}{\overline{m}(t)} \int_{m,\tau<\tilde{\tau}_m(s)} f_\tau(\tau)\pi(m|\tau)m(t)((1-c)s - \tau)\mathrm{d}\tau\mathrm{d}m + c\mathbb{E}(\tau|t),
\end{aligned}$$

where $\overline{m}(t) = \int_{\tau,m} f_\tau(\tau)\pi(m|\tau)m(t)\mathrm{d}\tau\mathrm{d}m$.

Proof. *See Appendix E.1.*

As mentioned before, the planner restricts his attention to the set of obedient signaling strategies. The definition of the obedience condition is stated below.

Definition 8 (Obedience Condition)**.** *The signaling strategy π is obedient if all of the agents prefer* *to arrive at the queue at the time they are recommended to do so. That is*

$$t \in \arg\min_{s} \overline{c}_{t,\pi}(s), \quad \forall t.$$

Definition 8 states that t must be a global minimizer of $\overline{c}_{t,\pi}(s)$ for π to be obedient. In the next lemma, we will show that for a signaling strategy π to be obedient, it is necessary and sufficient for t to be a local minimizer of $\overline{c}_{t,\pi}(s)$.

Lemma 17. *The signaling strategy π is obedient if and only if $\frac{\mathrm{d}}{\mathrm{d}s}\overline{c}_{t,\pi}(s)|_t = 0$ for all times t, which implies the following must hold for an obedient signaling strategy.*

$$(1-c)\int_m \int_{\tau=0}^{\tilde{\tau}_m(t)} f_\tau(\tau)\pi(m|\tau)m(t)\mathrm{d}\tau\mathrm{d}m + \frac{1}{\mu}\int_m \int_{\tilde{\tau}_m(t)}^{\infty} f_\tau(\tau)\pi(m|\tau)m(t)(m(t)-\mu c)\mathrm{d}\tau = 0.$$

Proof. *See Appendix E.2.*

6.4 Full Information and No Information Extremes

In this section, we characterize the full information (all agents know the value of τ) and the no information (there is no signal sent to the agents about the value of τ) equilibrium arrival processes.

Theorem 11 (Full Information)**.** *The full information equilibrium arrival process for the service time τ is as follows.*

$$m(t) = \mu c, \quad t \in (\tau - \frac{1-c}{\mu c}, \tau + \frac{1}{\mu}),$$

and $m(t) = 0$ elsewhere.

Proof. *See Appendix E.3.*

Note that the full information equilibrium arrival process induces a single queue throughout the whole time horizon and the queue is cleared out at the same time the arrival process is ended.

Next, we investigate the equilibrium when the agents have no information about τ, other than its prior distribution, $f_\tau(\cdot)$. For this part, we restrict our attention to the set of arrival processes that satisfy assumption (d). Also, we assume $f_\tau(\cdot)$ is an exponential distribution with parameter λ. Note that this assumption is not critical in finding the no information equilibrium and one can search the equilibrium arrival process for a different $f_\tau(\cdot)$. As we will see in the proof of Theorem 12, having a different $f_\tau(\cdot)$ will induce a different differential equation to be solved than the one we solve in this work.

Before stating the equilibrium, we present the following lemma that will enable us to restrict our attention to a smaller set of arrival processes for the no information equilibrium.

Lemma 18. *For the no information equilibrium arrival process, m, we have $m(t) \leq \mu$ for all t and m can not include a delta function.*

Proof. *See Appendix E.4.*

Using Lemma 18, we can characterize the no information equilibrium arrival process.

Theorem 12 (No Information)**.** *The no information equilibrium arrival process, if it exists, is as follows.*

$$m(t) = \mu - \frac{\mu}{\beta - \lambda t}, \quad t \in [t_1, t_2],$$

where $\beta = -\ln(1-c) + \frac{\lambda}{\mu} + 1$, $t_2 = \frac{-\ln(1-c)}{\lambda} + \frac{1}{\mu}$, and t_1 is derived from either of the following equations (or possibly both, which results in two solutions for the equilibrium).

$$\begin{aligned}
&\ln(1-c) + \lambda t_1 + \ln(\frac{\lambda}{\mu} - \ln(1-c) - \lambda t_1 + 1) = 0, \ t_1 \geq 0 \\
&t_1 = \frac{1-c}{\lambda c}\ln(1-c) - \frac{1-c}{\mu c} + \frac{1}{\lambda}, \quad t_1 < 0.
\end{aligned}$$

Proof. *See Appendix E.5.*

6.5 Structural Results

In this section, we assume that the planner restricts her attention to a set of arrival processes that satisfy assumption (b) and for such strategies, we present a structural property in the next theorem.

Theorem 13. *If a signaling strategy $\pi(\cdot|\tau)$ that satisfies assumption (b) is obedient and if we assume $\bar{t}_\tau - \underline{t}_\tau \leq \frac{1}{\mu c}$ and $c \leq 0.5$, then, $\pi(\cdot|\tau)$ is supported only over the full information equilibrium arrival process characterized in Theorem 11.*

Proof. *See Appendix E.6.*

Note that the interval $\frac{1}{\mu c}$ is the time span of the equilibrium arrival process in the full information case. Theorem 13 indicates that if the planner wants to induce a lower social cost than the full information equilibrium social cost, he should expand the time span of the arrival processes to intervals longer than $\frac{1}{\mu c}$.

6.6 GPM Formulation

In this section, we formulate our problem as a generalized problem of moments (GPM). A GPM is an optimization problem over finite probability measures that minimizes a cost that is linear in moments w.r.t. those measures, subject to constraints that are linear w.r.t. those moments. The GPM formulation will allow us to utilize the computation tools available for such problems to do numerical analysis for our model. In order to express our problem as a GPM, we impose two assumptions (b) and (c) on the set of arrival processes.

We define $\underline{\tau}(t)$ and $\overline{\tau}(t)$ to be the inverse of $\underline{t}_\tau$ and $\overline{t}_\tau$, respectively. That is, for $\tau < \underline{\tau}(t)$ or $\tau > \overline{\tau}(t)$, we have $m(t) = 0$ for all m with $\pi(m|\tau) > 0$.

The obedience condition is simplified in the next lemma.

Lemma 19. *A signaling strategy π that satisfies assumptions (b) and (c) is obedient iff the following holds.*

$$\int_{\underline{\tau}(t)}^{\overline{\tau}(t)} f_\tau(\tau) R_{m,\tau}(t,t)\mathrm{d}\tau = \mu c \int_{\underline{\tau}(t)}^{\overline{\tau}(t)} f_\tau(\tau)\overline{m}_\tau(t)\mathrm{d}\tau,$$

where we denote $\overline{m}_\tau(t) = \int_m \pi(m|\tau)m(t)\mathrm{d}m$ and $R_{m,\tau}(t,s) = \int_m \pi(m|\tau)m(t)m(s)\mathrm{d}m$.

Proof. *See Appendix E.7.*

One can easily see that the full information signaling strategy, i.e., $\pi(m|\tau) = \mathbf{1}(m = m_\tau^F)$, where m_τ^F is the full information equilibrium characterized in Theorem 11, satisfies the above obedience constraints.

We can also simplify the planner's objective as follows.

Lemma 20. *The planner's objective is given below if he restricts his attentions to the signalling strategies that satisfy assumptions (b) and (c).*

$$\overline{s}(\pi) = \frac{1}{\mu}\int_\tau f_\tau(\tau)(\int_{t=\underline{t}_\tau}^{\overline{t}_\tau}\int_{s=\underline{t}_\tau}^{t}(R_{m,\tau}(t,s) - \mu c\overline{m}_\tau(t))\mathrm{d}s\mathrm{d}t + \mu c(\tau - \underline{t}_\tau))\mathrm{d}\tau$$

Proof. *See Appendix E.8.*

According to lemmas 19 and 20, the planner's objective is linear in moments of m with respect to the measure $\pi(m|\tau)$. Also, the obedience condition is linear in moments of m. However, m is supported over real numbers and therefore, the measure π is not a finite measure. But for a problem to be a GPM, we must have finite probability measures. In order to have a finite measure, we need to discretize the time and consider a discretized version of the optimization problem, which is a GPM. Therefore, we can numerically solve it using the computation tools available for these types of problems such as Gloptipoly [56]. In the next section, we will present these numerical results.

Next lemma presents a result similar to one presented in Theorem 13 for the signaling strategies that satisfy assumptions (b) and (c).

Theorem 14. *If a signaling strategy $\pi(\cdot|\tau)$ that satisfies assumptions (b) and (c) is obedient and if we assume $\bar{t}_\tau - \underline{t}_\tau \leq \frac{1}{\mu c}$ then $\pi(\cdot|\tau)$ is supported only over the full information equilibrium arrival process characterized in Theorem 11.*

Proof. *See Appendix E.9.*

Note that the result of Theorem 14 holds regardless of the value of c, while in Theorem 13, we must have $c \leq 0.5$ for the result to hold.

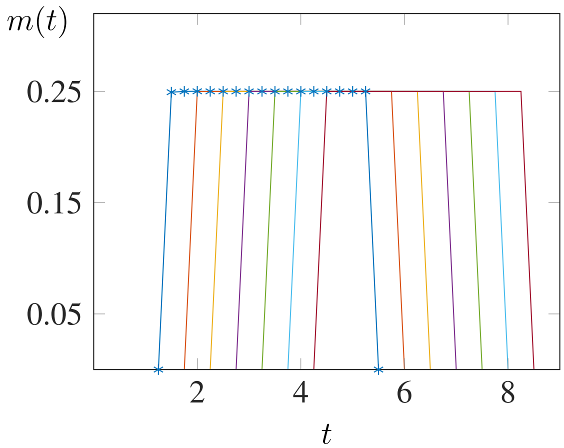
6.7 Numerical Analysis

In this section, based on the GPM formulation of our problem, we use Gloptipoly to solve the problem numerically. In this chapter, we consider uniform discretization of time.

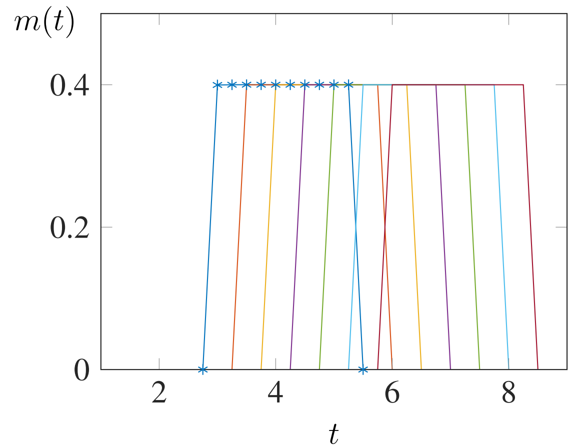
As showed in the previous section, if we restrict our attention to the signaling strategies that satisfy assumptions (b) and (c), and if $\bar{t}_\tau - \underline{t}_\tau \leq \frac{1}{\mu c}$, then the solution is known to have support only on the full information equilibrium of Theorem 11. This result is numerically confirmed as it is shown in Fig. 6.1a and 6.1b for $c = 0.5$ and $c = 0.8$, respectively, and for $\mu = 0.5$ and a bounded discrete interval of $\tau \in \{3, 3.5, 4, 4.5, 5, 5.5, 6\}$ with uniform distribution.

In order to investigate solutions other than the full information equilibrium, we allow the interval of $\bar{t}_\tau - \underline{t}_\tau$ to be longer than $\frac{1}{\mu c}$. We set $\bar{t}_\tau - \underline{t}_\tau = \frac{1}{\mu c} + 0.75$. The optimal signaling strategy for each τ turns out to have support on a singleton arrival process and the different arrival processes corresponding to each τ are represented in Fig. 6.1c and 6.1d for $c = 0.5$ and $c = 0.8$, respectively.

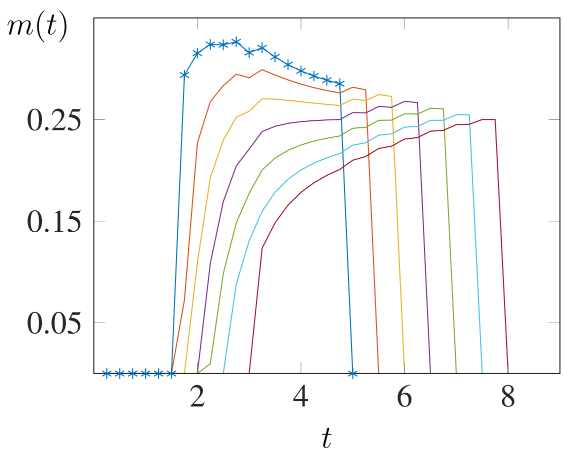
An intuitive explanation about why the solution looks like what we see in Fig. 6.1c, is that the planner decides to put smaller values of τ in higher priority compared to larger values. We can see that the arrival processes associated with smaller values of τ result in smaller social cost. However, they do not satisfy the obedience condition and are indeed far from it. This has been compensated with the arrival processes associated with larger values of τ that result in higher social cost but help with the obedience condition.



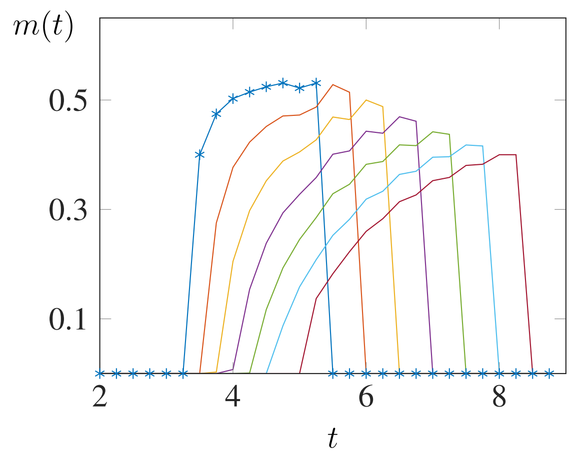
(a) $c = 0.5$ and $\bar{t}_\tau - \underline{t}_\tau = \frac{1}{\mu c}$



(b) $c = 0.8$ and $\bar{t}_\tau - \underline{t}_\tau = \frac{1}{\mu c}$.



(c) $c = 0.5$ and $\bar{t}_\tau - \underline{t}_\tau > \frac{1}{\mu c}$.



(d) $c = 0.8$ and $\bar{t}_\tau - \underline{t}_\tau > \frac{1}{\mu c}$.

Figure 6.1: $m(t)$ for different values of $\tau \in \{3, 3.5, 4, 4.5, 5, 5.5, 6\}$. The stared plots corresponds to $\tau = 3.5$.

6.8 Conclusion

In this chapter, we formulated and studied an information design problem for a non-atomic service scheduling game. We characterized the two extremes of full information and the no information equilibrium and investigated the conditions in which the planner should reveal the full information to the agents. We also formulated the information design problem as a GPM by imposing some assumptions on the model and then numerically solved some examples of the problem.

APPENDIX A

Proofs for Chapter 2

A.1 Proof of Lemma 1

$$\begin{aligned}
\pi_t(\xi_t|v) &= \mathbb{P}(\xi_t|v, a_{1:t-1}) = \frac{\int_{x_{1:t}} \mathbb{P}(x_{1:t}, \xi_t, a_{1:t-1}|v)}{\int_{x_{1:t}} \mathbb{P}(x_{1:t}, a_{1:t-1}|v)} \\
&= \frac{\displaystyle\int_{x_{1:t}} \prod_{i\in\mathcal{N}} \prod_{s=1}^{t-1} Q_X^i(x_s^i|v, a_{s-1})\mathbb{P}(a_s^i|x_{1:s}^i, a_{1:s-1}) Q_X^i(x_t^i|v, a_{t-1})\mathbb{P}(\xi_t^i|x_{1:t}^i, a_{1:t-1})}{\displaystyle\int_{x_{1:t}} \prod_{i\in\mathcal{N}} \prod_{s=1}^{t-1} Q_X^i(x_s^i|v, a_{s-1})\mathbb{P}(a_s^i|x_{1:s}^i, a_{1:s-1}) Q_X^i(x_t^i|v, a_{t-1})} \\
&= \prod_{i\in\mathcal{N}} \frac{\displaystyle\int_{x_{1:t}^i} \prod_{s=1}^{t-1} Q_X^i(x_s^i|v, a_{s-1})\mathbb{P}(a_s^i|x_{1:s}^i, a_{1:s-1}) Q_X^i(x_t^i|v, a_{t-1})\mathbb{P}(\xi_t^i|x_{1:t}^i, a_{1:t-1})}{\displaystyle\int_{x_{1:t}^i} \prod_{s=1}^{t-1} Q_X^i(x_s^i|v, a_{s-1})\mathbb{P}(a_s^i|x_{1:s}^i, a_{1:s-1}) Q_X^i(x_t^i|v, a_{t-1})} \\
&= \prod_{i\in\mathcal{N}} \frac{\mathbb{P}(\xi_t^i, a_{1:t-1}|v)}{\mathbb{P}(a_{1:t-1}|v)} = \prod_{i\in\mathcal{N}} \mathbb{P}(\xi_t^i|v, a_{1:t-1}) = \prod_{i\in\mathcal{N}} \pi_t(\xi_t^i|v).
\end{aligned} \tag{A.1}$$

The second part of the theorem is similarly proved as follows.

$$\begin{aligned}
\mathbb{P}(x_{1:t}|v, a_{1:t-1}) &= \frac{\mathbb{P}(x_{1:t}, a_{1:t-1}|v)}{\mathbb{P}(a_{1:t-1}|v)} \\
&= \frac{\displaystyle\prod_{i\in\mathcal{N}} \prod_{s=1}^{t-1} Q_X^i(x_s^i|v, a_{s-1})\mathbb{P}(a_s^i|x_{1:s}^i, a_{1:s-1}) Q_X^i(x_t^i|v, a_{t-1})}{\displaystyle\int_{x_{1:t}} \prod_{i\in\mathcal{N}} \prod_{s=1}^{t-1} Q_X^i(x_s^i|v, a_{s-1})\mathbb{P}(a_s^i|x_{1:s}^i, a_{1:s-1}) Q_X^i(x_t^i|v, a_{t-1})}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\prod_{i\in\mathcal{N}} \prod_{s=1}^{t-1} Q_X^i(x_s^i|v,a_{s-1})\mathbb{P}(a_s^i|x_{1:s}^i,a_{1:s-1})Q_X^i(x_t^i|v,a_{t-1})}{\prod_{i\in\mathcal{N}} \int_{x_{1:t}^i} \prod_{s=1}^{t-1} Q_X^i(x_s^i|v,a_{s-1})\mathbb{P}(a_s^i|x_{1:s}^i,a_{1:s-1})Q_X^i(x_t^i|v,a_{t-1})} \\
&= \prod_{i\in\mathcal{N}} \frac{\prod_{s=1}^{t-1} Q_X^i(x_s^i|v,a_{s-1})\mathbb{P}(a_s^i|x_{1:s}^i,a_{1:s-1})Q_X^i(x_t^i|v,a_{t-1})}{\int_{x_{1:t}^i} \prod_{s=1}^{t-1} Q_X^i(x_s^i|v,a_{s-1})\mathbb{P}(a_s^i|x_{1:s}^i,a_{1:s-1})Q_X^i(x_t^i|v,a_{t-1})} \\
&= \prod_{i\in\mathcal{N}} \frac{\mathbb{P}(x_{1:t}^i,a_{1:t-1}|v)}{\mathbb{P}(a_{1:t-1}|v)} = \prod_{i\in\mathcal{N}} \mathbb{P}(x_{1:t}^i|a_{1:t-1},v).
\end{aligned} \tag{A.2}
$$

A.2 Proof of Lemma 2

Using Bayes rule we have

$$
\begin{aligned}
\xi_{t+1}^i(v) &= \mathbb{P}(v|x_{1:t+1}^i,a_{1:t}) = \frac{\mathbb{P}(v,x_{t+1}^i,a_t|x_{1:t}^i,a_{1:t-1})}{\mathbb{P}(x_{t+1}^i,a_t|x_{1:t}^i,a_{1:t-1})} \\
&= \frac{\int_{\xi_t^{-i}} \mathbb{P}(v,x_{t+1}^i,a_t,\xi_t^{-i}|x_{1:t}^i,a_{1:t-1})}{\int_{\xi_t^{-i},\tilde{v}} \mathbb{P}(\tilde{v},x_{t+1}^i,a_t,\xi_t^{-i}|x_{1:t}^i,a_{1:t-1})} \\
&= \frac{\int_{\xi_t^{-i}} \mathbb{P}(v|x_{1:t}^i,a_{1:t-1})\mathbb{P}(\xi_t^{-i}|v,a_{1:t-1})\mathbb{P}(a_t|\xi_t^{-i},v,x_{1:t}^i,a_{1:t-1})Q_X^i(x_{t+1}^i|v,a_t)}{\int_{\xi_t^{-i},\tilde{v}} \mathbb{P}(\tilde{v}|x_{1:t}^i,a_{1:t-1})\mathbb{P}(\xi_t^{-i}|\tilde{v},a_{1:t-1})\mathbb{P}(a_t|\xi_t^{-i},\tilde{v},x_{1:t}^i,a_{1:t-1})Q_X^i(x_{t+1}^i|\tilde{v},a_t)} \\
&= \frac{\int_{\xi_t^{-i}} \xi_t^i(v)\pi_t^{-i}(\xi_t^{-i}|v)\prod_{j\in\mathcal{N}}\gamma_t^j(a_t^j|\xi_t^j)Q_X^i(x_{t+1}^i|v,a_t)}{\int_{\xi_t^{-i},\tilde{v}} \xi_t^i(\tilde{v})\pi_t^{-i}(\xi_t^{-i}|\tilde{v})\prod_{j\in\mathcal{N}}\gamma_t^j(a_t^j|\xi_t^j)Q_X^i(x_{t+1}^i|\tilde{v},a_t)} \\
&= \frac{\int_{\xi_t^{-i}} \xi_t^i(v)\prod_{j\in -i}\pi_t^j(\xi_t^j|v)\gamma_t^j(a_t^j|\xi_t^j)Q_X^i(x_{t+1}^i|v,a_t)}{\int_{\xi_t^{-i},\tilde{v}} \xi^i(\tilde{v})\prod_{j\in -i}\pi_t^j(\xi_t^j|\tilde{v})\gamma_t^j(a_t^j|\xi_t^j)Q_X^i(x_{t+1}^i|\tilde{v},a_t)}.
\end{aligned} \tag{A.3}
$$

A.3 Proof of Lemma 3

Using Bayes rule we have

$$
\begin{aligned}
\pi^i_{t+1}(\xi^i_{t+1}|v) &= \mathbb{P}(\xi^i_{t+1}|v, a_{1:t}) \\
&= \frac{\int_{\xi_t, x^i_{t+1}} \mathbb{P}(\xi^i_{t+1}, \xi_t, x^i_{t+1}, a_t|v, a_{1:t-1})}{\int_{\xi_t} \mathbb{P}(\xi_t, a_t|v, a_{1:t-1})} \\
&= \frac{\displaystyle\int_{\xi_t, x^i_{t+1}} \mathbb{P}(\xi_t|v, a_{1:t-1})\mathbb{P}(a_t|\xi_t, a_{1:t-1})\mathbb{P}(x^i_{t+1}|v, a_t)\delta(\xi^i_{t+1} - F^i[\xi^i_t, \pi^{-i}_t, \gamma^{-i}_t, a_t, x^i_{t+1}])}{\displaystyle\int_{\xi_t} \mathbb{P}(\xi_t|v, a_{1:t-1})\mathbb{P}(a_t|\xi_t, a_{1:t-1})} \\
&= \frac{\displaystyle\int_{\xi_t, x^i_{t+1}} \prod_{j\in\mathcal{N}} \pi^j_t(\xi^j_t|v)\gamma^j_t(a^j_t|\xi^j_t) Q^i_X(x^i_{t+1}|v, a_t)\delta(\xi^i_{t+1} - F^i[\xi^i_t, \pi^{-i}_t, \gamma^{-i}_t, a_t, x^i_{t+1}])}{\displaystyle\int_{\xi_t} \prod_{j\in\mathcal{N}} \pi^j_t(\xi^j_t|v)\gamma^j_t(a^j_t|\xi^j_t)} \\
&= \frac{\displaystyle\prod_{j\neq i}\int_{\xi^j_t} \pi^j_t(\xi^j_t|v)\gamma^j_t(a^j_t|\xi^j_t) \int_{\xi^i_t, x^i_{t+1}} \pi^i_t(\xi^i_t|v)\gamma^i_t(a^i_t|\xi^i_t) Q^i_X(x^i_{t+1}|v, a_t)\delta(\xi^i_{t+1} - F^i[\xi^i_t, \pi^{-i}_t, \gamma^{-i}_t, a_t, x^i_{t+1}])}{\displaystyle\prod_{j\in\mathcal{N}}\int_{\xi^j_t} \pi^j_t(\xi^j_t|v)\gamma^j_t(a^j_t|\xi^j_t)} \\
&= \frac{\displaystyle\int_{\xi^i_t, x^i_{t+1}} \pi^i_t(\xi^i_t|v)\gamma^i_t(a^i_t|\xi^i_t) Q^i_X(x^i_{t+1}|v, a_t)\delta(\xi^i_{t+1} - F^i[\xi^i_t, \pi^{-i}_t, \gamma^{-i}_t, a_t, x^i_{t+1}])}{\displaystyle\int_{\xi^i_t} \pi^i_t(\xi^i_t|v)\gamma^i_t(a^i_t|\xi^i_t)}.
\end{aligned}
\tag{A.4}
$$

A.4 Proof of Theorem 1

To prove the theorem, we show that if every player in $-i$ plays according to strategy $\gamma^{*,-i}_t = \theta^{-i}_t[\pi_t]$, the best response of player i is of the form $\gamma^{*,i}_t = \theta^i_t[\pi_t]$ and it is derived from the given fixed point equation. We show that if we fix the update rule of π_t to $\pi_{t+1} = F_\pi[\pi_t, \gamma^*_t, a_t] = F_\pi[\pi_t, \theta_t[\pi_t], a_t]$ and assume that player i is forced to use these beliefs as her true beliefs, then she faces an MDP with state (π_t, ξ_t^i), action a_t^i and instantaneous reward

$$\widehat{r}_t^i(\pi_t, \xi_t^i, a_t^i) = \mathbb{E}\left[r_t^i(V, A_t)|\pi_t, \xi_t^i, a_t^i\right].$$

We first need to prove that the state (π_t, ξ_t^i) evolves according to a controlled Markov process. Indeed,

$$\begin{aligned}
\mathbb{P}(\pi_{t+1}, \xi_{t+1}^i | \pi_{1:t}, \xi_{1:t}^i, a_{1:t}^i) &= \int_{v, \xi_t^{-i}, a_t^{-i}, x_{t+1}^i} \pi_t^{-i}(\xi_t^{-i}|v)\xi_t^i(v)\theta_t^{-i}[\pi_t](a_t^{-i}|\xi_t^{-i})Q(x_{t+1}^i|v, a_t) \\
&\qquad \delta(\pi_{t+1} - F_\pi[\pi_t, \theta_t[\pi_t], a_t])\delta(\xi_{t+1}^i - F^i[\xi_t^i, \pi_t^{-i}, \theta_t^{-i}[\pi_t], a_t, x_{t+1}^i]) \\
&= \mathbb{P}(\pi_{t+1}, \xi_{t+1}^i | \pi_t, \xi_t^i, a_t^i). \qquad \text{(A.5)}
\end{aligned}$$

The average instantaneous reward can now be written as $\mathbb{E}[r_t^i(V, A_t)] = \mathbb{E}[\mathbb{E}[r_t^i(V, A_t)|\Pi_t, \Xi_t^i, A_t^i]]$, where

$$\begin{aligned}
\mathbb{E}[r^i(V, A_t)|\pi_t, \xi_t^i, a_t^i] &= \int_{v, a_t^{-i}} r^i(v, a_t) \int_{\xi_t^{-i}} \mathbb{P}(v, a_t^{-i}, \xi_t^{-i} | \pi_t, \xi_t^i, a_t^i) \\
&= \int_{v, a_t^{-i}} r^i(v, a_t) \int_{\xi_t^{-i}} \theta_t^{-i}[\pi_t](a_t^{-i}|\xi_t^{-i})\pi_t^{-i}(\xi_t^{-i}|v)\xi_t^i(v) \\
&= \int_{v, a_t^{-i}} r^i(v, a_t) \int_{\xi_t^{-i}} \theta_t^{-i}[\pi_t](a_t^{-i}|\xi_t^{-i})\pi_t^{-i}(\xi_t^{-i}|v)\xi_t^i(v) \\
&=: \widehat{r}_t^i(\pi_t, \xi_t^i, a_t^i). \qquad \text{(A.6)}
\end{aligned}$$

Based on the above, it is now clear that user i faces an MDP and her best response strategy is the solution of the following backward dynamic program. We have $A_t^i \sim \gamma^{*,i}(\cdot|\xi_t^i)$, where

$$Supp(\gamma_t^{*,i}(\cdot|\xi_t^i)) \subset \arg\max_a \mathbb{E}[\widehat{r}_t^i(\pi_t, \xi_t^i, a) + J_{t+1}^i(\Pi_{t+1}, \Xi_{t+1}^i)|\pi_t, \xi_t^i, a] \qquad \text{(A.7a)}$$

$$J_t^i(\pi_t, \xi_t^i) = \max_a \mathbb{E}[\widehat{r}_t^i(\pi_t, \xi_t^i, a) + J_{t+1}^i(\Pi_{t+1}, \Xi_{t+1}^i)|\pi_t, \xi_t^i, a] \qquad \text{(A.7b)}$$

where expectation is w.r.t. γ_t^i and the conditional distribution in (A.5). Consequently the best response of user i is of the form $A_t^{*,i} \sim \psi_t^i(\cdot|\xi_t^i, \pi_t)$. Note that in the standard MDP formulation, it is sufficient to only consider the pure strategies. However, in equation (A.7), we see randomized strategies. The reason of this modification is that in a PBE, the beliefs have to be consistent with the equilibrium strategies and we need $\psi_t^i(\cdot|\xi_t^i, \pi_t) = \gamma_t^{*,i}(\cdot|\xi_t^i) = \theta_t^i[\pi_t](\cdot|\xi_t^i)$. Hence, the best responses satisfy the following fixed point equation at each time t. For all i and all ξ_t^i we have

$$\gamma^{*,i}(\cdot|\xi_t^i) \in \arg\max_{\gamma^i(\cdot|\xi_t^i)} \mathbb{E}[\widehat{r}_t^i(\pi_t, \xi_t^i, A_t^i) + J_{t+1}^i(F_\pi(\pi_t, \gamma_t^*, A_t), F^i(\xi_t^i, \pi_t, \gamma_t^{*,-i}, A_t, X_{t+1}^i)))|\pi_t, \xi_t^i], \tag{A.8}$$

where expectation is w.r.t. the distribution

$$\mathbb{P}(a_t, x_{t+1}^i|\pi_t, \xi_t^i) = \int_{\xi_t^{-i}, v} \gamma_t^i(a_t^i|\xi_t^i)\gamma_t^{*,-i}(a_t^{-i}|\xi_t^{-i})\pi_t^{-i}(\xi_t^{-i}|v)\xi_t^i(v)Q_X^i(x_{t+1}^i|v, a_t). \tag{A.9}$$

The above fixed point might not have a solution in pure strategies and therefore, we had to consider randomized strategies in equation (A.7).

A.5 Proof of Theorem 2

Throughout this proof, the submatrices that are not explicitly specified are all zero matrices with appropriate dimensions.

In order to prove the theorem we will define a dynamical system from the viewpoint of a specific user i and show inductively that it is a Gauss Markov model. Gaussianity of both private and conditional public beliefs follows from KF-type arguments.

For each player $i \in \mathcal{N}$, we define an unobserved state vector as

$$s_t^i = \left[\begin{array}{cc} v\ ; & \widehat{v}_{t-1}^{-i} \end{array}\right], \tag{A.10a}$$

and an observation vector

$$y_t^i = \left[\begin{array}{cc} a_{t-1}^{-i} - m_{t-1}^{-i}\ ; & x_t^i \end{array}\right]. \tag{A.10b}$$

We will show that the random vector s_t^i evolves according to a linear Gaussian process,

$$s_{t+1}^i = \mathbf{A}_t^i s_t^i + \begin{bmatrix} \mathbf{0} \\ \mathbf{D}_t^i \end{bmatrix} a_{t-1}^i + \begin{bmatrix} \mathbf{0} \\ \mathbf{H}_t^i \end{bmatrix} w_t^{-i} + \begin{bmatrix} 0 \\ d_t^i \end{bmatrix} \tag{A.11a}$$

$$y_t^i = \mathbf{C}_t^i s_t^i + \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} w_t^i, \tag{A.11b}$$

where

$$\mathbf{A}_t^i = \left[\begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline \multicolumn{2}{c}{\mathbf{G}_t^{-i}} \end{array} \right] \tag{A.11c}$$

Note that $(y_{1:t}^i, a_{1:t-1}^i)$ is a shifted version of h_t^i. We prove the validity of (A.11) and the claim of the theorem using induction. In particular, Lemma 21 below is the induction basis and the subsequent Lemma 22 is the induction step. This concludes the proof of the theorem.

Lemma 21. *The following are true.*
(a) ξ_1^i is Gaussian $N(\widehat{v}_1^i, \mathbf{\Sigma}_1^i)$, with $\widehat{v}_1^i = \mathbf{\Sigma}(\mathbf{\Sigma} + \mathbf{Q}^i)^{-1} x_1^i$ and $\mathbf{\Sigma}_1^i = \mathbf{\Sigma} - \mathbf{\Sigma}(\mathbf{\Sigma} + \mathbf{Q}^i)^{-1}\mathbf{\Sigma}$. Consequently the public belief $\pi_1^i(\xi_1^i|v)$ reduces to $\pi_1^i(\widehat{v}_1^i|v)$.
(b) (A.11) *holds for $t = 1$.*
(c) The public belief $\pi_1^i(\widehat{v}_1^i|v)$ is Gaussian with mean $\mathbb{E}[\widehat{V}_1^i|v] = \mathbf{E}_1^i v + f_1^i$, with $\mathbf{E}_1^i = \mathbf{\Sigma}(\mathbf{\Sigma} + \mathbf{Q}^i)^{-1}$, $f_1^i = 0$, and covariance matrix $\mathbf{\Sigma}(\mathbf{\Sigma} + \mathbf{Q}^i)^{-1}\mathbf{Q}^i(\mathbf{\Sigma} + \mathbf{Q}^i)^{-1}\mathbf{\Sigma}$.

Proof. *(a) We have $x_1^i = v + w_1^i$ and $\xi_1^i(v) = \mathbb{P}(v|x_1^i)$, so due to joint Gaussianity of V and X_1^i we have that ξ_1^i is $N(\widehat{v}_1^i, \mathbf{\Sigma}_1^i)$, with mean*

$$\begin{aligned} \widehat{v}_1^i &= \mathbb{E}[V|x_1^i] \\ &= \mathbb{E}[V] + \mathbb{E}[V X_1^{i\prime}]\mathbb{E}[X_1^i X_1^{i\prime}]^{-1}(x_1^i - \mathbb{E}[X_1^i]) \\ &= \mathbf{\Sigma}(\mathbf{\Sigma} + \mathbf{Q}^i)^{-1} x_1^i, \end{aligned} \tag{A.12}$$

and covariance matrix

$$\mathbf{\Sigma}_1^i = \mathbf{\Sigma} - \mathbf{\Sigma}(\mathbf{\Sigma} + \mathbf{Q}^i)^{-1}\mathbf{\Sigma}. \tag{A.13}$$

As a result the only private information of user i relevant to other users is $\widehat{v}_1^i$ and the public belief $\pi_1^i(\xi_1^i|v)$ can be reduced to $\pi_1^i(\widehat{v}_1^i|v)$.
(b) We have $s_1^i = \begin{bmatrix} v \,; & 0 \end{bmatrix}$ and $s_2^i = \begin{bmatrix} v \,; & \widehat{v_1^{-i}} \end{bmatrix}$. The first row of (A.11a) *is evidently true. For the second row, using the result (from part (a)) $\widehat{v}_1^j = \mathbf{\Sigma}(\mathbf{\Sigma} + \mathbf{Q}^j)^{-1}(v + w_1^j)$, we can derive $\mathbf{G}_1^{-i}$, $\mathbf{H}_1^i$, $\mathbf{D}_1^i$ and d_1^i as*

$$\mathbf{G}_1^{-i} = \begin{bmatrix} \mathbf{\Sigma}(\mathbf{\Sigma} + \mathbf{Q}^{-i})^{-1} & \mathbf{0} \end{bmatrix} \tag{A.14a}$$

$$\mathbf{H}_1^i = \mathfrak{D}(\mathbf{\Sigma}(\mathbf{\Sigma} + \mathbf{Q}^{-i})^{-1}) \tag{A.14b}$$

$$\mathbf{D}_1^i = \mathbf{0} \tag{A.14c}$$

$$d_1^i = \mathbf{0}, \tag{A.14d}$$

where $\mathbf{\Sigma}(\mathbf{\Sigma}+\mathbf{Q}^{-i})^{-1}$ *is the vertical concatenation of the matrices* $\mathbf{\Sigma}(\mathbf{\Sigma}+\mathbf{Q}^{j})^{-1}$ *for* $j \in -i$.
(c) Since $\widehat{v}_1^i = \mathbf{\Sigma}(\mathbf{\Sigma}+\mathbf{Q}^{i})^{-1}(v + w_1^i)$ *we deduce that* $\pi_1^i(\widehat{v}_1^i|v)$ *is Gaussian with mean* $\mathbb{E}[\widehat{V}_1^i|v] = \mathbf{\Sigma}(\mathbf{\Sigma}+\mathbf{Q}^{i})^{-1}v$ *and covariance matrix* $\tilde{\mathbf{\Sigma}}_1^i = \mathbf{\Sigma}(\mathbf{\Sigma}+\mathbf{Q}^{i})^{-1}\mathbf{Q}^i(\mathbf{\Sigma}+\mathbf{Q}^{i})^{-1}\mathbf{\Sigma}$.

Lemma 22. *Assuming pure linear strategies of the form* $\gamma_t^j(a_t^j|\widehat{v}_t^j) = \delta(a_t^j - \mathbf{L}_t^j\widehat{v}_t^j - m_t^j)$ *for all* $j \in \mathcal{N}$, *and assuming that* (A.11) *holds for* $t \le k$ *and* $\mathbb{E}[\widehat{V}_k^j|v, a_{1:k-1}] = \mathbf{E}_k^j v + f_k^j$, *the following are true.*
(a) ξ_{k+1}^i *is* $N(\widehat{v}_{k+1}^i, \mathbf{\Sigma}_{k+1}^i)$ *with*

$$\widehat{v}_{k+1}^i = \mathbf{G}_{k+1}^{i,i}\begin{bmatrix} \widehat{v}_k^i \\ x_{k+1}^i \end{bmatrix} + d_{k+1}^{i,i}, \tag{A.15}$$

where $\mathbf{G}_{k+1}^{i,i}$, $d_{k+1}^{i,i}$ *and* $\mathbf{\Sigma}_{k+1}^i$ *can be publicly evaluated. Consequently, the public belief* $\pi_{k+1}^i(\xi_{k+1}^i|v)$ *can be reduced to a belief* $\pi_{k+1}^i(\widehat{v}_{k+1}^i|v)$.
(b) (A.11) *holds for* $t = k+1$.
(c) The conditional public belief, $\pi_{k+1}^i(\widehat{v}_{k+1}^i|v)$, *are Gaussian with mean* $\mathbb{E}[\widehat{V}_{k+1}^i|V, a_{1:k}] = \mathbf{E}_{k+1}^i V + f_{k+1}^i$ *and covariance matrix* $\tilde{\mathbf{\Sigma}}_{k+1}^i$, *where matrices* $\mathbf{E}_{k+1}^i$ *and* $\tilde{\mathbf{\Sigma}}_{k+1}^i$ *and vector* f_{k+1}^i *can be publicly evaluated.*

Proof. *(a) We first show one important result from the lemma assumptions. Notice that due to conditional independence of* x_k^j*'s given* v *across time and players, and since* $\widehat{v}_k^j$ *is a function of* $x_{1:k}^j$ *and* $a_{1:k-1}$, *we have*

$$\begin{aligned}
\tilde{v}_k^{i,j} &= \mathbb{E}[\widehat{V}_k^j|x_{1:k}^i, a_{1:k-1}] \\
&= \mathbb{E}_V[\mathbb{E}[\widehat{V}_k^j|V, x_{1:k}^i, a_{1:k-1}]|x_{1:k}^i, a_{1:k-1}] \\
&= \mathbb{E}_V[\mathbb{E}[\widehat{V}_k^j|V, a_{1:k-1}]|x_{1:k}^i, a_{1:k-1}] \\
&= \mathbb{E}_V[\mathbf{E}_k^j V + f_k^j|x_{1:k}^i, a_{1:k-1}] \\
&= \mathbf{E}_k^j\mathbb{E}[V|x_{1:k}^i, a_{1:k-1}] + f_k^j \\
&= \mathbf{E}_k^j\widehat{v}_k^i + f_k^j.
\end{aligned} \tag{A.16}$$

By using the assumption that (A.11) *holds for* $t = k$, *we form a linear Gaussian model with partial observations and use Kalman filter results [67, Ch.7]. Consider equation* (A.11) *for* $t = k$. *By using standard Kalman filter results [67, Ch.7], we know that the belief over the system states given the observations is Gaussian and therefore, the private belief* ξ_k^i *is* $N(\widehat{v}_k^i, \mathbf{\Sigma}_k^i)$. *We denote*

$\mathbb{E}[S^i_{k+1}|y^i_{1:k+1}, a^i_{1:k}]$ and $\mathbb{E}[S^i_{k+1}|y^i_{1:k}, a^i_{1:k-1}]$ by $s^i_{k+1|k+1}$ and $s^i_{k+1|k}$, respectively. We have

$$\begin{aligned} s^i_{k+1|k+1} &= \mathbb{E}[S^i_{k+1}|x^i_{1:k+1}, a_{1:k}] \\ &= \begin{bmatrix} \widehat{v}^i_{k+1} \\ \mathbb{E}[\widehat{V}^{-i}_k|x^i_{1:k+1}, a_{1:k}] \end{bmatrix} \\ &= \mathbf{A}^i_k s^i_{k|k} + \begin{bmatrix} \mathbf{0} \\ \mathbf{D}^i_k \end{bmatrix} a^i_{k-1} + \mathbf{J}^i_{k+1}(y^i_{k+1} - \mathbf{C}^i_{k+1} s^i_{k+1|k}) + \begin{bmatrix} 0 \\ d^i_k \end{bmatrix}. \end{aligned} \tag{A.17}$$

Therefore,

$$\begin{aligned} \widehat{v}^i_{k+1} &= \widehat{v}^i_k + (\mathbf{J}^i_{k+1})_{\widehat{v}^i,:}(y^i_{k+1} - \mathbf{C}^i_{k+1} s^i_{k+1|k}) \\ &= \widehat{v}^i_k + (\mathbf{J}^i_{k+1})_{\widehat{v}^i,:} \begin{bmatrix} a^{-i}_k - m^{-i}_k - \mathfrak{D}(\mathbf{L}^{-i}_k)\tilde{v}^{i,-i}_k \\ x^i_{k+1} - \widehat{v}^i_k \end{bmatrix} \\ &= \widehat{v}^i_k + (\mathbf{J}^i_{k+1})_{\widehat{v}^i,:} \begin{bmatrix} -\mathfrak{D}(\mathbf{L}^{-i}_k)\mathbf{E}^{-i}_k \widehat{v}^i_k \\ x^i_{k+1} - \widehat{v}^i_k \end{bmatrix} + (\mathbf{J}^i_{k+1})_{\widehat{v}^i,a^{-i}}(a^{-i}_k - m^{-i}_k - \mathfrak{D}(\mathbf{L}^{-i}_k) f^{-i}_k) \\ &= \mathbf{G}^{i,i}_{k+1} \begin{bmatrix} \widehat{v}^i_k \\ x^i_{k+1} \end{bmatrix} + d^{i,i}_{k+1}, \end{aligned} \tag{A.18}$$

where

$$(\mathbf{G}^{i,i}_{k+1})_{:,x^i} = (\mathbf{J}^i_{k+1})_{\widehat{v}^i,x^i} \tag{A.19a}$$

$$(\mathbf{G}^{i,i}_{k+1})_{:,\widehat{v}^i} = \mathbf{I} - (\mathbf{J}^i_{k+1})_{\widehat{v}^i,a^{-i}}\mathfrak{D}(\mathbf{L}^{-i}_k)\mathbf{E}^{-i}_k - (\mathbf{J}^i_{k+1})_{\widehat{v}^i,x^i} \tag{A.19b}$$

$$d^{i,i}_{k+1} = (\mathbf{J}^i_{k+1})_{\widehat{v}^i,a^{-i}}(a^{-i}_k - m^{-i}_k - \mathfrak{D}(\mathbf{L}^{-i}_k) f^{-i}_k). \tag{A.19c}$$

The matrix $\mathbf{J}^i_{k+1}$ and the covariance matrix of S^i_{k+1} conditioned on $y^i_{1:k+1}$ and $y^i_{1:k}$, denoted by $\mathbf{\Sigma}^i_{k+1|k+1}$ and $\mathbf{\Sigma}^i_{k+1|k}$, respectively, can be derived from the standard Kalman filter equations as follows

$$\mathbf{\Sigma}^i_{k+1|k} = \mathbf{A}^i_k \mathbf{\Sigma}^i_{k|k} \mathbf{A}^{i\prime}_k + \begin{bmatrix} \mathbf{0} \\ \mathbf{H}^i_k \end{bmatrix} \mathfrak{D}(\mathbf{Q}^{-i}) \begin{bmatrix} \mathbf{0} \\ \mathbf{H}^i_k \end{bmatrix}' \tag{A.20a}$$

$$\mathbf{J}^i_{k+1} = \mathbf{\Sigma}^i_{k+1|k} \mathbf{C}^{i\prime}_{k+1} (\mathbf{C}^i_{k+1} \mathbf{\Sigma}^i_{k+1|k} \mathbf{C}^{i\prime}_{k+1} + \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \mathbf{Q}^i \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix}')^{-1} \tag{A.20b}$$

$$\mathbf{\Sigma}^i_{k+1|k+1} = (\mathbf{I} - \mathbf{J}^i_{k+1} \mathbf{C}^i_{k+1}) \mathbf{\Sigma}^i_{k+1|k} \tag{A.20c}$$

$$\mathbf{\Sigma}^i_{1|1} = \mathbb{E}[S^i_1 S^{i\prime}_1] - \mathbb{E}[S^i_1 X^{i\prime}_1](\mathbb{E}[X^i_1 X^{i\prime}_1])^{-1}\mathbb{E}[S^i_1 X^{i\prime}_1]'$$

$$
= \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{bmatrix} (\mathbf{\Sigma} + \mathbf{Q}^i)^{-1} \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \end{bmatrix}
$$

$$
= \begin{bmatrix} \mathbf{\Sigma} - \mathbf{\Sigma}(\mathbf{\Sigma} + \mathbf{Q}^i)^{-1}\mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}. \tag{A.20d}
$$

Note that, for notational simplicity, we remove the time subscripts from submatrix notation, so that $(\mathbf{J}^i_{k+1})_{\widehat{v}^i,x^i}$ *denotes* $(\mathbf{J}^i_{k+1})_{\widehat{v}^i_{k+1},x^i_k}$.

Finally, we have $\mathbf{\Sigma}^i_t = (\mathbf{\Sigma}^i_{t+1|t})_{v,v}$. *Unlike* $\widehat{v}^i_t$, *which is part of the private information of player* i, *the matrix* $\mathbf{\Sigma}^i_t$ *is a public quantity due to the independence of equation* (A.20c) *to the private observations of player* i.

(b) Equation (A.11a) *is obvious for the first part of the state,* v. *In order to prove the other parts of equation* (A.11a) *for* $t = k+1$, *we consider the dynamic system* (A.11) *for each of the players* $-i$ *for* $t = k$ *and we write* (A.18) *for players* $-i$. *Since* x^{-i}_{k+1} *is not part of* y^i_{k+1}, *we can substitute it by* $v + w^{-i}_{k+1}$ *and derive* $\mathbf{G}^j_{k+1}$, $\mathbf{D}^i_{k+1}$, $\mathbf{H}^i_{k+1}$, *and* d^i_{k+1} *for all* $j \in -i$ *as*

$$
(\mathbf{G}^j_{k+1})_{:,v} = (\mathbf{J}^j_{k+1})_{\widehat{v}^j,x^j} \tag{A.21a}
$$

$$
(\mathbf{G}^j_{k+1})_{:,\widehat{v}^j} = \mathbf{I} - (\mathbf{J}^j_{k+1})_{\widehat{v}^j,a^{-j}} \mathfrak{D}(\mathbf{L}^{-j}_k)\mathbf{E}^{-j}_k - (\mathbf{J}^j_{k+1})_{\widehat{v}^j,x^j} \tag{A.21b}
$$

$$
(\mathbf{D}^i_{k+1})_{\widehat{v}^j,:} = (\mathbf{J}^j_{k+1})_{\widehat{v}^j,a^i} \tag{A.21c}
$$

$$
(d^i_{k+1})_{\widehat{v}^j} = (\mathbf{J}^j_{k+1})_{\widehat{v}^j,a^{-ij}}(a^{-ij}_k - m^{-ij}_k - \mathfrak{D}(\mathbf{L}^{-ij}_k) f^{-ij}_k) + (\mathbf{J}^j_{k+1})_{\widehat{v}^j,a^i}(-m^i_k - \mathbf{L}^i_k f^i_k) \tag{A.21d}
$$

$$
\mathbf{H}^i_{k+1} = \mathfrak{D}((\mathbf{J}^{-i}_{k+1})_{\widehat{v}^{-i},x^{-i}}). \tag{A.21e}
$$

The notation $-ij$ *means all of the players except* i *and* j. *We have derived the matrices* $\mathbf{A}^i_{k+1}$, $\mathbf{D}^i_{k+1}$, $\mathbf{H}^i_{k+1}$, *and vector* d^i_{k+1} *and so* (A.11) *holds for* $t = k+1$.

(c) In order to show that the conditional public belief $\pi^i_{k+1}(\widehat{v}^i_{k+1}|v)$ *is Gaussian, we consider a conditional Gauss Markov model. Note that the conditional public belief is publicly measurable conditioned on* V. *We use this fact to form a conditional model, where the observations are the conditions in the conditional public belief and we derive conditional Kalman filters. Using* (A.11) *for* $t \leq k+1$, *we can construct the following linear Gaussian model for* $t \leq k+1$,

State:

$$
\tilde{s}_t = \begin{bmatrix} v \\ \widehat{v}_{t-1} \end{bmatrix}, \tag{A.22a}
$$

State Evolution:

$$
\tilde{s}_{t+1} = \tilde{\mathbf{A}}_t \tilde{s}_t + \tilde{\mathbf{H}}_t w_t + \tilde{d}_t, \tag{A.22b}
$$

Observation:

$$\tilde{y}_t = \begin{bmatrix} v \\ a_{t-1} - m_{t-1} \end{bmatrix} = \tilde{\mathbf{C}}_t s_t, \tag{A.22c}$$

where

$$\tilde{\mathbf{A}}_t = \left[\begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline \multicolumn{2}{c}{\tilde{\mathbf{G}}_t} \end{array}\right] \tag{A.23a}$$

$$(\tilde{\mathbf{G}}_t)_{\widehat{v}^i, v\widehat{v}^i} = (\mathbf{G}_t^i)_{:, v\widehat{v}^i}, \quad \forall i \in \mathcal{N} \tag{A.23b}$$

$$(\tilde{\mathbf{H}}_t)_{\widehat{v}^i, w^i} = (\mathbf{J}_t^i)_{\widehat{v}^i, x^i}, \quad \forall i \in \mathcal{N} \tag{A.23c}$$

$$(\tilde{d}_t)_{\widehat{v}^i} = d_t^{i,i}, \quad \forall i \in \mathcal{N} \tag{A.23d}$$

$$\tilde{\mathbf{C}}_t = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathfrak{D}(\mathbf{L}_{t-1}) \end{bmatrix}. \tag{A.23e}$$

Using this conditional Gauss Markov model, we can conclude that the conditional public beliefs $\pi_{k+1}^j(\widehat{v}_{k+1}^j|v)$ are Gaussian and by using Kalman filter results for $t = k+1$, we can write

$$\begin{aligned} \tilde{s}_{k+2|k+1} &= \mathbb{E}[\tilde{S}_{k+2}|\tilde{y}_{1:k+1}] = \mathbb{E}[\tilde{S}_{k+2}|v, a_{1:k}] \\ &= \tilde{\mathbf{A}}_{k+1}\tilde{s}_{k+1|k} + \tilde{\mathbf{A}}_{k+1}\tilde{\mathbf{J}}_{k+1}(\tilde{y}_{k+1} - \tilde{\mathbf{C}}_{k+1}\tilde{s}_{k+1|k}) + \tilde{d}_{k+1}. \end{aligned} \tag{A.24}$$

Therefore,

$$\begin{aligned} \mathbb{E}[\widehat{V}_{k+1}|v, a_{1:k}] = {} & (\tilde{\mathbf{G}}_{k+1})_{:,v} v + (\tilde{\mathbf{G}}_{k+1})_{:,\widehat{v}} \mathbb{E}[\widehat{V}_k|v, a_{1:k-1}] - (\tilde{\mathbf{A}}_{k+1}\tilde{\mathbf{J}}_{k+1})_{\widehat{v},a}\mathfrak{D}(\mathbf{L}_k)\mathbb{E}[\widehat{V}_k|v, a_{1:k-1}] \\ & + (\tilde{\mathbf{A}}_{k+1}\tilde{\mathbf{J}}_{k+1})_{\widehat{v},a}(a_k - m_k) + (\tilde{d}_{k+1})_{\widehat{v}}. \end{aligned} \tag{A.25}$$

Using the assumption of $\mathbb{E}[\widehat{V}_k|v, a_{1:k-1}] = \mathbf{E}_k v + f_k$, we have the following

$$\begin{aligned} \mathbb{E}[\widehat{V}_{k+1}|v, a_{1:k}] &= (\tilde{\mathbf{G}}_{k+1})_{:,v} v + (\tilde{\mathbf{G}}_{k+1})_{:,\widehat{v}}(\mathbf{E}_k v + f_k - (\tilde{\mathbf{A}}_{k+1}\tilde{\mathbf{J}}_{k+1})_{\widehat{v},a}\mathfrak{D}(\mathbf{L}_k)(\mathbf{E}_k v + f_k) \\ &\quad + (\tilde{\mathbf{A}}_{k+1}\tilde{\mathbf{J}}_{k+1})_{\widehat{v},a}(a_k - m_k) + (\tilde{d}_{k+1})_{\widehat{v}} \\ &= \mathbf{E}_{k+1} v + f_{k+1}, \end{aligned} \tag{A.26}$$

where

$$\mathbf{E}_{k+1} = (\tilde{\mathbf{G}}_{k+1})_{:,v} + ((\tilde{\mathbf{G}}_{k+1})_{:,\widehat{v}} - (\tilde{\mathbf{A}}_{k+1}\tilde{\mathbf{J}}_{k+1})_{\widehat{v},a}\mathfrak{D}(\mathbf{L}_k))\mathbf{E}_k \tag{A.27a}$$

$$f_{k+1} = ((\tilde{\mathbf{G}}_{k+1})_{:,\widehat{v}} - (\tilde{\mathbf{A}}_{k+1}\tilde{\mathbf{J}}_{k+1})_{\widehat{v},a}\mathfrak{D}(\mathbf{L}_k))f_k + (\tilde{\mathbf{A}}_{k+1}\tilde{\mathbf{J}}_{k+1})_{\widehat{v},a}(a_k - m_k) + (\tilde{d}_{k+1})_{\widehat{v}}, \quad \text{(A.27b)}$$

and similar to part (a) of the proof, the covariance matrix of $\tilde{S}_{k+1}$ conditioned on $\tilde{y}_{1:k+1}$ and $\tilde{y}_{1:k}$, denoted by $\tilde{\mathbf{\Sigma}}_{k+1|k+1}$ and $\tilde{\mathbf{\Sigma}}_{k+1|k}$, respectively, and the matrix $\tilde{\mathbf{J}}_{k+1}$ are derived from the following Kalman filter equations.

$$\tilde{\mathbf{\Sigma}}_{k+1|k} = \tilde{\mathbf{A}}_k\tilde{\mathbf{\Sigma}}_{k|k}\tilde{\mathbf{A}}_k' + \tilde{\mathbf{H}}_k\mathfrak{D}(\mathbf{Q})\tilde{\mathbf{H}}_k' \quad \text{(A.28a)}$$

$$\tilde{\mathbf{J}}_{k+1} = \tilde{\mathbf{\Sigma}}_{k+1|k}\tilde{\mathbf{C}}_{k+1}'(\tilde{\mathbf{C}}_{k+1}\tilde{\mathbf{\Sigma}}_{k+1|k}\tilde{\mathbf{C}}_{k+1}')^{-1} \quad \text{(A.28b)}$$

$$\tilde{\mathbf{\Sigma}}_{k+1|k+1} = (\mathbf{I} - \tilde{\mathbf{J}}_{k+1}\tilde{\mathbf{C}}_{k+1})\tilde{\mathbf{\Sigma}}_{k+1|k} \quad \text{(A.28c)}$$

$$\tilde{\mathbf{\Sigma}}_{1|1} = \mathbb{E}[\tilde{S}_1\tilde{S}_1'] - \mathbb{E}[\tilde{S}_1V'](\mathbb{E}[VV'])^{-1}\mathbb{E}[\tilde{S}_1V']' = \mathbf{0}. \quad \text{(A.28d)}$$

Note that if we know $\mathbf{\Sigma}_{k+1|k}$, $\tilde{\mathbf{\Sigma}}_{k+1|k}$, $\mathbf{E}_k$ and f_k, we can publicly evaluate all of the other quantities defined in this proof for $k+1$ for a given strategy matrices $\mathbf{L}_k$ and vectors m_k and therefore, we can find $\mathbf{\Sigma}_{k+2|k+1}$, $\tilde{\mathbf{\Sigma}}_{k+2|k+1}$, $\mathbf{E}_{k+1}$ and f_{k+1}. We can also find $\mathbf{G}_{k+1}^{i,i}$ and $d_{k+1}^{i,i}$, which are used to update $\widehat{v}_k^i$ to $\widehat{v}_{k+1}^i$.

A.6 Proof of Theorem 3

We show that for any $t \in \mathcal{T}$, if all players $-i$ play according to the strategy $\gamma_t^{-i}(a_t^{-i}|\widehat{v}_t^{-i}) = \delta(a_t^{-i} - \mathfrak{D}(\mathbf{L}_t^{-i})\widehat{v}_t^{-i} - m_t^{-i})$, where $m_t^{-i} = \mathbf{M}_t^{-i}f_t + \overline{m}_t^{-i}$, and the strategies of players are linear in $\widehat{v}_k$ for $k < t$, player i faces an MDP with state $(\widehat{v}_t^i, \mathbf{\Sigma}_t, \mathbf{E}_t, f_t)$ and her best response is of the form $\gamma_t^i(a_t^i|\widehat{v}_t^i) = \delta(a_t^i - \mathfrak{D}(\mathbf{L}_t^i)\widehat{v}_t^i - m_t^i)$, where $m_t^i = \mathbf{M}_t^i f_t + \overline{m}_t^i$.

By using the results from Theorem 2, given the strategy profile γ_t, $(\widehat{v}_t^i, \mathbf{\Sigma}_t, \mathbf{E}_t, f_t)$ forms a Markov chain. Notice that $\widehat{V}_{t+1}^i, \mathbf{\Sigma}_{t+1}, \mathbf{E}_{t+1}, f_{t+1}$ are updated by γ_t which is linear and therefore, all results from Theorem 2 hold.

Lemma 23. *One can write the expected value of the instantaneous reward $\overline{R}_t^i$ as*

$$\overline{R}_t^i = \mathbb{E}[r_t^i(V, A_t)|a_t^i, \widehat{v}_t^i, \mathbf{\Sigma}_t^i, \mathbf{E}_t, f_t] = \mathrm{qd}(\overline{\mathbf{R}}_t^i; \begin{bmatrix} \widehat{v}_t^i \\ a_t^i \\ f_t \end{bmatrix}) + \overline{b}_t^{i\prime} \begin{bmatrix} \widehat{v}_t^i \\ a_t^i \\ f_t \end{bmatrix} + \overline{c}_t^i, \quad \text{(A.29)}$$

where $\overline{\mathbf{R}}_t^i$, $\overline{b}_t^i$ and $\overline{c}_t^i$ are constructed in the proof.

Proof. *Since we assume all players $-i$ play according to γ_t^{-i}, we have*

$$a_t^{-i} = \mathfrak{D}(\mathbf{L}_t^{-i})\widehat{v}_t^{-i} + \mathbf{M}_t^{-i}f_t + \overline{m}_t^{-i},$$

and so the instantaneous reward can be rewritten as follows.

$$r_t^i(v, a_t) = \mathrm{qd}(\mathbf{R}_t^i; \begin{bmatrix} v \\ a_t \end{bmatrix})$$
$$= \mathrm{qd}(\tilde{\mathbf{R}}_t^i; \begin{bmatrix} v \\ a_t^i \\ \widehat{v_t}^{-i} \\ f_t \end{bmatrix}) + \tilde{b}_t^{i\prime} \begin{bmatrix} v \\ a_t^i \\ \widehat{v_t}^{-i} \\ f_t \end{bmatrix} + \tilde{c}_t^i, \tag{A.30}$$

where

$$\tilde{\mathbf{R}}_t^i = \begin{bmatrix} \mathbf{I}_{N_v+N_a} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathfrak{D}(\mathbf{L}_t^{-i}) & \mathbf{M}_t^{-i} \end{bmatrix}' \tilde{\mathbf{I}}'_{2,i+1} \mathbf{R}_t^i \tilde{\mathbf{I}}_{2,i+1} \begin{bmatrix} \mathbf{I}_{N_v+N_a} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathfrak{D}(\mathbf{L}_t^{-i}) & \mathbf{M}_t^{-i} \end{bmatrix} \tag{A.31a}$$

$$\tilde{\mathbf{I}}_{2,i+1} = \begin{bmatrix} \mathbf{I}_{N_v} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{(i-1)N_a} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N_a} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{(N-i)N_a} \end{bmatrix}, \tag{A.31b}$$

and $\mathbf{I}_k$ *is the identity matrix with size* $k \times k$*.*

$$\tilde{b}_t^{i\prime} = 2 \begin{bmatrix} 0 \\ \overline{m}_t^{-i} \end{bmatrix}' \tilde{\mathbf{I}}'_{2,i+1} \mathbf{R}_t^i \tilde{\mathbf{I}}_{2,i+1} \begin{bmatrix} \mathbf{I}_{N_v+N_a} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathfrak{D}(\mathbf{L}_t^{-i}) & \mathbf{M}_t^{-i} \end{bmatrix} \tag{A.31c}$$

$$\tilde{c}_t^i = \begin{bmatrix} 0 \\ \overline{m}_t^{-i} \end{bmatrix}' \tilde{\mathbf{I}}'_{2,i+1} \mathbf{R}_t^i \tilde{\mathbf{I}}_{2,i+1} \begin{bmatrix} 0 \\ \overline{m}_t^{-i} \end{bmatrix}. \tag{A.31d}$$

We can now calculate the expected value of R^i *as follows.*

$$\overline{R}_t^i = \mathrm{qd}(\tilde{\mathbf{R}}_t^i; \begin{bmatrix} \widehat{v}_t^i \\ a_t^i \\ \tilde{v}_t^{i,-i} \\ f_t \end{bmatrix}) + \mathrm{tr}(\tilde{\mathbf{R}}_t^i \overline{\boldsymbol{\Sigma}}_t^i) + \tilde{b}_t^{i\prime} \begin{bmatrix} \widehat{v}_t^i \\ a_t^i \\ \tilde{v}_t^{i,-i} \\ f_t \end{bmatrix} + \tilde{c}_t^i, \tag{A.32}$$

where

$$\overline{\boldsymbol{\Sigma}}_t^i = \begin{bmatrix} \boldsymbol{\Sigma}_t^i & \mathbf{0} & \boldsymbol{\Sigma}_t^i \mathbf{E}_t^{-i\prime} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{E}_t^{-i}\boldsymbol{\Sigma}_t^i & \mathbf{0} & (\boldsymbol{\Sigma}_{t+1|t}^i)_{\widehat{v}^{-i},\widehat{v}^{-i}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}. \tag{A.33}$$

By using $\tilde{v}_t^{i,-i} = \mathbf{E}_t^{-i}\widehat{v}_t^i + f_t^{-i}$*, we can derive the equations for* $\overline{\mathbf{R}}_t^i$*,* $\overline{b}_t^i$ *and* $\overline{c}_t^i$*.*

$$\overline{\mathbf{R}}_t^i = \begin{bmatrix} \mathbf{I}_{N_v} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N_a} & \mathbf{0} \\ \mathbf{E}_t^{-i} & \mathbf{0} & \widehat{\mathbf{I}}_{-i} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{N_v N} \end{bmatrix}' \tilde{\mathbf{R}}_t^i \begin{bmatrix} \mathbf{I}_{N_v} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N_a} & \mathbf{0} \\ \mathbf{E}_t^{-i} & \mathbf{0} & \widehat{\mathbf{I}}_{-i} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{N_v N} \end{bmatrix} \tag{A.34a}$$

$$\overline{b}_t^{i\prime} = \tilde{b}_t^{i\prime} \begin{bmatrix} \mathbf{I}_{N_v} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N_a} & \mathbf{0} \\ \mathbf{E}_t^{-i} & \mathbf{0} & \widehat{\mathbf{I}}_{-i} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{N_v N} \end{bmatrix} \tag{A.34b}$$

$$(\widehat{\mathbf{I}}_{-i})_{:,f^{-i}} = \mathbf{I}_{(N-1)N_v} \tag{A.34c}$$

$$\overline{c}_t^i = \mathrm{tr}(\tilde{\mathbf{R}}_t^i \overline{\boldsymbol{\Sigma}}_t^i) + \tilde{c}_t^i, \tag{A.34d}$$

In the next lemma, we show that the reward-to-go at time t is a quadratic functions of $\begin{bmatrix} \widehat{v}_t^i \\ f_t \end{bmatrix}$ and we will construct the strategy matrix and vector $\mathbf{L}_t^i$ and m_t^i.

Lemma 24. *We have the following equation for the reward-to-go function,* $J_t^i(\widehat{v}_t^i, \boldsymbol{\Sigma}_t, \mathbf{E}_t, f_t) = \mathrm{qd}(\mathbf{Z}_t^i; \begin{bmatrix} \widehat{v}_t^i \\ f_t \end{bmatrix}) + z_t^{i\prime} \begin{bmatrix} \widehat{v}_t^i \\ f_t \end{bmatrix} + o_t^i.$

Note that the above equation only highlights the functionality of the reward-to-go with respect to $\widehat{v}_t^i$ and f_t. We do not care about its functionality with respect to $\boldsymbol{\Sigma}_t$ and $\mathbf{E}_t$ due to two reasons. First, they are part of the public part of the history and are not parameters of the partial strategies γ. Second, they are not controlled by the actions. As we will see in the proof of this lemma, $\mathbf{Z}_t^i$, z_t^i and o_t^i are functions of $\boldsymbol{\Sigma}_t$ and $\mathbf{E}_t$.

Proof. *We prove the lemma by backward induction. For* $T+1$*, we have*

$$J_{T+1}^i(\widehat{v}_{T+1}^i, \boldsymbol{\Sigma}_{T+1}, \mathbf{E}_{T+1}, f_{T+1}) = 0$$

and by setting $\mathbf{Z}^i_{T+1} = \mathbf{0}$, $z^i_{T+1} = 0$, $o^i_{T+1} = 0$, the equation holds.

Assume that the lemma holds for $t+1$. We will show that it will also hold for t.

$$
\begin{aligned}
J^i_t(\widehat{v}^i_t, \mathbf{\Sigma}_t, \mathbf{E}_t, f_t) &= \max_{a^i_t} \mathbb{E}^{\gamma_t^{-i}}[r^i_t(V, A_t) + J^i_{t+1}(\widehat{V}^i_{t+1}, \mathbf{\Sigma}_{t+1}, \mathbf{E}_{t+1}, f_{t+1})|a^i_t, \widehat{v}^i_t, \mathbf{\Sigma}_t, \mathbf{E}_t, f_t] \\
&= \max_{a^i_t}\{\mathrm{qd}(\overline{\mathbf{R}}^i_t; \begin{bmatrix} \widehat{v}^i_t \\ a^i_t \\ f_t \end{bmatrix}) + \overline{b}^{i\prime}_t \begin{bmatrix} \widehat{v}^i_t \\ a^i_t \\ f_t \end{bmatrix} + \overline{c}^i_t + \mathbb{E}^{\gamma_t^{-i}}[\mathrm{qd}(\mathbf{Z}^i_{t+1}; \begin{bmatrix} \widehat{V}^i_{t+1} \\ f_{t+1} \end{bmatrix}) \\
&\qquad + z^{i\prime}_{t+1} \begin{bmatrix} \widehat{V}^i_{t+1} \\ f_{t+1} \end{bmatrix} + o^i_{t+1}|a^i_t, \widehat{v}^i_t, \mathbf{\Sigma}_t, \mathbf{E}_t, f_t]\}.
\end{aligned} \tag{A.35}
$$

First consider the J^i_{t+1} part.

$$
\begin{aligned}
&\mathbb{E}^{\gamma_t^{-i}}[\mathrm{qd}(\mathbf{Z}^i_{t+1}; \begin{bmatrix} \widehat{V}^i_{t+1} \\ f_{t+1} \end{bmatrix}) + z^{i\prime}_{t+1} \begin{bmatrix} \widehat{V}^i_{t+1} \\ f_{t+1} \end{bmatrix} + o^i_{t+1}|a^i_t, \widehat{v}^i_t, \mathbf{\Sigma}_t, \mathbf{E}_t, f_t] \\
&= \mathbb{E}^{\gamma_t^{-i}}[\mathrm{qd}(\mathbf{Z}^i_{t+1}; \widehat{\mathbf{G}}^i_{t+1} \begin{bmatrix} \widehat{v}^i_t \\ a^i_t \\ \widehat{V}^{-i}_t \\ X^i_{t+1} \\ f_t \end{bmatrix} + \widehat{g}^i_{t+1}) + z^{i\prime}_{t+1}(\widehat{\mathbf{G}}^i_{t+1} \begin{bmatrix} \widehat{v}^i_t \\ a^i_t \\ \widehat{V}^{-i}_t \\ X^i_{t+1} \\ f_t \end{bmatrix} + \widehat{g}^i_{t+1}) \\
&\qquad + o^i_{t+1}|a^i_t, \widehat{v}^i_t, \mathbf{\Sigma}_t, \mathbf{E}_t, f_t] \\
&= \mathrm{qd}(\overline{\mathbf{Z}}^i_{t+1}; \begin{bmatrix} \widehat{v}^i_t \\ a^i_t \\ f_t \end{bmatrix}) + \overline{z}^{i\prime}_{t+1} \begin{bmatrix} \widehat{v}^i_t \\ a^i_t \\ f_t \end{bmatrix} + \overline{o}^i_{t+1},
\end{aligned} \tag{A.36}
$$

where

$$(\widehat{\mathbf{G}}^i_{t+1})_{\widehat{v}^i, \widehat{v}^i} = (\mathbf{G}^{i,i}_{t+1})_{:,\widehat{v}^i} \tag{A.37a}$$

$$(\widehat{\mathbf{G}}^i_{t+1})_{\widehat{v}^i, \widehat{v}^{-i}} = (\mathbf{J}^i_{t+1})_{\widehat{v}^i, a^{-i}} \mathfrak{D}(\mathbf{L}^{-i}_t) \tag{A.37b}$$

$$(\widehat{\mathbf{G}}^i_{t+1})_{\widehat{v}^i, x^i} = (\mathbf{G}^{i,i}_{t+1})_{:,x^i} \tag{A.37c}$$

$$(\widehat{\mathbf{G}}^i_{t+1})_{\widehat{v}^i, f^{-i}} = (\mathbf{J}^i_{t+1})_{\widehat{v}^i, a^{-i}} \mathfrak{D}(\mathbf{L}^{-i}_t) \tag{A.37d}$$

$$
\begin{aligned}
(\widehat{\mathbf{G}}^i_{t+1})_{f^j, f^{-j}} &= ((\tilde{\mathbf{G}}_{t+1})_{:,\widehat{v}} - (\tilde{\mathbf{A}}_{t+1}\tilde{\mathbf{J}}_{t+1})_{\widehat{v},a} \mathfrak{D}(\mathbf{L}_t))_{f^j, f^{-j}} \\
&\quad - (\mathbf{J}^j_{t+1})_{\widehat{v}^j, a^{-j}} \mathfrak{D}(\mathbf{L}^{-j}_t) - (\tilde{\mathbf{A}}_{t+1}\tilde{\mathbf{J}}_{t+1})_{\widehat{v}^j, a^i} (\mathbf{M}^i_t)_{:,f^{-j}} - (\mathbf{J}^j_{t+1})_{\widehat{v}^j, a^i} (\mathbf{M}^i_t)_{:,f^{-j}}, \ \forall j \neq i
\end{aligned} \tag{A.37e}
$$

$$
\begin{aligned}
(\widehat{\mathbf{G}}^i_{t+1})_{f^j,f^j} &= ((\tilde{\mathbf{G}}_{t+1})_{:,\widehat{v}} - (\tilde{\mathbf{A}}_{t+1}\tilde{\mathbf{J}}_{t+1})_{\widehat{v},a}\mathfrak{D}(\mathbf{L}_t))_{f^j,f^j} \\
&\quad - (\tilde{\mathbf{A}}_{t+1}\tilde{\mathbf{J}}_{t+1})_{\widehat{v}^j,a^i}(\mathbf{M}^i_t)_{:,f^j} - (\mathbf{J}^j_{t+1})_{\widehat{v}^j,a^i}(\mathbf{M}^i_t)_{:,f^j}, \quad \forall j \neq i \qquad \text{(A.37f)} \\
(\widehat{\mathbf{G}}^i_{t+1})_{f^i,f^{-i}} &= ((\tilde{\mathbf{G}}_{t+1})_{:,\widehat{v}} - (\tilde{\mathbf{A}}_{t+1}\tilde{\mathbf{J}}_{t+1})_{\widehat{v},a}\mathfrak{D}(\mathbf{L}_t))_{f^i,f^{-i}} \\
&\quad - (\mathbf{J}^i_{t+1})_{\widehat{v}^i,a^{-i}}\mathfrak{D}(\mathbf{L}^{-i}_t) - (\tilde{\mathbf{A}}_{t+1}\tilde{\mathbf{J}}_{t+1})_{\widehat{v}^i,a^i}(\mathbf{M}^i_t)_{:,f^{-i}} \qquad \text{(A.37g)} \\
(\widehat{\mathbf{G}}^i_{t+1})_{f^i,f^i} &= ((\tilde{\mathbf{G}}_{t+1})_{:,\widehat{v}} - (\tilde{\mathbf{A}}_{t+1}\tilde{\mathbf{J}}_{t+1})_{\widehat{v},a}\mathfrak{D}(\mathbf{L}_t))_{f^i,f^i} - (\tilde{\mathbf{A}}_{t+1}\tilde{\mathbf{J}}_{t+1})_{\widehat{v}^i,a^i}(\mathbf{M}^i_t)_{:,f^i}, \quad \forall j \neq i \\
&\qquad \text{(A.37h)} \\
(\widehat{\mathbf{G}}^i_{t+1})_{f^j,a^i} &= (\mathbf{J}^j_{t+1})_{\widehat{v}^j,a^i} + (\tilde{\mathbf{A}}_{t+1}\tilde{\mathbf{J}}_{t+1})_{\widehat{v}^j,a^i}, \quad \forall j \neq i \qquad \text{(A.37i)} \\
(\widehat{\mathbf{G}}^i_{t+1})_{f^i,a^i} &= (\tilde{\mathbf{A}}_{t+1}\tilde{\mathbf{J}}_{t+1})_{\widehat{v}^i,a^i} \qquad \text{(A.37j)} \\
(\widehat{\mathbf{G}}^i_{t+1})_{f^k,\widehat{v}^j} &= (\mathbf{J}^k_{t+1})_{\widehat{v}^k,a^j}\mathbf{L}^j_t + (\tilde{\mathbf{A}}_{t+1}\tilde{\mathbf{J}}_{t+1})_{\widehat{v}^k,a^j}\mathbf{L}^j_t, \quad \forall j \neq i, \forall k \neq j \qquad \text{(A.37k)} \\
(\widehat{\mathbf{G}}^i_{t+1})_{f^j,\widehat{v}^j} &= (\tilde{\mathbf{A}}_{t+1}\tilde{\mathbf{J}}_{t+1})_{\widehat{v}^j,a^j}\mathbf{L}^j_t, \quad \forall j \neq i \qquad \text{(A.37l)} \\
(\widehat{g}^i_{t+1})_{f^i} &= -(\tilde{\mathbf{A}}_{t+1}\tilde{\mathbf{J}}_{t+1})_{\widehat{v}^i,a^i}\overline{m}^i_t \qquad \text{(A.37m)} \\
(\widehat{g}^i_{t+1})_{f^j} &= -(\tilde{\mathbf{A}}_{t+1}\tilde{\mathbf{J}}_{t+1})_{\widehat{v}^j,a^i}\overline{m}^i_t - (\mathbf{J}^j_{t+1})_{\widehat{v}^j,a^i}\overline{m}^i_t, \ \forall j \neq i, \qquad \text{(A.37n)}
\end{aligned}
$$

and we have

$$
\begin{aligned}
\overline{\mathbf{Z}}^i_{t+1} &= \mathbf{T}^{i\prime}_{t+1}\widehat{\mathbf{G}}^{i\prime}_{t+1}\mathbf{Z}^i_{t+1}\widehat{\mathbf{G}}^i_{t+1}\mathbf{T}^i_{t+1} \qquad \text{(A.38a)} \\
\mathbf{T}^i_{t+1} &= \begin{bmatrix} \mathbf{I}_{N_v} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N_a} & \mathbf{0} \\ \mathbf{E}^{-i}_t & \mathbf{0} & \widehat{\mathbf{I}}_{-i} \\ \mathbf{I}_{N_v} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{N_v N} \end{bmatrix} \qquad \text{(A.38b)} \\
\overline{z}^{i\prime}_{t+1} &= (2\widehat{g}^{i\prime}_{t+1}\mathbf{Z}^i_{t+1}\widehat{\mathbf{G}}^i_{t+1} + z^{i\prime}_{t+1}\widehat{\mathbf{G}}^i_{t+1})\mathbf{T}^i_{t+1} \qquad \text{(A.38c)} \\
\overline{o}^i_{t+1} &= \widehat{g}^{i\prime}_{t+1}\mathbf{Z}^i_{t+1}\widehat{g}^i_{t+1} + \mathrm{tr}(\widehat{\mathbf{G}}^{i\prime}_{t+1}\mathbf{Z}^i_{t+1}\widehat{\mathbf{G}}^i_{t+1}\widehat{\mathbf{\Sigma}}^i_{t+1}) + z^{i\prime}_{t+1}\widehat{g}^i_{t+1} + o^i_{t+1} \qquad \text{(A.38d)} \\
\widehat{\mathbf{\Sigma}}^i_{t+1} &= Cov(\begin{bmatrix} \widehat{v}^i_t \\ a^i_t \\ \widehat{V}^{-i}_t \\ X^i_{t+1} \\ f_t \end{bmatrix} | a^i_t, \widehat{v}^i_t, \mathbf{\Sigma}_t, \mathbf{E}_t, f_t) \qquad \text{(A.38e)} \\
(\widehat{\mathbf{\Sigma}}^i_{t+1})_{\widehat{v}^{-i}x^i,\widehat{v}^{-i}x^i} &= \begin{bmatrix} (\mathbf{\Sigma}^i_{t+1|t})_{\widehat{v}^{-i},\widehat{v}^{-i}} & \mathbf{E}^{-i}_t\mathbf{\Sigma}^i_t \\ \mathbf{\Sigma}^i_t\mathbf{E}^{-i\prime}_t & \mathbf{\Sigma}^i_t + \mathbf{Q}^i \end{bmatrix}. \qquad \text{(A.38f)}
\end{aligned}
$$

Therefore, one can write the expected reward-to-go as follows.

$$
\begin{aligned}
J_t^i(\widehat{v}_t^i, \boldsymbol{\Sigma}_t^i, &\mathbf{E}_t, f_t) \\
&= \max_{a_t^i}\{\mathrm{qd}(\overline{\mathbf{R}}_t^i; \begin{bmatrix} \widehat{v}_t^i \\ a_t^i \\ f_t \end{bmatrix}) + \overline{b}_t^{i\prime} \begin{bmatrix} \widehat{v}_t^i \\ a_t^i \\ f_t \end{bmatrix} + \overline{c}_t^i + \mathrm{qd}(\overline{\mathbf{Z}}_{t+1}^i; \begin{bmatrix} \widehat{v}_t^i \\ a_t^i \\ f_t \end{bmatrix}) + \overline{z}_{t+1}^{i\prime} \begin{bmatrix} \widehat{v}_t^i \\ a_t^i \\ f_t \end{bmatrix} + \overline{o}_{t+1}^i\} \\
&= \max_{a_t^i}\{\mathrm{qd}(\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i; \begin{bmatrix} \widehat{v}_t^i \\ a_t^i \\ f_t \end{bmatrix}) + (\overline{b}_t^{i\prime} + \overline{z}_{t+1}^{i\prime}) \begin{bmatrix} \widehat{v}_t^i \\ a_t^i \\ f_t \end{bmatrix} + \overline{c}_t^i + \overline{o}_{t+1}^i\}. \qquad \text{(A.39)}
\end{aligned}
$$

The above equation is quadratic with respect to a_t^i and therefore, if $(\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)_{a^i,a^i}$ is negative definite, the maximum value is achieved when the gradient of the above equation with respect to a_t^i is zero.

$$
2(\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)_{a^i,a^i} a_t^i + 2(\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)_{a^i,\widehat{v}^i f} \begin{bmatrix} \widehat{v}_t^i \\ f_t \end{bmatrix} + (\overline{b}_t^i + \overline{z}_{t+1}^i)_{a^i} = 0 \qquad \text{(A.40a)}
$$

$$
\Rightarrow a_t^i = -(\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)_{a^i,a^i}^{-1}((\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)_{a^i,\widehat{v}^i f} \begin{bmatrix} \widehat{v}_t^i \\ f_t \end{bmatrix} + \frac{1}{2}(\overline{b}_t^i + \overline{z}_{t+1}^i)_{a^i}). \qquad \text{(A.40b)}
$$

Finally, we can derive the best response strategy of player i to be $\gamma_t^i(\cdot|\widehat{v}_t^i) = \delta(a_t^i - \mathbf{L}_t^i \widehat{v}_t^i - m_t^i)$ where

$$
\mathbf{L}_t^i = -(\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)_{a^i,a^i}^{-1}(\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)_{a^i,\widehat{v}^i} \qquad \text{(A.41a)}
$$

$$
m_t^i = -(\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)_{a^i,a^i}^{-1}((\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)_{a^i,f} f_t + \frac{1}{2}(\overline{b}_t^i + \overline{z}_{t+1}^i)_{a^i}). \qquad \text{(A.41b)}
$$

Note that we have $m_t^i = \mathbf{M}_t^i f_t + \overline{m}_t^i$, where

$$
\mathbf{M}_t^i = -(\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)_{a^i,a^i}^{-1}(\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)_{a^i,f} \qquad \text{(A.41c)}
$$

$$
\overline{m}_t^i = -\frac{1}{2}(\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)_{a^i,a^i}^{-1}(\overline{b}_t^i + \overline{z}_{t+1}^i)_{a^i}. \qquad \text{(A.41d)}
$$

By substituting the best response action in the reward-to-go equation (A.39)*, we have the following final step of the proof.*

$$
J_t^i(\widehat{v}_t^i, \boldsymbol{\Sigma}_t, \mathbf{E}_t, f_t) = \mathrm{qd}(\mathbf{Z}_t^i; \begin{bmatrix} \widehat{v}_t^i \\ f_t \end{bmatrix}) + z_t^{i\prime} \begin{bmatrix} \widehat{v}_t^i \\ f_t \end{bmatrix} + o_t^i, \qquad \text{(A.42)}
$$

where

$$\mathbf{Z}_t^i = \widehat{\mathbf{T}}_t^{i\prime}(\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)\widehat{\mathbf{T}}_t^i \tag{A.43a}$$

$$\widehat{\mathbf{T}}_t^i = \begin{bmatrix} \mathbf{I}_{N_v} & \mathbf{0} \\ \mathbf{L}_t^i & \mathbf{M}_t^i \\ \mathbf{0} & \mathbf{I}_{N_v N} \end{bmatrix} \tag{A.43b}$$

$$z_t^{i\prime} = 2\widehat{m}_t^{i\prime}(\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)\widehat{\mathbf{T}}_t^i + (\overline{b}_t^{i\prime} + \overline{z}_{t+1}^{i\prime})\widehat{\mathbf{T}}_t^i \tag{A.43c}$$

$$\widehat{m}_t^i = \begin{bmatrix} \mathbf{0} \\ \overline{m}_t^i \\ \mathbf{0} \end{bmatrix} \tag{A.43d}$$

$$o_t^i = \widehat{m}_t^{i\prime}(\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)\widehat{m}_t^i + (\overline{b}_t^{i\prime} + \overline{z}_{t+1}^{i\prime})\widehat{m}_t^i + \overline{c}_t^i + \overline{o}_{t+1}^i. \tag{A.43e}$$

Note that in order to derive the γ_t^i strategy matrix and vector, $\mathbf{L}_t^i$ and m_t^i, we need to know $\mathbf{L}_t^{-i}$ and m_t^{-i}. Clearly, the same is true for calculating $\mathbf{L}_t^{-i}$ and m_t^{-i}. On the other hand, some of the quantities used in the proof, like $\widehat{\mathbf{G}}_{t+1}^i$, require $\mathbf{L}_t^i$ and m_t^i to be evaluated. Therefore, we have a fixed point equation over $\mathbf{L}_t$ and m_t.

Note that we have such linear solution only if the matrix $(\overline{\mathbf{R}}_t^i + \overline{\mathbf{Z}}_{t+1}^i)_{a^i,a^i}$ is invertible and negative semidefinite for all $i \in \mathcal{N}$.

We conclude the proof of the theorem by noting that in Lemma 24, we proved that the reward to go is a quadratic function of $\begin{bmatrix} \widehat{v}_t^i \\ f_t \end{bmatrix}$ and as a result and throughout the proof, we showed that the strategies that are linear in terms of $\begin{bmatrix} \widehat{v}_t^i \\ f_t \end{bmatrix}$ form equilibria of the game and the theorem is proved.

APPENDIX B

Proofs for Chapter 3

B.1 Proof of Lemma 2

Proof. *The proof follows by induction. For $t = 0$ we have*

$$\pi_0(\boldsymbol{x}, v) = \mathbb{P}^s(\boldsymbol{x}, v|n_0) = Q(v) \prod_{m=1}^{N} Q(x^m|v).$$

Assuming that $\pi_{t-1}(\boldsymbol{x}, v) = \pi_{t-1}(v) \prod_{m=1}^{N} \pi_{t-1}(x^m|v)$ we have

$$\begin{aligned}
\pi_t(\boldsymbol{x}, v) &= \mathbb{P}^s(\boldsymbol{x}, v|\boldsymbol{a}_{0:t-1}, n_{0:t}) && \text{(B.1a)}\\
&= \frac{\mathbb{P}^s(\boldsymbol{x}, v, \boldsymbol{a}_{t-1}, n_t|\boldsymbol{a}_{0:t-2}, n_{0:t-1})}{\mathbb{P}^s(\boldsymbol{a}_{t-1}, n_t|\boldsymbol{a}_{0:t-2}, n_{0:t-1})} && \text{(B.1b)}\\
&= \frac{(1/N)\mathbb{P}^s(\boldsymbol{a}_{t-1}|\boldsymbol{x}, v, \boldsymbol{a}_{0:t-2}, n_{0:t-1})\mathbb{P}^s(\boldsymbol{x}, v|\boldsymbol{a}_{0:t-2}, n_{0:t-1})}{\mathbb{P}^s(\boldsymbol{a}_{t-1}, n_t|\boldsymbol{a}_{0:t-2}, n_{0:t-1})} && \text{(B.1c)}\\
&= \frac{(1/N)\left(\prod_{m=1}^{N} \boldsymbol{1}_{\gamma_{t-1}^m(x^m)}(a_{t-1}^m)\right) \pi_{t-1}(\boldsymbol{x}, v)}{\sum_{x,v}(1/N)\left(\prod_{m=1}^{N} \boldsymbol{1}_{\gamma_{t-1}^m(x^m)}(a_{t-1}^m)\right) \pi_{t-1}(\boldsymbol{x}, v)} && \text{(B.1d)}\\
&= \frac{\left(\prod_{m=1}^{N} \boldsymbol{1}_{\gamma_{t-1}^m(x^m)}(a_{t-1}^m)\right) \pi_{t-1}(v) \prod_{m=1}^{N} \pi_{t-1}(x^m|v)}{\sum_{x,v}\left(\prod_{m=1}^{N} \boldsymbol{1}_{\gamma_{t-1}^m(x^m)}(a_{t-1}^m)\right) \pi_{t-1}(v) \prod_{m=1}^{N} \pi_{t-1}(x^m|v)} && \text{(B.1e)}\\
&= \frac{\left(\prod_{m=1}^{N} \boldsymbol{1}_{\gamma_{t-1}^m(x^m)}(a_{t-1}^m)\pi_{t-1}(x^m|v)\right) \pi_{t-1}(v)}{\sum_{v}\left(\prod_{m=1}^{N} \sum_{x^m} \boldsymbol{1}_{\gamma_{t-1}^m(x^m)}(a_{t-1}^m)\pi_{t-1}(x^m|v)\right) \pi_{t-1}(v)}. && \text{(B.1f)}
\end{aligned}$$

The conditional distribution of X given V and $\boldsymbol{h}_t^c$ can now be written as

$$\pi_t(\boldsymbol{x}|v) = \frac{\prod_{m=1}^{N} \boldsymbol{1}_{\gamma_{t-1}^m(x^m)}(a_{t-1}^m)\pi_{t-1}(x^m|v)}{\sum_{x}\left(\prod_{m=1}^{N} \boldsymbol{1}_{\gamma_{t-1}^m(x^m)}(a_{t-1}^m)\pi_{t-1}(x^m|v)\right)} \quad \text{(B.2a)}$$

$$= \prod_{m=1}^{N} \frac{\boldsymbol{1}_{\gamma_{t-1}^m(x^m)}(a_{t-1}^m)\pi_{t-1}(x^m|v)}{\sum_{x^m} \boldsymbol{1}_{\gamma_{t-1}^m(x^m)}(a_{t-1}^m)\pi_{t-1}(x^m|v)} \tag{B.2b}$$

$$= \prod_{m=1}^{N} \pi_t(x^m|v), \tag{B.2c}$$

where the second equality follows since given V, $\{x^m\}$ are independent, so the expectation of the product is the product of the expectations. This completes the induction step proving that $X^1, \ldots, X^N$ are conditionally independent given $v, \boldsymbol{h}_t^c$, which gives (3.17)*. Furthermore,* (B.2c) *provides an update equation for the conditional beliefs as*

$$\pi_t(x^m|v) = \frac{\boldsymbol{1}_{\gamma_{t-1}^m(x^m)}(a_{t-1}^m)\pi_{t-1}(x^m|v)}{\sum_{x^m} \boldsymbol{1}_{\gamma_{t-1}^m(x^m)}(a_{t-1}^m)\pi_{t-1}(x^m|v)}$$

$$= \begin{cases} \pi_{t-1}(x^m|v), & m \neq n_{t-1} \textit{ or } \gamma_{t-1}^m \neq \boldsymbol{I} \\ \boldsymbol{1}_{\frac{x^m+1}{2}}(a_{t-1}^m), & m = n_{t-1} \textit{ and } \gamma_{t-1}^m = \boldsymbol{I}. \end{cases} \tag{B.3a}$$

Consequently, if player m has not yet revealed her information up to time t, then $\pi_t(x^m|v) = \cdots = \pi_0(x^m|v) = Q(x^m|v)$. Alternatively, if player m has revealed her information before time t, we have $\pi_t(x^m|v) = \boldsymbol{1}_{\tilde{x}^m}(x^m)$, thus proving (3.18)*.*

Now, marginalizing (B.1a) *w.r.t. x we have*

$$\frac{\pi_{t+1}(1)}{\pi_{t+1}(-1)} = \frac{\prod_{m=1}^{N} \sum_{x^m} \boldsymbol{1}_{\gamma_t^m(x^m)}(a_t^m)\pi_t(x^m|1)}{\prod_{m=1}^{N} \sum_{x^m} \boldsymbol{1}_{\gamma_t^m(x^m)}(a_t^m)\pi_t(x^m|-1)} \frac{\pi_t(1)}{\pi_t(-1)} \tag{B.4a}$$

$$= \frac{\sum_{x^{n_t}} \boldsymbol{1}_{\gamma_t(x^{n_t})}(a_t^{n_t})\pi_t(x^{n_t}|1)}{\sum_{x^{n_t}} \boldsymbol{1}_{\gamma_t(x^{n_t})}(a_t^{n_t})\pi_t(x^{n_t}|-1)} \frac{\pi_t(1)}{\pi_t(-1)}, \tag{B.4b}$$

where the last equality is due to the fact that for all non-acting players $m \neq n_t$ we have $\gamma_t^m = \boldsymbol{0}$. Hence, for $m \neq n_t$ we always have $\boldsymbol{1}_{\gamma_t(x^m)}(a_t^m) = 1$, and then $\sum \pi_t\left(x^m \mid \pm 1\right) = 1$ so these terms have no effect on the products in the numerator and denominator. Furthermore, if $\gamma_t \neq \boldsymbol{I}$ or the acting player has already revealed her information, the multiplicative factor reduces to 1. Else, the factor becomes

$$\frac{\sum_{x^{n_t}} \boldsymbol{1}_{\frac{x^{n_t}+1}{2}}(a_t^{n_t})Q(x^{n_t}|1)}{\sum_{x^{n_t}} \boldsymbol{1}_{\frac{x^{n_t}+1}{2}}(a_t^{n_t})Q(x^{n_t}|-1)} = \frac{Q(2a_t^{n_t}-1|1)}{Q(2a_t^{n_t}-1|-1)} = q^{2a_t^{n_t}-1}, \tag{B.4c}$$

which gives (3.19)*. We derive* (3.20) *by repeating the substitution of* (B.4c) *for all n, and using $Q(1) = Q(-1) = \frac{1}{2}$.*

B.2 Proof of Theorem 1

Proof. *For clarity, we first prove the result with π replacing $\tilde{\boldsymbol{x}}$ all throughout FPE 1, and then use the fact that π can be computed from $\tilde{\boldsymbol{x}}$. Let us assume that all players other than player n play according to $\gamma_t^* = \theta[n_t, \pi_t, \boldsymbol{b}_t]$, i.e., so if at time t we have $n_t \neq n$ then $a_t^{n_t} = \gamma_t^*(x^{n_t}) = \theta[n_t, \pi_t, \boldsymbol{b}_t](x^{n_t})$. Let us further assume that the update of the belief π_t is fixed to $\pi_{t+1} = F(\pi_t, \gamma_t^*, a_t^{n_t}, n_t) = F(\pi_t, \theta[n_t, \pi_t, \boldsymbol{b}_t], a_t^{n_t}, n_t) =: F^\theta(\pi_t, n_t, a_t^{n_t}, \boldsymbol{b}_t)$. We will show that the optimization problem faced by player n can be formulated as a Markov decision process (MDP). For this we will define a state, action, and instantaneous reward of a dynamical system as follows. The state of the system is defined as $\lambda_t = (x^n, n_t, \pi_t, \boldsymbol{b}_t)$. Further, the action space is defined according to* (3.7)*, where at each time t, player n takes the action $a_t^n \in \mathcal{A}^n(b_t^n, n_t)$ and receives an expected instantaneous reward of $R(\lambda_t, a_t^n) = a_t^n \sum_v v\mu^n(v)$.*

We first show that $(\Lambda_t)_t$ is a controlled Markov process with actions a_t^n, i.e.,

$$\mathbb{P}(\Lambda_{t+1}|\Lambda_{1:t}, a_{1:t}^n) = \mathbb{P}(\Lambda_{t+1}|\Lambda_t, a_t^n). \tag{B.5}$$

Indeed,

$$\mathbb{P}(\Lambda_{t+1}|\Lambda_{1:t}, a_{1:t}^n) = \mathbb{P}(\overline{x}^n, n_{t+1}, \pi_{t+1}, \boldsymbol{b}_{t+1}|x^n, n_{1:t}, \pi_{1:t}, \boldsymbol{b}_{1:t}, a_{1:t}^n) \tag{B.6a}$$

$$= \boldsymbol{1}_{x^n}(\overline{x}^n)\frac{1}{N}Q_b(\boldsymbol{b}_{t+1}|x^n, n_t, \pi_t, \boldsymbol{b}_t, a_t^n)Q_\pi(\pi_{t+1}|x^n, n_t, \pi_t, \boldsymbol{b}_t, a_t^n), \tag{B.6b}$$

where the kernels Q_b and Q_π are defined through

$$Q_b(\boldsymbol{b}_{t+1}|x^n, n_t, \pi_t, \boldsymbol{b}_t, a_t^n) = Q_{b^n}(b_{t+1}^n|b_t^n, a_t^n) \prod_{m=1, m\neq n}^{N} Q_{\boldsymbol{b}^{-n}}(b_{t+1}^m|x^n, n_t, \pi_t, \boldsymbol{b}_t) \tag{B.7a}$$

with

$$Q_{b^n}(b_{t+1}^n = 1|b_t^n, a_t^n) = \begin{cases} 1, & b_t^n = 1, \text{ or } a_t^n = 1 \\ 0, & \text{else} \end{cases} \tag{B.7b}$$

$$Q_{\boldsymbol{b}^{-n}}(b_{t+1}^m = 1|x^n, n_t, \pi_t, \boldsymbol{b}_t) = \begin{cases} \boldsymbol{1}_m(n_t) \sum_{x^m} \mu^n(x^m) \boldsymbol{1}_{\theta[n_t, \pi_t, \boldsymbol{b}_t](x^m)}(1), & b_t^m = 0 \\ 1, & b_t^m = 1 \end{cases} \tag{B.7c}$$

and

$$Q_\pi(\pi_{t+1}|x^n, n_t, \pi_t, \boldsymbol{b}_t, a_t^n) = \begin{cases} \sum_{x^{n_t}} \mu^n(x^{n_t}) \boldsymbol{1}_{F^\theta(\pi_t, n_t, \theta[n_t, \pi_t, \boldsymbol{b}_t](x^{n_t}), \boldsymbol{b}_t)}(\pi_{t+1}), & n_t \neq n \\ \boldsymbol{1}_{F^\theta(\pi_t, n_t, a_t^n, \boldsymbol{b}_t)}(\pi_{t+1}), & n_t = n. \end{cases} \tag{B.7d}$$

It is exactly the above equation that reveals why the belief update has to be fixed in order to prove that player n faces an MDP. If that were not the case, the above equation would require that the belief is updated through an expression of the form $\pi_{t+1} = F(\pi_t, \gamma_t, a_t^{n_t}, n_t)$ which would require to include the partial function γ_t in the action space for the case $n_t = n$ as opposed to only including the action a_t^n. We have now proved (B.5). *Hence, the state process $(\Lambda_t)_t$ with the reward $R(\Lambda_t, a_t^n)$ form an infinite horizon MDP and so the optimal pure strategy can be derived from the following FPE for the state $\Lambda = (x^n, n_a, \pi, \boldsymbol{b})$,*

$$a^{*n} = \gamma^*(x^n) = \underset{a^n \in \mathcal{A}^n(b^n, n_a)}{\arg\max} \left\{ a^n \sum_v v\mu^n(v) + \delta \mathbb{E}\left[V^n(x^n, N_a', \Pi, B)|x^n, n_a, \pi, b, a^n\right] \right\}, \tag{B.8a}$$

where n_a denotes the acting player and N_a', Π and B are random variables for the next state elements and the expectation is according to the transition kernels (B.7). *Furthermore,*

$$V^n(x^n, n_a, \pi, b) = \max_{a^n \in \mathcal{A}^n(b^n, n_a)} \left\{ a^n \sum_v v\mu^n(v) + \delta \mathbb{E}\left[V^n(x^n, N_a', \Pi, B)|x^n, n_a, \pi, b, a^n\right] \right\}. \tag{B.8b}$$

Next, we need to show that the above FPE is equivalent to FPE 1. We first show that for all $x^n, n, \pi, \boldsymbol{b}^{-n}$,

$$V^n(x^n, n_a, \pi, b^n = 1, \boldsymbol{b}^{-n}) = 0.$$

According to the action space defined in (3.7)*, if $b^n = 1$, $\mathcal{A}^n(b^n, n_a) = \{0\}$. This means that the instantaneous reward at this state is* 0*. On the other hand, according to the transition kernel of $\boldsymbol{b}$ in* (B.7)*, this state is absorbing in terms of b^n, which means that $b^n = 1$ for all future states too. This will cause player n to have 0 rewards in all of the upcoming states and so $V^n(x^n, n_a, \pi, b^n = 1, \boldsymbol{b}^{-n}) = 0$. The above implies that player n faces a stopping time problem.*

If n is the acting player ($n = n_a$), then FPE (B.8) *is indeed choosing between buying and* *getting the instantaneous reward* $\sum_v v\mu^n(v)$*, or waiting and getting*

$$\delta\mathbb{E}\left[V^n(x^n, N_a', \Pi, B)|x^n, n_a, \pi, \boldsymbol{b}, a^n\right] = \frac{\delta}{N}\sum_{n_a'=1}^{N} V^n\left(x^n, n_a', F\left(\pi, \gamma^*, 0, n\right), \boldsymbol{b}\right)$$

using the transition kernels in (B.7)*. Hence, for* $n = n_a$*, FPE* (B.8) *is equivalent to* (3.21a) *and the first three cases of* (3.21b)*.*

If n *is not the acting player (*$n \neq n_a$*), since* $\mathcal{A}^n(b^n, n_a) = \{0\}$ *then*

$$V^n(x^n, n_a, \pi, \boldsymbol{b}) = \delta\mathbb{E}\left[V^n(x^n, N_a', \Pi, B)|x^n, n_a, \pi, \boldsymbol{b}, a^n\right].$$

According the transition kernels (B.7)*,*

$$\delta\mathbb{E}\left[V^n(x^n, N_a', \Pi, B)|x^n, n_a, \pi, \boldsymbol{b}, a^n\right] = \frac{\delta}{N}\sum_{n_a'=1}^{N}\mathbb{E}\left[V^n(x^n, n_a', \Pi, (B^{n_a}, \boldsymbol{b}^{-n_a}))|x^n, n_a, \pi, \boldsymbol{b}, a^n\right],$$

and $\Pi = F(\pi, \gamma^*, \gamma^*(X^{n_a}), n_a)$ *with probability 1. Thus,*

$$V^n(x^n, n_a, \pi, \boldsymbol{b}) = \frac{\delta}{N}\sum_{n_a'=1}^{N}\mathbb{E}\{V^n(x^n, n_a', F(\pi, \gamma^*, \gamma^*(X^{n_a}), n_a), B^{n_a}\boldsymbol{b}^{-n_a})|x^n, n_a, \pi, \boldsymbol{b}, a^n\},$$

which is the fourth case of (3.21b)*. Next, note that the transition kernel of* B^{n_a} *in* (B.7b) *is the same is in* (3.21c)*. It is now a simple task to construct the PBE by the forward algorithm in* (3.22) *following each information set recursively (we are also using the fact that the private variables* $X^1, \ldots, X^N$ *are independent conditioned on* V*, as shown in Lemma 2). The proof is completed by showing that* π *can be computed using* $\widetilde{\boldsymbol{x}}$*. In particular, using* (3.20) *in Lemma 2, and* (3.16) *in* (3.21a) *we substitute*

$$\pi(1\,|\,x^n) = \frac{q^{\sum_m \widetilde{x}^m - \widetilde{x}^n + x^n}}{1 + q^{\sum_m \widetilde{x}^m - \widetilde{x}^n + x^n}}. \tag{B.9}$$

Similarly, using (3.18) *in Lemma 2, in* (3.21e) *we get* (3.21d)*.*

B.3 Computing a PBE though a polynomial-dimensional FPE

Owing to the symmetry of the problem we define the set $\mathcal{K} = \{00, -10, 01, -11, +11\}$ where the elements of this set are all possible values that the pair $\tilde{x}^i b^i$ can take for each player i. Note that $+10$ can never happen under any strategy so it is not included in the set. So players are grouped into 5 groups according to their value of the pair $\tilde{x}^i b^i$. We define the joint type (scaled empirical distribution), $t_{\tilde{\boldsymbol{x}}\boldsymbol{b}}$ of the sequence $(\tilde{\boldsymbol{x}}, \boldsymbol{b})$ as

$$t_{\tilde{\boldsymbol{x}}\boldsymbol{b}}(k) = \sum_{i=1}^{N} \mathbf{1}_{\tilde{x}^i b^i}(k), \qquad \forall k \in \mathcal{K}. \tag{B.10}$$

Clearly for every type $\boldsymbol{t}$, $t(k) \geq 0$ and $\sum_{k\in\mathcal{K}} t(k) = N$, so there are exactly $\binom{N+4}{4} \sim N^4$ such possible types.

Note that with the above definition, the aggregate state information $y = \sum_{i=1}^{N} \tilde{x}_i$ equals to $y = t(+11) - t(-10) - t(-11)$.

We define the following functions $U_a : \mathcal{X} \times \mathcal{K} \times \mathcal{T} \to \mathbb{R}$, and $U_{na}^l : \mathcal{X} \times \mathcal{K} \times \mathcal{T} \to \mathbb{R}$ for all $l \in \mathcal{K}$. The meaning of these functions is as follows. $U_a(x, k, \boldsymbol{t})$ denotes the value function of the acting player n whose private information $x^n = x$, her pair $\tilde{x}^n b^n = k$ (and so she belongs to group k) and the joint type of the sequence $(\tilde{\boldsymbol{x}}, \boldsymbol{b})$ is t. Similarly, $U_{na}^l(x, k, \boldsymbol{t})$ denotes the value function of a non-acting player m whose private information $x^m = x$, her pair $\tilde{x}^m b^m = l$ (and so she belongs to group l), with an acting player n whose pair $\tilde{x}^n b^n = k$ (i.e., belonging to group k), and the joint type of the sequence $(\tilde{\boldsymbol{x}}, \boldsymbol{b})$ is $\boldsymbol{t}$.

Finally we define the update functions g^x, g^b, and g^t as follows

$$g^x(k_x, \gamma, a) = \begin{cases} 2a - 1 & \text{, if } k_x = 0 \text{ and } \gamma = \boldsymbol{I} \\ k_x & \text{, else} \end{cases}, \tag{B.11a}$$

$$g^b(k_b, a) = \begin{cases} a & \text{, if } k_b = 0 \\ k_b & \text{, else} \end{cases}, \tag{B.11b}$$

$$g^{\boldsymbol{t}}(k, \boldsymbol{t}, \gamma, a)(k') = \begin{cases} t(k') - 1 & \text{, if } k' = k \text{ and } g^{xb}(k, \gamma, a) \neq k \\ t(k') + 1 & \text{, if } k' = g^{xb}(k, \gamma, a) \text{ and } g^{xb}(k, \gamma, a) \neq k \\ t(k') & \text{, else} \end{cases} \tag{B.11c}$$

where we use the notation $k = k_x k_b$ to decompose the two parts of the k index, and with the understanding that we also use the notation g^{efg} to denote (g^e, g^f, g^g) for any $e, f, g \in \{x, b, t\}$.

We consider the following FP equation in FPE 3.

Fixed-Point Equation 3 (Polynomial dimension)**.** *For every $k = k_x k_b \in \mathcal{K}$, $\boldsymbol{t} \in \mathcal{T}$ we evaluate $\gamma^* = \phi[k, \boldsymbol{t}]$ as follows.*

- *If $k_b = 1$ then $\gamma^* = \boldsymbol{0}$.*

- *If* $k_b = 0$ *then* γ^* *is the solution of the following system of equations*

$$\gamma^*(x) = \arg\max\{ \underbrace{A}_{0=\textit{don't buy}}, \underbrace{\frac{q^{y+x\boldsymbol{I}_0(k_x)}-1}{q^{y+x\boldsymbol{I}_0(k_x)}+1}}_{1=\textit{buy}} \} \quad \forall x \in \mathcal{X}, \tag{B.12a}$$

where

$$A = \frac{\delta}{N} U_a(x, g^{xb\boldsymbol{t}}(k, \boldsymbol{t}, \gamma^*, 0)) + \frac{\delta}{N} \sum_{k' \in \mathcal{K}} [t(k') - \boldsymbol{I}_k(k')] U_{na}^{g^{xb}(k,\gamma^*,0)}(x, k', g^{\boldsymbol{t}}(k, \boldsymbol{t}, \gamma^*, 0))] \tag{B.12b}$$

where the value functions satisfy

$$U_a(x, k, \boldsymbol{t}) = \begin{cases} 0, & \textit{if } k_b = 1 \\ A, & \textit{if } k_b = 0, \gamma^*(x) = 0 \\ \frac{q^{y+x\boldsymbol{I}_0(k_x)}-1}{q^{y+x\boldsymbol{I}_0(k_x)}+1}, & \textit{if } k_b = 0, \gamma^*(x) = 1 \end{cases}, \tag{B.12c}$$

and for all $l = l_x l_b \in \mathcal{K}$

$$\begin{aligned} &U_{na}^l(x, k, \boldsymbol{t}) \\ &= \begin{cases} 0, & \textit{if } l_b = 1 \\ \frac{\delta}{N}\mathbb{E}[U_a(x, l, g^{\boldsymbol{t}}(k, \boldsymbol{t}, \gamma^*, \gamma^*(X^n)))] + \frac{\delta}{N}\mathbb{E}[U_{na}^l(x, g^{xb\boldsymbol{t}}(k, \boldsymbol{t}, \gamma^*, \gamma^*(X^n)))] & \\ + \frac{\delta}{N} \sum_{k' \in \mathcal{K}} [t(k') - \boldsymbol{I}_k(k') - \boldsymbol{I}_l(k')] \mathbb{E}[U_{na}^l(x, k', g^{\boldsymbol{t}}(k, \boldsymbol{t}, \gamma^*, \gamma^*(X^n)))], & \textit{if } l_b = 0, \end{cases} \end{aligned} \tag{B.12d}$$

where expectation in the last equation is w.r.t. the RV X^n *where*

$$P(X^n = x^n | l, x, k, \boldsymbol{t}) = \begin{cases} \boldsymbol{I}_{k_x}(x^n), & \textit{if } k_x \neq 0 \\ \frac{Q(x^n|-1) + Q(x^n|1) q^{y+x\boldsymbol{I}_0(l_x)}}{q^{y+x\boldsymbol{I}_0(l_x)}+1}, & \textit{else.} \end{cases} \tag{B.12e}$$

■

We will now show that if the above FPE has a solution U^*, then the original FPE has a solution V^* where V^* can be readily derived from U^*.

Given the solution U^* of the above FP equation (together with the strategy ϕ) we construct the following strategies and value functions.

$$\gamma^* = \theta[n, \tilde{\boldsymbol{x}}, \boldsymbol{b}] = \phi[\tilde{x}^n b^n, \boldsymbol{t}_{\tilde{\boldsymbol{x}},\boldsymbol{b}}] \tag{B.13a}$$

$$\tilde{V}^m(\cdot, n, \tilde{\boldsymbol{x}}, \boldsymbol{b}) = \begin{cases} U_a(\cdot, \tilde{x}^n b^n, \boldsymbol{t}_{\tilde{x},b}), & \text{if } m = n \\ U_{na}^{\tilde{x}^m b^m}(\cdot, \tilde{x}^n b^n, \boldsymbol{t}_{\tilde{\boldsymbol{x}},\boldsymbol{b}}), & \text{if } m \neq n \end{cases}. \tag{B.13b}$$

We will show that these value functions are solutions of the original FPE 1.

Theorem 9. *The value functions $(\tilde{V}^m)_{m\in\mathcal{N}}$ together with the strategy mapping $\gamma^* = \phi[\cdot]$ satisfy FPE 1.*

Proof. *Fix n, $\tilde{\boldsymbol{x}}$, and $\boldsymbol{b}$ that result in a type $\boldsymbol{t}$ with accumulated state y. The acting player n belongs to a group $k = k_x k_b = \tilde{x}^n b^n$. If $b_n = 1$ then $k_b = 1$ and $\gamma^* = \boldsymbol{0}$. If $b_n = 0$ then it is clear that the second term in* (3.21a) *becomes $\frac{q^{y+x^n \boldsymbol{1}_0(k_x)}-1}{q^{y+x^n \boldsymbol{1}_0(k_x)}+1}$, which is exactly the same as the second term in* (B.12b) *(with $x^n = x$). Consider the first term in* (3.21a)*. The new group of the acting player n is $\widehat{k} = (f(\tilde{x}^n, \gamma^*, 0), 0) = g^{xb}(k, \gamma^*, 0)$ and the new value for the overall type will change to $\widehat{\boldsymbol{t}} = g^{\boldsymbol{t}}(k, \boldsymbol{t}, \gamma^*, 0)$. The implication of the above is that the first term in* (3.21a) *will be*

$$\begin{aligned}
&\sum_{n'=1}^{N} \tilde{V}^n(x^n, n', \left(\tilde{\boldsymbol{x}}^{-n}, f(\tilde{x}^n, \gamma^*, 0)\right) (\boldsymbol{b}^{-n}, 0)) \\
&= \tilde{V}^n(x^n, n, \left(\tilde{\boldsymbol{x}}^{-n}, f(\tilde{x}^n, \gamma^*, 0)\right), (\boldsymbol{b}^{-n}, 0)) + \sum_{n'=1, n'\neq n}^{N} \tilde{V}^n(x^n, n', \left(\tilde{\boldsymbol{x}}^{-\boldsymbol{n}}, f(\tilde{x}^n, \gamma^*, 0)\right), (\boldsymbol{b}^{-n}, 0)) \\
&= U_a(x^n, \widehat{k}, \widehat{\boldsymbol{t}}) + \sum_{n'=1, n'\neq n}^{N} U_{na}^{\widehat{k}}(x^n, \tilde{x}^{n'} b^{n'}, \widehat{\boldsymbol{t}}) \\
&= U_a(x^n, \widehat{k}, \widehat{\boldsymbol{t}}) + \sum_{k'\in\mathcal{K}} \sum_{n'=1, n'\neq n, \tilde{x}^{n'} b^{n'} = k'}^{N} U_{na}^{\widehat{k}}(x^n, \tilde{x}^{n'} b^{n'}, \widehat{\boldsymbol{t}}) \\
&= U_a(x^n, g^{xb\boldsymbol{t}}(k, \boldsymbol{t}, \gamma^*, 0)) + \sum_{k'\in\mathcal{K}} [t(k') - \boldsymbol{1}_k(k')] U_{na}^{g^{xb}(k,\gamma^*,0)}(x^n, k', g^{\boldsymbol{t}}(k, \boldsymbol{t}, \gamma^*, 0)),
\end{aligned} \tag{B.14}$$

where the term $t(k') - \boldsymbol{1}_k(k')$ enumerates all players $n' \neq n$ in the vector $(\tilde{\boldsymbol{x}}^{-\boldsymbol{n}}, f(\tilde{x}^n, \gamma^*, 0)), (\boldsymbol{b}^{-n}, 0)$ which are given by the original type $\boldsymbol{t}$ subtracting one from the group of the acting player. This is exactly the expression in (B.12b) *and thus* (3.21a) *is satisfied.*

Now consider (3.21b)*. Fix m and denote the group of the m-th player by $l = l_x l_b = \tilde{x}^m b^m$. The first three branches of this equation are obviously satisfied. Regarding the fourth branch we know* *that the new group of the acting player n will be $\widehat{K} = f(\tilde{x}^n, \gamma^*, \gamma^*(X^n))$, $B'^n = g^b(k_b, \gamma^*(X^n))$ and the new type will be $\widehat{T} = g^{\boldsymbol{t}}(k, \boldsymbol{t}, \gamma^*, \gamma^*(X^n))$. The left-hand side of* (3.21b) *becomes* $U^l_{na}(x^m, k, \boldsymbol{t})$ *with $l_b = 0$. The right-hand side becomes*

$$\begin{aligned}
&\sum_{n'=1}^{N}\mathbb{E}\left[V^m\left(x^m, n', \left(\tilde{\boldsymbol{x}}^{-\boldsymbol{n}}, f\left(\tilde{x}^n, \gamma^*, \gamma^*\left(X^n\right)\right)\right), \left(\boldsymbol{b}^{-n}, B'^n\right)\right)\right]\\
&\quad = \mathbb{E}\left[V^m\left(x^m, m, \left(\tilde{\boldsymbol{x}}^{-\boldsymbol{n}}, f\left(\tilde{x}^n, \gamma^*, \gamma^*\left(X^n\right)\right)\right), \left(\boldsymbol{b}^{-n}, B'^n\right)\right)\right]\\
&\quad + \mathbb{E}\left[V^m\left(x^m, n, \left(\tilde{\boldsymbol{x}}^{-\boldsymbol{n}}, f\left(\tilde{x}^n, \gamma^*, \gamma^*\left(X^n\right)\right)\right), \left(\boldsymbol{b}^{-n}, B'^n\right)\right)\right]\\
&\quad + \sum_{n'=1, n'\neq m,n}^{N} \mathbb{E}\left[V^m\left(x^m, n', \left(\tilde{\boldsymbol{x}}^{-\boldsymbol{n}}, f\left(\tilde{x}^n, \gamma^*, \gamma^*\left(X^n\right)\right)\right), \left(\boldsymbol{b}^{-n}, B'^n\right)\right)\right]\\
&\quad = \mathbb{E}[U_a(x^m, l, \widehat{T})] + \mathbb{E}[U^l_{na}(x^m, \widehat{K}, \widehat{T})] + \sum_{k'\in\mathcal{K}} \sum_{n'=1, n'\neq n,m,\ \tilde{x}^{n'}b^{n'}=k'}^{N} \mathbb{E}[U^l_{na}(x^m, \tilde{x}^{n'}b^{n'}, \widehat{T})]\\
&\quad = \mathbb{E}[U_a(x^m, l, g^{\boldsymbol{t}}(k, \boldsymbol{t}, \gamma^*, \gamma^*(X^n)))] + \mathbb{E}[U^l_{na}(x^m, g^{xb\boldsymbol{t}}(k, \boldsymbol{t}, \gamma^*, \gamma^*(X^n)))]\\
&\quad + \sum_{k'\in\mathcal{K}} [t(k') - \boldsymbol{1}_k(k') - \boldsymbol{1}_l(k')]\mathbb{E}[U^l_{na}(x^m, k', g^{\boldsymbol{t}}(k, \boldsymbol{t}, \gamma^*, \gamma^*(X^n)))].
\end{aligned} \tag{B.15}$$

This is exactly the expression in (B.12d) *and thus* (3.21b) *is satisfied.*

We remark at this point that this method can be generalized for heterogeneous players with different values of δ. All is needed is to consider joint types of the vectors $\tilde{\boldsymbol{x}}, \boldsymbol{b}, \delta$. The corresponding dimensionality of the FP equation will be $\sim N^{4K_\delta}$ where K_δ is the number of different types of δ.

B.4 Proof of Theorem 2

Proof. *Fix n, $\tilde{\boldsymbol{x}}$ that results in population parameters y and w. The acting player has either not revealed her information ($\tilde{x}^n = 0$) or she has revealed a bad signal ($\tilde{x}^n = -1$), since otherwise she would have already bought the product and must play $a^n = 0$. This implies that $\tilde{x}^n = -r$. It is then clear that the first term in* (3.21a) *becomes $\frac{q^{y+r+x^n}-1}{q^{y+r+x^n}+1}$, which is exactly the same as the first term in* (3.26a) *(with $x^n = x$). Consider the second term in* (3.21a)*. The new parameter of the acting player n is $\widehat{r} = |f(\tilde{x}^n, \gamma^*, 0)| = G^r(r, \gamma^*)$. Define the new population parameters by $\widehat{y} = G^y\left(r, y, \gamma^*, 0\right)$ and $\widehat{w} = G^w\left(r, w, \gamma^*, 0\right)$. The implication of the above is that the second term in* (3.21a) *will be (apart for the δ/N factor)*

$$\sum_{n'=1}^{N} \tilde{V}^n\left(x^n, n', \tilde{\boldsymbol{x}}^{-n} f(\tilde{x}^n, \gamma^*, 0), (\boldsymbol{b}^{-n}, 0)\right) \tag{B.16a}$$

$$= \tilde{V}^n \left(x^n, n, \tilde{\boldsymbol{x}}^{-n} f\left(\tilde{x}^n, \gamma^*, 0\right), \left(\boldsymbol{b}^{-n}, 0\right)\right) + \sum_{n'=1, n'\neq n}^{N} \tilde{V}^n \left(x^n, n', \tilde{\boldsymbol{x}}^{-n} f\left(\tilde{x}^n, \gamma^*, 0\right), \left(\boldsymbol{b}^{-n}, 0\right)\right) \tag{B.16b}$$

$$= U_a\left(x^n, \widehat{r}, \widehat{y}, \widehat{w}\right) + \sum_{n'=1, n'\neq n}^{N} U_{na}^{\widehat{r}} \left(x^n, z^{n'}, \widehat{y}, \widehat{w}\right) \tag{B.16c}$$

$$= U_a\left(x^n, \widehat{r}, \widehat{y}, \widehat{w}\right) + \sum_{n'=1, n'\neq n, z^{n'}=0}^{N} U_{na}^{\widehat{r}}\left(x^n, 0, \widehat{y}, \widehat{w}\right) + \sum_{n'=1, n'\neq n, z^{n'}=1}^{N} U_{na}^{\widehat{r}}\left(x^n, 1, \widehat{y}, \widehat{w}\right) \tag{B.16d}$$

$$\begin{aligned}
&= U_a\left(x^n, \widehat{r}, \widehat{y}, \widehat{w}\right) + \left(N - w - 1 + r\right) U_{na}^{\widehat{r}}\left(x^n, 0, \widehat{y}, \widehat{w}\right) + \left(w - r\right) U_{na}^{\widehat{r}}\left(x^n, 1, \widehat{y}, \widehat{w}\right) \\
&= U_a\left(x^n, G^{ryw}(r, y, w, \gamma^*, 0)\right) + \left(N - w - 1 + r\right) U_{na}^{G^r(r,\gamma^*)}\left(x^n, 0, G^{yw}(r, y, w, \gamma^*, 0)\right) \\
&\quad + \left(w - r\right) U_{na}^{G^r(r,\gamma^*)}\left(x^n, 1, G^{yw}(r, y, w, \gamma^*, 0)\right).
\end{aligned} \tag{B.16e}$$

This is exactly the expression in (3.26b) *so* (3.21a) *is satisfied. Now consider* (3.21b)*. Fix* m *and denote the parameter of the* m*-th player by* $\tilde{r} = |\tilde{x}^m|$*. The first three branches of this equation are obviously satisfied. Regarding the fourth branch we know that the new parameter of the acting player* n *will be* $\widehat{Z} = G^z(z, \gamma^*, \gamma^*(X^n))$ *and the new population parameters will be* $(\widehat{Y}, \widehat{W}) = G^{yw}(z, y, w, \gamma^*, \gamma^*(X^n))$*. The left-hand side of* (3.21b) *becomes* $U_{na}^{\tilde{r}}(x^m, z, y, w)$*. The right-hand side is as follows.*

$$\begin{aligned}
&\sum_{n'=1}^{N} \mathbb{E}\left[V^m\left(x^m, n', \tilde{x}^{-n} f\left(\tilde{x}^n, \gamma^*, \gamma^*\left(X^n\right)\right), \left(\boldsymbol{b}^{-n}, B'^n\right)\right)\right] \\
&\quad = \mathbb{E}\left[V^m\left(x^m, m, \tilde{x}^{-n} f\left(\tilde{x}^n, \gamma^*, \gamma^*\left(X^n\right)\right), \left(\boldsymbol{b}^{-n}, B'^n\right)\right)\right] \\
&\qquad + \mathbb{E}\left[V^m\left(x^m, n, \tilde{x}^{-n} f\left(\tilde{x}^n, \gamma^*, \gamma^*\left(X^n\right)\right), \boldsymbol{b}^{-n} B'^n\right)\right] \\
&\qquad + \sum_{n'=1, n'\neq m, n}^{N} \mathbb{E}\left[V^m\left(x^m, n', \tilde{x}^{-n} f\left(\tilde{x}^n, \gamma^*, \gamma^*\left(X^n\right)\right), \left(\boldsymbol{b}^{-n}, B'^n\right)\right)\right] \\
&\quad = \mathbb{E}[U_a(x^m, \tilde{r}, \widehat{Y}, \widehat{W})] + \mathbb{E}[U_{na}^{\tilde{r}}(x^m, \widehat{Z}, \widehat{Y}, \widehat{W})] + \sum_{n'=1, n'\neq n, m,\ z^{n'}=1}^{N} \mathbb{E}[U_{na}^{\tilde{r}}(x^m, 1, \widehat{Y}, \widehat{W})] \\
&\qquad + \sum_{n'=1, n'\neq n, m,\ z^{n'}=0}^{N} \mathbb{E}[U_{na}^{\tilde{r}}(x^m, 0, \widehat{Y}, \widehat{W})] \\
&\quad = \mathbb{E}[U_a(x^m, \tilde{r}, G^{yw}(z, y, w, \gamma^*, \gamma^*(X^n)))] + \mathbb{E}[U_{na}^{\tilde{r}}(x^m, G^{zyw}(z, y, w, \gamma^*, \gamma^*(X^n)))] \\
&\qquad + (w - z - \tilde{r})\mathbb{E}[U_{na}^{\tilde{r}}(x^m, 1, G^{yw}(z, y, w, \gamma^*, \gamma^*(X^n)))]
\end{aligned}$$

$$+ (N - w - 2 + z + \tilde{r})\mathbb{E}[U_{na}^{\tilde{r}}(x^m, 0, G^{yw}(z, y, w, \gamma^*, \gamma^*(X^n)))]. \tag{B.17}$$

This is exactly the expression in (3.26g) *and thus* (3.21b) *is satisfied.*

B.5 Proof of Lemma 3

First, we show that whenever $\gamma^* = \phi[0, y, w] = \mathbf{0}$, the valuation functions are all 0 and we must have $\gamma^* = \phi[1, y, w] = \mathbf{0}$. According to FPE 2, at the state (x, r, y, w) we have

$$A = \frac{\delta}{N} U_a (x, r, y, w) + \frac{\delta}{N} (N - w - 1 + z) U_{na}^r (x, 0, y, w) + \frac{\delta}{N} (w - z) U_{na}^r (x, 1, y, w),$$

where for both $\tilde{z} = 0, 1$,

$$\begin{aligned} U_{na}^r (x, \tilde{z}, y, w) = {} & \frac{\delta}{N} U_a (x, r, y, w) + \frac{\delta}{N} (N - w - 1 + z) U_{na}^r (x, 0, y, w) \\ & + \frac{\delta}{N} (w - z) U_{na}^r (x, 1, y, w). \end{aligned}$$

and since $\gamma^* = \phi[r, y, w] = \mathbf{0}$, we should have $U_a (x, r, y, w) = A$. Therefore, we can solve for $U_a (x, r, y, w)$, $U_{na}^r (x, 0, y, w)$, $U_{na}^r (x, 1, y, w)$ and A in above equations. It is easy to see that the solution for all of these quantities is 0 and hence, $A = 0$. Therefore, $U_a (x, r, y, w) = 0$ and it is obvious that for $y < -2$, players strictly prefer to wait since the expected value of instantaneous reward is negative and they prefer to get A, which is 0. Further, for $y = -2$, players with $r = 0$ or $r = 1$ and $x = -1$ strictly prefer to wait. A player with $r = 1$ and $x = 1$ is indifferent between buying and not buying. The reason is that a player with $r = 1$ that gets to play again, must have revealed $x = -1$. Therefore, if $x = 1$, the player is at an off-equilibrium point and according to equation (3.16), she forms her true belief by canceling out what she has revealed and then augmenting the belief by her private information. In terms of y, this is translated into using $y - \tilde{x} + x$ to form the belief over V. For $y = -2$, a player with $r = 1$ and $x = 1$ uses $y - (-1) + 1 = 0$ to form her belief over V. Thus, the expected value of her instantaneous reward is 0. It completes the proof of the third part of the theorem.

Now consider $\delta = 1$. Assume that $\gamma^* = \phi[r, y, w]$ is a solution of FPE 2. According to γ^*, define player n's terminating states to be the (r, y, w) values for which player n decides to either buy the product (and leave the game) or to not buy it ever after, i.e., playing $\gamma^* = \mathbf{0}$ when everyone else is (which means that the player practically leaves the game). At each state of the game, γ^* imposes a probability distribution on the future terminating states of the game that are reached by not buying decision of the acting player. So at each state, a player compares the expected value of

V with the expected valuation she can get in future by not buying, which is an average between the value of V at the terminating states where she buys the product, and zero, corresponding to the terminating states she decides not to buy the product ever.

Lemma 5. *Assume that according to $\gamma^* = \phi[r, y, w]$, the acting player will buy the product in all of the future terminating states (the ones with positive probability to happen if the acting player decided not to buy). Then, the player is indifferent between buying and not buying the product for $\delta = 1$. Otherwise, she strictly prefers to wait for $\delta = 1$.*

Proof. *According to FPE 2, the acting player compares the expected value of v with the average of expected values of v in future terminating states with positive probability to happen if the acting player decided not to buy. More formally, if we denote the current state by s and the future terminating states that happen with positive probability if the player does not buy with $s_1, ..., s_k$, then we have*

$$\gamma^*(x) = \arg\max \left\{ \sum_{j=1}^{k} \mathbb{E}[v|s_j]p(s_j|s), \mathbb{E}[v|s] \right\} \tag{B.18}$$

By the law of total expectation, we know that the above terms are always equal to each other, no matter what the states $s_1, ...s_k$ are and with what probability they happen. The only requirement is that at all of the states $s_1, ...s_k$ the player decides to buy the product. Therefore, if a player finds herself in a state s that could lead to terminating states $s_1, ...s_k$ (by not buying), in all of which she will decide to buy the product (according to γ^*), she is in fact indifferent between buying and not buying at the current state for $\delta = 1$.

Next, assume that at one of the terminating states $s_1, ...s_k$, let's say s_j, the player strictly prefer not to buy the product, she will receive zero valuation at s_j and therefore, the expected value of v should have been negative at s_j. Hence, by substituting zero instead of $\mathbb{E}[v|s_j]$ in equation (B.18)*, we get a term that is greater than $\mathbb{E}[v|s]$. It means that the expected valuation of not buying is greater than the expected value of v at the current state s which implies that the player strictly prefers not to buy the product. The same argument holds if in more than one future terminating states the player strictly prefers not to buy.*

Lemma 6. *Assume that according to γ^*, we know that for state s, there exists at least one future terminating state s_j (with positive probability to happen if the acting player does not buy the product), at which the acting player strictly prefers not to buy the product, then she strictly prefers to wait at the state s for large enough $\delta \leq 1$.*

Proof. *According to the proof of the second part of Lemma 5, for $\delta = 1$, the acting player strictly prefers to wait. Hence, there exists large enough $\delta < 1$ for which the acting player still strictly prefers to wait.*

Next, we prove the first two parts of the theorem. We first characterize the equilibrium strategies for $w = N$. It is evident what the equilibrium is for $w = N$, because all of the states are absorbing and players act based on the expected instantaneous reward. For any value of δ, we have $\gamma^* = \phi[1, y, N] = \mathbf{0}$ for[1] $y \leq -2$, $\gamma^* = \phi[1, y, N] = \boldsymbol{I}$ for $y = -1$ and $\gamma^* = \phi[1, y, N] = \mathbf{1}$ for $y \geq 0$. We also know that $\gamma^* = \phi[r, y, w] = \mathbf{0}$ for $y \leq -2$. In order to prove the theorem, we investigate the terminating states for all states of the game with $y \geq -1$. Since $\gamma^* = \mathbf{0}$ is played at $y = -2$, all states with $y = -2$ are absorbing. Hence, no state with $y < -2$ is reachable from states with $y \geq -1$. Therefore, all of the terminating states have $y \geq -2$. On the other hand, if $\gamma^* = \boldsymbol{I}$ at the current state, the acting player with $x = 1$ can reach her terminating states only when she has $r = 1$. A player with $x = 1$ and $r = 1$ is indifferent between buying and not buying at $y = -2$ (she is on an off-equilibrium path) and prefers to buy at all states with $y \geq -1$. Therefore, at all of the terminating states the player prefers to buy and according to Lemma 5, she is indifferent between buying and not buying at the current state. Furthermore, if $\gamma^* = \mathbf{1}$ at the current state, the acting player should be indifferent between buying and not buying according to Lemma 5 (the player should either be indifferent or strictly prefer to wait. The latter is impossible due to the strategy $\gamma^* = \mathbf{1}$). It means that a player with $x = 1$ is indifferent between buying and not buying for all states with $y \geq -1$ and all w. It implies that for $\delta < 1$ a player with $x = 1$ strictly prefers to buy if her instantaneous reward is positive, i.e., $y \geq 0$ and is indifferent if her instantaneous reward is 0, i.e., $y = -1$.

Next consider the players with $x = -1$. If at all of the terminating states of a player with $x = -1$ we have $y \geq 1$ or $y = 0$ and $w = N$ (the states in which she prefers to buy the product), then this player should be indifferent between buying and not buying at the current state. Assume that we have $\gamma^* = \phi[r, y, w] = \boldsymbol{I}$ for every r, y, w (this strategy profile shows us the biggest set of approachable terminating states from each state s, although it may not be the solution). It is evident that for $y + w \geq N$ and all r, all of the terminating states that are approachable have $y \geq 1$ or $y = 0$ and $w = N$ (at each state (r, y, w), the player can move to $y - 1$ and $w + 1$ by playing $\gamma^* = \boldsymbol{I}$). Hence, for $\delta = 1$, a player with $x = -1$ is indifferent between buying and not buying for $y + w \geq N$ and all r. It implies that for $\delta < 1$, a player with $x = -1$ strictly prefers to buy for $y + w \geq N$ if her instantaneous reward is positive, i.e., $y \geq 2$ or $y = 1$ and $r = 1$, and is indifferent if her instantaneous reward is 0, i.e., $y = 1$ and $r = 0$ or $y = 0$ and $r = 1$.

[1] Note that we can also have $\gamma^* = \phi[1, y, N] = \boldsymbol{I}$ for $y = -2$ due to the tie for the player with $x = 1$.

B.6 Proof of Theorem 3

We prove this theorem by referring to Lemma 3. The strategy profile proposed for $y \leq -2$ is an evident solution of FPE 2 due to the fact that both types of players with $r = 0$ prefer not to buy and hence they play $\gamma^* = \mathbf{0}$. This implies that both types of players with $r = 1$ will also play $\gamma^* = \mathbf{0}$. For $y \geq -1$, a player with $x = 1$ is either indifferent or prefers to buy for all $\delta \leq 1$ and therefore, she can decide to buy for $-1 \leq y \leq 1$. Furthermore, a player with $x = -1$ and $r = 0$ always prefers to wait or is indifferent for $-1 \leq y \leq 1$ (her expected instantaneous reward is either negative or zero) and so she can decide to wait for $-1 \leq y \leq 1$. The same argument holds for a player with $x = -1$ and $r = 1$ for $-1 \leq y \leq 0$. For $y = 1$, a player with $x = -1$ and $r = 1$ has positive expected instantaneous reward and so whether she prefers to wait or to buy depends on δ. Since we know that the action of a player with $x = 1$ and $r = 1$ at $y = 1$ is buying, the strategy at $y = 1$ and $r = 1$ should be either of $\mathbf{1}$ or $\boldsymbol{I}$. Notice that this strategy does not affect the decision of players at other states (it can not be reached from states with $y > 1$ and the solution for $y < 1$ does not depend on what is played at $y = 1$, as we just proved). Hence, it can be determined independently based on FPE 2. We next have to prove that the strategy $\gamma^* = \phi[r, y, w] = \mathbf{1}$ is a solution for $y \geq 2$ and all w and r. According to Lemma 3, if the strategy profile is $\gamma^* = \phi[r, y, w] = \mathbf{1}$ for some state of the game $s = (r, y, w)$ and we have $\gamma^* = \phi[r, y, w'] = \mathbf{1}$ for all $w' > w$ (which is the case in the suggested strategy profile in this theorem), then in all of the terminating states (see the proof of Lemma 3) that are reachable from s, the acting player buys the product. Hence, the player is either indifferent ($\delta = 1$) or strictly prefers to buy ($\delta < 1$) and it completes the proof of the strategy $\gamma^* = \phi[r, y, w] = \mathbf{1}$ being a solution for $y \geq 2$, all w, r and all $\delta \leq 1$.

B.7 Proof of Theorem 4

According to Lemma 3, the solution is evident for $y \leq -2$, for both $\delta = 1$ and $\delta < 1$. We also know the solution for $w = N$ and all δ according to the proof of Lemma 3. For $w = N$, which implies that $r = 1$, we must have $\gamma^* = \phi[r, y, w] = \mathbf{1}$ for $y \geq 1$. Further, since for $y = 0, 1$, the expected instantaneous reward for players with $x = 1$ and $x = -1$ is positive and non-positive, respectively, we can have $\gamma^* = \phi[r, y, w] = \boldsymbol{I}$ as the solution. This proves the first, third and fourth part of $\delta = 1$ case and first and fifth part ($w = N$) of $\delta < 1$ case.

Next, consider the second part of $\delta = 1$ case. According to Lemma 3, a player with $x = 1$ and all r and $w < N$ is indifferent between buying and not buying for $y \geq -1$ and hence, she can decide to buy. On the other hand, a player with $x = -1$ is indifferent for $y + w \geq N$ and so she can decide not to buy for these states. If the proposed strategy is the solution of FPE 2, then from the states with $y + w < N$, a player with $x = -1$ can reach the terminating states with negative y (-1 or -2), in which she strictly prefers not to buy (this is evident by tracing the states that can be reached by going from y, w to $y-1, w+1$ by each revelation through the strategy $\gamma^* = \phi[r, y, w] = \boldsymbol{I}$). Therefore, for $\delta = 1$ a player with $x = -1$ strictly prefers to wait for $y + w < N$ and so strategy $\gamma^* = \phi[r, y, w] = \boldsymbol{I}$ can be a solution for $y \geq -1$ and $w < N$. This completes the proof of the $\delta = 1$ case.

Now consider $\delta < 1$. With the same arguments as in the $\delta = 1$ case, since a player with $x = 1$ is indifferent for $y \geq -1$ and $\delta = 1$, she strictly prefers to buy if $\delta < 1$ (she is losing valuation by waiting and is not gaining anything). In the same manner, a player with $x = -1$ strictly prefers to buy for $y + w \geq N$ and $w < N$. Therefore, the strategy $\gamma^* = \phi[r, y, w] = \mathbf{1}$ could be a solution for $y + w \geq N$ and $w < N$. For the rest of the states which are $y \geq -1$, $w < N$ and $y + w < N$, a player with $x = -1$ strictly prefers to wait for $\delta = 1$ and hence, there exists large enough $\delta < 1$ such that this player still prefers to wait and therefore, $\gamma^* = \phi[r, y, w] = \boldsymbol{I}$ can be a solution for $y \geq -1$, $w < N$ and $y + w < N$ when $\delta < 1$ is large enough. Further, for $w = N - 1$, $y = 1$ and $r = 1$, both types prefer buying over waiting therefore $\gamma^* = \phi[r, y, w] = \mathbf{1}$ can be a solution. This completes the proof of this theorem.

B.8 Proof of Theorem 5

We first prove the fourth part of the theorem. For $y \leq -2$, the instantaneous reward is negative for $r = 0$ and both $x = 1$ and $x = -1$. On the other hand, according to the proof of Lemma 3, the value functions are 0 when $\gamma^* = \phi(r, y, w) = \mathbf{0}$. Therefore, the equilibrium strategy is not buying for both values of x and so $\gamma^* = \phi(r, y, w) = \mathbf{0}$ is the only solution for $y \leq -2$. For $y \geq 0$, the instantaneous reward is positive for $x = 1$ and therefore, $\gamma^* = \phi(0, y, w) = \mathbf{0}$ (so that the value functions are all 0) can not be an equilibrium strategy.

The fifth part is obvious due to the fact that at $y = 0$ the reward is negative for $x = -1$ and it is positive for $x = 1$. Hence, neither $\gamma^* = \phi[0, 0, w] = \mathbf{0}$ nor $\gamma^* = \phi[0, 0, w] = \mathbf{1}$ can be solution of FPE 2. Therefore, if a solution exists, which we know it does, we must have $\gamma^* = \phi(0, 0, w) = \boldsymbol{I}$.

Now we prove the sixth part. If for some equilibrium strategy and some w and y, $\gamma^* = \phi[0, y, w] = \boldsymbol{I}$ or $\gamma^* = \phi[0, y, w] = \mathbf{1}$, we can not have $\gamma^* = \phi[0, y, w] = \mathbf{0}$ for $w' \neq w$ and $y \neq -1$. The reason is that if for $w' \neq w$, we have $\gamma^* = \phi[r, y, w'] = \mathbf{0}$, the valuation function $U_a(x, 0, y, w) = 0$ as proved in the proof of Lemma 3. On the other hand, since $\gamma^* = \phi[0, y, w] = \boldsymbol{I}$ or $\gamma^* = \phi[0, y, w] = \mathbf{1}$, we know that $\frac{q^{y+1}-1}{q^{y+1}+1} > 0$ for $y \neq -1$. Hence the instantaneous reward for a player with $x = 1$ at $r = 0, y, w'$ is positive and therefore, $\gamma^* = \phi[0, y, w'] = \mathbf{0}$ can not be an equilibrium strategy. Hence, $\gamma^* = \phi[0, y, w] = \boldsymbol{I}$ or $\gamma^* = \phi[0, y, w] = \mathbf{1}$ can not happen with

$\gamma^* = \phi[0, y, w'] = \mathbf{0}$ for the same y. Therefore, for a fixed y, we either have $\gamma^* = \phi[0, y, w] = \mathbf{0}$ for all w or a combination of $\gamma^* = \phi[0, y, w] = \boldsymbol{I}$ or $\gamma^* = \phi[0, y, w] = \mathbf{1}$ for different w.

The seventh part is evident by using the fourth part and Lemma 3. As we saw in Lemma 3, a player with $x = 1$ is indifferent between buying and waiting for $y \geq -1$, which includes $y = -1$. It means that she can always decide to buy for $y \geq -1$. On the other hand, a player with $x = -1$ has negative instantaneous reward and she should not buy at $y = -1$. Hence, since the expected reward of the player with $x = 1$ is 0, both $\gamma^* = \phi[0, -1, w] = \boldsymbol{I}$ and $\gamma^* = \phi[0, -1, w] = \mathbf{0}$ can be solutions for all w.

We can prove the eighth part in a similar way. A player with $x = 1$ and $r = 1$ is always indifferent between buying and not buying at $y = -2$ since the instantaneous reward is 0 (she is on an off-equilibrium path). On the other hand, a player with $x = -1$ and $r = 1$ prefers to wait at $y = -2$ since her instantaneous reward is negative and hence, both $\gamma^* = \phi[1, -2, w] = \boldsymbol{I}$ and $\gamma^* = \phi[1, -2, w] = \mathbf{0}$ are the solutions.

The third part is a direct consequence of fourth and seventh parts.

In order to prove the first part, it is sufficient to show that if $\gamma^* = \phi[r, y, w] = \boldsymbol{I}$ and the solution is a threshold policy w.r.t. w for $y' < y$, then $\gamma^* = \phi[r, y, w'] = \boldsymbol{I}$ is a solution for $w' < w$ (Note that it might not be the only case, and we are arguing about existence. So if a solution is not of this type, we can construct a solution of this type, as we explain later on).

Assume that for the state $s = (x, r, y, w)$, we have $\gamma^* = \phi[r, y, w] = \boldsymbol{I}$. It means that the instantaneous reward for $x = -1$ at y has been no more than the expected valuation of not buying, which is the average of rewards, at those terminating states that the player will buy, and 0's, for those in which the player decides not to buy the product (see the proof of Lemma 3). The more likely the final states with not buying decision are, the larger the difference between the instantaneous reward and expected valuation of not buying is. So for two different states with the same instantaneous reward, i.e., the same y, we can compare their terminating states to get a sense of how the player decides in these two states. Consider $s' = (r, y, w')$ for $w' < w$. Since the solution is a threshold policy w.r.t. w for $y' < y$, it is clear that for each terminating state s_j for the state s that the player decides not to buy the product, there is a corresponding state s'_j for the state s' which is at least as likely to happen as s_j (there are more players that can reveal their private signal and change the state). At s'_j, the player has the opportunity to decide not to buy the product, or decide later on if it is beneficial for her (which implies that s'_j may not be a terminating state for s'). In both cases, the valuation of not buying at s' is at least as good as s and hence, if the valuation of not buying at s is not less than the instantaneous reward, it has to be true for s' too. Therefore, if $\gamma^* = \phi[r, y, w] = \boldsymbol{I}$, we can have $\gamma^* = \phi[r, y, w'] = \boldsymbol{I}$ for $w' < w$. If a solution is not of this type, we can construct such strategy as follows. According to the other parts of this theorem, we know the solution can be $\gamma^* = \phi[r, y, w] = \mathbf{0}$ for $y \le -2$, all r and w, and also $\gamma^* = \phi[r, y, w] = \boldsymbol{I}$ for $r = 0$, $-1 \le y \le 0$ and all w and for $r = 1$, $-1 \le y \le 0$ and all w. So whatever else is the solution, we can change it to the mentioned strategy profile. Next, we start at $y = 1$ and we know that the solution is a threshold policy for $y' < y$. Starting at $w = N$ and going back step by step for both $r = 0$ and $r = 1$, we can change all the solutions $\gamma^* = \phi[r, y, w'] = \mathbf{1}$ to $\gamma^* = \phi[r, y, w'] = \boldsymbol{I}$ for all $w' < w$ such that the solution is $\gamma^* = \phi[r, y, w] = \boldsymbol{I}$. In this way we construct a strategy profile that is a threshold policy w.r.t. w and is solution to FPE 2.

Now we restrict our attention to the equilibrium strategies that are threshold policies w.r.t. w and prove that whenever $\gamma^* = \phi[0, y, w] = \mathbf{1}$, then we must have $\gamma^* = \phi[0, y', w] = \mathbf{1}$ for all $y' > y$ and whenever $\gamma^* = \phi[0, y, w] = \boldsymbol{I}$ then we must have $\gamma^* = \phi[0, y', w] \neq \mathbf{0}$ for all $y' > y$.

Similar to the arguments in the proof of sixth part of this theorem, whenever $\gamma^* = \phi[0, y, w] = \boldsymbol{I}$, the instantaneous reward is positive for $x = 1$ and if we have $\gamma^* = \phi[0, y', w] = \mathbf{0}$, the valuation will be 0 while the instantaneous reward for y' is greater than y and so it is positive. Hence, we can not have $\gamma^* = \phi[0, y', w] = \mathbf{0}$ as a solution.

In order to prove that whenever $\gamma^* = \phi[0, y, w] = \mathbf{1}$, then we must have $\gamma^* = \phi[0, y', w] = \mathbf{1}$ for all $y' > y$, we assume this is not true and hence, we have a case where $\gamma^* = \phi[0, y, w] = \mathbf{1}$ and $\gamma^* = \phi[0, y+1, w] = \boldsymbol{I}$. In this case the player with $x = -1$ at $y + 1$ is choosing not buying over buying which means that

$$\frac{q^y - 1}{q^y + 1} \le \frac{\delta}{N} U_a\left(-1, 1, y, w+1\right) + \frac{\delta}{N}\left(N - w - 1\right) U_{na}^1\left(-1, 0, y, w+1\right) + \frac{\delta}{N} w U_{na}^1\left(-1, 1, y, w+1\right), \quad \text{(B.19)}$$

since $\gamma^* = \phi[0, y, w] = \mathbf{1}$, we know that $\gamma^* = \phi[0, y, w+1] = \mathbf{1}$ and hence, according to the proof of Theorem 3, $U_a\left(-1, 1, y, w+1\right) = \frac{q^y-1}{q^y+1}$, $U_{na}^1\left(-1, 0, y, w+1\right) \le \frac{q^y-1}{q^y+1}$ and $U_{na}^1\left(-1, 1, y, w+1\right) \le \frac{q^y-1}{q^y+1}$ which means that for $\delta < 1$, $\frac{q^y-1}{q^y+1} < \frac{q^y-1}{q^y+1}$ and it is a contradiction.

Next consider $r = 1$. We first prove a relation between a player's decision in states $s = (x, 0, y, w)$ and $s' = (x, 1, y-1, w+1)$. Assume that $\gamma^* = \phi[0, y, w] = \boldsymbol{I}$. It means that

$$\frac{q^{y-1} - 1}{q^{y-1} + 1} \le \frac{\delta}{N} U_a\left(-1, 1, y-1, w+1\right) + \frac{\delta}{N}\left(N - w - 1\right) U_{na}^1\left(-1, 0, y-1, w+1\right) + \frac{\delta}{N} w U_{na}^1\left(-1, 1, y-1, w+1\right), \quad \text{(B.20)}$$

on the other hand, if we write the fixed point equation for $x=-1,\ r=1,\ y-1,\ w+1$, we have

$$\gamma^*(-1) = \arg\max\left\{\frac{\delta}{N}U_a\left(-1,1,y-1,w+1\right)+\frac{\delta}{N}\left(N-w-1\right)U_{na}^1\left(-1,0,y-1,w+1\right)\right.$$
$$\left.+\frac{\delta}{N}wU_{na}^1\left(-1,1,y-1,w+1\right),\frac{q^{y-1}-1}{q^{y-1}+1}\right\}. \quad \text{(B.21)}$$

According to (B.20), we can say that the solution of the above fixed point equation can be not buy. Hence, whenever $\gamma^*=\phi[0,y,w]=\boldsymbol{I}$, we can have $\gamma^*=\phi[1,y-1,w+1]=\boldsymbol{I}$. Also, according to Lemma 3, whenever $\gamma^*=\phi[0,y,w]=\mathbf{1}$ and $\gamma^*=\phi[0,y,w']=\mathbf{1}$ for $w'>w$, we must have $\gamma^*=\phi[1,y,w]=\mathbf{1}$. This all means that whenever we have a solution that is a threshold policy w.r.t. y for $r=0$, the solution can also be a threshold policy w.r.t. y for $r=1$.

B.9 Proof of Theorem 6

If Y_t remains constant with probability one for all $t'>t$, then it is an informational cascade by definition since y_t sums all the revealed private information. Hence, the absorbing states of $\overline{Y}_i$ are informational cascades. We have shown that some $\overline{Y}_i=y_L\geq Y_{\min}$ and $\overline{Y}_i=y_R\leq Y_{\max}$ are absorbing. The values of both y_L,y_R are independent of N. The transition probabilities of $\overline{Y}_i$ are $\frac{p+(1-p)q^y}{q^y+1}$ for moving right and $\frac{1-p+pq^y}{q^y+1}$ for moving left, so they are also independent of N. We conclude that the distribution (specifically, expectation and variance) of the absorption time is independent of N. Hence, for large enough N, the probability that the absorption time is larger than M_N vanishes to zero. This absorption time is counted in the number of revealings i. We conclude that the probability that a cascade occurs before M_N revealings occur approaches 1 as $N\to\infty$.

Now assume that $\phi\left[r,y,w\right]=\mathbf{1}$ implies that $\phi\left[r,y,\widehat{w}\right]=\mathbf{1}$ for $\widehat{w}>w$. Denote the number of turns up to turn $t=M_N$ where the acting player n_t has $r^{n_t}=1$ or $b^{n_t}=1$ by $\overline{R}\left(M_N\right)$, which is stochastically dominated by a binomial distributed variable with $p=\frac{M_N}{N}$ and M_N trials since $\frac{w_t}{N}\leq\frac{M_N}{N}$.

Hence, for all $N>0$,

$$\mathbb{P}\left(\overline{R}\left(M_N\right)\geq 1\right)\leq 1-\left(1-\frac{M_N}{N}\right)^{M_N}. \quad \text{(B.22)}$$

Since by assumption $\frac{M_N^2}{N}\to 0$ as $N\to\infty$, then $1-\left(1-\frac{M_N}{N}\right)^{M_N}\to 1-e^{-\frac{M_N^2}{N}}\to 0$ and we conclude that with high probability, at least M_N-1 of the first turns are of players with $z^n=0$. Assume that in turn $t<M_N-1$ the acting player did not reveal her private information.

- If she waited, then $w_{t+1} = w_t$ and $y_{t+1} = y_t$. The next player with $z^n = 0$ will also wait since she uses the same strategy $\gamma^* = \phi[r, y, w]$.
- If she bought, then $w_{t+1} = w_t + 1$ and $y_{t+1} = y_t$. The next acting player with $z^n = 0$ will also buy for $x^n = -1, 1$ (and not reveal) since $w_{t+1} > w_t$.

The same argument applies to all subsequent players with $z^n = 0$, and by definition to players with $z^n = 1$, so a cascade occurred.

APPENDIX C

Proofs for Chapter 4

C.1 Proof of Lemma 4

Since functions $\widehat{u}_i^{\mathtt{u}}(m_i^{\mathtt{u}}, m_{-i}^{\mathtt{u}})$ and $\widehat{u}_i^{\mathtt{m}}(m_i^{\mathtt{m}}, m_{-i}^{\mathtt{m}})$ are twice differentiable w.r.t $m_i^{\mathtt{u}}$ and $m_i^{\mathtt{m}}$, respectively, we can prove strict concavity by showing that their Hessian matrices, $\mathrm{H}^{\mathtt{u}}$ and $\mathrm{H}^{\mathtt{m}}$, w.r.t. $m_i^{\mathtt{u}}$ and $m_i^{\mathtt{m}}$, respectively, are negative definite. The cross derivatives of $\widehat{u}_i^{\mathtt{u}}(m_i^{\mathtt{u}}, m_{-i}^{\mathtt{u}})$ and $\widehat{u}_i^{\mathtt{m}}(m_i^{\mathtt{m}}, m_{-i}^{\mathtt{m}})$ w.r.t. different components of $m_i^{\mathtt{u}}$ and $m_i^{\mathtt{m}}$, which are the non-diagonal elements of $\mathrm{H}^{\mathtt{u}}$ and $\mathrm{H}^{\mathtt{m}}$, respectively, are zero. Hence, we consider the diagonal elements and show that they are all negative. The second partial derivative of $\widehat{u}_i^{\mathtt{u}}(m^{\mathtt{u}})$ w.r.t. all elements of messages n_i, q_i, p_i is equal to -2. Also, the second partial derivative of $\widehat{u}_i^{\mathtt{m}}(m^{\mathtt{m}})$ w.r.t. all elements of messages n_i, q_i, p_i, w_i, z_i, a_i and s_i is equal to -2. The only message element left is y_i. The second partial derivative of $\widehat{u}_i^{\mathtt{u}}(m^{\mathtt{u}})$ w.r.t. y_i is $\partial^2 \widehat{u}_i^{\mathtt{u}}(m^{\mathtt{u}})/(\partial y_i)^2 = (r_i^{\mathtt{u}})^2 \partial^2 v_i(\widehat{x}_i^{\mathtt{u}})/(\partial \widehat{x}_i^{\mathtt{u}})^2$. Since $v_i(\widehat{x}_i^{\mathtt{u}})$ is strictly concave w.r.t. $\widehat{x}_i^{\mathtt{u}}$, $\partial^2 \widehat{u}_i^{\mathtt{u}}(m^{\mathtt{u}})/(\partial y_i)^2 < 0$. Similarly, the second partial derivative of $\widehat{u}_i^{\mathtt{m}}(m^{\mathtt{m}})$ w.r.t. y_i is $\partial^2 \widehat{u}_i^{\mathtt{m}}(m^{\mathtt{m}})/(\partial y_i)^2 = (r_i^{\mathtt{m}})^2 \partial^2 v_i(\widehat{x}_i^{\mathtt{m}})/(\partial \widehat{x}_i^{\mathtt{m}})^2 < 0$. Note that $r_i^{\mathtt{u}}$ and $r_i^{\mathtt{m}}$ don't consist of any of agent i's messages and so they are constant factors.

Therefore, matrices $\mathrm{H}^{\mathtt{u}}$ and $\mathrm{H}^{\mathtt{m}}$ are negative definite because all of their diagonal elements are negative and non-diagonal elements are zero.

C.2 Proof of Lemma 5

At NE, every agent is best responding to other agents' messages. Each of the results in this lemma corresponds to one of agent i's messages and its best response to other agents messages. Therefore, all of the results can be directly derived by setting each of their corresponding element of gradient to zero. For all $i \in \mathcal{N}$ we have

$$\frac{\partial \widehat{u}_i^{\mathtt{u}}(m_i^{\mathtt{u}}, m_{-i}^{\mathtt{u}})}{\partial q_{i,j}} = 0 \Rightarrow q_{i,j} = y_j, \ \forall j \in \mathcal{I}_i. \tag{C.1}$$

Since $y_j \geq 0$ the above equation can always hold.

$$\frac{\partial \widehat{u}_i^{\mathtt{u}}(m_i^{\mathtt{u}}, m_{-i}^{\mathtt{u}})}{\partial n_{i,j}^l} = 0 \Rightarrow n_{i,j}^l = y_j^l + \sum_{h \in \mathcal{N}(j), h \neq i} n_{j,h}^l \ \ \forall j \in \mathcal{N}(i), l \in \mathcal{L}. \tag{C.2}$$

Using a similar argument as the one used in [113, p. 131] all above equations can be combined to show that

$$n_{i,j}^l = \sum_{h \in \mathcal{N}, n(i,h)=j} y_h^l, \tag{C.3}$$

and equivalently,

$$\sum_{j \in \mathcal{N}(i)} n_{i,j}^l = \sum_{h \in \mathcal{N}, h \neq i} y_h^l. \tag{C.4}$$

In order to verify this conclusion, we mention that the message graph is a tree, and hence we can form an induction on the level of nodes from the leaf nodes. If $j \in \mathcal{N}(i)$ is a leaf node, there is no message components $n_{j,h}^l$ for $h \in \mathcal{N}(j), h \neq i$. because the only neighbor of j is i. As a result, $n_{i,j}^l = y_j^l$ and therefore, the induction basis holds. Suppose for all $j \in \mathcal{N}(i)$, $n_{j,h}^l = \sum_{k \in \mathcal{N}, n(j,k)=h} y_k^l$, $\forall h \in \mathcal{N}(j)$. Substituting this to (C.2) we have

$$n_{i,j}^l = y_j^l + \sum_{\substack{h \in \mathcal{N}(j) \\ h \neq i}} \sum_{\substack{k \in \mathcal{N} \\ n(j,k)=h}} y_k^l \ \ \forall j \in \mathcal{N}(i), l \in \mathcal{L}. \tag{C.5}$$

We need to check whether the set of nodes covered in the summations above is equal to the set of nodes in the summation of (C.3). In (C.5), we are summing over all nodes k that can be reached to node j by nodes $h \in \mathcal{N}(j),\ h \neq i$. Since $j \in \mathcal{N}(i)$ and there is only one path between any two nodes in the graph, we conclude that all of these nodes reach i through j and therefore, $n(i,k) = j$. This means that the summations in (C.3) and (C.5) include the same set of nodes and the result is proved.

C.3 Proof of Lemma 6

According to the explanations at the beginning of the proof of Lemma 5, by setting each of the corresponding element of gradients of $\widehat{u}_i^{\mathtt{m}}(m_i^{\mathtt{m}}, m_{-i}^{\mathtt{m}})$ to zero for all $i \in \mathcal{N}$ we have

$$\frac{\partial \widehat{u}_i^{\mathtt{m}}(m_i^{\mathtt{m}}, m_{-i}^{\mathtt{m}})}{\partial q_{i,j}} = 0 \Rightarrow q_{i,j} = y_j, \ \ \forall j \in \mathcal{I}_i \tag{C.6a}$$

$$\frac{\partial \widehat{u}_i^{\mathfrak{m}}(m_i^{\mathfrak{m}}, m_{-i}^{\mathfrak{m}})}{\partial (s_i^l)} = 0 \Rightarrow s_i^l = \frac{q_{\phi(i),i}\mathbf{1}_{\{q_{\phi(i),i}\}}(\overline{z}_i^{1,l})}{\overline{z}_i^{2,l}}, \ \ \forall l \in \mathcal{L}_i \tag{C.6b}$$

$$\frac{\partial \widehat{u}_i^{\mathfrak{m}}(m_i^{\mathfrak{m}}, m_{-i}^{\mathfrak{m}})}{\partial (p_{i,j}^{2,l})} = 0 \Rightarrow p_{i,j}^{2,l} = p_j^{1,l}, \ \ \forall j \in I_i, l \in \mathcal{L}_j \tag{C.6c}$$

$$\frac{\partial \widehat{u}_i^{\mathfrak{m}}(m_i^{\mathfrak{m}}, m_{-i}^{\mathfrak{m}})}{\partial w_i^l} = 0 \Rightarrow w_i^l = w_{c(k(i),l)}^l, \ \forall l \in \mathcal{L}_i, l \notin \mathcal{C}_i \tag{C.6d}$$

$$\frac{\partial \widehat{u}_i^{\mathfrak{m}}(m_i^{\mathfrak{m}}, m_{-i}^{\mathfrak{m}})}{\partial w_i^l} = 0 \Rightarrow \quad w_i^l = p_{\phi(i),i}^{2,l} + \sum_{j \in \mathcal{G}_{k(i)}^l, j \neq i} p_j^{1,l}, \ \ \forall l \in \mathcal{L}_i, l \in \mathcal{C}_i \tag{C.6e}$$

$$\frac{\partial \widehat{u}_i^{\mathfrak{m}}(m_i^{\mathfrak{m}}, m_{-i}^{\mathfrak{m}})}{\partial n_{i,j}^l} = 0 \Rightarrow n_{i,j}^l = y_j^l + \sum_{h \in \mathcal{N}(j), h \neq i} n_{j,h}^l, \ \ \forall j \in \mathcal{N}(i), l \in \mathcal{L}. \tag{C.6f}$$

Using a similar argument as the one used in [113, p. 131] we prove that

$$n_{i,j}^l = \sum_{h \in \mathcal{N}, n(i,h) = j} y_h^l, \tag{C.7}$$

and consequently,

$$\sum_{j \in \mathcal{N}(i)} n_{i,j}^l = \sum_{h \in \mathcal{N}, h \neq i} y_h^l. \tag{C.8}$$

The remaining results are related to message elements z_i and ${}^2a_i$.

$$\frac{\partial \widehat{u}_i^{\mathfrak{m}}(m_i^{\mathfrak{m}}, m_{-i}^{\mathfrak{m}})}{\partial (z_i^{1,l})} = 0 \Rightarrow z_i^{1,l} = \overline{z}_i^{1,l}, \quad \forall l \in \mathcal{L}_i, l \in \mathcal{C}_i \tag{C.9a}$$

$$\frac{\partial \widehat{u}_i^{\mathfrak{m}}(m_i^{\mathfrak{m}}, m_{-i}^{\mathfrak{m}})}{\partial (z_i^{2,l})} = 0 \Rightarrow z_i^{2,l} = \overline{z}_i^{2,l}, \quad \forall l \in \mathcal{L}_i, l \in \mathcal{C}_i \tag{C.9b}$$

$$\frac{\partial \widehat{u}_i^{\mathfrak{m}}(m_i^{\mathfrak{m}}, m_{-i}^{\mathfrak{m}})}{\partial (a_{i,j}^{2,l})} = 0 \Rightarrow a_{i,j}^{2,l} = a_j^{1,l}, \quad \forall j \in \mathcal{I}_i, l \in \mathcal{L}_j. \tag{C.9c}$$

C.4 Proof of Lemma 7

According to Lemmas 5 and 6, the following relation holds at NE,

$$f_i^{\mathfrak{u},l} = \sum_{j \in \mathcal{N}} y_j^l, \tag{C.10}$$

and similarly,

$$f_i^{\mathtt{m},l} = \sum_{j \in \mathcal{N}} y_j^l. \tag{C.11}$$

This implies that all of the agents $i \in \mathcal{N}$ have the same $f_i^{\mathtt{u},l}$ and $f_i^{\mathtt{m},l}$ and consequently, have the same $r_i^{\mathtt{u}}$ and $r_i^{\mathtt{m}}$ at NE of the games $\mathfrak{G}^{\mathtt{u}}$ and $\mathfrak{G}^{\mathtt{m}}$, respectively. Further, due to Lemma 6, for the game $\mathfrak{G}^{\mathtt{m}}$ we can write for every $j \in \mathcal{G}_k^l$

$$y_j^l = \begin{cases} \frac{y_j}{n\_\max_k^l} & \text{if } l \in \mathcal{L}_j,\ y_j = \max_{i \in \mathcal{G}_k^l}\{y_i\} \\ 0 & \text{oth.}, \end{cases} \tag{C.12}$$

where $n\_\max_k^l$ is the number of agents $j \in \mathcal{G}_k^l$ with $y_j = \max_{i \in \mathcal{G}_k^l}\{y_i\}$. Consequently, $\forall l \in \mathcal{L}$ we can write

$$\sum_{k \in \mathcal{K}^l} \max_{i \in \mathcal{G}_k^l}\{\widehat{x}_i^{\mathtt{m}}\} = \sum_{k \in \mathcal{K}^l} \max_{i \in \mathcal{G}_k^l}\{r_i^{\mathtt{m}} y_i\} \leq \frac{c^l}{\sum_{j \in \mathcal{N}} y_j^l} \sum_{k \in \mathcal{K}^l} \max_{i \in \mathcal{G}_k^l}\{y_i\} = \frac{c^l}{\sum_{j \in \mathcal{N}} y_j^l} \sum_{i \in \mathcal{N}} y_i^l = c^l, \tag{C.13}$$

which proves that the allocation $\widehat{x}^{\mathtt{m}}$ is feasible at NE.

By using similar steps, we can show the feasibility of allocation of the game $\mathfrak{G}^{\mathtt{u}}$ at NE.

C.5 Proof of Lemma 8

In order to prove the first result, we first derive the following

$$p_i^l = \overline{p}_{-i}^l \quad \forall i \in \mathcal{N}, l \in \mathcal{L}_i. \tag{C.14}$$

Suppose the above equation does not hold, i.e.,

$$\exists i \in \mathcal{N}, l \in \mathcal{L}_i : \quad p_i^l \neq \overline{p}_{-i}^l. \tag{C.15}$$

Then there exists an agent $j \in \mathcal{N}^l : p_j^l > \overline{p}_{-j}^l$ (as an example, we could consider the agent j with the highest p_j^l over all of the agents and if we have multiple choices, at least one of them will work). We can show that agent j has a profitable deviation to $p_j^{l'} = \overline{p}_{-j}^l = p_j^l - \epsilon$. Indeed, we can write

$$\widehat{u}_j^{\mathtt{u}}(., p_j^{l'}) - \widehat{u}_j^{\mathtt{u}}(., p_j^l) = \epsilon^2 + \epsilon \overline{p}_{-j}^l (c^l - r_j^{\mathtt{u}} f_j^{\mathtt{u},l})^2 = \epsilon(\underbrace{\epsilon}_{>0} + \underbrace{\overline{p}_{-j}^l (c^l - r_j^{\mathtt{u}} f_j^{\mathtt{u},l})^2}_{\geq 0}) > 0, \tag{C.16}$$

therefore, we must have $p_i^l = \overline{p}_{-i}^l$.

As a result of this equality and because of Assumption 1, it is obvious that $p_i^l = p_j^l,\ \forall i, j \in \mathcal{N}^l$ and we denote this common price by p^l.

For the second result, we set the derivative of the utility function w.r.t. p_i^l to zero,

$$\begin{aligned}\frac{\partial \widehat{u}_i^{\mathtt{u}}(m_i^{\mathtt{u}}, m_{-i}^{\mathtt{u}})}{\partial p_i^l} = 0 \Rightarrow \underbrace{2(p_i^l - \overline{p}_{-i}^l)}_{=0,\text{Due to (4.23a)}} + \overline{p}_{-i}^l (c^l - r_i^{\mathtt{u}} f_i^{\mathtt{u},l})^2 = 0 \\ \Rightarrow p^l (c^l - r_i^{\mathtt{u}} f_i^{\mathtt{u},l})^2 = 0 \Rightarrow p^l (c^l - \sum_{i \in \mathcal{N}^l} \widehat{x}_i^{\mathtt{u}}) = 0.\end{aligned} \tag{C.17}$$

C.6 Proof of Lemma 9

We first prove result (4.24a). This result is equivalent with $\widehat{w}_i^l = \widehat{w}_j^l,\ \forall i, j \in \mathcal{N}^l$. Assume $\exists i, j \in \mathcal{N}^l,\ \widehat{w}_i^l \neq \widehat{w}_j^l$. Since $w_i^l = \widehat{w}_i^l$ at NE and due to Assumption 1, there exists an agent $h \in \mathcal{N}^l$ for which $\widehat{w}_h^l > \overline{w}_{-h}^l$ or equivalently, $\widehat{w}_h^l = \overline{w}_{-h}^l + \epsilon$ for some $\epsilon > 0$. We will show that agent h has a profitable deviation by decreasing his message $a_h^{1,l}$ to $a_h^{1,l'} = a_h^{1,l} - \epsilon' > 0$ for some $0 < \epsilon' < \epsilon$. Consequently, $\widehat{w}_h^{l'} = \widehat{w}_h^l - \epsilon' = \overline{w}_{-h}^l + \epsilon - \epsilon' = \overline{w}_{-h}^l + \epsilon''$. We can write

$$\begin{aligned}\widehat{u}_h^{\mathtt{m}}(., a_h^{1,l'}) - \widehat{u}_h^{\mathtt{m}}(., a_h^{1,l}) &= -\epsilon''^2 - \overline{w}_{-h}^l \epsilon'' (c^l - r_h^{\mathtt{m}} f_h^{\mathtt{m},l})^2 + \epsilon^2 + \overline{w}_{-h}^l \epsilon (c^l - r_h^{\mathtt{m}} f_h^{\mathtt{m},l})^2 \\ &= \underbrace{\epsilon^2 - \epsilon''^2}_{>0} + \underbrace{\overline{w}_{-h}^l (\epsilon - \epsilon'')(c^l - r_h^{\mathtt{m}} f_h^{\mathtt{m},l})^2}_{\geq 0} > 0,\end{aligned} \tag{C.18}$$

and we conclude that at any NE, $\widehat{w}_i^l = \widehat{w}_j^l,\ \forall l \in \mathcal{L}, i, j \in \mathcal{N}^l$. Therefore, we can denote this common value for each link l by w^l and we arrive at the result $\widehat{w}_i^l = w^l,\ \forall i \in \mathcal{N},\ l \in \mathcal{L}_i$.

Now we prove result (4.24b). Suppose $\exists i \in \mathcal{N}, l \in \mathcal{L}_i$ so that $w^l (c^l - r_i^{\mathtt{m}} \sum_{i \in \mathcal{N}} y_i^l) \neq 0$. This implies $\overline{w}_{-i}^l (c^l - r_i^{\mathtt{m}} f_i^{\mathtt{m},l})^2 > 0$. We show that agent i benefits from deviating to $a_i^{1,l'} = a_i^{1,l} - \epsilon > 0$, for some $\epsilon > 0$. According to the first result of this lemma, $\widehat{w}_i^l = \overline{w}_{-i}^l$ and we have

$$\begin{aligned}\widehat{u}_i^{\mathtt{m}}(., a_i^{1,l'}) - \widehat{u}_i^{\mathtt{m}}(., a_i^{1,l}) &= -\epsilon^2 + \overline{w}_{-i}^l \epsilon (c^l - r_i^{\mathtt{m}} f_i^{\mathtt{m},l})^2 \\ &= \epsilon(-\epsilon + \underbrace{\overline{w}_{-i}^l (c^l - r_i^{\mathtt{m}} f_i^{\mathtt{m},l})^2}_{>0,\ \text{Due to assumption}}) = \epsilon(-\epsilon + \alpha) > 0,\ \text{for } \epsilon < \alpha.\end{aligned} \tag{C.19}$$

Since $\alpha > 0$, profitable deviation by a positive ϵ is possible and the result is proved.

Proving result (4.24c) is similar to result (4.24b). Assume $\exists i \in \mathcal{N}, l \in \mathcal{L}_i$ so that $p_i^{1,l}(y_i - \overline{z}_i^{1,l}) \neq 0$. This implies that $p_{\phi(i),i}^{2,l}(\overline{z}_i^{1,l} - q_{\phi(i),i})^2 > 0$ and $p_i^{1,l} > 0$. We prove agent i has a profitable deviation to $p_i^{1,l'} = p_i^{1,l} - \epsilon > 0$, for some $\epsilon > 0$. Indeed,

$$\begin{aligned}\widehat{u}_i^{\mathtt{m}}(., p_i^{1,l'}) - \widehat{u}_i^{\mathtt{m}}(., p_i^{1,l}) &= -\epsilon^2 + p_{\phi(i),i}^{2,l}\epsilon(\overline{z}_i^{1,l} - q_{\phi(i),i})^2 \\ &= \epsilon(-\epsilon + \underbrace{p_{\phi(i),i}^{2,l}(\overline{z}_i^{1,l} - q_{\phi(i),i})^2}_{>0,\ \text{Due to assumption}}) = \epsilon(-\epsilon + \alpha) > 0, \text{for } \epsilon < \alpha. \quad \text{(C.20)}\end{aligned}$$

Since $\alpha > 0$, agent i can profit by deviating with a positive ϵ and the result is proved.

C.7 Proof of Lemma 10

If $\widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}}) > 0$, then $y_i > 0$ and hence, the partial derivative of $\widehat{u}_i^{\mathtt{u}}(m_i^{\mathtt{u}}, m_{-i}^{\mathtt{u}})$ w.r.t. y_i must be zero at NE. Therefore,

$$\begin{aligned}\frac{\partial \widehat{u}_i^{\mathtt{u}}(m_i^{\mathtt{u}}, m_{-i}^{\mathtt{u}})}{\partial y_i} &= 0 \Rightarrow \left(\frac{\partial \widehat{u}_i^{\mathtt{u}}(m_i^{\mathtt{u}}, m_{-i}^{\mathtt{u}})}{\partial \widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}})}\right)\frac{d\widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}})}{dy_i} = 0 \\ &= (v_i'(\widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}})) - \sum_{l\in\mathcal{L}_i} p^l)r_i^{\mathtt{u}} \Rightarrow v_i'(\widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}})) = \sum_{l\in\mathcal{L}_i} p^l, \quad \text{(C.21)}\end{aligned}$$

and if $\widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}}) = 0$, then $y_i = 0$ and therefore, the partial derivative of $\widehat{u}_i^{\mathtt{u}}(m_i^{\mathtt{u}}, m_{-i}^{\mathtt{u}})$ w.r.t. y_i must not be positive at NE. Hence,

$$\begin{aligned}\frac{\partial \widehat{u}_i^{\mathtt{u}}(m_i^{\mathtt{u}}, m_{-i}^{\mathtt{u}})}{\partial y_i} &\le 0 \Rightarrow \left(\frac{\partial \widehat{u}_i^{\mathtt{u}}(m_i^{\mathtt{u}}, m_{-i}^{\mathtt{u}})}{\partial \widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}})}\right)\frac{d\widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}})}{dy_i} \\ &= (v_i'(\widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}})) - \sum_{l\in\mathcal{L}_i} p^l)r_i^{\mathtt{u}} \le 0 \Rightarrow v_i'(\widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}})) \le \sum_{l\in\mathcal{L}_i} p^l. \quad \text{(C.22)}\end{aligned}$$

C.8 Proof of Lemma 11

Similar to Lemma 10, if $\widehat{x}_i^{\mathtt{m}}(m^{\mathtt{m}}) > 0$,

$$\begin{aligned}\frac{\partial \widehat{u}_i^{\mathtt{m}}(m_i^{\mathtt{m}}, m_{-i}^{\mathtt{m}})}{\partial y_i} &= 0 \Rightarrow \left(\frac{\partial \widehat{u}_i^{\mathtt{m}}(m_i^{\mathtt{m}}, m_{-i}^{\mathtt{m}})}{\partial \widehat{x}_i^{\mathtt{m}}(m^{\mathtt{m}})}\right)\frac{d\widehat{x}_i^{\mathtt{m}}(m^{\mathtt{m}})}{dy_i} = 0 \\ &\Rightarrow (v_i'(\widehat{x}_i^{\mathtt{m}}(m^{\mathtt{m}})) - \sum_{l\in\mathcal{L}_i} p_{\phi(i),i}^{2,l})r_i^{\mathtt{m}} = 0 \\ &\Rightarrow v_i'(\widehat{x}_i^{\mathtt{m}}(m^{\mathtt{m}})) = \sum_{l\in\mathcal{L}_i} p_i^{1,l}. \quad \text{(C.23)}\end{aligned}$$

Note that $r_i^{\mathtt{m}} > 0$. If $\widehat{x}_i^{\mathtt{m}}(m^{\mathtt{m}}) = 0$,

$$\begin{aligned}\frac{\partial \widehat{u}_i^{\mathtt{m}}(m_i^{\mathtt{m}}, m_{-i}^{\mathtt{m}})}{\partial y_i} \leq 0 &\Rightarrow (\frac{\partial \widehat{u}_i^{\mathtt{m}}(m_i^{\mathtt{m}}, m_{-i}^{\mathtt{m}})}{\partial \widehat{x}_i^{\mathtt{m}}(m^{\mathtt{m}})})\frac{d\widehat{x}_i^{\mathtt{m}}(m^{\mathtt{m}})}{dy_i} \leq 0 \\ &\Rightarrow (v_i'(\widehat{x}_i^{\mathtt{m}}(m^{\mathtt{m}})) - \sum_{l\in\mathcal{L}_i} p_{\phi(i),i}^{2,l})r_i^{\mathtt{m}} \leq 0 \\ &\Rightarrow v_i'(\widehat{x}_i^{\mathtt{m}}(m^{\mathtt{m}})) \leq \sum_{l\in\mathcal{L}_i} p_i^{1,l}. \end{aligned} \tag{C.24}$$

C.9 Proof of Lemma 12

To prove the existence of a NE, we show that a suggested valid message is a NE. For each of the games $\mathfrak{G}^{\mathtt{u}}$ and $\mathfrak{G}^{\mathtt{m}}$, the suggested message is generated based on the solution of problems (4.1) and (4.3), respectively, which we know exist and is unique. We notice that because of the monotonicity of valuation functions, the solution of problems (4.1) and (4.3) always lies in the Pareto optimal region of the feasible set which, in our case, is the upper boundary of feasible set in both UTP and MMTP. First consider the game $\mathfrak{G}^{\mathtt{u}}$. Suppose (x^*, λ^*) is the solution of problem (4.1). We generate $m^{\mathtt{u}}$ as follows. First assume $m^{\mathtt{u}}$ satisfies all of the constraints in Lemma 5. Further, y is set to be any scaled version of x^* and since x^* is on the boundary of feasible region, $\widehat{x}^{\mathtt{u}}(m^{\mathtt{u}}) = x^*$. In addition, p_i^l is set to be equal to λ^{l*} and this is valid since $\lambda^{l*} \geq 0$. Hence, Lemma 8 is satisfied for $m^{\mathtt{u}}$. Also, due to stationarity condition, Lemma 10 is also satisfied for $m^{\mathtt{u}}$. Overall, since Lemmas 5, 8 and 10 are satisfied, we know that the elements of the gradient vector of agent i's utility function w.r.t. $m_i^{\mathtt{u}}$ is either zero (positive messages) or not positive (zero messages) which implies that each agent is best responding to other agents' messages and therefore, $m^{\mathtt{u}}$ is a NE of the game $\mathfrak{G}^{\mathtt{u}}$.

Similar steps are taken for the proof of existence of NE for the game $\mathfrak{G}^{\mathtt{m}}$. Let $(x^*, b^*, \lambda^*, \mu^*)$ be the solution of problem (4.4). We generate $m^{\mathtt{m}}$ as following. First assume $m^{\mathtt{m}}$ satisfies all of the constraints in Lemma 6. Further, y is set to be any scaled version of x^* and since x^* is on the boundary of feasible region, $\widehat{x}^{\mathtt{m}}(m^{\mathtt{m}}) = x^*$. In addition, $p_i^{1,l}$ is set to be equal to μ_i^{l*} and this is valid because $\mu_i^{l*} \geq 0$. Then, $w^l = \sum_{j\in\mathcal{G}_k^l} p_j^{1,l} = \sum_{j\in\mathcal{G}_k^l} \mu_j^{l*} = \lambda^{l*}$. Also $r_i^{\mathtt{m}}\bar{z}_i^{1,l} = \max_{j\in\mathcal{G}_{k(i)}^l}\{r_i^{\mathtt{m}} y_j\} = \max_{j\in\mathcal{G}_{k(i)}^l}\{\widehat{x}_i^{\mathtt{m}}\} = \max_{j\in\mathcal{G}_{k(i)}^l}\{x_i^*\} = b_{k(i)}^{l*}$. Hence Lemma 9 is satisfied for $m^{\mathtt{m}}$. Also, due to stationarity condition, Lemma 11 is satisfied for $m^{\mathtt{m}}$. Overall, since Lemmas 6, 9 and 11 are satisfied, we know that the elements of the gradient vector of utility function of each agent i w.r.t. $m_i^{\mathtt{m}}$ is either zero (positive messages) or not positive (zero messages) which implies that each agent is best responding to other agents' messages and therefore, $m^{\mathtt{m}}$ is a NE.

Notice that the dual variables in the solution of optimization problems (4.1) and (4.4) are not unique, even though the primal solution (x) is unique. For each game and each value of dual variables there is a suggested message that is a Nash equilibrium for that game. Further, the y messages at Nash equilibria of these games have infinitely many options as it was mentioned in its construction. This means that the Nash equilibria of these games are not unique and in fact there are infinitely many Nash equilibria for these games.

C.10 Proof of Lemma 13

First consider the weak budget balance equations. At NE, we can write $\widehat{t}_i^{\mathtt{u}} = \widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}}) \sum_{l \in \mathcal{L}_i} p^l$ and hence $\sum_{i \in \mathcal{N}} \widehat{t}_i^{\mathtt{u}} \geq 0$. Similarly, $\widehat{t}_i^{\mathtt{m}} = \widehat{x}_i^{\mathtt{m}}(m^{\mathtt{m}}) \sum_{l \in \mathcal{L}_i} p_i^{1,l}$ and hence, $\sum_{i \in \mathcal{N}} \widehat{t}_i^{\mathtt{m}} \geq 0$ and both mechanisms are weak budget balanced.

Next, consider the individual rationality part for UTP mechanism (the MMTP version is almost identical). For $\widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}}) = 0$, the result is obvious. For $\widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}}) > 0$, we define the function u_i as

$$u_i(x) = v_i(x) - x \sum_{l \in \mathcal{L}_i} p^l. \tag{C.25}$$

Since $u_i(x)$ is concave w.r.t. x and $u_i'(\widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}})) = 0$, then $u_i'(y) \geq 0$ for $0 \leq y \leq \widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}})$, we can conclude $u_i(y) \geq u_i(0)$ and since $u_i(0) = v_i(0)$ and $u_i(\widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}})) = v_i(\widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}})) - \widehat{t}_i^{\mathtt{u}}(m^{\mathtt{u}})$, it follows that $v_i(\widehat{x}_i^{\mathtt{u}}(m^{\mathtt{u}})) - \widehat{t}_i^{\mathtt{u}}(m^{\mathtt{u}}) \geq v_i(0)$ and the result is proved.

APPENDIX D

Proofs for Chapter 5

D.1 Proof of Lemma 14

We have a continuous time M/M/1 queue with transition rates $g_{i,i+1} = \lambda\mathbb{P}(S = 1|x = i)$, i.e., rate of transiting from state i to state $i+1$, and $g_{i,i-1} = 1$ for $i > 0$. Therefore, we have $g_{i,i} = -1 - \lambda\mathbb{P}(S = 1|x = i)$ for $i > 0$ and $g_{0,0} = -\lambda\mathbb{P}(S = 1|x = 0)$. We can calculate the stationary distribution of the queue as follows.

$$\sum_j \mu(j) g_{j,i} = 0, \quad \forall i \geq 0 \tag{D.1}$$

$$\Rightarrow \mu(1) = \lambda\mathbb{P}(S = 1|x = 0)\mu(0), \tag{D.2}$$

$$\mu(i+1) + \lambda\mathbb{P}(S = 1|x = i-1)\mu(i-1) \tag{D.3}$$

$$= (1 + \lambda\mathbb{P}(S = 1|x = i))\mu(i) \tag{D.4}$$

$$\Rightarrow \mu(i+1) = \lambda\mathbb{P}(S = 1|x = i)\mu(i), \quad \forall i \geq 0 \tag{D.5}$$

D.2 Proof of Lemma 15

The proof for α_i is evident by comparing the utility of action $e = 1$ at its decision point, i.e., $u(i, d = 0, e = 1) = i\overline{v} - p$ by the utility of $e = 0$, which is $u(i, d = 0, e = 0) = 0$.

To prove the equation of f, we need to maximize the average utility of the user at the decision point of function f, i.e., $iq(m) - t(m)$ w.r.t. m, which gives us the result.

Similarly, we can prove the equations for γ_i by comparing the average utility of actions $d = 1$ and $d = 0$ at the corresponding decision point. The average utility of action $d = 1$ is $\mathbb{E}[u(i, d = 1, m, S)] = iq(m) - t(m)$, where we have $m = f(i)$. The average utility of the action $d = 0$ is the utility of the outside option which is $(i\overline{v} - p)^+$, and this completes the proof.

D.3 Proof of Theorem 6

This theorem can be proved using Myerson's Lemma [155]. In order to see the connection, note given the tax function described in equation (5.13), one can write the following for $\mathbb{E}(u(i, d = 1, m, S))$.

$$\mathbb{E}(u(i, d = 1, m, S)) = iq(m) - mq(m) + \sum_{j=1}^{m-1} q(j) - t_0 \tag{D.6a}$$

$$= (i - m)q(m) + \sum_{j=1}^{m-1} q(j) - t_0 \tag{D.6b}$$

where $m = f(i)$ and if $m = f(i) = i$, we have

$$\mathbb{E}(u(i, d = 1, m = i, S)) = \sum_{j=1}^{i-1} q(j) - t_0. \tag{D.7}$$

It is now clear that Myerson's lemma can be used here and therefore we have DSIC.

In order to prove the second part of the theorem, note that because of the discrete type space, we do not have uniqueness for the tax function. In other words, revenue equivalence theorem does not hold. However, we can show that the tax function defined in equation (5.13) is an upper bound on all of the tax functions satisfying DSIC and therefore, it is the best the planner can do in terms of his revenue.

One can write the following for any tax function satisfying DSIC.

$$t(m + 1) - t(m) \geq m[q(m + 1) - q(m)] \tag{D.8}$$

$$t(m + 1) - t(m) \leq (m + 1)[q(m + 1) - q(m)] \tag{D.9}$$

Therefore, for $m \geq 2$ we have

$$t(m) \leq t(m - 1) + m[q(m) - q(m - 1)] \tag{D.10}$$

We will show that if tax function is defined according to equation (5.13), each $t(m)$ is equal to its upper bound. We can see it through induction. For $m = 2$, according to equation (5.13), we have $t(2) = t(1) + 2(q(2) - q(1))$ which is the upper bound in the above inequality. Assume we have $t(m) = t(m - 1) + m[q(m) - q(m - 1)]$. According to equation (5.13), we have $t(m + 1) = t(m)+(m+1)q(m+1)-mq(m)-q(m)$ and therefore, $t(m+1) = t(m)+(m+1)[q(m+1)-q(m)]$.

Therefore, we have proved that the planner can not gain any more revenue using other forms of tax functions.

D.4 Proof of Theorem 7

Since optimization problem (5.17) is linear in γ, we can use KKT conditions to characterize the solution. We have the following optimization problem for $N = 2$ and the corresponding dual variables for each constraints.

$$-2\sum_{x=0}^{\infty} v(x)\gamma(s=1,1,x) + 2\sum_{x=0}^{\infty} v(x)\sum_{s,i}\gamma(s,i,x) - p + t_0 \leq 0 : \ \epsilon_1 \tag{D.11}$$

$$\sum_{x=0}^{\infty} v(x)\sum_{s,i}\gamma(s,i,x) - p + t_0 \leq 0 : \ \epsilon_2 \tag{D.12}$$

$$-2\sum_{x=0}^{\infty} v(x)\gamma(s=1,1,x) + t_0 \leq 0 : \ \epsilon_3 \tag{D.13}$$

$$t_0 \leq 0 : \ \epsilon_4 \tag{D.14}$$

$$\sum_{x=0}^{\infty} v(x)\gamma(s=1,1,x) - \sum_{x=0}^{\infty} v(x)\gamma(s=1,2,x) \leq 0 : \eta \tag{D.15}$$

$$\sum_{s,i}\gamma(s,i,x+1) - \lambda\sum_{i}\gamma(1,i,x) = 0, \ \forall x \geq 0 : \alpha_x \tag{D.16}$$

$$2\sum_{s}\gamma(s,i,x) - \sum_{s,i}\gamma(s,i,x) = 0, \forall i \in \mathcal{I}, x \geq 0 : \nu_x^i \tag{D.17}$$

$$\sum_{s,i,x}\gamma(s,i,x) - 1 = 0 : \ \psi \tag{D.18}$$

$$-\gamma(s,i,x) \leq 0, \ \forall s \in \{0,1\}, i \in \mathcal{I}, x \geq 0 : \ \beta_{s,x}^i \tag{D.19}$$

where $\epsilon_1 \geq 0$, $\epsilon_2 \geq 0$, $\epsilon_3 \geq 0$, $\epsilon_4 \geq 0$, $\eta \geq 0$ and $\beta_{s,x}^i \geq 0$. By taking the derivative of the dual function with respect to $\gamma(1,1,x)$, $\gamma(1,2,x)$, $\gamma(0,1,x)$, and $\gamma(0,2,x)$ for $x > 0$, and also with respect to t_0 and setting them to zero, we have the following.

$$(\epsilon_2 - 2\epsilon_3 + \eta)v(x) + \alpha_{x-1} - \lambda\alpha_x + \nu_x^1 - \nu_x^2 + \psi - \beta_{1,x}^1 = 0 \tag{D.20a}$$

$$(-2 + 2\epsilon_1 + \epsilon_2 - \eta)v(x) + \alpha_{x-1} - \lambda\alpha_x + \nu_x^2 - \nu_x^1 + \psi - \beta_{1,x}^2 = 0 \tag{D.20b}$$

$$(2\epsilon_1 + \epsilon_2)v(x) + \alpha_{x-1} + \nu_x^1 - \nu_x^2 + \psi - \beta_{0,x}^1 = 0 \tag{D.20c}$$

$$(2\epsilon_1 + \epsilon_2)v(x) + \alpha_{x-1} + \nu_x^2 - \nu_x^1 + \psi - \beta_{0,x}^2 = 0 \tag{D.20d}$$

$$-1+\epsilon_1+\epsilon_2+\epsilon_3+\epsilon_4=0. \tag{D.20e}$$

Therefore, we have

$$v(x)=\frac{-\lambda\alpha_x+\alpha_{x-1}+\nu_x^1-\nu_x^2+\psi-\beta_{1,x}^1}{2\epsilon_3-\epsilon_2-\eta} \tag{D.21a}$$
$$=\frac{-\lambda\alpha_x+\alpha_{x-1}+\nu_x^2-\nu_x^1+\psi-\beta_{1,x}^2}{2-2\epsilon_1-\epsilon_2+\eta} \tag{D.21b}$$
$$=\frac{-\alpha_{x-1}-\nu_x^1+\nu_x^2-\psi+\beta_{0,x}^1}{2\epsilon_1+\epsilon_2} \tag{D.21c}$$
$$=\frac{-\alpha_{x-1}-\nu_x^2+\nu_x^1-\psi+\beta_{0,x}^2}{2\epsilon_1+\epsilon_2}. \tag{D.21d}$$

Based on the above equations, we can have the following lemma.

Lemma 25. *If there exists a $\tilde{x}>0$ for which we have $\gamma(0,1,\tilde{x})>0$, $\gamma(0,2,\tilde{x})>0$, $\gamma(1,1,\tilde{x})>0$, and $\gamma(1,2,\tilde{x})>0$, then $\eta=0$, $\epsilon_2=\epsilon_4=0$, and $\nu_x^1=\nu_x^2$ for all $x>0$. Further, we have $\beta_{0,x}^1=\beta_{0,x}^2$ and $\beta_{1,x}^1=\beta_{1,x}^2$.*

Proof. *Looking at equations* (D.21c) *and* (D.21d)*, we have $2\nu_x^1-2\nu_x^2+\beta_{0,x}^2=\beta_{0,x}^1$. Since $\gamma(0,1,\tilde{x})>0$ and $\gamma(0,2,\tilde{x})>0$ we have $\beta_{0,\tilde{x}}^2=\beta_{0,\tilde{x}}^1=0$ and it results in $\nu_{\tilde{x}}^1=\nu_{\tilde{x}}^2$. Also since $\gamma(1,1,\tilde{x})>0$ and $\gamma(1,2,\tilde{x})>0$, which results in $\beta_{1,\tilde{x}}^2=\beta_{1,\tilde{x}}^1=0$, we must have $2\epsilon_3-\epsilon_2-\eta=2-2\epsilon_1-\epsilon_2+\eta$ or equivalently, $\epsilon_1+\epsilon_3=1+\eta$. According to equation* (D.20e)*, we must have $\eta=0$ and $\epsilon_2=\epsilon_4=0$. Further, for any $x>0$ and each type $i\in\mathcal{I}$, we either have $\gamma(1,i,x)>0$ or $\gamma(0,i,x)>0$ or both. Therefore, either $\beta_{1,\tilde{x}}^i=0$ or $\beta_{0,\tilde{x}}^i=0$ or both. Assume $\beta_{0,\tilde{x}}^1=0$. Then if $\beta_{0,\tilde{x}}^2=0$, we must have $\nu_{\tilde{x}}^1=\nu_{\tilde{x}}^2$ and the result is proved. If $\beta_{1,\tilde{x}}^2=0$, then from equations* (D.21c) *and* (D.21d) *we must have $\nu_{\tilde{x}}^2\geq\nu_{\tilde{x}}^1$. From equations* (D.21a) *and* (D.21b)*, we must have $\nu_{\tilde{x}}^1\geq\nu_{\tilde{x}}^2$. Therefore, $\nu_{\tilde{x}}^1=\nu_{\tilde{x}}^2$. Therefore, due to equation* (D.21)*, we must have $\beta_{0,x}^1=\beta_{0,x}^2$ and $\beta_{1,x}^1=\beta_{1,x}^2$ for all $x>0$.*

Note that all of the results of Lemma 25 hold if we have $2\epsilon_3-\epsilon_2-\eta=2-2\epsilon_1-\epsilon_2+\eta$, i.e., the denominators of the coefficients in equations (D.21a) and (D.21b) are equal. In other words, Lemma 25 states a condition in which we must have $2\epsilon_3-\epsilon_2-\eta=2-2\epsilon_1-\epsilon_2+\eta$ and consequently, the rest of the results hold. However, these coefficients might be equal without having the condition stated in Lemma 25. In the next lemma, we show what the solution looks like if we have $2\epsilon_3-\epsilon_2-\eta=2-2\epsilon_1-\epsilon_2+\eta$.

Lemma 26. *If we have $2\epsilon_3-\epsilon_2-\eta=2-2\epsilon_1-\epsilon_2+\eta$, then there is a threshold $\tilde{x}$ such that for $x\geq\tilde{x}$ we have $\gamma(s=1,i,x)=0$ for all $i\in\mathcal{I}$, and for $x<\tilde{x}$, $\gamma(s=0,i,x)=0$ for all*

$i \in \mathcal{I}$ except for some points $\tilde{\mathcal{X}} = \{x_1, x_2, \ldots\},\ x_k < \tilde{x}$ for which we can have $\gamma(0,1,x_k) > 0$, or $\gamma(0,2,x_k) > 0$, where all $x_k \in \tilde{\mathcal{X}}$ satisfy the following condition. There exists $\epsilon_1 > 0$ and ψ such that

$$(2\sum_{x=0}^{x_k} \lambda^x v(x))\epsilon_1 + (\sum_{x=0}^{x_k} \lambda^x)\psi = \sum_{x=0}^{x_k-1} \lambda^x v(x),\ \forall x_k \in \tilde{\mathcal{X}}.$$

Proof. *If we have $2\epsilon_3 - \epsilon_2 - \eta = 2 - 2\epsilon_1 - \epsilon_2 + \eta$, then $\eta = 0$, $\epsilon_2 = \epsilon_4 = 0$, and $\nu_x^1 = \nu_x^2$ for all $x > 0$. Further, we have $\beta_{0,x}^1 = \beta_{0,x}^2$ and $\beta_{1,x}^1 = \beta_{1,x}^2$. This results in the following.*

$$v(x) = \frac{-\lambda\alpha_x + \alpha_{x-1} + \psi - \beta_{1,x}^1}{2 - 2\epsilon_1} \tag{D.22a}$$
$$= \frac{-\lambda\alpha_x + \alpha_{x-1} + \psi - \beta_{1,x}^2}{2 - 2\epsilon_1} \tag{D.22b}$$
$$= \frac{-\alpha_{x-1} - \psi + \beta_{0,x}^1}{2\epsilon_1} = \frac{-\alpha_{x-1} - \psi + \beta_{0,x}^2}{2\epsilon_1}. \tag{D.22c}$$

From the above equations, we can conclude that $\beta_{0,x}^1 = \beta_{0,x}^2$ and $\beta_{1,x}^1 = \beta_{1,x}^2$. It means that we if we have $\beta_{1,\tilde{x}}^1 = \beta_{1,\tilde{x}}^2 > 0$ which results in $\gamma(s=1,i,\tilde{x}) = 0$ for all $i \in \mathcal{I}$, the stationary distribution $\mu(x)$ is zero for all $x > \tilde{x}$ and therefore $\gamma(s=1,i,x) = 0$ for all $i \in \mathcal{I}$. For $x < \tilde{x}$, we have $\beta_{1,x}^1 = \beta_{1,x}^2 = 0$. We also either have $\beta_{0,x}^1 = \beta_{0,x}^2 > 0$, which results in $\gamma(s=0,i,x) = 0$ for all $i \in \mathcal{I}$ or we have $\beta_{0,x}^1 = \beta_{0,x}^2 = 0$ which can allow us to have $\gamma(0,1,x) > 0$, $\gamma(0,2,x) > 0$. Suppose either $\gamma(0,1,x) > 0$ or $\gamma(0,2,x) > 0$ for $x \in \{x_1, x_2, \ldots\}$. By writing the equation (D.22) *for $x = 0$ and due to the fact that at least one of $\beta_{1,0}^1$ or $\beta_{1,0}^2$ is zero, we have $v(0) = \frac{1}{2-2\epsilon_1}(-\lambda\alpha_0 + \psi)$.*

For $x \in \{x_1, x_2, \ldots\}$ we have

$$v(x) = \frac{-\lambda\alpha_x + \alpha_{x-1} + \psi}{2 - 2\epsilon_1} = \frac{-\alpha_{x-1} - \psi}{2\epsilon_1} \tag{D.23}$$

which results in $v(x) = -\frac{\lambda}{2}\alpha_x$. Using equations (D.22) *and*(D.23) *we have*

$$\sum_{x=0}^{x_1} \lambda^x v(x) = \frac{1}{2 - 2\epsilon_1}(-\lambda^{x_1+1}\alpha_{x_1} + \psi\sum_{x=0}^{x_1} \lambda^x) \tag{D.24}$$
$$\Rightarrow\ (2\sum_{x=0}^{x_1} \lambda^x v(x))\epsilon_1 + (\sum_{x=0}^{x_1} \lambda^x)\psi = \sum_{x=0}^{x_1-1} \lambda^x v(x). \tag{D.25}$$

having x_1, the above is an equation with respect to ϵ and ψ. In general, if we have x_1 and x_2, we should be able to determine ϵ and ψ and we can have x_k for $k \geq 3$ only if they result in linearly *dependent equations in* (D.25) *and this might not be true for general* $v(\cdot)$.

Note that if for all $x < \tilde{x}$ *we have* $\gamma(s = 0, i, x) = 0$ *for all* $i \in \mathcal{I}$*, then* $q(1) = q(2)$ *and the objective of the planner would be zero. Therefore, this is probably not the solution of the optimization problem. In order to create discrimination between users of type 1 and type 2, the planner can only consider different policies for these two types at* $x_k \in \tilde{\mathcal{X}}$.

In Lemma 26, we investigated the solution under the assumption of $2\epsilon_3 - \epsilon_2 - \eta = 2 - 2\epsilon_1 - \epsilon_2 + \eta$. Note that due to equation (D.20e), we can not have $2\epsilon_3 - \epsilon_2 - \eta > 2 - 2\epsilon_1 - \epsilon_2 + \eta$. Therefore if the equality does not hold, we have $2\epsilon_3 - \epsilon_2 - \eta < 2 - 2\epsilon_1 - \epsilon_2 + \eta$. In the next lemma we present some results under this inequality assumption.

Lemma 27. *If* $2\epsilon_3 - \epsilon_2 - \eta < 2 - 2\epsilon_1 - \epsilon_2 + \eta$*, then we have the following. If* $v(x) > 0$ *and* $\sigma(s = 1|1, x) > 0$*, we have* $\sigma(s = 1|2, x) = 1$*. Furthermore, if* $v(x) < 0$ *and* $\sigma(s = 1|2, x) > 0$*, we have* $\sigma(s = 1|1, x) = 1$.

Proof. *Looking at equation* (D.21)*, if* $v(x) > 0$*, since* $2\epsilon_3 - \epsilon_2 - \eta < 2 - 2\epsilon_1 - \epsilon_2 + \eta$*, we must have the following.*

$$v(x) = \frac{\alpha_{x-1} - \lambda\alpha_x + \nu_x^1 - \nu_x^2 + \psi - \beta_{1,x}^1}{2\epsilon_3 - \epsilon_2 - \eta} \tag{D.26}$$

$$= \frac{\alpha_{x-1} - \lambda\alpha_x + \nu_x^2 - \nu_x^1 + \psi - \beta_{1,x}^2}{2 - 2\epsilon_1 - \epsilon_2 + \eta} \tag{D.27}$$

$$\Rightarrow\ \alpha_{x-1} - \lambda\alpha_x + \nu_x^1 - \nu_x^2 + \psi - \beta_{1,x}^1 < \alpha_{x-1} - \lambda\alpha_x + \nu_x^2 - \nu_x^1 + \psi - \beta_{1,x}^2 \tag{D.28}$$

$$\Rightarrow\ \beta_{1,x}^2 + 2\nu_x^1 - 2\nu_x^2 < \beta_{1,x}^1 \tag{D.29}$$

We also have the following.

$$-\nu_x^1 + \nu_x^2 + \beta_{0,x}^1 = -\nu_x^2 + \nu_x^1 + \beta_{0,x}^2 \tag{D.30}$$

$$\Rightarrow\ \beta_{0,x}^1 = \beta_{0,x}^2 + 2\nu_x^1 - 2\nu_x^2 \tag{D.31}$$

We can have three cases for $\nu_x^1 - \nu_x^2$*. We either have* $\nu_x^1 > \nu_x^2$*, which results in* $\beta_{1,x}^2 < \beta_{1,x}^1$*, which means that* $\beta_{1,x}^1 > 0$ *and therefore,* $\gamma(s = 1, 1, x) = 0$*. On the other hand, we must have* $\beta_{0,x}^1 > 0$ *and therefore,* $\gamma(s = 0, 1, x) = 0$*. This is a contradiction for those* x*'s that* $\mu(x) > 0$*, which are the ones that we are interested in. If we have* $\nu_x^1 < \nu_x^2$*, we must have* $\beta_{0,x}^2 > 0$*, and therefore,* $\gamma(s = 0, 2, x) = 0$*. If we have* $\nu_x^1 = \nu_x^2$*, we must have* $\beta_{1,x}^1 > 0$ *and therefore,* $\gamma(s = 1, 1, x) = 0$*. Therefore, for* $v(x) > 0$*, we have* $\sigma(s = 1|2, x) = 1$ *if* $\sigma(s = 1|1, x) > 0$*. If* $v(x) < 0$*, then we must* *have*

$$\alpha_{x-1} - \lambda\alpha_x + \nu_x^1 - \nu_x^2 + \psi - \beta_{1,x}^1 > \alpha_{x-1} - \lambda\alpha_x + \nu_x^2 - \nu_x^1 + \psi - \beta_{1,x}^2 \tag{D.32}$$

$$\Rightarrow\ \beta_{1,x}^1 + 2\nu_x^2 - 2\nu_x^1 < \beta_{1,x}^2. \tag{D.33}$$

Consequently, if $\nu_x^2 > \nu_x^1$*, we have* $\beta_{1,x}^2 > 0$ *and therefore,* $\gamma(s = 1, 2, x) = 0$*. On the other hand, due to equation* (D.31)*, we have* $\beta_{0,x}^2 > 0$ *and therefore,* $\gamma(s = 0, 2, x) = 0$*, which is a contradiction. Hence,* $\nu_x^2 \leq \nu_x^1$*. If* $\nu_x^2 < \nu_x^1$*, we have* $\beta_{0,x}^1 > 0$ *and so* $\gamma(s = 0, 1, x) = 0$*. If* $\nu_x^2 = \nu_x^1$*, we have* $\beta_{1,x}^2 > 0$ *and therefore,* $\gamma(s = 1, 2, x) = 0$*. Hence, for* $v(x) < 0$ *we have* $\sigma(s = 1|1, x) = 1$ *if* $\sigma(s = 1|2, x) > 0$*.*

D.5 Proof of Theorem 8

Similar to the proof of Theorem 7, we use KKT conditions to find the properties for the solution of the optimization problem (5.17). We use the following dual variables for each constraints in (5.17).

$$t_0 - N\sum_{j=1}^{i-1}\sum_{x=0}^{\infty} v(x)\gamma(s = 1, j, x) - i\sum_{x=0}^{\infty} v(x)\sum_{s,j}\gamma(s, j, x) + p \leq 0,\ \forall i \in \mathcal{I}:\ \epsilon_1^i \tag{D.34a}$$

$$t_0 - N\sum_{j=1}^{i-1}\sum_{x=0}^{\infty} v(x)\gamma(s = 1, j, x) \leq 0,\ \forall i \in \mathcal{I}:\ \epsilon_2^i \tag{D.34b}$$

$$\sum_{x=0}^{\infty} v(x)\gamma(s = 1, i, x) - \sum_{x=0}^{\infty} v(x)\gamma(s = 1, i + 1, x) \leq 0,\ 1 \leq i < N:\ \eta^i \tag{D.34c}$$

$$\sum_{s,i}\gamma(s, i, x + 1) - \lambda\sum_{i}\gamma(1, i, x) = 0,\ \forall x \geq 0:\ \alpha_x \tag{D.34d}$$

$$\sum_{s}\gamma(s, i, x) - \frac{1}{N}\sum_{s,j}\gamma(s, j, x) = 0,\ \forall i \in \mathcal{I},\ x \geq 0:\ \nu_x^i \tag{D.34e}$$

$$\sum_{s,i,x}\gamma(s, i, x) - 1 = 0:\ \psi \tag{D.34f}$$

$$-\gamma(s, i, x) \leq 0,\ \forall s \in \{0, 1\}, i \in \mathcal{I}, x \geq 0:\ \beta_{s,x}^i. \tag{D.34g}$$

The KKT Stationarity conditions will result in the following.

$$v(x) = \frac{\lambda\alpha_x - \alpha_{x-1} - \nu_x^i + \frac{1}{N}\sum_{i\in\mathcal{I}}\nu_x^i - \psi + \beta_{1,x}^i}{\sum_{j\in\mathcal{I}} j\epsilon_1^j - (2i - N) - N\sum_{j=i+1}^{N}(\epsilon_1^j + \epsilon_2^j) + \eta^i - \eta^{i-1}} \tag{D.35a}$$

$$= \frac{-\alpha_{x-1} - \nu_x^i + \frac{1}{N}\sum_{i\in\mathcal{I}} \nu_x^i - \psi + \beta_{0,x}^i}{\sum_{j\in\mathcal{I}} j\epsilon_1^j} \tag{D.35b}$$

We denote the denominator in equation (D.35a) by c^i. By comparing c^i for different $i \in \mathcal{I}$, we can prove the results.

Lemma 28. *If $c^{i_1} > c^{i_2}$, then i_2 dominates i_1.*

Proof. *From equation* (D.35b)*, we denote*

$$\beta_{0,x}^i - \nu_x^i = k_x. \tag{D.36}$$

Notice that $\beta_{0,x}^i - \nu_x^i$ is fixed across types of users. Therefore, we have

$$v(x) = \frac{\lambda\alpha_x - \alpha_{x-1} + \frac{1}{N}\sum_{i\in\mathcal{I}} \nu_x^i - \psi + k_x + \beta_{1,x}^i - \beta_{0,x}^i}{\sum_{j\in\mathcal{I}} j\epsilon_1^j - (2i-N) - N\sum_{j=i+1}^N(\epsilon_1^j + \epsilon_2^j) + \eta^i - \eta^{i-1}}. \tag{D.37}$$

Hence, if $c^{i_1} > c^{i_2}$ and for $v(x) > 0$, we have $\beta_{1,x}^{i_1} - \beta_{0,x}^{i_1} > \beta_{1,x}^{i_2} - \beta_{0,x}^{i_2}$. It means that for $v(x) > 0$, whenever we have $\beta_{1,x}^{i_2} > 0$ (meaning $\gamma(s = 1, i_2, x) = 0$), we must have $\beta_{1,x}^{i_1} > 0$ (meaning $\gamma(s = 1, i_1, x) = 0$). Similarly, for $v(x) < 0$, we have $\beta_{1,x}^{i_1} - \beta_{0,x}^{i_1} < \beta_{1,x}^{i_2} - \beta_{0,x}^{i_2}$, which means whenever we have $\beta_{1,x}^{i_1} > 0$ (meaning $\gamma(s = 1, i_1, x) = 0$), we must have $\beta_{1,x}^{i_2} > 0$ (meaning $\gamma(s = 1, i_2, x) = 0$). Therefore, i_2 dominates i_1.

Assume $i_2 > i_1$. We either have $c^{i_1} > c^{i_2}$, which according to Lemma 28 indicates that i_2 dominates i_1, or we have $c^{i_1} \leq c^{i_2}$. It indicates the following.

$$\begin{aligned}
&\sum_{j\in\mathcal{I}} j\epsilon_1^j - (2i_1 - N) - N\sum_{j=i_1+1}^N(\epsilon_1^j + \epsilon_2^j) + \eta^{i_1} - \eta^{i_1-1} \\
&\qquad \leq \sum_{j\in\mathcal{I}} j\epsilon_1^j - (2i_2 - N) - N\sum_{j=i_2+1}^N(\epsilon_1^j + \epsilon_2^j) + \eta^{i_2} - \eta^{i_2-1} && \text{(D.38)} \\
&\qquad \Rightarrow 2(i_2 - i_1) - N\sum_{j=i_1+1}^{i_2}(\epsilon_1^j + \epsilon_2^j) + \eta^{i_1} - \eta^{i_1-1} - \eta^{i_2} + \eta^{i_2-1} \leq 0 && \text{(D.39)}
\end{aligned}$$

Since we have $\eta^i \geq 0$, $\epsilon_1^i \geq 0$, and $\epsilon_2^i \geq 0$, for the above inequality to hold, we must have

$$N\sum_{j=i_1+1}^{i_2}(\epsilon_1^j + \epsilon_2^j) + \eta^{i_1-1} + \eta^{i_2} \geq 2(i_2 - i_1).$$

Also note that using KKT conditions, we have

$$\sum_{i\in\mathcal{I}}(\epsilon_1^i+\epsilon_2^i)=1. \tag{D.40}$$

Therefore, for $i_2-i_1>\frac{N}{2}$, we must have $\eta^{i_1-1}+\eta^{i_2}>0$, which means we should either have $\eta^{i_2}>0$ (meaning $q(i_2)=q(i_2+1)$) or $\eta^{i_1-1}>0$ (meaning $q(i_1)=q(i_1-1)$). If $i_2-i_1<=\frac{N}{2}$, then we either have $\eta^{i_2}>0$ (meaning $q(i_2)=q(i_2+1)$) or $\eta^{i_1-1}>0$ (meaning $q(i_1)=q(i_1-1)$), or $\sum_{j=i_1+1}^{i_2}(\epsilon_1^j+\epsilon_2^j)>0$, which means that there exists an $\bar{i}$ where $i_1<\bar{i}\le i_2$, for which we have $\epsilon_1^{\bar{i}}>0$ or $\epsilon_2^{\bar{i}}>0$. It indicates that for $\bar{i}$, the utility of the mechanism is equal to the outside option utility. That is, we have $\bar{i}q(\bar{i})-t(\bar{i})=(\bar{i}\overline{v}-p)^+$.

D.6 Proof of Theorem 9

In the no information case, the objective of the planner is to maximize the following.

$$\lambda\sum_{i=1}^{N}\gamma_i P_I(i)(t_0+iq(i)-\sum_{j=1}^{i-1}q(j)) \tag{D.41}$$

If $\overline{v}\ge 0$, we have

$$\begin{aligned}\lambda\sum_{i=1}^{N}\gamma_i P_I(i)(t_0+iq(i)-\sum_{j=1}^{i-1}q(j)) =&\lambda\sum_{i=1}^{N}\gamma_i P_I(i)(t_0+i\sigma(1|i)\overline{v}-\sum_{j=1}^{i-1}q(j)) &\text{(D.42a)}\\ \le&\lambda\sum_{i=1}^{N}\gamma_i P_I(i)(t_0+i\overline{v}-\sum_{j=1}^{i-1}q(j))\le\lambda p &\text{(D.42b)}\end{aligned}$$

where the last inequality is due to the individual rationality constrains ($\gamma_i=1,\ \forall i\in\mathcal{I}$) that impose the condition $iq(i)-t(i)=\sum_{j=1}^{i-1}q(j)-t_0\ge(i\overline{v}-p)^+$.

If $\overline{v}<0$, we have $i\overline{v}-p<0$ and therefore, we must have $iq(i)-t(i)=\sum_{j=1}^{i-1}q(j)-t_0\ge 0$. Hence, we can write the following.

$$\begin{aligned}\lambda\sum_{i=1}^{N}\gamma_i P_I(i)(t_0+iq(i)-\sum_{j=1}^{i-1}q(j)) =&\lambda\sum_{i=1}^{N}\gamma_i P_I(i)(t_0+i\sigma(1|i)\overline{v}-\sum_{j=1}^{i-1}q(j)) &\text{(D.43a)}\\ \le&0\le\lambda p &\text{(D.43b)}\end{aligned}$$

D.7 Proof of Theorem 10

Due to irrelevance of message quoting in the full information scenario, the planner only needs to satisfy individual rationality constraints, which are given below.

$$i\sum_{x=0}^{\infty}(v(x))^{+}\mu(x)-t\geq(i\sum_{x=0}^{\infty}v(x)\mu(x)-p)^{+},\ \forall i\in\mathcal{I} \tag{D.44}$$

Therefore, we have

$$t=\min_{i}(i\sum_{x=0}^{\infty}(v(x))^{+}\mu(x)-(i\sum_{x=0}^{\infty}v(x)\mu(x)-p)^{+}) \tag{D.45}$$

$$\leq\min_{i}(i\sum_{x=0}^{\infty}(v(x))^{+}\mu(x)-(i\sum_{x=0}^{\infty}v(x)\mu(x)-p)) \tag{D.46}$$

$$=\min_{i}-iv(x_{neg})\mu(x_{neg})+p=-v(x_{neg})\mu(x_{neg})+p, \tag{D.47}$$

where x_{neg} is the smallest x for which we have $v(x)<0$. We have $\mu(x_{neg})=\frac{\lambda^{x_{neg}}}{\sum_{x=0}^{x_{neg}}\lambda^{x}}$. Therefore, the revenue of the planner is less than or equal to $-\lambda v(x_{neg})\frac{\lambda^{x_{neg}}}{\sum_{x=0}^{x_{neg}}\lambda^{x}}+\lambda p$.

APPENDIX E

Proofs for Chapter 6

E.1 Proof of Lemma 16

$$\begin{aligned}
\overline{c}_{t,\pi}(s) &= \mathbb{E}\{c_{\tau,m}(s)|t,\pi\} \\
&= \frac{\mathbb{E}\{q(s)|t,\pi\}}{\mu} + \mathbb{E}\{(s-\tau)^{+} - c(s-\tau)\} \\
&= \frac{1}{\mu\overline{m}(t)} \int_{m,\tau>\tilde{\tau}_m(s)} f_\tau(\tau)\pi(m|\tau)m(t)(\int_{l=-\infty}^{s} m(l)\mathrm{d}l - \mu(s-\tau)^{+} + \mu c(\tau - s) + \mu(s-\tau)^{+})\mathrm{d}\tau\mathrm{d}m \\
&\quad + \frac{1}{\overline{m}(t)} \int_{m,\tau<\tilde{\tau}_m(s)} f_\tau(\tau)\pi(m|\tau)m(t)(c(\tau-s) + (s-\tau)^{+})\,\mathrm{d}\tau\mathrm{d}m \\
&= \frac{1}{\mu\overline{m}(t)} \int_{m,\tau>\tilde{\tau}_m(s)} f_\tau(\tau)\pi(m|\tau)m(t)(\int_{l=-\infty}^{s} m(l)\mathrm{d}l - \mu cs + \mu c\tau)\,\mathrm{d}\tau\mathrm{d}m \\
&\quad + \frac{1}{\overline{m}(t)} \int_{m,\tau<\tilde{\tau}_m(s)} f_\tau(\tau)\pi(m|\tau)m(t) \\
&\qquad\qquad (c(\tau-s) + (s-\tau))\,\mathrm{d}\tau\mathrm{d}m \\
&= \frac{1}{\mu\overline{m}(t)} \int_{m,\tau>\tilde{\tau}_m(s)} f_\tau(\tau)\pi(m|\tau)m(t)(\int_{l=-\infty}^{s} m(l)\mathrm{d}l - \mu cs)\mathrm{d}\tau\mathrm{d}m \\
&\quad + \frac{1}{\overline{m}(t)} \int_{m,\tau<\tilde{\tau}_m(s)} f_\tau(\tau)\pi(m|\tau)m(t)((1-c)s - \tau)\,\mathrm{d}\tau\mathrm{d}m + c\mathbb{E}(\tau|t) \qquad \text{(E.1)}
\end{aligned}$$

where $\overline{m}(t) = \int_{\tau,m} f_\tau(\tau)\pi(m|\tau)m(t)\mathrm{d}\tau\mathrm{d}m$.

E.2 Proof of Lemma 17

In order for an agent to obey her suggestion, we must have $\overline{c}_{t,\pi}(t)$ to be the global minimizer of $\overline{c}_{t,\pi}(s)$. In the next lemma, we will show that $\overline{c}_{t,\pi}(s)$ is convex and therefore, any local minimizer is a global minimizer.

Lemma 29. *$\overline{c}_{t,\pi}(s)$ is convex with respect to s.*

Proof. In order to prove convexity of $\overline{c}_{t,\pi}(s)$, we prove that its derivative is increasing. But we first go over some preliminary results.

For $t \geq 0$, if $\int_{-\infty}^{t} m(s)\mathrm{d}s \leq \mu t$, we have

$$\int_{-\infty}^{t} m(s)\mathrm{d}s = \mu(t - \tilde{\tau}_m(t)) \tag{E.2}$$

Since we have $\tilde{\tau}_m(t) = 0$ for $t \leq 0$, the following holds.

$$\tilde{\tau}_m(t) = (t - \int_{-\infty}^{t} \frac{m(s)}{\mu}\mathrm{d}s)^{+} \tag{E.3}$$

Lemma 30. *If $m(t) \leq \mu$ for all t, then $\tilde{\tau}_m(t)$ is continuous and increasing with respect to t. Furthermore, $\tilde{\tau}_m(t)$ is differentiable for all t except possibly for $t = \tilde{t}$, where $\tilde{t}$ will be characterized in the proof.*

Proof. Since $m(t) \leq \mu$, there exists a time $\tilde{t}$, for which we have $t - \int_{-\infty}^{t} \frac{m(s)}{\mu}\mathrm{d}s \geq 0$ for $t \geq \tilde{t}$ and $t - \int_{-\infty}^{t} \frac{m(s)}{\mu}\mathrm{d}s < 0$ for $t < \tilde{t}$. It is clear that $\tilde{\tau}_m(t)$ is continuous, differentiable and increasing for $t < \tilde{t}$ and for $t > \tilde{t}$. Also, since the assumption of $m(t) \leq \mu$ eliminates the possibility of $m(t)$ including a delta function, $\tilde{\tau}_m(t)$ is continuous for all t.

□

Since we have assumed $m(t) \leq \mu$, and according to Lemma 30, we know $\tilde{\tau}_m(t)$ is continuous, increasing, and differentiable for $t \neq \tilde{t}$, we can write the following for $\frac{\mathrm{d}}{\mathrm{d}s}\overline{c}_{t,\pi}(s)$ for $s \neq \tilde{t}$.

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}\overline{c}_{t,\pi}(s) &= \frac{\mathrm{d}}{\mathrm{d}s}\frac{1}{\overline{m}(t)}\int_{\tau,m} c_{m,\tau}(s) f_\tau(\tau)\pi(m|\tau)m(t)\mathrm{d}\tau\mathrm{d}m \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\frac{1}{\overline{m}(t)}(\int_m \int_0^{\tilde{\tau}_m(s)} (1-c)(s-\tau) f_\tau(\tau)\pi(m|\tau)m(t)\mathrm{d}\tau\mathrm{d}m \\
&\quad + \int_m \int_{\tilde{\tau}_m(s)}^{\infty} (\int_{l=-\infty}^{s} \frac{m(l)}{\mu}\mathrm{d}l - c(s-\tau)) f_\tau(\tau)\pi(m|\tau)m(t)\mathrm{d}\tau\mathrm{d}m) \\
&= \frac{1}{\overline{m}(t)}(\int_m \int_0^{\tilde{\tau}_m(s)} (1-c) f_\tau(\tau)\pi(m|\tau)m(t)\mathrm{d}\tau\mathrm{d}m \\
&\quad + \int_m \tilde{\tau}_m'(s)(1-c)(s-\tilde{\tau}_m(s)) f_\tau(\tilde{\tau}_m(s))\pi(m|\tilde{\tau}_m(s))m(t)\mathrm{d}m \\
&\quad - \int_m \tilde{\tau}_m'(s)(\int_{l=-\infty}^{s} \frac{m(l)}{\mu}\mathrm{d}l - c(s-\tilde{\tau}_m(s))) f_\tau(\tilde{\tau}_m(s))\pi(m|\tilde{\tau}_m(s))m(t)\mathrm{d}m
\end{aligned}$$

$$
\begin{aligned}
&+\int_m \int_{\tilde{\tau}_m(s)}^{\infty} (\frac{m(s)}{\mu} - c) f_\tau(\tau)\pi(m|\tau)m(t)\mathrm{d}\tau\mathrm{d}m) \\
&= \frac{1}{\overline{m}(t)}(\int_m \int_{\tau=0}^{\tilde{\tau}_m(s)} f_\tau(\tau)\pi(m|\tau)m(t)\mathrm{d}\tau\mathrm{d}m \\
&+\int_m \int_{\tilde{\tau}_m(s)}^{\infty} f_\tau(\tau)\pi(m|\tau)m(t)\frac{m(s)}{\mu}\mathrm{d}\tau\mathrm{d}m) - c
\end{aligned} \tag{E.4}
$$

One can see that the left and right derivative of $\overline{c}_{t,\pi}(s)$ at $s = \tilde{t}$ are equal to equation (E.4) and therefore, equation (E.4) holds for all s. According to equation (E.4), since $m(t) \leq \mu$ for all t, the term in the first integral of $\frac{\mathrm{d}}{\mathrm{d}s}\overline{c}_{t,\pi}(s)$, i.e., $f_\tau(\tau)\pi(m|\tau)m(t)$, is greater than the term in the second integral, i.e., $f_\tau(\tau)\pi(m|\tau)m(t)\frac{m(s)}{\mu}$. Therefore, as we increase s and therefore we increase $\tilde{\tau}_m(s)$, we are increasing the range of the first integral and decreasing the range of the second, thus, increasing $\frac{\mathrm{d}}{\mathrm{d}s}\overline{c}_{t,\pi}(s)$. Hence, $\overline{c}_{t,\pi}(s)$ is convex with respect to s. □

According to Lemma 29, it is necessary and sufficient for t to be a local minimizer of $\overline{c}_{t,\pi}(s)$ to be its global minimizer. Therefore, we should have $\frac{\mathrm{d}}{\mathrm{d}s}\overline{c}_{t,\pi}(s)|_t = 0$ and we have the result by setting (E.4) at t to 0.

E.3 Proof of Theorem 11

Let us assume that we have a positive queue over the interval $[t_1, t_2]$, i.e., $q_{\tau,m}(t) > 0$ for $t \in (t_1, t_2)$ and $q_{\tau,m}(t_1) = 0$ and $q_{\tau,m}(t_2) = 0$. Note that we do not have any assumptions on the queue length at other times. In order not to have any profitable deviations for agents arriving in $t \in [t_1, t_2]$, we should have $c'_{\tau,m}(t) = 0$ for $t \in (t_1, t_2)$ to avoid profitable deviations by changing the position inside the queue. It implies the following.

$$
c'_{\tau,m}(t) = \frac{m(t)}{\mu} - c = 0 \Rightarrow m(t) = \mu c,\ t \in (t_1, t_2) \tag{E.5}
$$

Since the queue size is 0 at t_2, an agent arriving at t_2 does not have any incentives for arriving later. Furthermore, in order for an agent arriving at t_1 not to have profit by arriving earlier, it is sufficient to have $t_1 \leq \tau$. This condition implies that we can not have multiple queues in the full information equilibrium.

We can calculate the queue length at t as follows.

$$
q_{\tau,m}(t) = \int_{t_1}^{t} m(t)\mathrm{d}t - \mu(t-\tau)^+ = c\mu(t-t_1) - \mu(t-\tau)^+ \tag{E.6}
$$

Setting the queue at t_2 to 0 will give us the equation below that relates t_1 and t_2 to τ.

$$\tau = ct_1 + (1-c)t_2$$

Since all agents must arrive between $[t_1, t_2]$, we have $c\mu(t_2 - t_1) = 1$, and therefore, we have

$$t_1 = \tau - \frac{1-c}{c\mu}, \quad t_2 = \tau + \frac{1}{\mu} \tag{E.7}$$

E.4 Proof of Lemma 18

Assume we have a delta function of size a at some time t in the arrival process. We will show that the agent arriving at time t has a profit by arriving slightly before t at $s = t - \mathrm{d}t$. Note that we have $q_{\tau,m}(t - \mathrm{d}t) = q_{\tau,m}(t) - a$ for every τ. The average cost of arriving at time t is

$$\overline{c}(t) = \int_{\tau=0}^{\infty} (\frac{q_{\tau,m}(t)}{\mu} - c(t-\tau) + (t-\tau)^+) f_\tau(\tau)\mathrm{d}\tau \tag{E.8}$$

On the other hand, the average cost of arriving at time $s = t - \mathrm{d}t$ is

$$\overline{c}(t-\mathrm{d}t) = \int_{\tau=0}^{\infty} (\frac{q_{\tau,m}(t-\mathrm{d}t)}{\mu} - c(t-\mathrm{d}t-\tau) + (t-\mathrm{d}t-\tau)^+) f_\tau(\tau)\mathrm{d}\tau \tag{E.9}$$

Subtracting the two will result in the following.

$$\overline{c}(t) - \overline{c}(t-\mathrm{d}t) = \frac{a}{\mu} - c\mathrm{d}t + \mathrm{d}t\mathbf{1}(t \geq \tau) > 0, \text{ for } \mathrm{d}t \text{ small enough.}$$

Therefore, we can not have a delta function in the arrival process.

Next, assume $m(t) > \mu$ for some time t, which due to piecewise continuity of m, implies that $m(s) > \mu$ in a neighborhood of t. If m is not continuous at t, we consider a point in the neighborhood of t where m is continuous. Therefore, we have $q(t) > 0$ in a neighborhood of t. Let us denote t_0 to be the latest time before t that $q(t_0) = 0$. We have $q(t) = \int_{l=t_0}^{t} m(l)\mathrm{d}l - \mu(t - \max(t_0, \tau))\mathbf{1}(t \geq \tau)$. Therefore, we can write the derivative of the average value of the cost as follows.

$$\begin{aligned}\overline{c}'(t) = \frac{\mathrm{d}}{\mathrm{d}t}(\int_{\tau=0}^{\infty}(\int_{l=t_0}^{t} \frac{m(l)}{\mu}\mathrm{d}l - (t - \max(t_0,\tau))\mathbf{1}(t \geq \tau) \\ + (t-\tau)^+ - c(t-\tau)) f_\tau(\tau)\mathrm{d}\tau = \frac{m(t)}{\mu} - c\end{aligned} \tag{E.10}$$

Setting $\overline{c}'(t) = 0$ results in $m(t) = c\mu < \mu$. Therefore, we can not have $m(t) > \mu$.

E.5 Proof of Theorem 12

We define $\overline{c}(t)$ to be the average value of the cost $c_{\tau,m}(t)$ with respect to τ using $f_\tau(\cdot)$. In order to have an equilibrium, each agent arriving at time t should to be acting rationally by doing so. Therefore, we should have $\overline{c}'(t) = 0$ for every t that $m(t) > 0$ in a neighborhood of t. If $m(t) > 0$ in a right neighborhood of t, the right derivative of the expected cost should be zero and the left derivative should be non-positive. Similar rule applies for the left neighborhoods.

In Lemma 18, we proved that in order to satisfy incentive constraints in the no information case, we can never have $m(t) > \mu$, and $m(t)$ can never include a delta function. Therefore, we have $m(t) \le \mu$ for all t. Also, according to Lemma 30, we know $\tilde{\tau}_m(t)$ is continuous and differentiable. Therefore, the derivative of the average value of the cost, $\overline{c}'(t)$ is given as follows.

$$
\begin{aligned}
\overline{c}'(t) =& \frac{\mathrm{d}}{\mathrm{d}t}\int_\tau c_{\tau,m}(t) f_\tau(\tau) d\tau \\
=& \frac{\mathrm{d}}{\mathrm{d}t}\Big(\int_0^{\tilde{\tau}_m(t)} (1-c)(t-\tau) f_\tau(\tau)\mathrm{d}\tau + \int_{\tilde{\tau}_m(t)}^{\infty} \Big(\int_{l=-\infty}^{t} \frac{m(l)}{\mu}\mathrm{d}l - c(t-\tau)\Big) f_\tau(\tau)\mathrm{d}\tau\Big) \\
=& \tilde{\tau}_m'(t)(1-c)(t-\tilde{\tau}_m(t)) f_\tau(\tilde{\tau}_m(t)) - \tilde{\tau}_m'(t)\Big(\int_{l=-\infty}^{t} \frac{m(l)}{\mu}\mathrm{d}l - c(t-\tilde{\tau}_m(t))\Big) f_\tau(\tilde{\tau}_m(t)) \\
&+ \int_0^{\tilde{\tau}_m(t)} (1-c) f_\tau(\tau)\mathrm{d}\tau + \int_{\tilde{\tau}_m(t)}^{\infty} \Big(\frac{m(t)}{\mu} - c\Big) f_\tau(\tau)\mathrm{d}\tau \\
=& \int_0^{\tilde{\tau}_m(t)} (1-c) f_\tau(\tau)\mathrm{d}\tau + \int_{\tilde{\tau}_m(t)}^{\infty} \Big(\frac{m(t)}{\mu} - c\Big) f_\tau(\tau)\mathrm{d}\tau \\
=& 1 - e^{-\lambda\tilde{\tau}_m(t)} + \frac{m(t)}{\mu} e^{-\lambda\tilde{\tau}_m(t)} - c \\
=& 1 - c - e^{-\lambda(t - \int_{l=-\infty}^{t} \frac{m(l)}{\mu}\mathrm{d}l)^+}\Big(1 - \frac{m(t)}{\mu}\Big)
\end{aligned} \tag{E.11}
$$

Setting $\overline{c}'(t) = 0$ will result in the following.

$$
e^{-\lambda(t - \int_{l=-\infty}^{t} \frac{m(l)}{\mu}\mathrm{d}l)^+}\Big(1 - \frac{m(t)}{\mu}\Big) = 1 - c \tag{E.12}
$$

Equation (E.12) holds for all t such that we have $m(t) > 0$. Note that if $m(s) = 0$ for $s < t$ and $m(s) > 0$ for $s \ge t$, the left derivative of $\overline{c}(t)$ is non-positive given equation (E.12) holds for t. This implies that, as we increase t, we can have discontinuity in $m(t)$ from 0 to a non zero value. This is not the case for right neighborhoods with zero arrivals, i.e., $m(s) = 0$ for an interval of $s > t$ and $m(s) > 0$ for an interval of $s \leq t$. In this case, the right derivative will be non-positive if (E.12) holds for t. However, we need the right derivative to be positive for the agents to not have profitable deviations. Hence, whenever we have $m(s) = 0$ for an interval of $s > t$, we should have $m(t) = 0$, i.e., $m(t)$ must be continuous when transitioning to zero from non zero values. Also, note that the assumption of $m(t) \leq \mu$ clearly holds for any $m(t)$ satisfying equation (E.12). Therefore, we have the following.

If we take the derivative of equation (E.12) w.r.t. t for $t \geq \tilde{t}$ ($\tilde{t}$ is defined in Lemma 30), we have

$$\begin{aligned}
&e^{-\lambda(t-\int_{l=-\infty}^{t}\frac{m(l)}{\mu}\mathrm{d}l)}(\lambda(1-\frac{m(t)}{\mu})^2+\frac{m'(t)}{\mu})=0\\
&\Rightarrow\ -\lambda(1-\frac{m(t)}{\mu})^2-\frac{m'(t)}{\mu}=0\ \Rightarrow\ \frac{\mathrm{d}m}{(\mu-m)^2}=-\frac{\lambda}{\mu}\mathrm{d}t\\
&\Rightarrow\frac{1}{\mu-m}=\frac{-\lambda t+\beta}{\mu}\Rightarrow m(t)=\mu-\frac{\mu}{\beta-\lambda t}
\end{aligned}\tag{E.13}$$

In order to derive constant β, we assume that $m(t)$ is 0 outside of an interval of $[t_1, t_2]$. If $\tilde{t} > 0$ then we must have $t_1 < 0$. For now, we assume $\tilde{t_1} = 0$ and therefore, $t_1 \geq 0$. We must have $m(t_2) = 0$ as mentioned in the discussions above. Also, since $\int_0^{t_2} m(t)\mathrm{d}t = 1$, we have $\tilde{\tau}_m(t_2) = t_2 - \frac{1}{\mu}$. Therefore, according to equation (E.12), we have the following for t_2.

$$e^{-\lambda(t_2-\frac{1}{\mu})}=1-c\Rightarrow\ t_2=\frac{-\ln(1-c)}{\lambda}+\frac{1}{\mu}\tag{E.14}$$

and we know $m(t_2) = 0$ which will give us β as follows.

$$\mu-\frac{\mu}{\beta-\lambda t_2}=0\Rightarrow\ \beta=\lambda t_2+1\Rightarrow\beta=-\ln(1-c)+\frac{\lambda}{\mu}+1.\tag{E.15}$$

On the other hand, we must have $\int_{t_1}^{t_2} m(t) = 1$, which results in the following equation to derive t_1.

$$\begin{aligned}
&\int_{t_1}^{t_2}m(t)\mathrm{d}t=\int_{t_1}^{t_2}(\mu-\frac{\mu}{\beta-\lambda t})\mathrm{d}t=1\\
&\Rightarrow\mu(t_2-t_1)-\frac{\mu}{\lambda}\ln(\lambda(t_2-t_1)+1)=1\\
&\Rightarrow\ln(1-c)+\lambda t_1+\ln(\frac{\lambda}{\mu}-\ln(1-c)-\lambda t_1+1)=0
\end{aligned}\tag{E.16}$$

If t_1 derived from the above equation is non-negative, then the equilibrium is characterized. Next, we consider the possibility of $t_1 \leq 0$, which results in $\tilde{t} > 0$. For $t \leq \tilde{t}$, $\tilde{\tau}_m(t) = 0$ and according to (E.12) we have $1 - \frac{m(t)}{\mu} = 1 - c$ and therefore, we must have $m(t) = \mu c$ for $t_1 \leq t \leq \tilde{t}$. The queue size must be 0 at $\tilde{t}$ if $\tau = 0$, because for $t > \tilde{t}$, we have $\tilde{\tau}_m(t) > 0$. This results in the following.

$$\mu c(\tilde{t} - t_1) = \mu \tilde{t} \Rightarrow \tilde{t} = -\frac{c}{1-c} t_1 \tag{E.17}$$

On the other hand, since $\tilde{\tau}_m(t) > 0$ for $t > \tilde{t}$ and $\tilde{\tau}_m(\tilde{t}) = 0$, $m(t)$ follows equation (E.13) for $t \geq \tilde{t}$ and we have $m(\tilde{t}) = \mu - \frac{\mu}{\beta - \lambda \tilde{t}}$. Therefore, we have the following.

$$\begin{aligned}
&m(\tilde{t}) = \mu - \frac{\mu}{\beta - \lambda \tilde{t}} = \mu c \\
&\Rightarrow 1 - \frac{1}{\lambda(t_2 + \frac{c}{1-c} t_1) + 1} = c \\
&\Rightarrow \lambda((1-c)t_2 + ct_1) = c \\
&\Rightarrow -(1-c)\ln(1-c) + \frac{(1-c)\lambda}{\mu} + \lambda c t_1 = c \\
&\Rightarrow t_1 = \frac{1-c}{\lambda c}\ln(1-c) - \frac{1-c}{\mu c} + \frac{1}{\lambda}
\end{aligned} \tag{E.18}$$

If the value of t_1 above is negative, the no information equilibrium is characterized. Notice that we might have two types of no information equilibrium, one with negative t_1 and one with a positive one if the value of t_1 satisfying equations (E.16) and (E.18) is positive and negative, respectively.

E.6 Proof of Theorem 13

Consider any m in the support of $\pi(\cdot|\tau)$. We show $m(t)$ as $m(t) = \mu c + \delta(t)$, where $\delta(t)$ is defined over $[\underline{t}_\tau, \overline{t}_\tau]$. Since we have $\int_{\underline{t}_\tau}^{\overline{t}_\tau} m(t)\mathrm{d}t = 1$ and $\overline{t}_\tau - \underline{t}_\tau \leq \frac{1}{\mu c}$, we must have $\int_{\underline{t}_\tau}^{\overline{t}_\tau} \delta(t)\mathrm{d}t \geq 0$. Using Lemma 17 we have the following.

$$\begin{aligned}
(1-c)\int_{\tau,m} & f_\tau(\tau)\pi(m|\tau)m(t)\mathbf{1}(\tau \leq \tilde{\tau}_m(t))\mathrm{d}m \\
+ \frac{1}{\mu}\int_{\tau,m} & f_\tau(\tau)\pi(m|\tau)m(t)(m(t) - \mu c)\mathbf{1}(\tau > \tilde{\tau}_m(t))\mathrm{d}m = 0
\end{aligned} \tag{E.19}$$

Since $\tilde{\tau}_m(t)$ is increasing in t, we can define its inverse by $\tilde{t}_m(\tau)$, i.e., we have $q_{\tau,m}(t) > 0$ for $\underline{t}_\tau \leq t < \tilde{t}_m(\tau)$ and $q_{\tau,m}(t) = 0$ for $t \geq \tilde{t}_m(\tau)$. We have

$$
\begin{aligned}
&\frac{1}{\mu}\int_{\tau,m} f_\tau(\tau)\pi(m|\tau)\int_t m(t)(\mu(1-c)\mathbf{1}(t \geq \tilde{t}_m(\tau)) \\
&\quad + (m(t)-\mu c)\mathbf{1}(t < \tilde{t}_m(\tau)))\mathrm{d}\tau\mathrm{d}m\mathrm{d}t = 0 \\
&\frac{1}{\mu}\int_{\tau,m} f_\tau(\tau)\pi(m|\tau)\int_t (\delta(t)+\mu c)(\mu(1-2c+c) \\
&\qquad\quad \mathbf{1}(t \geq \tilde{t}_m(\tau)) + \delta(t)\mathbf{1}(t < \tilde{t}_m(\tau)))\mathrm{d}\tau\mathrm{d}m\mathrm{d}t = 0 \\
&\frac{1}{\mu}\int_{\tau,m} f_\tau(\tau)\pi(m|\tau)\int_{\underline{t}_\tau}^{\bar{t}_\tau} (\mu c\delta(t) + \mu^2c^2\mathbf{1}(t \geq \tilde{t}_m(\tau)) \\
&\quad + \mu(1-2c)m(t)\mathbf{1}(t \geq \tilde{t}_m(\tau)) + \delta(t)^2\mathbf{1}(t < \tilde{t}_m(\tau)))\mathrm{d}\tau\mathrm{d}m\mathrm{d}t = 0 \qquad \text{(E.20)}
\end{aligned}
$$

We notice that all of the elements of the above integral are greater than or equal to zero. Therefore, they must all be zero for the sum to be zero. Hence, we have

$$
\begin{aligned}
&\int_{\underline{t}_\tau}^{\bar{t}_\tau} \mu^2c^2\mathbf{1}(t \geq \tilde{t}_m(\tau))\mathrm{d}t = 0 \\
&\int_{\underline{t}_\tau}^{\bar{t}_\tau} \mu(1-2c)m(t)\mathbf{1}(t \geq \tilde{t}_m(\tau))\mathrm{d}t = 0 \\
&\int_{\underline{t}_\tau}^{\bar{t}_\tau} \delta(t)^2\mathbf{1}(t < \tilde{t}_m(\tau))\mathrm{d}t = 0 \qquad \text{(E.21)}
\end{aligned}
$$

Therefore, we must have $\delta(t) = 0$ for all $t \in [\underline{t}_\tau, \bar{t}_\tau], m$. Hence, $m(t) = \mu c$ and thus, $\bar{t}_\tau - \underline{t}_\tau = \frac{1}{\mu c}$, i.e., the time span of the arrival processes are equal to the one in the full information equilibrium. We must also have $\mathbf{1}(t \geq \tilde{t}_m(\tau)) = 0$ for all $t \in [\underline{t}_\tau, \bar{t}_\tau], m, \tau$, which is consistent with assumption (c). Therefore, we must have $\underline{t}_\tau = \tau - \frac{1-c}{c\mu}$ and $\bar{t}_\tau = \tau + \frac{1}{\mu}$. Hence, $\pi(\cdot|\tau)$ is supported only over the full information equilibrium arrival process and the theorem is proved.

E.7 Proof of Lemma 19

If the planner restricts his attention to the set of signaling strategies that satisfy assumptions (b) and (c), we have $q_{\tau,m}(t) = \int_{l=-\infty}^{t} m(l)\mathrm{d}l - \mu(\tau - t)^+$. Therefore, we have the following for $\bar{c}_{t,\pi}(s)$

and its derivative.

$$
\begin{aligned}
\overline{c}_{t,\pi}(s) &= \frac{1}{\mu\overline{m}(t)}\int_m\int_{\tau=\underline{\tau}(t)}^{\overline{\tau}(t)}(\int_{l=-\infty}^{s} m(l)\mathrm{d}l - \mu(s-\tau)^+ \\
&\quad + \mu c(\tau - s) + \mu(s-\tau)^+)f_\tau(\tau)\pi(m|\tau)m(t)\mathrm{d}\tau\mathrm{d}mc \\
&= \frac{1}{\mu\overline{m}(t)}\int_m\int_{\tau=\underline{\tau}(t)}^{\overline{\tau}(t)} f_\tau(\tau)\pi(m|\tau)m(t)(\int_{l=-\infty}^{s} m(l)\mathrm{d}l - \mu cs)\,\mathrm{d}\tau\mathrm{d}m + c\mathbb{E}(\tau|t) \quad \text{(E.22)}
\end{aligned}
$$

$$
\frac{\mathrm{d}}{\mathrm{d}s}\overline{c}_{t,\pi}(s) = \frac{1}{\mu\overline{m}(t)}\int_m\int_{\tau=\underline{\tau}(t)}^{\overline{\tau}(t)} f_\tau(\tau)\pi(m|\tau)m(t)(m(s)-\mu c)\,\mathrm{d}\tau\mathrm{d}m \quad \text{(E.23)}
$$

According to Lemma 17, if we set $\frac{\mathrm{d}}{\mathrm{d}s}\overline{c}_{t,\pi}(s)|_t = 0$ we get the result.

E.8 Proof of Lemma 20

$$
\begin{aligned}
\overline{s}(\pi) &= \int_t\int_{\tau,m} f_\tau(\tau)\pi(m|\tau)m(t)c_{\tau,m}(t)\mathrm{d}\tau\mathrm{d}m\mathrm{d}t \\
&= \int_t\int_{\tau,m} f_\tau(\tau)\pi(m|\tau)m(t)(\frac{q(t)}{\mu} + c(\tau-t)^+ + (1-c)(t-\tau)^+)\mathrm{d}\tau\mathrm{d}m\mathrm{d}t \\
&= \int_t\int_{\tau,m} f_\tau(\tau)\pi(m|\tau)m(t)(\frac{\int_{l=-\infty}^{t} m(l)dl - \mu(t-\tau)^+}{\mu} + c(\tau-t) + (t-\tau)^+)\mathrm{d}\tau\mathrm{d}m\mathrm{d}t \\
&= \int_t\int_{\tau,m} f_\tau(\tau)\pi(m|\tau)m(t)(\frac{\int_{l=-\infty}^{t} m(l)\mathrm{d}l}{\mu} + c(\tau-t))\mathrm{d}\tau\mathrm{d}m\mathrm{d}t \\
&= \frac{1}{\mu}\int_\tau f_\tau(\tau)(\int_{t=\underline{t}_\tau}^{\overline{t}_\tau}\int_{s=\underline{t}_\tau}^{t}(R_{m,\tau}(t,s) - \mu c\overline{m}_\tau(t))\mathrm{d}s\mathrm{d}t + \mu c(\tau - \underline{t}_\tau))\mathrm{d}\tau \quad \text{(E.24)}
\end{aligned}
$$

E.9 Proof of Theorem 14

Suppose $m(t)$ is in the support of $\pi(\cdot|\tau)$. Assume $\overline{t}_\tau - \underline{t}_\tau = T$. We show $m(t)$ as $m(t) = \mu c + \delta(t)$. Since we have $\int_{\underline{t}_\tau}^{\overline{t}_\tau} m(t)\mathrm{d}t = 1$, we must have $\int_{\underline{t}_\tau}^{\overline{t}_\tau}\delta(t)\mathrm{d}t = 1 - \mu cT \geq 0$. Lemma 19 results in the following.

$$
\begin{aligned}
&\int_{\tau,m} f_\tau(\tau)\pi(m|\tau)\int_t(\mu^2c^2 + 2\mu c\delta(t) + \delta(t)^2)\mathrm{d}t\mathrm{d}m\mathrm{d}\tau = \mu c \\
&\Rightarrow \mu c(1-\mu cT) + \int_{\tau,m} f_\tau(\tau)\pi(m|\tau)\int_t\delta(t)^2\mathrm{d}t\mathrm{d}m\mathrm{d}\tau = 0
\end{aligned}
$$

$$\Rightarrow \mathbb{E}[\int_t \delta^2(t)\mathrm{d}t] = 0 \;\Rightarrow \delta(t) = 0 \quad wp.\ 1$$

$$\mu c T = 1 \Rightarrow T = \frac{1}{\mu c} \tag{E.25}$$

Therefore, we have $m(t) = \mu c$ with probability one and $\bar{t}_\tau - \underline{t}_\tau = \frac{1}{\mu c}$. Therefore, we must have $m(t)$ to be the full information equilibrium, i.e., $\underline{t}_\tau = \tau - \frac{1-c}{c\mu}$ and $\bar{t}_\tau = \tau + \frac{1}{\mu}$. Therefore, $\pi(\cdot|\tau)$ is supported only over the full information equilibrium arrival process and the result is proved.

Bibliography

[1] M. J. Osborne and A. Rubinstein, *A course in game theory*. MIT press, 1994.

[2] D. Fudenberg and J. Tirole, *Game theory*. MIT press, 1991.

[3] J. Watson, "A general, practicable definition of perfect bayesian equilibrium," *unpublished draft*, 2017.

[4] Y. Ouyang, H. Tavafoghi, and D. Teneketzis, "Dynamic games with asymmetric information: Common information based perfect bayesian equilibria and sequential decomposition," *IEEE Trans. Automatic Control*, vol. 62, no. 1, pp. 222–237, Jan 2017.

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

[6] A. V. Banerjee, "A simple model of herd behavior," *The Quarterly Journal of Economics*, vol. 107, no. 3, pp. 797–817, 1992.

[7] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, vol. 100, no. 5, pp. 992–1026, 1992.

[8] L. Smith and P. Sørensen, "Pathological outcomes of observational learning," *Econometrica*, vol. 68, no. 2, pp. 371–398, 2000.

[9] T. N. Le, V. G. Subramanian, and R. A. Berry, "Information cascades with noise," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 2, pp. 239–251, 2017.

[10] S.-T. Su, V. G. Subramanian, and G. Schoenebeck, “Social learning with questions,” in *Proceedings of the 14th Workshop on the Economics of Networks, Systems and Computation*, 2019, pp. 1–1.

[11] L. Hurwicz and S. Reiter, *Designing economic mechanisms*. Cambridge University Press, 2006.

[12] T. Börgers, D. Krahmer, and R. Strausz, *An introduction to the theory of mechanism design*. Oxford University Press, 2015.

[13] T. Groves and J. Ledyard, “Optimal allocation of public goods: A solution to the “free rider” problem,” *Econometrica: Journal of the Econometric Society*, pp. 783–809, 1977.

[14] L. Hurwicz, “Outcome functions yielding Walrasian and Lindahl allocations at Nash equilibrium points,” *The Review of Economic Studies*, vol. 46, no. 2, pp. 217–225, 1979.

[15] S. Yang and B. Hajek, “Revenue and stability of a mechanism for efficient allocation of a divisible good,” University of Illinois Urbana-Champaign, Tech. Rep., 2005. [Online]. Available: http://www.ifp.illinois.edu/~hajek/Papers/YangHajek06.pdf

[16] R. Srikant and L. Ying, *Communication networks: an optimization, control, and stochastic networks perspective*. Cambridge University Press, 2013.

[17] R. T. Maheswaran and T. Başar, “Social welfare of selfish agents: motivating efficiency for divisible resources,” in *43rd IEEE Conference on Decision and Control, 2004. CDC.*, vol. 2. IEEE, 2004, pp. 1550–1555.

[18] R. Jain and J. Walrand, “An efficient Nash-implementation mechanism for network resource allocation,” *Automatica*, vol. 46, no. 8, pp. 1276–1283, 2010.

[19] A. Kakhbod and D. Teneketzis, “Correction to “An efficient game form for multi-rate multicast service provisioning”,” *IEEE J. Select. Areas Commun.*, vol. 31, no. 7, pp. 1355–1356, July 2013.

[20] ——, “Correction to “An efficient game form for unicast service provisioning” [feb 12 392-404],” *IEEE Transactions on Automatic Control*, vol. 60, no. 2, pp. 584–585, Feb 2015.

[21] A. Sinha and A. Anastasopoulos, “Mechanism design for resource allocation in networks with intergroup competition and intragroup sharing,” *IEEE Trans. on Control of Network Systems*, vol. 5, no. 3, pp. 1098–1109, Sept. 2018.

[22] J. Huang, R. A. Berry, and M. L. Honig, "Auction-based spectrum sharing," *Mobile Networks and Applications*, vol. 11, no. 3, pp. 405–418, 2006.

[23] B. Wang, Y. Wu, Z. Ji, K. J. R. Liu, and T. C. Clancy, "Game theoretical mechanism design methods," *IEEE Signal Processing Magazine*, vol. 25, no. 6, pp. 74–84, November 2008.

[24] S. Wang, P. Xu, X. Xu, S. Tang, X. Li, and X. Liu, "TODA: truthful online double auction for spectrum allocation in wireless networks," in *IEEE Symposium on New Frontiers in Dynamic Spectrum*, 2010, pp. 1–10.

[25] M. M. Khalili, P. Naghizadeh, and M. Liu, "Designing cyber insurance policies: Mitigating moral hazard through security pre-screening," in *International Conference on Game Theory for Networks*. Springer, 2017, pp. 63–73.

[26] F. Fu and M. van der Schaar, "Noncollaborative resource management for wireless multimedia applications using mechanism design," *IEEE Transactions on Multimedia*, vol. 9, no. 4, pp. 851–868, June 2007.

[27] S. Sharma and D. Teneketzis, "Local public good provisioning in networks: A Nash implementation mechanism," *IEEE Journal on Selected Areas in Communication*, vol. 30, no. 11, pp. 2105–2116, 2012.

[28] M. Fahrioglu and F. L. Alvarado, "Designing incentive compatible contracts for effective demand management," *IEEE Transactions on Power Systems*, vol. 15, no. 4, pp. 1255–1260, Nov 2000.

[29] S. Caron and G. Kesidis, "Incentive-based energy consumption scheduling algorithms for the smart grid," in *First IEEE International Conference on Smart Grid Communications (SmartGridComm), 2010*, Oct 2010, pp. 391–396.

[30] P. Samadi, H. Mohsenian-Rad, R. Schober, and V. W. S. Won, "Advanced demand side management for the future smart grid using mechanism design," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1170–1180, Sept 2012.

[31] H. Everett III, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources," *Operations research*, vol. 11, no. 3, pp. 399–417, 1963.

[32] D. Gabay and B. Mercier, *A dual algorithm for the solution of non linear variational problems via finite element approximation*. Institut de recherche d'informatique et d'automatique, 1975.

[33] J. Eckstein and D. P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 1-3, pp. 293–318, 1992.

[34] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[35] M. Kraning, E. Chu, J. Lavaei, S. Boyd *et al.*, "Dynamic network energy management via proximal message passing," *Foundations and Trends® in Optimization*, vol. 1, no. 2, pp. 73–126, 2014.

[36] M. Zhu and S. Martínez, *Distributed optimization-based control of multi-agent networks in complex environments*. Springer, 2015.

[37] G. Dantzig, *Linear programming and extensions*. Princeton university press, 2016.

[38] E. Kamenica and M. Gentzkow, "Bayesian persuasion," *American Economic Review*, vol. 101, no. 6, pp. 2590–2615, 2011.

[39] D. Bergemann, A. Bonatti, and A. Smolin, "The design and price of information," *American economic review*, vol. 108, no. 1, pp. 1–48, 2018.

[40] N. Heydaribeni and A. Anastasopoulos, "Structured equilibria for dynamic games with asymmetric information and dependent types," *arXiv preprint arXiv:2009.04253*, 2020.

[41] ——, "Linear equilibria for dynamic LQG games with asymmetric information and dependent types," *IEEE Conference on Decision and Control (CDC)*, 2019.

[42] ——, "Structured equilibria for dynamic games with asymmetric information and dependent types," *Submitted to IEEE Transactions on Automatic Control*, 2020.

[43] D. Vasal and A. Anastasopoulos, "Signaling equilibria for dynamic LQG games with asymmetric information," in *Proc. IEEE Conf. on Decision and Control*, Las Vegas, NV, Dec. 2016, pp. 6901–6908.

[44] D. Vasal, A. Sinha, and A. Anastasopoulos, "A systematic process for evaluating structured perfect bayesian equilibria in dynamic games with asymmetric information," *IEEE Transactions on Automatic Control*, vol. 64, no. 1, pp. 81–96, 2018.

[45] D. Abreu, D. Pearce, and E. Stacchetti, "Toward a theory of discounted repeated games with imperfect monitoring," *Econometrica: Journal of the Econometric Society*, pp. 1041–1063, 1990.

[46] I. Bistritz, N. Heydaribeni, and A. Anastasopoulos, "Do Informational Cascades Happen with Non-myopic Agents?" *arXiv preprint arXiv:1905.01327*, 2019.

[47] N. Heydaribeni, I. Bistritz, and A. Anastasopoulos, "Informational cascades can be avoided with non-myopic agents," in *Proc. Allerton Conf. Commun., Control, Comp.*, Monticello, Illinois, Sept. 2019, pp. 655–662.

[48] I. Bistritz, N. Heydaribeni, and A. Anastasopoulos, "Do Informational Cascades Happen with Non-myopic Agents?" *Submitted to IEEE Transactions on Automatic Control*, 2019.

[49] I. Bistritz and A. Anastasopoulos, "Characterizing non-myopic information cascades in Bayesian learning," in *Proc. IEEE Conf. on Decision and Control*, Miami Beach, FL, Dec. 2018, pp. 2716–2721.

[50] N. Heydaribeni and A. Anastasopoulos, "Distributed mechanism design for network resource allocation problems," *IEEE Transactions on Network Science and Engineering*, 2019.

[51] ——, "Distributed mechanism design for multicast transmission," in *Proc. IEEE Conf. on Decision and Control*, Miami Beach, FL, Dec. 2018, pp. 4200–4205.

[52] ——, "Distributed mechanism design for unicast transmission," in *Information Theory and Applications*, San Diego, CA, Feb. 2018, pp. 1–6.

[53] ——, "Joint information and mechanism design for a queuing system," in *Proc. IEEE Conf. on Decision and Control*, Austin, TX, USA, Dec. 2021.

[54] N. Heydaribeni and K. Savla, "Information design for a non-atomic service scheduling game," in *2021 IEEE Conference on Decision and Control (CDC)*. Austin, TX, USA: IEEE, 2021.

[55] J. B. Lasserre, "A semidefinite programming approach to the generalized problem of moments," *Mathematical Programming*, vol. 112, no. 1, pp. 65–92, 2008.

[56] D. Henrion, J.-B. Lasserre, and J. Löfberg, "Gloptipoly 3: moments, optimization and semidefinite programming," *Optimization Methods & Software*, vol. 24, no. 4-5, pp. 761–779, 2009.

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

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

[59] D. Vasal, A. Sinha, and A. Anastasopoulos, "A systematic process for evaluating structured perfect Bayesian equilibria in dynamic games with asymmetric information," *IEEE Trans. Automatic Control*, vol. 64, no. 1, pp. 81–96, Jan. 2019.

[60] A. Mahajan and A. Nayyar, "Sufficient statistics for linear control strategies in decentralized systems with partial history sharing," *IEEE Transactions on Automatic Control*, vol. 60, no. 8, pp. 2046–2056, 2015.

[61] S. Yuksel, "Stochastic nestedness and the belief sharing information pattern," *IEEE Transactions on Automatic Control*, vol. 54, no. 12, pp. 2773–2786, 2009.

[62] H. Tavafoghi, Y. Ouyang, and D. Teneketzis, "A unified approach to dynamic decision problems with asymmetric information-part ii: Strategic agents," *arXiv preprint arXiv:1812.01132*, 2018.

[63] D. Tang, H. Tavafoghi, V. Subramanian, A. Nayyar, and D. Teneketzis, "Dynamic games among teams with delayed intra-team information sharing," *arXiv preprint arXiv:2102.11920*, 2021.

[64] D. Vasal and A. Anastasopoulos, "Signaling equilibria for dynamic LQG games with asymmetric information," *IEEE Trans. on Control of Network Systems*, vol. 8, no. 3, pp. 1177–1188, 2021.

[65] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, pp. 992–1026, 1992.

[66] Y.-C. Ho *et al.*, "Team decision theory and information structures in optimal control problems–part i," *IEEE Transactions on Automatic control*, vol. 17, no. 1, pp. 15–22, 1972.

[67] P. R. Kumar and P. Varaiya, *Stochastic systems: estimation, identification, and adaptive control*. Englewood Cliffs, NJ: Prentice-Hall, 1986.

[68] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968.

[69] T. Başar, "Two-criteria LQG decision problems with one-step delay observation sharing pattern," *Information and Control*, vol. 38, no. 1, pp. 21–50, 1978.

[70] E. Altman, V. Kambley, and A. Silva, "Stochastic games with one step delay sharing information pattern with application to power control," in *2009 International Conference on Game Theory for Networks*. IEEE, 2009, pp. 124–129.

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

[72] V. P. Crawford and J. Sobel, "Strategic information transmission," *Econometrica: Journal of the Econometric Society*, pp. 1431–1451, 1982.

[73] F. Farokhi, A. M. Teixeira, and C. Langbort, "Gaussian cheap talk game with quadratic cost functions: When herding between strategic senders is a virtue," in *2014 American Control Conference*. IEEE, 2014, pp. 2267–2272.

[74] M. O. Sayin and T. Başar, "Dynamic information disclosure for deception," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 1110–1117.

[75] B. Gharesifard and J. Cortés, "Evolution of players' misperceptions in hypergames under perfect observations," *IEEE Transactions on Automatic Control*, vol. 57, no. 7, pp. 1627–1640, 2011.

[76] D. Fudenberg and J. Tirole, "Perfect bayesian equilibrium and sequential equilibrium," *journal of Economic Theory*, vol. 53, no. 2, pp. 236–260, 1991.

[77] D. M. Kreps and R. Wilson, "Sequential equilibria," *Econometrica: Journal of the Econometric Society*, pp. 863–894, 1982.

[78] R. S. Bielefeld, "Reexamination of the perfectness concept for equilibrium points in extensive games," in *Models of Strategic Rationality*. Springer, 1988, pp. 1–31.

[79] D. Bergemann and S. Morris, "Correlated equilibrium in games with incomplete information," *Cowles Foundation Discussion Paper No. 1822, Economic Theory Center*

Working Paper No. 024-2011, Available at SSRN: https://ssrn.com/abstract=1941708 or http://dx.doi.org/10.2139/ssrn.1941708, 2011.

[80] B. Von Stengel and F. Forges, "Extensive-form correlated equilibrium: Definition and computational complexity," *Mathematics of Operations Research*, vol. 33, no. 4, pp. 1002–1022, 2008.

[81] H. T. Jahormi, "On design and analysis of cyber-physical systems with strategic agents," Ph.D. dissertation, University of Michigan, Ann Arbor, 2017.

[82] D. Acemoglu, M. A. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian learning in social networks," *The Review of Economic Studies*, vol. 78, no. 4, pp. 1201–1236, 2011.

[83] T. N. Le, V. G. Subramanian, and R. A. Berry, "Information cascades with noise," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 2, pp. 239–251, 2017.

[84] B. Çelen and S. Kariv, "Observational learning under imperfect information," *Games and Economic Behavior*, vol. 47, no. 1, pp. 72–86, 2004.

[85] G. Schoenebeck, S.-T. Su, and V. Subramanian, "Social learning with questions," *arXiv preprint arXiv:1811.00226*, 2018.

[86] V. Bala and S. Goyal, "Learning from neighbours," *The review of economic studies*, vol. 65, no. 3, pp. 595–621, 1998.

[87] E. Mossel and O. Tamuz, "Making consensus tractable," *ACM Transactions on Economics and Computation (TEAC)*, vol. 1, no. 4, p. 20, 2013.

[88] E. Mossel, A. Sly, and O. Tamuz, "Asymptotic learning on bayesian social networks," *Probability Theory and Related Fields*, vol. 158, no. 1-2, pp. 127–157, 2014.

[89] ——, "Strategic learning and the topology of social networks," *Econometrica*, vol. 83, no. 5, pp. 1755–1794, 2015.

[90] M. Harel, E. Mossel, P. Strack, and O. Tamuz, "On the speed of social learning," arXiv, Tech. Rep., 2014.

[91] A. Jadbabaie, E. Mossel, and M. A. Rahimian, "Bayesian group decisions: Algorithms and complexity," arXiv, Tech. Rep., 2017.

[92] K. Dasaratha, B. Golub, and N. Hak, "Social learning in a dynamic environment," arXiv, Tech. Rep., 2018.

[93] B. Çelen and S. Kariv, "Distinguishing informational cascades from herd behavior in the laboratory," *American Economic Review*, vol. 94, no. 3, pp. 484–498, 2004.

[94] A. Guarino, H. Harmgart, and S. Huck, "Aggregate information cascades," *Games and Economic Behavior*, vol. 73, no. 1, pp. 167–185, 2011.

[95] H. Herrera and J. Hörner, "Biased social learning," *Games and Economic Behavior*, vol. 80, pp. 131–146, 2013.

[96] A. Guarino and P. Jehiel, "Social learning with coarse inference," *American Economic Journal: Microeconomics*, vol. 5, no. 1, pp. 147–74, 2013.

[97] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

[98] G. Ellison and D. Fudenberg, "Rules of thumb for social learning," *Journal of political Economy*, vol. 101, no. 4, pp. 612–643, 1993.

[99] ——, "Word-of-mouth communication and social learning," *The Quarterly Journal of Economics*, vol. 110, no. 1, pp. 93–125, 1995.

[100] P. Molavi, A. Tahbaz-Salehi, and A. Jadbabaie, "Foundations of non-bayesian social learning," Columbia Business School Research Paper No. 15-95, Tech. Rep., 2017. [Online]. Available: https://ssrn.com/abstract=2683607

[101] A. Jadbabaie, P. Molavi, A. Sandroni, and A. Tahbaz-Salehi, "Non-bayesian social learning," *Games and Economic Behavior*, vol. 76, no. 1, pp. 210–225, 2012.

[102] Y. Peres, M. Z. Racz, A. Sly, and I. Stuhl, "How fragile are information cascades?" arXiv, Tech. Rep., 2017.

[103] A. Nedić, A. Olshevsky, and C. A. Uribe, "A tutorial on distributed (non-bayesian) learning: Problem, algorithms and results," in *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 6795–6801.

[104] I. H. Lee, "On the convergence of informational cascades," *Journal of Economic theory*, vol. 61, no. 2, pp. 395–411, 1993.

[105] W. Hann-Caruthers, V. V. Martynov, and O. Tamuz, “The speed of sequential asymptotic learning,” *Journal of Economic Theory*, vol. 173, pp. 383–409, 2018.

[106] D. Vasal and A. Anastasopoulos, “Decentralized Bayesian learning in dynamic games,” Arxiv, Tech. Rep., 2016. [Online]. Available: http://arxiv.org/abs/1607.06847

[107] ——, “A systematic process for evaluating structured perfect Bayesian equilibria in dynamic games with asymmetric information,” in *American Control Conference*, Boston, MA, July 2016, pp. 3378–3385.

[108] Y. Ouyang, H. Tavafoghi, and D. Teneketzis, “Dynamic games with asymmetric information: Common information based perfect bayesian equilibria and sequential decomposition,” *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 222–237, 2016.

[109] D. Vasal and A. Anastasopoulos, “Signaling equilibria for dynamic LQG games with asymmetric information,” *IEEE Transactions on Control of Network Systems*, 2021.

[110] D. Fudenberg and J. Tirole, *Game theory*. Cambridge, Massachusetts, 1991.

[111] A. Sinha and A. Anastasopoulos, “A practical mechanism for network utility maximization for unicast flows on the Internet,” in *Proc. International Conf. Communications*, June 2015, pp. 5679–5684.

[112] ——, “Distributed mechanism design with learning guarantees,” in *Proc. IEEE Conf. on Decision and Control*, Melbourne, Australia, Dec. 2017, pp. 5000–5007.

[113] A. Sinha, “Mechanism design with allocative, informational and learning constraints,” Ph.D. dissertation, University of Michigan, 2017.

[114] J. Feigenbaum, C. H. Papadimitriou, and S. Shenker, “Sharing the cost of multicast transmissions,” *Journal of Computer and System Sciences*, vol. 63, no. 1, pp. 21–41, 2001.

[115] J. Feigenbaum and S. Shenker, “Distributed algorithmic mechanism design: recent results and future directions,” in *Proceedings of the 6th international workshop on Discrete algorithms and methods for mobile computing and communications*. ACM, 2002, pp. 1–13.

[116] R. Sami, “Distributed algorithmic mechanism design,” Ph.D. dissertation, Yale University, 2003.

[117] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic game theory*. Cambridge University Press, 2007.

[118] Z. Huang, S. Mitra, and N. Vaidya, "Differentially private distributed optimization," in *Proceedings of the 2015 International Conference on Distributed Computing and Networking*, ser. ICDCN '15. New York, NY, USA: ACM, 2015, pp. 4:1–4:10. [Online]. Available: http://doi.acm.org/10.1145/2684464.2684480

[119] E. Nozari, P. Tallapragada, and J. Cortés, "Differentially private distributed convex optimization via objective perturbation," in *American Control Conference (ACC), 2016*. IEEE, 2016, pp. 2061–2066.

[120] J. Cortés, G. E. Dullerud, S. Han, J. Le Ny, S. Mitra, and G. J. Pappas, "Differential privacy in control and network systems," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*. IEEE, 2016, pp. 4252–4272.

[121] S. Han, U. Topcu, and G. J. Pappas, "Differentially private distributed constrained optimization," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 50–64, 2017.

[122] M. Kearns, M. L. Littman, and S. Singh, "Graphical models for game theory," in *Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 2001, pp. 253–260.

[123] M. O. Jackson, *Social and economic networks*. Princeton university press, 2010.

[124] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic theory*. Oxford university press New York, 1995, vol. 1.

[125] P. J. Healy and L. Mathevet, "Designing stable mechanisms for economic environments," *Theoretical Economics*, vol. 7, no. 3, pp. 609–661, 2012.

[126] T. Börgers and D. Krahmer, *An introduction to the theory of mechanism design*. Oxford University Press, USA, 2015.

[127] M. M. Khalili, X. Zhang, and M. Liu, "Incentivizing effort in interdependent security games using resource pooling," in *Proceedings of the 14th Workshop on the Economics of Networks, Systems and Computation*, 2019, pp. 1–6.

[128] N. Heydaribeni and A. Anastasopoulos, “Distributed mechanism design for network resource allocation problems,” *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 2, pp. 621–636, 2020.

[129] A. Sinha and A. Anastasopoulos, “Distributed mechanism design with learning guarantees for private and public goods problems,” *IEEE Trans. Automatic Control*, vol. 65, no. 10, pp. 4106–4121, 2020.

[130] M. Rasouli and D. Teneketzis, “Electricity pooling markets with elastic demand: A mechanism design approach,” in *2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2014, pp. 70–75.

[131] C. Chen, R. A. Berry, M. L. Honig, and V. G. Subramanian, “Pricing, bandwidth allocation, and service competition in heterogeneous wireless networks,” *IEEE/ACM Transactions on Networking*, vol. 28, no. 5, pp. 2299–2308, 2020.

[132] M. Zhang, A. Arafa, E. Wei, and R. A. Berry, “Optimal and quantized mechanism design for fresh data acquisition,” *arXiv preprint arXiv:2006.15751*, 2020.

[133] A. Moradipari and M. Alizadeh, “Pricing and routing mechanisms for differentiated services in an electric vehicle public charging station network,” *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1489–1499, 2020.

[134] D. Bergemann and S. Morris, “Information design: A unified perspective,” *Journal of Economic Literature*, vol. 57, no. 1, pp. 44–95, 2019.

[135] F. Farhadi and D. Teneketzis, “Dynamic information design: A simple problem on optimal sequential information disclosure,” *Available at SSRN 3554960*, 2020.

[136] M. O. Sayin and T. Basar, “Bayesian persuasion with state-dependent quadratic cost measures,” *IEEE Transactions on Automatic Control*, 2021.

[137] J. Anunrojwong, K. Iyer, and V. Manshadi, “Information design for congested social services: Optimal need-based persuasion,” *arXiv preprint arXiv:2005.07253*, 2020.

[138] D. Bergemann and S. Morris, “Bayes correlated equilibrium and the comparison of information structures in games,” *Theoretical Economics*, vol. 11, no. 2, pp. 487–522, 2016.

[139] A. Kolotilin, T. Mylovanov, A. Zapechelnyuk, and M. Li, “Persuasion of a privately informed receiver,” *Econometrica*, vol. 85, no. 6, pp. 1949–1964, 2017.

[140] C. Daskalakis, C. Papadimitriou, and C. Tzamos, “Does information revelation improve revenue?” in *Proceedings of the 2016 ACM Conference on Economics and Computation*, 2016, pp. 233–250.

[141] J. Hörner and A. Skrzypacz, “Selling information,” *Journal of Political Economy*, vol. 124, no. 6, pp. 1515–1562, 2016.

[142] C. Li, “A model of bayesian persuasion with transfers,” *Economics Letters*, vol. 161, pp. 93–95, 2017.

[143] T. Yamashita, “Optimal public information disclosure by mechanism designer,” 2018.

[144] Y. Cai, F. Echenique, H. Fu, K. Ligett, A. Wierman, and J. Ziani, “Third-party data providers ruin simple mechanisms,” *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 4, no. 1, pp. 1–31, 2020.

[145] D. Lingenbrink and K. Iyer, “Optimal signaling mechanisms in unobservable queues,” *Operations research*, vol. 67, no. 5, pp. 1397–1416, 2019.

[146] D. Bergemann and S. Morris, “Information design, bayesian persuasion, and bayes correlated equilibrium,” *American Economic Review*, vol. 106, no. 5, pp. 586–91, 2016.

[147] S. Das, E. Kamenica, and R. Mirka, “Reducing congestion through information design,” in *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2017, pp. 1279–1284.

[148] H. Tavafoghi and D. Teneketzis, “Informational incentives for congestion games,” in *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2017, pp. 1285–1292.

[149] Y. Zhu and K. Savla, “On the stability of optimal bayesian persuasion strategy under a mistrust dynamics in routing games,” in *2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2018, pp. 92–99.

[150] J. C. Ely, “Beeps,” *American Economic Review*, vol. 107, no. 1, pp. 31–53, 2017.

[151] L. Doval and J. C. Ely, “Sequential information design,” *Econometrica*, vol. 88, no. 6, pp. 2575–2608, 2020.

[152] M. J. Smith, "The existence of a time-dependent equilibrium distribution of arrivals at a single bottleneck," *Transportation science*, vol. 18, no. 4, pp. 385–394, 1984.

[153] C. F. Daganzo, "The uniqueness of a time-dependent equilibrium distribution of arrivals at a single bottleneck," *Transportation science*, vol. 19, no. 1, pp. 29–37, 1985.

[154] Y. Zhu and K. Savla, "A semidefinite approach to information design in non-atomic routing games," *arXiv preprint arXiv:2005.03000*, 2020.

[155] R. B. Myerson, "Optimal auction design," *Mathematics of operations research*, vol. 6, no. 1, pp. 58–73, 1981.