On Design and Analysis of Cyber-Physical Systems with Strategic Agents

by

Hamidreza Tavafoghi Jahormi

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering: Systems)
in the University of Michigan
2017

Doctoral Committee:

Professor Demosthenis Teneketzis, Chair
Associate Professor Suarabh Amin, Massachusetts Institute of Technology
Professor Tamer Başar, University of Illinois at Urbana-Champaign
Professor Mingyan Liu
Associate Professor David Miller
Professor Asuman Ozdaglar, Massachusetts Institute of Technology

یک چند به کودکی باستاد شدیم یک چند به استادی خود شاد شدیم
پایان سخن شنو که ما را چه رسید از خاک در آمدیم و بر باد شدیم

خیام

For a while, to a master we have gone
For a while, in our mastery we have flown
In the end, listen to what became of us
From dust we came, with wind have gone

Khayyam, Persian poet

Hamidreza Tavafoghi Jahromi
tavaf@umich.edu
ORCID iD: 0000-0002-9583-9727

To Maman and Baba

Acknowledgements

I wish to express my sincere gratitude to people without whom none of this would be possible. I am especially thankful to my advisor, Professor Demosthenis Teneketzis for his continued guidance and unconditional support. You have been a patient mentor, a knowledgeable colleague, and a great friend. Thank you and Barbara for many evenings we spent together.

I am also deeply grateful to all my committee members, Professor Mingyan Liu, Professor David Miller, Professor Saurabh Amin, Professor Tamer Başar, and Professor Asuman Ozdaglar. Their insightful comments and constructive suggestions have been essential for me to complete and improve this work.

I would like to extend my gratitude to Professor Tilman Börgers, a fantastic teacher, for everything I have learnt from him in his terrific classes. I am also thankful to Ali Kakhbod for being a great mentor and an honest friend since day one of my doctoral studies. I would also like to thank the members of our research group: Farzaneh, Erik, Mohammad, and Ouyang. I am grateful to them for their valuable suggestions and many fruitful discussions.

The best part of my Ph.D. life was to be surrounded with many amazing friends who were my family in Ann Arbor. I cannot imagine how my life would have been without your friendship and all the wonderful moments and memories we shared together. My special thanks go to Parinaz Naghizadeh for being an incredible colleague and the best friend.

Above all, I want to express my most sincere gratitude to my family: Maman, Baba, Erfaneh, Alireza, and Laleh. The hardest part of my Ph.D. was not being able to see you over the last six years. Nevertheless, I have always felt your unconditional

support and endless love despite being so far away. I am forever indebted to you for everything you gave me and for putting my happiness before yours. I miss you and I love you.

Table of Contents

List of Figures

Figure

List of Tables

List of Appendices

Appendix

Abstract

In contrast to traditional CPS where a designer can specify an action plan for each agent, in CPS with strategic agents, every agent acts selfishly and chooses his strategy privately so as to maximize his own objective. In this dissertation, we study problems arising in the design and analysis of CPSs with strategic agents.

We consider two classes of design problems. In the first class, the designer utilizes her control over decisions and resources in the system to incentivize the agents via monetary incentive mechanisms to reveal their private information that is crucial for the efficient operation of the system. In particular, we consider market mechanism design for the integration of renewable energy and flexible loads into power grids. We consider a model that captures the dynamic and intermittent nature of these resources, and demonstrate the advantage of dynamic market mechanism over static market mechanisms that underly the existing architecture of the electricity markets. In the second class of design problems, the designer utilizes her informational advantage over the agents and employ informational incentive mechanisms to disclose selectively information to the agents so as to influence the agents' decisions. Specifically, we consider the design of public and private information disclosure mechanisms in a transportation system so as to improve the overall congestion.

We also study the analysis of CPS with strategic agents as a stochastic dynamic game of asymmetric information. We present a set of conditions sufficient to characterize an information state for each agent that effectively compresses his private and common information over time. This information state provides a sufficient statistic for decision-making purposes in strategic and non-strategic settings. Accordingly, we provide a sequential decomposition of the dynamic game over time, and formulate a dynamic program that enables us to determine a set of equilibria of the game. The proposed approach generalizes and unifies the existing results for

dynamic teams with non-classical information structure and dynamic games with asymmetric information.

Chapter 1

Introduction

1.1 Motivation

In recent years, societal and environmental changes along with advances in communication and information technologies have led to the emergence of new dynamic multi-agent systems in which a group of autonomous selfish decision makers (DMs) interact with one another, as well as their surrounding environment, over time. For instance, the traditional government-regulated electricity markets have been transformed to competitive electricity markets where revenue-maximizing generators compete with one another to sell electricity to loadserving entities over the power network [57, 58]. Another example is the advent of navigation applications (*e.g.* Google map, Waze, *etc.*), which have been developed following the commercialization of GPS technology and smart phones, and provide traffic information and routing suggestions to drivers in transportation networks [10].

The above-emerging systems can be modeled as cyber-physical systems (CPSs) with strategic DMs. By cyber-physical system, we refer to the integration of physical and cyber components of a dynamic system, each operating in different spatial and temporal scales. In a traditional CPS where decisions and/or information are decentralized only due to the limitation/cost in communications and processing, a designer has control over the operation of every local component of the system, and thus, can design (in principle) its components using techniques from control and optimization theories. However, in designing a CPS with strategic DMs, a designer cannot dictate the behavior of DMs. In a CPS with strategic DMs, each DM acts autonomously considering his[1] knowledge about the overall operation of the CPS as well as other DMs' behavior over time. Each DM makes private (imperfect) observations about the current state of the CPS as well as other DMs' behavior over time. Combining this information along with his anticipation about the other DMs' behavior, each agent makes decisions in *real time* trying to maximize *his own objective*. The agents' decisions, in turn, affect the evolution of the CPS over time, and thus, determine its overall performance. Therefore, it is important to understand how strategic DMs interact with a CPS and develop analysis and design approaches to CPSs that incorporate the DMs' selfish behavior in real time.

The main focus of this dissertation is to work towards developing analysis and design approaches to CPSs with strategic DMs. We study specific problems that are motivated by theoretical challenges in the study of CPSs with strategic agents as well as particular applications where these systems are prevailing. We discuss below the general framework that underlies the problems we study in this dissertation.

1.2 Research Framework

Throughout this dissertation, we assume that DMs who interact with a CPS are *rational Bayesian agents* [50]. We call a DM a Bayesian agent if he forms his belief about the current state of the CPS by using Bayesian inference. We call a DM a rational agent if whenever he makes a decision he is not limited by the complexity of the decision problem he faces, a cognitive limitation (e.g. imperfect recall), or time available to make a decision.

A CPS with strategic agents can be described by a game that has the following two key components:

(i) A *decision tree* $\mathcal{G}$ that determines agents' feasible actions at any time, system dynamics given the agents' actions, and agents' utility along each path of the CPS's

[1] Throughout this dissertation we refer to the designer/principal as "she" and to agent/DM as "he".

evolution. As seen in Figure 1.1, the primitive random variables that appear in the evolution of the CPS can be modeled as *nature*'s actions.

(ii) An *information structure* $\mathcal{S}$ that determines the information that each agent knows about the current state of the CPS, as well as other agents' information, at any time (see Figure 1.2).

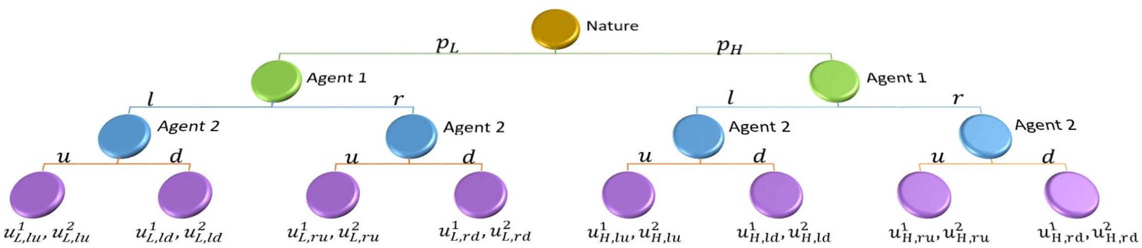


Figure 1.1: An example of decision tree $\mathcal{G}$ with two agents.

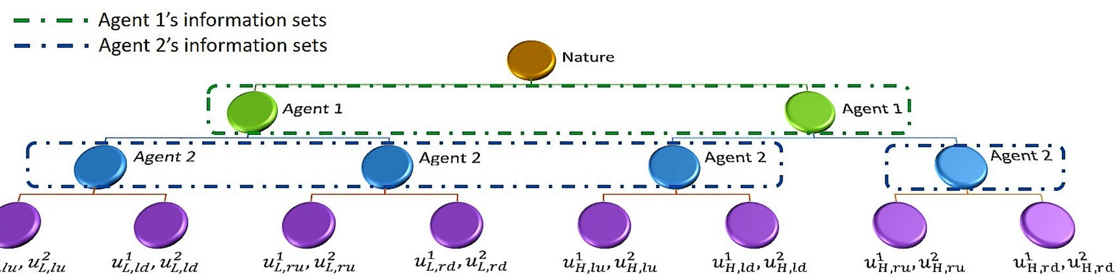


Figure 1.2: An example of information structure $\mathcal{S}$ with two agents.

The above components of a CPS capture the two main elements present in any decision making problem: *decisions* and *information* (see Figure 1.3). It is the interplay between decisions and information that determines the outcome/evolution of a CPS with strategic agents. The interplay between decisions and information is more subtle and essential in a CPS with strategic agents and asymmetric information. In CPS with strategic agents and asymmetric information, agents' decisions influence the information that each agent has about the current state of the system and other agents' information and decisions over time. This is known as *signaling*. Moreover, when agents are strategic with misaligned objectives, each agent has an incentive to make decisions so as to manipulate the information of other agents and influence their behavior, in an attempt to maximize his own objective.

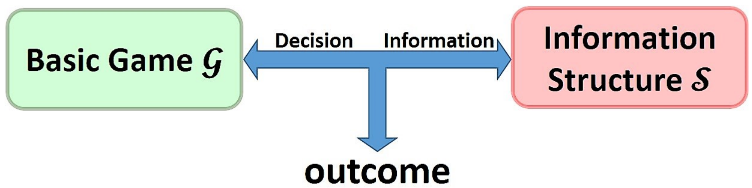


Figure 1.3: The main two components of a CPS with strategic agents

1.3 Problem Formulations and Thesis Outline

In this dissertation, we study three classes of problems concerning the analysis and design of CPS with strategic agents using the framework described above. To design a CPS, a designer can alter its decision tree $\mathcal{G}$ and/or its information structure $\mathcal{S}$ depending on his resources and control over various components of the CPS in the specific application of interest. In the first class of problems, we consider monetary incentive mechanisms where a designer alters the decision tree $\mathcal{G}$ by modifying the agents' payoffs through payments. In the second class of problems, we consider informational incentive mechanisms where a designer alters the information structure $\mathcal{S}$ by disclosing information to the agents. In the third class of problems, we consider the analysis of a CPS when the decision tree $\mathcal{G}$ and information structure $\mathcal{S}$ are fixed, and we want to evaluate the overall outcome of the strategic agents' interaction with the CPS.

1.3.1 Monetary Incentive Mechanisms

There are many instances of dynamic CPSs where strategic agents possess private information that is crucial for the efficient operation of the system. In these instances, the designer (principal) has authority/control over the decisions and resources but does not know all information that is necessary to evaluate the performance of every decision available to her over time. For instance, in a power grid, the independent system operator needs to know the private production cost of all generators so as to solve the optimal power flow problem and determine the energy generation for each of them. In these instances, the designer (principal) can utilize his authority over the decisions and resources in the CPS and alter the decision tree $\mathcal{G}$ so as to induce the agents to reveal (directly or indirectly) the information she needs to know for the operation of CPS.

We note that when agents are non-strategic, a system designer faces a problem that is equivalent to a resource allocation problem. Therefore, the mechanism design problem described above can be interpreted as the strategic analogue of resource allocation problems when the agents are strategic.

In a strategic setting where the agents' objectives are different from that of the designer, they do not necessarily follow the strategy that is prescribed by her and reveal their private information. Therefore, the designer needs to provide incentives so that the agents are willing to follow her proposed strategy prescriptions.

Providing monetary incentive payments is a common instrument the designer can utilize to incentivize the agents to reveal their information truthfully and follow her proposed strategies. However, not all strategy prescriptions can be incentivized through monetary payments. Moreover, when determining the monetary payments, the designer may further need to consider a tradeoff between the desired outcome and the resources required to achieve it. Therefore, to determine an optimal strategy prescription for the agents the designer needs to take into account the feasibility of the prescribed strategies and the corresponding incentive payments.

In the first two chapters of this dissertation, we study the application of monetary incentive mechanisms in the design of electricity markets for the integration of renewable energy and flexible load into power grids.

1.3.1.1 Market Mechanisms for Renewable Energy and Flexible Loads

The current electricity market architecture is mainly designed for conventional generators with slowly varying cost and static information structure, and assume that the major uncertainty in balancing the demand and supply in a power grid is due to the mismatch between the load forecast and load realization in real time. In contrast to the production from conventional generators, the energy production from renewable resources and the availability of flexible loads participating in demand response programs depend on variables that are intermittent and become available dynamically over time. For example, an accurate prediction of wind energy generation is only feasible within a few minutes of the generation time [56]. Moreover, the flexibility of thermal loads participating in demand response programs depends on variables such as temperature that can only be predicted within a few hours [82].

The current practices for the integration of intermittent renewable energy and flexible loads into the electricity markets are to incorporate them into the existing two-settlement market architecture (a forward market followed by a real-time market) along with various subsidies and credits that encourage the investment in renewable energy generation and flexible loads. However, as the share of renewable energy generation and flexible loads increases and the supportive program and subsidies phase out the current practices cannot be sustained [56, 30, 126], and we need to investigate market mechanisms that are appropriate for the integration of these new resources and revisit the electricity market structure. In Chapter 2 and Chapter 3 of this dissertation, we study two mechanism design problems for electricity markets that aim to provide insight into an appropriate market architecture for the integration of intermittent renewable energy and flexible loads to the power grid.

We consider a dynamic model that accounts for the dynamic and uncertain nature of energy production from renewable resources and the availability of flexible loads. In this model, a renewable generator/flexible load receives information about its generation capacity/flexibility level over time. Using this dynamic setting, we study the problem of dynamic market mechanism design for the integration of renewable generation and flexible loads into the power grid. We take a principal-agent view point and adopt a Bayesian mechanism design framework [25] where the independent system operator (principal) designs the market rules in advance and announces them to the generators; the generators then voluntarily participate in the market assuming that the independent system operator is committed to implementing the market rules she has announced.

In the current electricity market architecture, pooling markets along with forward bilateral trades between generators and utility companies are the main two forms of market mechanisms that are used to determine energy trades over the power network. Therefore, using the general dynamic environment described above, we study the problem of mechanism design within the context of forward contracts and pooling markets for the integration of renewable energy and flexible loads in Chapter 2 and Chapter 3, respectively.

1.3.1.2 Forward Contracts for Uncertain Electricity Resources

In electricity markets with conventional generators, it has been shown that having forward contracts, alongside the pooling-based electricity markets, lowers the market price, hedges pricing risks, and increases the reliability of the market operation [3, 24]. Motivated by these results, we study the problem of optimal forward contract design for energy procurement (resp. direct control of flexible loads [82]) from a diversified electricity producer with renewable generation (resp. set of flexible loads) in Chapter 2. Due to the seller's multiple energy sources (resp. demand constraints), the seller has multidimensional private information. Moreover, the seller has incomplete information about his renewable generation (resp. energy consumption) at $T = 1$; this information becomes complete at $T = 2$ with the realization of wind (resp. temperature).

Assuming the wind (resp. temperature) realizations can be monitored, we characterize the optimal forward contract in which the buyer can accept random energy delivery (resp. partial direct control of the load) from the diversified energy producer (resp. flexible loads). We assume that the buyer needs to make a decision over the amount of flexible and inflexible loads (resp. amount of renewable and conventional generations) she schedules based on her agreement with the renewable generator (resp. flexible load). We show that an optimal procurement mechanism is a contingent contract. The seller chooses a contingent contract at $T = 1$ based on his available private information at that time. At $T = 2$, the specific allocation and payment are chosen from the selected contract based on the realization of the wind speed (resp. temperature).

We illustrate through examples that the contingent contract described above provides flexibility to a renewable generator in his generation according to the new information he receives over time, and enables the load serving-entities to utilize more effectively the set of potential flexible loads that are available to him.

1.3.1.3 Dynamic Market Mechanisms for Wind Energy

We study the problem of market mechanism design for wind energy in Chapter 3. We first consider a dynamic two-step model with one strategic seller with wind generation and one buyer that captures the essential elements that appear in the design of a pooling market for renewable energy. The seller has private information about his generation capability which he learns dynamically over time. At $T = 1$ the seller has imperfect information about his generation capability. At $T = 2$, the seller learns more accurate information about the realization of wind speed, and thus, has perfect information about his generation capability.

We consider (static) forward and real-time mechanisms that take place at time $T = 1$ and $T = 2$, respectively. The formulation of these mechanisms is motivated by current practices that are in place in Europe and the U.S., respectively [20, 62]. We further propose a dynamic market mechanism that provides a coupling between the real-time and forward markets over time. We show that the proposed dynamic mechanism outperforms the forward and real-time mechanisms; thus, we demonstrate the advantage of adopting the dynamic mechanism over a sequence of static mechanisms for wind energy. On one hand, in contrast to the forward mechanism, the dynamic mechanism incorporates the additional information arriving over time and provides flexibility for intermittent wind generation. On the other hand, in contrast to the real-time mechanism, the dynamic market mechanism provides early information about wind generation which is necessary to maintain the reliability of power grids and scheduling of adequate reserves. Moreover, by requiring the seller to reveal his private information sequentially over time instead of simultaneously when he has all his information, the dynamic mechanism restricts the seller's market manipulation power. We show that the main advantage of the dynamic mechanism over the real-time mechanism is due to the fact that in the dynamic mechanism it is possible to (i) price discriminate different types of generators based on the uncertainty level in their generation, and (ii) expose generators to the risk of penalty charges.

We further consider two variants of the dynamic mechanism. First, we investigate the dynamic mechanism that guarantees no penalty risk for the seller. We characterize the additional incentive payments the designer needs to provide so as to guarantee no penalty risk for the seller. We show that the performance of the dynamic mechanism with no penalty risk is in general inferior to the dynamic mechanism with penalty risk. Second, we study the dynamic mechanism with wind monitoring. We show that when the wind condition is monitored, the outcome of the dynamic mechanism improves. This happens because the required incentive payments the designer needs to provide decrease as the seller cannot manipulate the outcome of the mechanism by misrepresenting the wind condition. However, we show that the benefit of wind monitoring vanishes as the number of possible technologies for wind generation increases.

We discuss how our results generalize to settings with many sellers using a handicap auction mechanism [38]. In a handicap auction, sellers bid for a set of quantity-payment options in the forward market. Next, in the real time, sellers receive more information and bid for modifications in their generation. The allocations are determined based on the seller's bids in the real-time market and the quantity-payment options they choose in the forward market. Moreover, the payment that each seller receives depends on the outcome of the real-time market as well as the quantity-payment option he chooses at the forward market.

1.3.2 Informational Incentive Mechanisms

In contrast to the situation described in Section 1.3.1, there are instances of CPS where the designer (principal) does not have perfect control over the decisions made in the CPS but has *superior information* about the current state of the CPS that is of value to the agents who make the decisions. In many of these systems, the implementation of monetary incentive mechanisms in not feasible or desirable. For instance, in a transportation network where every driver makes his routing decision individually, the effect of toll payments on traffic flow is limited, and it is not possible to track the routing decision of every driver and charge them accordingly. In these instances, the designer can utilize her superior information and provide informational incentives to the agents by a selective disclosure of information that alters the information structure $\mathcal{S}$ so as to influence the agents' decisions, and maximize her own objective.

We note that when the agents are non-strategic, the designer's problem is equivalent to a real-time source coding problem with a noiseless channel [77, 123, 131, 136]. Therefore, the design of informational incentive mechanisms can be viewed as the analogue of the real-time source coding problem when agents are strategic.

In a strategic setting, each agent utilizes the information he receives from the designer to his own advantage and does necessarily follow the actions the designer suggests. Therefore, the designer needs to reveal her information strategically so that the agents' best response to the information they receive maximizes her objective.

There are two general approaches to the study of information design problems. In the first approach, the designer announces a recommendation policy in advance, and for every realization of her information about the current state of the system, she recommends a specific action to every agent according to the recommendation policy announced before [16, 15]. Therefore, the principal must choose an action recommendation policy so that it is a best response for each agent to follow/obey the action recommendation. Optimizing over the set of recommendation policies which the agents obey, the designer chooses the one that maximizes her objective.

In the second approach, the designer directly works with the agents' beliefs about the current state of the CPS and attempts to modify them by disclosing selective information conditioned on the realization of her information [59, 35]. Assuming that it is possible to determine the agents' best response for every realization of their beliefs, the principal then chooses from all possible modifications of the agents' beliefs so as to maximize her objective.

In this proposal, using the two approaches discussed above, we study the application of informational incentive mechanisms in the design of advanced traveler information systems in transportation networks (*e.g.* variable (dynamic) message signs on roadsides and/or routing recommendation on GPS-enabled devices).

1.3.2.1 Informational Incentives in Congestion Games

In recent years, the development of navigation applications (*e.g.* Waze, Google map) and increasing utilization of variable message signs on roadsides have enabled the drivers to have better information about the congestion level and the condition of every link in a transportation network, and thus, make informed routing decisions. The provision of real-time data about road conditions to drivers creates new opportunities to alleviate congestion in transportation networks.

Several studies have examined the effect of information provision to drivers on the social welfare, and identified instances where provision of information to drivers can be harmful and reduce the social welfare as well as the drivers' utility [78, 4, 14, 33, 73, 69, 124, 1, 74]. This is because in a transportation network every driver makes his routing decision individually trying to maximize his own utility and does not consider the negative externality that he creates by increasing the traffic along the route that he takes.

In Chapter 4, we study the problem of designing an information disclosure mechanism in a transportation network. In contrast to the works of [78, 4, 14, 33, 73, 69, 124, 1, 74] that analyze the effect of specific information disclosure policies, we investigate the problem of optimal design of information disclosure mechanisms in transportation networks. We consider a congestion game over a parallel two-link network, where drivers choose their route/link individually trying to minimize their travel time. The travel time through every route/link in the network depends on the route's condition as well as the number of cars traveling through it. We assume that the condition of one route (safe route) is known to all drivers while the condition of the other route (risky route) is random and only known to the designer (principal). The principal wants to design an information disclosure mechanism so as to minimize the overall traveling time (social welfare). We consider two cases: (i) when the principal can only disclose information publicly to all drivers (*e.g.* variable message signs), and (ii) when the principal can disclose information privately to each driver (*e.g.* navigation applications). We investigate these two cases using the two general approaches to information design described above in Section 1.3.2.

We show that when the principal employs a public information disclosure mechanism her optimal mechanism depends on the second derivative of the social welfare function with respect to the risky route's condition. If the social welfare is a convex (resp. concave) function of the risky route's condition then it is socially optimal to disclose no information (resp. perfect information) about the risky routes' condition to the drivers. However, if the social welfare is neither convex nor concave, there may exist a probability distribution over the possible risky route's condition such that an optimal mechanism is a partial information disclosure mechanism.

When the principal can employ private information disclosure mechanisms, we show that the principal can improve the social welfare by coordinating the individual recommendation she makes to the drivers based on the realization of the risky route's condition. When the uncertainty about the risky routes' condition is high relative to the *ex-ante* difference in the routes' conditions (*i.e.* the value of information is high), the principal can achieve the socially efficient routing outcome using an optimal private information disclosure mechanism. When the uncertainty about the risky routes' condition is low relative to the *ex-ante* difference in the routes' conditions (*i.e.* the value of information is low), the principal does not have enough power to persuade the drivers to change their routing decision so as to achieve the socially efficient routing outcome; nevertheless, by disclosing private information to each driver, the principal can improve the social welfare compared to the one under the "no information disclosure" mechanism.

In Section 4.6, we investigate the problem of dynamic information disclosure mechanism design in a dynamic setting with time horizon $T = 2$, where the risky route's condition has uncontrolled Markovian dynamics and the drivers learn from their experience at $t = 1$. We consider the following three scenarios for what drivers learn at $t = 1$: (i) the drivers only learn from the information they receive from the principal and do not make any additional observation about the risky route's condition and/or the number of cars on each route, (ii) in addition to the information they receive from the principal at $t = 1$, the drivers who take the risky route at $t = 1$ learn its condition perfectly, and (iii) in addition to the information they receive from the principal at $t = 1$, the drivers observe the number of cars on the route they take at $t = 1$. Due to privacy constraints and practicality issues, we assume that the information the principal discloses to every driver at $t = 2$ does not depend on his actual routing decision at $t = 1$. Using numerical simulations, we conjecture that in scenario (i) the principal can achieve the same performance per time step in a dynamic setting as in the static setting even though the drivers learn from the information they receive at $t = 1$. However, in scenarios (ii) and (iii), where the drivers make additional observation than in scenario (i), the performance of an optimal dynamic information mechanism per time step decreases as the correlation between the risky route's conditions at $t = 1$ and $t = 2$ increases. In particular, for scenario (ii) we identify instances where the principal's optimal information mechanism is to commit to revealing the risky route's condition at $t = 2$ so that the drivers do not have an incentive to experiment by taking the risky route at $t = 1$. Moreover, for scenario (iii), we identified instances where the principal's optimal information mechanism is to not utilize all her information about the risky route's condition at $t = 1$ so as to have a higher information superiority at $t = 2$.

1.3.3 Dynamic Games with Asymmetric Information

Many CPSs can be modeled as a dynamical system with controlled Markovian dynamics. Therefore, given a fixed decision tree $\mathcal{G}$ and information structure $\mathcal{S}$, we can analyze CPSs with strategic agents as stochastic dynamic games with asymmetric information and Markovian dynamics. In dynamic games with asymmetric information, agents have different observations of the game evolution (*i.e.* current state of the CPS), and thus, different information histories. Every agent plays a strategy which is a function of his information history. In order to anticipate other agents' strategies over time, an agent needs to form a belief about other agents' information histories so as to predict other agents' decisions. Therefore, to characterize the outcome of a dynamic game with asymmetric information we need to define an *assessment* that consists of a set of strategies for all agents as well as a set of beliefs for all agents at every information state. In our work, we adopt Perfect Bayesian Equilibrium (PBE) as the solution concept to study the outcome of dynamic games with asymmetric information. A PBE is an assessment that satisfies the *sequential rationality* and *consistency* conditions [43]. Sequential rationality requires that each agent's strategy is optimal at each of his information sets given his belief and other agents' strategies. Consistency requires that each agent's belief at each of his information sets complies with Bayes' rule given all agents' strategies.

While the definition of PBE provides a formalization of the connections between the agents' strategies and information/beliefs in dynamic games with asymmetric information, it does not provide a tractable methodology for their analysis. In a dynamic game with asymmetric information, an agent's belief about the current state of the game at any time t depends on the strategy of all other agents up to time t; this dependency is captured by the consistency condition, and it is known as *signaling*. Moreover, at any time t, an agent chooses his strategy from time t onward according to his belief about the current state of the game at time t; this dependency is captured by the sequential rationality condition. Therefore, there exists a circular dependency between all agents' strategies and beliefs over time. As a result, one needs to determine the agents' strategies and beliefs simultaneously for the whole time horizon so as to satisfy the consistency and sequential rationality conditions. Furthermore, as an agent gathers more information over time, the domains of his strategies grow. Consequently, the determination of an agent's strategy at any time has a complexity that grows exponentially over time. As a result, the existing literature has only studied special instances of dynamic games with asymmetric information (see [79, 42, 138, 44, 89, 46, 98, 97, 127, 108]).

In Chapter 5 of this dissertation, we aim to develop a general approach to the study of dynamic games with asymmetric information and propose a tractable approach to determine a specific set of PBE of these games.

1.3.3.1 A Common Information Approach to Dynamic Games with Asymmetric information

Dynamic games with asymmetric information can be considered as the analogue of decentralized stochastic control problems [53, 100] when agents are strategic and have different objectives. Alternatively, dynamic games with asymmetric information can be considered as a generalization of dynamic games with symmetric information [43, 41, 9] when agents possess private information in addition to the common information they share. The authors of [91, 90] and [81] propose a tractable methodology to study decentralized stochastic control problems and dynamic games with symmetric information, respectively. In Chapter 5, we provide an analogue of the results of [91, 90, 81] for dynamic games with asymmetric information.

The authors of [81] consider Markovian dynamic games with symmetric information. They propose the common information based belief about the system state as the information state for all agents, and characterize a class of *subgame perfect equilibria (SPE)* for dynamic games with symmetric information called *Markov Perfect Equilibrium (MPE)*. In an MPE, agents play strategies that are only functions of the common information based belief (information state). Using the notion of MPE, they propose a sequential decomposition of Markovian dynamic games with symmetric information, and formulate a dynamic program that can be used to determine the set of MPE of the game. The approach proposed in [81] does not apply to dynamic games with asymmetric information simply because it does not consider the agents' private information and beliefs about other agents' private information over time.

The authors of [91, 90] study dynamic decentralized stochastic control problems with non-classical information structure. Using the common information among the agents, they present a centralized stochastic control problem that is equivalent to the original problem as follows. For every local controller, they consider a *fictitious* controller who has access to the local controller's common information but not his private information. Every fictitious controller has to determine a sequence of prescriptions that determine the corresponding local controller's action at every time for every possible realization of his private information. The problem of determining the optimal strategies for the fictitious controllers is a centralized stochastic control problem since they have symmetric information. Using standard results in centralized stochastic control literature [67], they characterize a sufficient statistic/information state for every agent, sequentially decompose the problem over time, and formulate a dynamic program that can be used to determine the optimal strategy of the original problem.

The works of [89, 46] have utilized the common information approach proposed in [91, 90] to study a class of dynamic games with asymmetric informations where there is no signaling among agents, *i.e.* agents' beliefs are strategy independent. In such situations, the circular coupling between strategies and beliefs, discussed above, does not exist; this feature is a significant simplification assumed in the solution methodology for the dynamic games considered in [89, 46]. Such a methodology does not work for games where there is signaling among agents (*i.e.* agents' beliefs are strategy dependent).

In this chapter, we propose a general approach for the study of dynamic games with asymmetric information when signaling occurs. We present a set of conditions sufficient to characterize information states where the agents' common and private information are effectively compressed in a mutually consistent manner. We identify instances of dynamic games with asymmetric information where we can characterize an information state for every agent that has a time-invariant domain.

When the agents are non-strategic, we show that for every arbitrary but fixed agents' strategies, there exists an equivalent set of strategies that utilize only the above-mentioned information state and results in the same expected flow of utility over time (Theorem 5.4). This result generalizes the results of [90, 91] for decentralized stochastic control problems in two aspects. First, the set of conditions sufficient to characterize an information state presented in this dissertation are more general than those of [90, 91] and includes them as special cases. In contrast to the approach presented in [90, 91] that requires the agents to use all of their private information (or use perfectly a predetermined stored memory of it), in our approach the agents' private information can be effectively compressed according to the sufficient conditions presented in Section 5.6.

		Strategic DMs	Non-strategic DMS
Design	Control over resources/decisions	Mechanism design (Chapters 2 and 3)	Resources Allocation
	Possess superior information	Information design (Chapter 4)	Real-time communication system
Analysis		Dynamic games with asymmetric information (Chapter 5)	Decentralized stochastic control (Chapter 5)

Table 1.1: An overall view of the dissertation

Based on the information state characterized in Section 5.6, we introduce the notion of Common Information Based Perfect Bayesian Equilibrium (CIB-PBE) that characterizes a set of outcomes for dynamic games. Using the notion of CIB-PBE, we provide a sequential decomposition of the dynamic games over time and formulate a dynamic program that enables us to compute the set of CIB-PBEs via backward induction. The results appearing in Chapter 5 generalize the results of [97, 98, 116] that consider special instances of dynamic games with asymmetric information when signaling occurs.

We discuss the connection between the sets of PBEs and CIB-PBEs in dynamic games with asymmetric information and argue that when the underlying system is highly dynamic and there exists a significant information asymmetry among the agents, the notion of CIB-PBE provides a plausible prediction of the outcome in practice. We provide conditions under which we can guarantee the existence of CIB-PBEs. Using these conditions, we prove the existence of CIB-PBEs for zero-sum dynamic games and special instances of non-zero-sum dynamic games.

The information state characterized in Chapter 5 provides a sufficient statistic for decision making purposes in strategic and non-strategic settings. Therefore, we propose a universal approach to dynamic decision problems in CPSs with strategic and non-strategic agents that can be used to study dynamic games among teams.

1.4 Contributions of the Thesis

An overall view of the dissertation can be seen in Table 1.1. The main contribution of this dissertation can be summarized as follows:

- Market mechanisms for renewable energy and flexible loads (Chapter 2)
    - We propose a two-time step model that captures the dynamic and uncertain nature of energy generation from renewable resources and allows for scheduling of flexible/inflexible loads based on the availability of renewable generation.
    - We investigate the problem of forward contract design for uncertain electricity resources (*e.g.* wind generators, flexible loads) between a buyer and seller with general cost/utility functions and multi-dimensional private information. We show that the optimal contract is a contingent contract that is signed at $t = 1$ and determines the payment and the energy quantity for every realization of the uncertain variable (*e.g.* wind, temperature) at $t = 2$.
    - We present a modified payment function contingent on the realization of the uncertainty that achieves any arbitrary risk sharing between the buyer and seller without changing the energy allocation function. In particular, we show that it is possible to ensure the stronger notion of ex-post individual rationality instead of interim individual rationality at no performance loss.
- Dynamic market mechanisms for wind energy (Chapter 3)
    - We demonstrate the advantage of dynamic market mechanisms over static mechanisms (*e.g.* day-ahead and real-time markets) for a general designer's objective. A dynamic market mechanism couples dynamically the outcomes/payments of different markets over time. Compared to a forward pooling market (*e.g.* day-ahead markets), a dynamic market incorporates the additional information that arrives after the market closes, and

thus, provides flexibility in generation from renewable resources. Compared to a real-time market, a dynamic market gives less manipulation power to a seller to misreport his generation cost, guarantees a certain level of commitment by the seller for energy generation, and reveals in advance the information necessary to schedule the adequate reserve generators/flexible loads.

  - We show that the advantage of the dynamic market mechanism is mainly due to (i) the designer's power to price discriminate sellers with different generation uncertainty (non-uniform pricing) and (ii) the designer's ability to expose sellers to the risk of penalty charges.
  - We characterize the benefit of wind monitoring on the performance of the dynamic market mechanism, and show that the value of wind monitoring decreases as the number of possible generation technologies increases.

- Informational incentives in congestion games (Chapter 4)
  - We investigate the problems of optimal public and private information disclosure mechanism design in a transportation network so as to improve the social welfare. We show that perfect disclosure of information about the routes' conditions is not an optimal mechanism . Therefore, our results propose a solution to the concern raised in [78, 4, 14, 33, 73, 69, 124, 1, 74] about the potential negative impact of information provision on the overall congestion in transportation networks.
  - When the principal can employ a private information disclosure mechanism, we show that she can implement the socially efficient routing outcome by providing coordinated routing recommendation to the drivers if the value (variance) of her information about the routes' condition is high relative to the ex-ante differences in routes' conditions.
  - We investigate the problem of optimal dynamic private information disclosure mechanism design in a dynamic two-time step setting where the routes' conditions have uncontrolled Markovian dynamics. We identify

the following three scenarios in which the drivers learn from their experience at $t = 1$: (i) they learn from the routing recommendations they receive at $t = 1$ (ii) in addition to the routing recommendation they receive, they observe the condition of the route they take at $t = 1$ (iii) in addition to the routing recommendation they receive, they observe the number of cars (traffic) on the route they take at $t = 1$. Using numerical simulation, we conjecture that in scenario (i) the performance of the optimal dynamic private information disclosure mechanism per time-step is the same as that of the optimal static private information disclosure mechanism. However, in scenarios (ii) and (iii) the performance of the optimal dynamic private information disclosure mechanism decreases as the correlation among the routes' conditions at $t = 1$ and $t = 2$ increases.

- A common information approach to dynamic games with asymmetric information (Chapter 5)
  - We present a set of conditions sufficient to characterize an information state where the agents' private and common information are effectively compressed in a mutually consistent manner. We identify instances of dynamic games with asymmetric information where we can characterize an information state with a time-invariant domain.
  - When agents are non-strategic, we show that for any arbitrary agents' strategy profile there exists an equivalent set of strategies for the agents that depend on the above-mentioned information state that results in the same flow of utility for all agents over time. Therefore, we propose a general methodology for the study and analysis of dynamic team problems with asymmetric information and generalize the existing results in [90, 91].
  - We introduce a subclass of PBE of dynamic games with asymmetric information, called CIB-PBE, that utilizes the above-mentioned information state for every agent. Using the notion of CIB-PBE, we provide a sequential decomposition of the dynamic game over time. Accordingly, we

formulate a dynamic program that enables us to compute the set of CIB-PBE of a dynamic game via backward induction.

- We provide conditions under which we can guarantee the existence of CIB-PBEs in a dynamic game. Using these conditions, we prove the existence of CIB-PBE for dynamic zero-sum games and specific instances of dynamic non-zero-sum games.

1.5 Notation

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

Chapter 2

Forward Contracts under Uncertainty for Electricity Markets

2.1 Introduction

2.1.1 Background and Motivation

In recent years, electricity markets have undergone profound structural changes in both the generation and the demand side. The traditionally monopolistic governm-ent-regulated markets reformed toward liberalized electricity markets in order to introduce competition and increase efficiency in generation [134]. Privately-owned generators and utility companies possess private information about their cost/utility, behave strategically, and seek to maximize their profits. Moreover, the developing network of smart grids aims to utilize the available flexibility on the demand side to increase the efficiency of the grid. To involve the demand side actively into the operation of the grid, one needs to design appropriate mechanisms that incentivize the demand to exercise flexibility in its consumption behavior.

Long-term contracts, as an agreement between strategic parties with private information, is one of the main trading mechanisms used in electricity markets. Generators and utility companies sign long-term contracts to hedge themselves against the risk of pooling markets. In fact, it has been suggested that long-term contracts are necessary along with the existing pooling markets to ensure the stability and reliability of electricity markets [24].

Contracts have been considered as one of the main mechanisms to induce a desired behavior on the demand side of smart grids. In comparison to real time pricing or direct market participation, contracts with incentive payments result in a direct control of resources, and thus, give reliability and stability guarantees [82]. Furthermore, contracts with incentive payments are simpler to implement and more appealing to smaller market participants (*e.g.* households) [125].

In this chapter, we study a general contract design problem for electricity markets in a principal-agent (buyer-seller(s)) setup. We assume that both the buyer and the seller sides have multi-dimensional private information and general utility/cost functions. Furthermore, we explicitly consider a general uncertainty in our problem formulation which is becoming a critical issue in the operation of electricity markets.

As the share of intermittent generation from renewable generation increases, the uncertainty in the available generation will increase. Furthermore, the added flexibility on the demand side in smart grids also means a higher uncertainty on demand; such uncertainty should be properly managed through appropriately designed incentives. In general, both the buyer and the seller may have uncertainty, either in their cost/utility functions, or the availability of the resources being traded between them. By explicitly including uncertainty into our problem formulation we capture these facts and can address the problem of commitment (ex-post voluntary participation), risk sharing, and forward contracts with random allocation.

The problems formulated in this chapter enable us to capture and analyze interaction between energy consumers and renewable energy generators, as well as interactions between an aggregator and a network of a demand population participating in the demand response program. We provide examples for each of these scenarios so as to illustrate our results.

2.1.2 Related Literature

There is a growing literature on contract design for electricity with information asymmetry and strategic behavior. A contract design problem for demand management with one-dimensional private information and linear utility has been studied in [39]. The work in [21] addresses the problem of contract design for deferrable demands with constant marginal utility for demand. The work in [28] considers a mechanism design problem for the forward reserved market assuming that the participants have constant marginal cost and no market power. Although the private information in [21], and [28] is multi-dimensional, the simplifying assumption of constant marginal cost/utility enables the authors to rank different types, and is critical to the solution approaches they provide. The specific structures of utility/cost functions assumed in [21], [28], and [39] enable the authors to provide solutions that are inspired by the solution methodology of the one-dimensional screening problem. Contract design problem for demand response with quadratic cost functions is investigated in [49] by numerical methods. The work in [114] considers a mechanism design problem for energy procurement with a general utility/cost function and uncertainty and applies a Vickery-Clacks-Gloves (VCG) based mechanism. However, the VCG mechanism is suboptimal for the problem formulated in [114] when the cost function cannot be parameterized by only a one-dimensional type (see [66], Ch. 14).

From the economics point of view, the problem we formulate in this chapter belongs to the class of screening problems. In economics, the one-dimensional screening problem has been well-studied with both linear and nonlinear utility functions [25]. However, the extension to the multi-dimensional screening problem is not straightforward and no general solution is available. The authors in [75] study a general framework for a deterministic multi-dimensional screening problem with linear utilities. They discuss two general approaches, the parametric-utility approach and the demand-profile approach. The methodology we use to solve the problem formulated in this chapter is similar to the demand-profile approach. We consider a multi-dimensional screening problem under uncertainty with nonlinear utilities. The presence of nonlinearities and uncertainty results in additional complications that are not present in [75] where the utilities are linear and there is uncertainty[1].

[1]When a problem is linear, expectation of any random variable can be replaced by its expected value and reduce the problem to a deterministic one.

2.1.3 Contribution

The contribution of this chapter is two-fold. First, we consider an optimal contract design problem for electricity markets with utility/cost functions that are more general than those considered in the literature ([21],[28],[39],[49]). The nature of utility/cost functions with multi-dimensional private information is such that the solution methodology presented in [21],[28], and [39] does not extend to our problem. The generality of our model enables us to capture many instances of problems arising in electricity markets. Two such instances are discussed in Sections 2.4 and 2.6.

Second, we explicitly incorporate a general uncertainty in the realized cost/utility of the buyer and the seller. The presence of uncertainty along with the nonlinearity of the utilities result in problems where the methodology used in previous works ([21],[28],[39]) cannot be applied, as in these works the utilities are linear and any uncertainty can be replaced by its expected value. The inclusion of uncertainty is crucial in the modeling and analysis of emerging electricity markets because: (1) the share of renewable generation increases; (2) the existing demand becomes less shielded from the market outcome and more elastic; and (3) new resources/loads (e.g. storage, plug-in electric vehicles) enter the market. Due to uncertainty, firm forward contracts (a priori fixed allocation and fixed payment) do not appear to be an appropriate form of contract for emerging electricity markets. Moreover, in the presence of uncertainty, interim voluntary participation (defined in Section 2.3) of the seller does not necessarily imply ex-post voluntary participation of the seller (defined in Section 2.5). Therefore, additional considerations are needed to ensure the commitment of the agents to the contract for every realization of the uncertainty. We show that, in general, the optimal mechanism for the problem formulated in this chapter is a menu of nonlinear pricing schemes. We prove that by allowing the payment to depend on the uncertainty, we can achieve ex-post voluntary participation of the seller, and a desired risk-sharing (associated with the uncertainty) between the buyer and the seller. To the best of our knowledge, our results present the first optimal forward contract under uncertainty for electricity markets where the buyer and the seller have general utility/cost functions parameterized by multi-dimensional private information. We illustrate our results by providing two examples from electricity markets: an optimal demand response contract for ancillary service; and a bilateral trade between a buyer and a renewable energy generator.

2.1.4 Organization

The rest of this chapter is organized as follows. We introduce the model in Section 2.2. In Section 2.3, we formulate and analyze an optimal forward contract with deterministic allocation, and address the problem of risk sharing between the buyer and the seller. We illustrate the result via an example for a contract design problem for demand response program in Section 2.4. In Section 2.5, we formulate and analyze an optimal forward contract with random allocation that depends on the uncertainty, and address the problem of the seller's imperfect commitment (ex-post voluntary participation). We provide an example of a bilateral trade between a buyer and a renewable energy generator in Section 2.6. We discuss our results in Section 2.7. We conclude in Section 4.7. The proofs of the lemmas and corollaries appearing in this chapter can be found in Appendix A.

2.2 Model

A buyer wants to design a mechanism to procure energy/resource from a seller.[2] Let q be the amount of energy/resource the buyer procures, and t be his payment to the seller. The buyer's total profit is given by $\mathcal{V}(q) - t$, where $\mathcal{V}(q)$ is his utility by receiving q amount of energy/resource. The function $\mathcal{V}(\cdot)$ is the buyer's private information and $\mathcal{V}(0) = 0$.

The seller's provision cost is given by $C(q, \mathbf{x}, w)$, convex and increasing in q, $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in \chi \subseteq \mathbf{R}^n$ is the seller's type, and w denotes the realization of a random variable W (uncertainty) with a probability distribution $F_W(w)$ that is common knowledge. We assume that $C(0, \mathbf{x}, w)$ (zero-provision cost) does not depend on the realization of random variable w and is equal to x_1, *i.e.* $C(0, \mathbf{x}, w) = C(0, \mathbf{x}) = x_1$.

[2]From now on, we refer to the buyer as "he" and to the seller as "she".

The seller's utility is given by her total expected revenue $\mathbb{E}_W \{t - C(q, \mathbf{x}, W)\}$. The seller's type $\mathbf{x}$ is her private information, the set χ is common knowledge, and there is a prior probability distribution $F_{\mathbf{x}}$ over χ which is common knowledge between the buyer and the seller.

Let $c(q,\mathbf{x}) := \frac{\partial \mathbb{E}_W\{C(q,\mathbf{x},W)\}}{\partial q}$ denote the expected marginal cost for the seller's type $\mathbf{x}$. We assume that $\exists m$, $1 < m \leq n$, such that $c(q,\mathbf{x})$ is increasing in x_i for $1 \leq i \leq m$, and decreasing in x_i for $m < i \leq n$.[3] Moreover, there exists $\underline{\mathbf{x}} \in \chi$ (the seller's worst type) such that $\underline{x}_i \leq x_i$ and $\underline{x_j} \geq x_j$ for all $\mathbf{x} \in \chi$, $1 \leq i \leq m$ and $m < j \leq n$.

Definition 2.1. *We say the seller's type $\mathbf{x}$ is better (resp. worse) than the seller's type $\hat{\mathbf{x}}$ if $c(q,\mathbf{x}) \leq c(q,\hat{\mathbf{x}})$ for all $q \geq 0$ (resp. $c(q,\mathbf{x}) \geq c(q,\hat{\mathbf{x}})$) with strict inequality for some q.*

Therefore, the seller's type $\mathbf{x}$ is better than the seller's type $\hat{\mathbf{x}}$ if and only if $x_i \leq \hat{x}_i$ for $1 \leq i \leq m$, and $x_i \geq \hat{x}_i$ for $m < i \leq n$ with strict inequality for some i. The following example illustrates such ordering.

Example 2.1. *Consider an energy seller with a wind turbine and a gas generator. The generation from the wind turbine is free and given by γw^3, where γ is the turbine's technology and w is the realized weather. The gas generator has a fixed marginal cost θ_c. There is a fixed cost c_0 which includes the start-up cost for both plants and the capital cost for the seller. Therefore, the seller's type has $n = 3$ dimensions. The generation cost for the seller is given by*

$$C(q, w, \mathbf{x}) = c_0 + \theta_c \max\left\{q - \gamma w^3, 0\right\}. \tag{2.1}$$

The seller's type $\mathbf{x} = (c_0, \theta_c, \gamma)$ is better than the seller's type $\hat{\mathbf{x}} = (\hat{c}_0, \hat{\theta}_c, \hat{\gamma})$ if and only if $c_0 \leq \hat{c}_0$, $\theta_c \leq \hat{\theta}_c$, and $\gamma \geq \hat{\gamma}$, with one of the above inequalities being strict.

Note that in the one-dimensional screening problem, the cost of production induces a complete order among the seller's types, which is crucial to the solution of

[3] Note that for a general cost function $C(q, \mathbf{x}, W)$ if the corresponding $c(q, \mathbf{x})$ changes sign for only finite number of times, one can expand the type space χ and reorder its dimensions so that it satisfies the assumption on the existence of m.

the optimal mechanism design problem. However, in multi-dimensional screening problems, the expected cost of production induces, in general, only a partial order among the seller's types.

We assume that the buyer has all the bargaining power; thus, he can design the mechanism/set of rules that determines the agreement for the procurement quantity q, and the payment t. After the buyer announces the mechanism for procurement and the seller accepts it, both the buyer and the seller are fully committed to following the rules of the mechanism.

As a consequence of the assumption on the buyer's bargaining power and the fact that the seller's utility does not directly depend on the buyer's private information (private value), the solution of the problem formulated in this chapter does not depend on whether the buyer's utility $\mathcal{V}(\cdot)$ is private information or common knowledge[4].

In the rest of this chapter we formulate two contract design problems. In Section 2.3, we assume that the buyer can only accept an a priori fixed energy delivery and formulate a forward contract design problem with deterministic allocation. In Section 2.5, we assume that the buyer can tolerate intermittency in the delivered energy by utilizing his existing storage/reserve resources, and formulate a forward contract design with random allocation.

2.3 Forward Contracts with Deterministic Allocation

In this section we consider a problem of forward contract design where the allocation q is deterministic and is decided in advance at the time of contract signing. Bilateral trades with conventional generators and demand response (DR) contracts for direct load control are forms of such a contract.

[4] This becomes more clear by looking at the result of Theorem 1.

2.3.1 Problem Formulation

Let $(\mathcal{M}, h)$ be the mechanism/game form (see [80], Ch. 23) for energy procurement designed by the buyer. In this game form, $\mathcal{M}$ describes the message/strategy space for the buyer and the seller, respectively, and h determines the outcome function; $h : \mathcal{M} \rightarrow \mathbf{R}_+ \times \mathbf{R}$. For every message $m \in \mathcal{M}$ the outcome function h specifies the amount q of the procured energy/resource and the payment t made to the seller, i.e. $h(m) = (q(m), t(m))$.[5]

The objective is to determine a mechanism $(\mathcal{M}, h)$ so as to

$$\underset{(\mathcal{M},(q(\cdot),t(\cdot)))}{\text{maximize}} \quad \mathbb{E}_{\mathbf{x},W} \left\{ \mathcal{V}(q(m^*)) - t(m^*) \right\}, \tag{2.2}$$

where $m^* \in \mathcal{M}$ is a Bayesian Nash equilibrium (BNE) of the game induced by the mechanism $(\mathcal{M}, h)$. We want the seller to voluntarily participate in the procurement process. The voluntary participation (VP) (or individual rationality) for each type of the seller can be written as

$$\text{interim VP: } \mathbb{E}_W\{t(m^*) - C(q(m^*), \mathbf{x}, W)\} \geq 0, \forall \mathbf{x} \in \chi \tag{2.3}$$

That is, at equilibrium m^* of the induced game the mechanism the seller must have an expected (with respect to the uncertainty W) non-negative payoff. We call the requirement expressed by (2.3) an *interim voluntary participation* constraint.

We call the above problem **(P1)**.

2.3.2 Analysis & Results

We prove that the optimal procurement mechanism is a pricing scheme that the buyer offers to the seller and the seller chooses a quantity according to her type. In such a pricing scheme we have $\mathcal{M} = \chi$, $q : \chi \rightarrow \mathbf{R}_+$, and the payment function $t(\cdot)$ can be defined indirectly as a function of the quantity $q(\mathbf{x})$, *i.e.* $t(q(\mathbf{x}))$. We characterize the optimal procurement mechanism by the following theorem, which

[5]Note that we use q (resp. t) to denote both the quantity value (resp. payment value) and the quantity outcome function (resp. payment function) of mechanism $(\mathcal{M}, h)$.

reduces the original functional maximization problem (P1) to a set of equivalent pointwise maximization problems.

Theorem 2.1. *Under a certain concavity condition, stated in Lemma 2.3 below, the optimal mechanism $(q(\cdot), t(\cdot))$ for the buyer is a nonlinear pricing scheme given by*

$$p(q) = \arg\max_{\hat{p}} \{P[\mathbf{x} \in \chi | \hat{p} \geq c(q, \mathbf{x})] (\mathcal{V}'(q) - \hat{p})\}, \tag{2.4}$$

$$t(q) = \int_0^q p(l)dl + C(0, \underline{\mathbf{x}}), \tag{2.5}$$

$$q(\mathbf{x}) = \arg\max_{l \in \mathbf{R}_+} \{t(l) - \mathbb{E}_W \{C(l, \mathbf{x}, W)\}\} \tag{2.6}$$

where $\mathcal{V}'(q) := \frac{d\mathcal{V}(q)}{dq}$ *and* $\mathcal{M} = \chi$.

The assertion of Theorem 2.1 is established via several steps. Below we present these steps and the key ideas behind each step. The proofs of the lemmas and corollaries appearing in theses steps can be found in the appendix. In the sequel, we omit the argument of the functions $q(\cdot)$ and $t(\cdot)$ whenever such an omission causes no confusion.

Step 1. We set message space $\mathcal{M} = \chi$ and formulate the following problem **(P2)** that is equivalent to problem **(P1):**

$$\underset{(q(\cdot),t(\cdot))}{\text{maximize}} \quad \mathbb{E}_{\mathbf{x},W} \{\mathcal{V}(q(\mathbf{x})) - t(\mathbf{x})\} \tag{2.7}$$

subject to

$$IC: \mathbf{x} = \arg\max_{\mathbf{x}'} \mathbb{E}_W[t(\mathbf{x}') - C(q(\mathbf{x}'), \mathbf{x}, W)], \forall \mathbf{x} \in \chi \tag{2.8}$$

$$interim\ VP: \mathbb{E}_W[t(\mathbf{x}) - C(q(\mathbf{x}), \mathbf{x}, W)] \geq 0, \forall \mathbf{x} \in \chi, \tag{2.9}$$

where $q : \chi \to \mathbf{R}_+$ and $t : \chi \to \mathbf{R}$.

The equivalence follows from the revelation principle [32]. By invoking the revelation principle, without loss of optimality, we restrict attention to direct mechanisms (where $\mathcal{M} = \chi$) that are incentive compatible and individually rational. Incentive compatibility (IC) for a direct mechanism requires that truth-telling must be an optimal strategy for the seller.

Step 2. We show that for any incentive compatible mechanism (q,t) the seller's worst type $\underline{\mathbf{x}}$ gets the minimum utility among all of the seller's types. We utilize the partial order among the seller's different types to rank her utility for her different types (Lemma 2.1), and reduce the VP constraint (2.13) for all the seller's types to the VP constraint only for the seller's worst type (Corollary 2.1).

Lemma 2.1. *For a given incentive compatible mechanism (q,t), a better type of the seller gets a higher utility. That is, let $U(\mathbf{x}) := \mathbb{E}_W \{t(\mathbf{x}) - C(q(\mathbf{x}), \mathbf{x}, W)\}$ denote the expected profit of the seller with type $\mathbf{x}$. Then,*

1. $\frac{\partial U}{\partial x_i} \leq 0, 1 \leq i \leq m,$
2. $\frac{\partial U}{\partial x_i} \geq 0, m < i \leq n.$

A direct consequence of Lemma 2.1, is that the seller's worst type $\underline{\mathbf{x}}$ receives the minimum utility among all the seller's types.

Corollary 2.1. *The voluntary participation constraint is only binding for the worst type $\underline{\mathbf{x}}$. That is, the general VP constraint (2.13) can be reduced to*

$$U(\underline{\mathbf{x}}) := \mathbb{E}_W \{t(\underline{\mathbf{x}}) - C(q(\underline{\mathbf{x}}), \underline{\mathbf{x}}, W)\} \geq 0. \tag{2.10}$$

Step 3. We show, via Lemma 2.2 below, that the optimal mechanism (q,t) is a pricing scheme. That is the payment function $t(\mathbf{x})$ can be defined indirectly as a function of q as $t(q(\mathbf{x}))$.

Lemma 2.2. *For any pair of functions (q,t) that satisfies the IC constraint, we can rewrite $t(\mathbf{x}')$ as $t(q(\mathbf{x}'))$.*

With some abuse of notation we assume that the payment function $t : \mathbf{R} \to \mathbf{R}$ refers to the indirectly defined function $t(q(\mathbf{x}))$ (non-linear pricing scheme) and we denote $t(q(\mathbf{x}))$ by $t(q)$.

Lemma 2.2 implies that the VP constraint (2.10) can be written as

$$U(\underline{\mathbf{x}}) := \mathbb{E}_W \left\{ t(q(\underline{\mathbf{x}})) - C(q(\underline{\mathbf{x}}), \underline{\mathbf{x}}, W) \right\} \geq 0. \tag{2.11}$$

.

Step 4. We show that under a certain quasi-concavity condition, stated in Lemma 2.3 below, we can define indirectly the allocation function $q(\mathbf{x})$ as a function of the payment function $t(l)$ by utilizing the incentive compatibility constraint. We define the following problem **(P3)**, that is equivalent to problem **(P2)**, in terms of the marginal price $p(l) = \frac{dt(l)}{dl}$ and the minimum payment $t(0)$:

$$\max_{p(\cdot),t(0)} \int_0^\infty P[\mathbf{x}\in\chi | p(l) \geq c(l,\mathbf{x})](\mathcal{V}'(l) - p(l))\, dl - t(0) \tag{2.12}$$

subject to

$$\text{interim VP:} \, \mathbb{E}_W \left\{ t(0) + \int_0^{q(\underline{\mathbf{x}})} p(l) dl - C(q(\underline{\mathbf{x}}), \underline{\mathbf{x}}, W) \right\} \geq 0. \tag{2.13}$$

The equivalence is established in two steps. First, consider an arbitrary incentive compatible mechanism (q, t). The optimal quantity $q^*(\mathbf{x})$ for each type $\mathbf{x}$ of the seller is given by

$$q^*(\mathbf{x}) = \arg\max_l \mathbb{E}_W \left\{ t(l) - C(l, \mathbf{x}, W) \right\}. \tag{2.14}$$

Incentive compatibility then requires that the seller must tell the truth to achieve this optimal value, and cannot do better by lying, *i.e.* $q(\mathbf{x}) = q^*(\mathbf{x})$ for all $\mathbf{x} \in \chi$. For any function $t(\cdot)$, this last equality can be taken as the definition for the associated function $q(\cdot)$. Thus, the IC constraint can be eliminated by defining $q(\cdot) := q^*(\cdot)$ and the problem of designing the optimal direct revelation mechanism (q, t) can be reduced to an equivalent problem where we determine only the optimal payment function $t(\cdot)$ subject to the voluntary participation constraint for the worst type.

Next, using Lemma 2.3, stated below, we rewrite the buyer's expected utility in terms of the marginal price $p(q) := \frac{\partial t(q)}{\partial q}$ and the minimum payment $t(0)$ (which along with $p(\cdot)$ uniquely determines the payment function $t(\cdot)$).

Lemma 2.3. *Assume that the seller's problem defined by (2.14) is continuous and quasi-concave*[6]*. Then, the buyer's expected utility can be expressed in terms of* $p(.)$ *and* $t(0)$ *as*

$$\begin{aligned}\mathbb{E}_{\mathbf{x}}[\mathcal{V}(q^*(\mathbf{x}))]-\mathbb{E}_{\mathbf{x}}[t(q^*(\mathbf{x}))]= &\\ &\textstyle\int_0^\infty P\left(\mathbf{x}\in\chi|q^*(\mathbf{x})\geq l\right)\mathcal{V}'(l)dl\\ &-t(0)-\textstyle\int_0^\infty P\left(\mathbf{x}\in\chi|q^*(\mathbf{x})\geq l\right)p(l)dl,\end{aligned} \tag{2.15}$$

where

$$P\left(\mathbf{x}\in\chi|q^*(\mathbf{x})\geq l\right)=P\left[\mathbf{x}\in\chi|p(l)\geq c(l,\mathbf{x})\right]. \tag{2.16}$$

Using (2.15) and (2.16), we can rewrite the objective of problem **(P2)** and obtain the equivalent problem **(P3)** given by (2.12) and (2.13)

Equation (2.16) states that the seller is willing to produce the marginal quantity at l if the resulting expected marginal profit is positive, *i.e.* the marginal price $p(l)$ exceeds the marginal expected cost of generation $c(l,\mathbf{x})$. Equation (2.15) expresses the buyer's total expected utility in term of an integral of his total marginal utility $\mathcal{V}'(l)-p(l)$ at quantity l, times the probability that the seller's production exceeds l, minus the minimum payment $t(0)$.

Step 5. We prove that the seller's worst type produces the minimum quantity among all the seller's types, *i.e.* $q(\underline{\mathbf{x}})=\min_{x\in\chi}q(\mathbf{x})$. As a result, we show that

[6] This is a standard assumption in economics literature, e.g. see [75] and [133]. Basically, it can be seen as a situation where the seller can decide for each marginal unit of production independently. Thus, in general, there is no guarantee that the seller's independent decisions about each marginal unit of production results in a continuous and plausible total production quantity q. Therefore, the continuity of the result must be checked a posteriori for each type of the seller.

problem **(P3)** is equivalent to the following problem **(P4):**

$$\max_{p(\cdot)} \int_0^\infty P\left[\mathbf{x} \in \chi | p(l) \geq c(l, \mathbf{x})\right] \left(\mathcal{V}'(l) - p(l)\right) dl \tag{2.17}$$

subject to

$$\text{iterim VP:} C(0, \underline{\mathbf{x}}) + \int_0^{q^*(\underline{\mathbf{x}})} p(l) dl \geq \mathbb{E}_W\left[C(q^*(\underline{\mathbf{x}}), \mathbf{x}, W)\right]. \tag{2.18}$$

We establish the equivalence by providing a ranking for the seller's optimal decision $q^*(\mathbf{x})$ based on the partial order among the seller's types.

Lemma 2.4. *For a given mechanism specified by* $(t(\cdot), q(\cdot))$*, a better type of the seller produces more. That is, the optimal quantity* $q^*(\mathbf{x})$ *that the seller with true type* $\mathbf{x}$ *wishes to produce satisfies the following properties:*

a) $\frac{\partial q^*(\mathbf{x})}{\partial x_i} \leq 0, 1 \leq i \leq m,$

b) $\frac{\partial q^*(\mathbf{x})}{\partial x_i} \geq 0, m < i \leq n.$

As a consequence of Corollary 2.1 and Lemma 2.4 we can then simplify the VP constraint (2.13) as follows.

Corollary 2.2. *The interim VP constraint is satisfied if* $t(0) = C(0, \underline{\mathbf{x}})$ *and the seller's worst type payment is equal to her expected production cost, i.e.* $t(q^*(\underline{\mathbf{x}})) = \mathbb{E}_W\{C(q^*(\underline{\mathbf{x}}), \underline{\mathbf{x}}, W)\}$.

The equivalence of problems **(P3)** and **(P4)** follows from Corollary 2.2 and by replacing the VP constraint (2.13) by (2.18). Note that we also dropped the constant term $t(0) = C(0, \underline{\mathbf{x}})$ (from Corollary 2.2) in the objective of problem **P4** given by (2.17).

Problem **(P4)** is in terms of the marginal price $p(l)$ and requires that the payment the seller's worst type receives is equal to her cost of production.

Step 6. We show that the solution of problem **(P4)** is given by

$$p(l) = \arg\max_{\hat{p}} \left\{P\left[\mathbf{x} \in \chi | \hat{p} \geq c(l, \mathbf{x})\right] (\mathcal{V}'(l) - \hat{p})\right\}$$

To prove the claim of Step 6 we consider a relaxed version of (P4) without the VP constraint (2.18). The unconstrained problem can be solved pointwise at each quantity l to determine the optimal $p(l)$ as

$$p(l) = \arg\max_{\hat{p}} \left\{ P\left[\mathbf{x} \in \chi | \hat{p} \geq c(l, \mathbf{x})\right] (\mathcal{V}'(l) - \hat{p}) \right\}, \tag{2.19}$$

which is the same as (2.4). From Corollary 2.2 and the fact that the worst type has the highest expected marginal cost, we can simplify (2.19), for $l \leq q^*(\underline{\mathbf{x}})$, as

$$p(l) = c(l, \underline{\mathbf{x}}), \quad \text{for } l \leq q^*(\underline{\mathbf{x}}). \tag{2.20}$$

Note that for $l \leq q^*(\underline{\mathbf{x}})$ we have $P\left[\mathbf{x} \in \chi | \hat{p} \geq c(l, \mathbf{x})\right] = 1$ from Lemma 2.4. Therefore, the minimum marginal price $p(l)$ that ensures all the seller's type are willing to produce more than $q^*(\underline{\mathbf{x}})$ is equal to the marginal expected cost for the seller's worst type $c(l, \underline{\mathbf{x}})$. Therefore, the solution to the unconstrained version of problem (P4) satisfies constraint (2.18) of problem (P4), and therefore, (2.19) is also the optimal solution of problem (P4).

We complete now the proof of Theorem 2.1. Using claim of Step 4 along with Corollary 2.2, the optimal payment function (nonlinear pricing) can be written as

$$t(q) = \int_0^q p(l) dl + C(0, \underline{\mathbf{x}})$$

which is the same as (2.5). From (2.14) we determine the optimal energy procurement function,

$$q(\mathbf{x}) = \arg\max_{l} \mathbb{E}_W \left\{ t\left(l\right) - C(l, \mathbf{x}, W) \right\}$$

which is the same as (2.6). The specification of $t(\cdot)$ and $q(\cdot)$ completes the proof of Theorem 2.1 and the solution to problem (P1).

In essence, Theorem 2.1 states that at each quantity l, the optimal marginal price $p(l)$ is chosen so as to maximize the expected total marginal utility at l, which is given by the total marginal utility $(\mathcal{V}'(l) - p(l))$ times the probability that the seller generates at least l.

Remark 2.1. *In problem (P1), we assume that there exists a seller's worst type which has the highest cost at any quantity among all the seller's types, and we reduce the VP constraint for all of the seller's type to only the VP constraint for this worst type. As a result, we pin down the optimal payment function by setting* $t(0) = C(0, \underline{\mathbf{x}})$ *to ensure the voluntary participation of the worst type, which consequently implies the voluntary participation for all of the seller's types. In absence of the assumption on the existence of the seller's worst type, the argument used to reduce the VP constraint is not valid anymore and we cannot pin down the payment function and specify* $t(0)$ *a priori. Assuming that all types of the seller participate in the contract, their decision on the optimal quantity* q^* *only depends on the marginal price* $p(q)$*, and therefore, the optimal marginal price* $p(q)$*, given by (2.19), is still valid without the assumption on the existence of the worst type. To pin down the payment function* $t(\cdot)$*, we find the minimum payment* $t(0)$ *a posteriori so that all types of the seller voluntarily participate. That is,*

$$t(0) = \max_{\mathbf{x} \in \chi} \left[\mathbb{E}_W \left\{ C(q(\mathbf{x}), \mathbf{x}, W) \right\} - \int_0^{q^*(\mathbf{x})} p(\hat{l}) d\hat{l} \right], \tag{2.21}$$

where the optimal decision of type $\mathbf{x}$ *is given by*

$$q^*(\mathbf{x}) = \arg\max_l \left[\int_0^l p(\hat{l}) d\hat{l} - \mathbb{E}_W \left\{ C(l, \mathbf{x}, W) \right\} \right]. \tag{2.22}$$

Remark 2.2. *In a setup with a positive zero-provision cost for the seller, it might not be optimal for the buyer to require all the seller's types to voluntarily participate in the procurement process, since* $t(0)$ *depends on the zero-provision cost of the seller's worst type* $C(0, \underline{\boldsymbol{x}})$*. In such cases, it might be optimal for the buyer to exclude some "less efficient" types of the seller from the contract, select an admissible set of the seller's types, and then design the optimal contract for this admissible set of the seller's types*[7]*. Note that this is not the case for setups without a zero-provision cost.*

[7] To find the optimal admissible set, the optimal contract can be computed for different potential

In such setups, if it is not optimal for some type $\mathbf{x}$ *to be included in the optimal contract, it is equivalent to set* $q(\mathbf{x})=0$ *in a contract menu that considers all types of the seller.*

2.3.3 Risk Allocation

In the optimal mechanism/contract menu presented by Theorem 2.1, the buyer faces no uncertainty, and he is guaranteed to receive quantity $q(\mathbf{x})$, and all the risk associated with the realization of W is taken by the seller. We wish to modify the mechanism to reallocate the above-mentioned risk between the buyer and the seller. To do so, we modify the payment function so that the risk is reallocated between the buyer and the seller. Consider the following modified payment function with $\alpha \in [0,1]$,

$$\begin{aligned} \hat{t}(\mathbf{x}, w) = \;& t(q(\mathbf{x})) + \alpha \left[C(q(\mathbf{x}), \mathbf{x}, w) \right. \\ & \left. - \mathbb{E}_W \left\{ C(q(\mathbf{x}), \mathbf{x}, W) \right\} \right]. \end{aligned} \tag{2.23}$$

From (2.23) it follows that $\mathbb{E}_W\left\{\hat{t}(\mathbf{x}, W)\right\} = t(q(\mathbf{x}))$. Therefore, the strategic behavior of the seller does not change and the seller chooses the same quantity under the modified payment function $\hat{t}(\cdot)$ as under the original payment function $t(q)$ given by (2.5). Note that for $\alpha = 0$ we have the same payment as $t(q)$. For $\alpha = 1$, the seller is completely insured against any risk and all the risk is taken by the buyer. The parameter α determines the allocation of the risk between the buyer and the seller; the buyer undertakes α and the seller undertakes $(1-\alpha)$ share of the risk.

We illustrate the result of Theorem 2.1 by an example below.

2.4 Example - Demand Response (DR)

We consider a contract design for DR program. There is a load aggregator that offers contracts with incentive payments to a heterogeneous population of loads who are willing to yield the direct control of their load to the aggregator given that they

admissible sets. Then, the resulting utilities can be compared to find the optimal admissible set.

are offered an appropriate incentive payment. The aggregator participates in an ancillary service market and sells the aggregated resources to the reserve market at exogenous marginal price p_r.[8] Formally, there are I types of loads with a population distribution f over different types. Each load of type i has a maximum controllable load L_i. Let $q_i \leq L_i$ denote the quantity that each load of type i yields its control to the aggregator to be dispatched. We assume that each load of type i has a quadratic cost (increasing marginal cost) given by

$$C_i = \alpha_i^0 + \alpha_i^1 q_i + \alpha_i^2 q_i^2. \tag{2.24}$$

Therefore, the load's type is $\mathbf{x} = (L_i, \alpha_i^0, \alpha_i^1, \alpha_i^2)$. Let t_i denotes the incentive payment to each load of type i for yielding the control of load q_i. Then, the total utility of each load of type i is given by

$$U_i = t_i - \left(\alpha_i^0 + \alpha_i^1 q_i + \alpha_i^2 q_i^2\right). \tag{2.25}$$

The aggregator participates in the ancillary service market and provides capacity $q = \sum f_i q_i$ at a given uniform price p_r. Therefore, the aggregator's revenue is given by

$$p_r \sum (f_i q_i) - \sum (f_i t_i). \tag{2.26}$$

We consider $I = 5$ types of loads described in Table 2.1 along with a normalized population distribution f with $\sum f_i = 1$, and set $p_r = 2$ ¢$/kWh$. Note that no complete ordering can be defined based on their marginal cost and there exists no worst type; at lower quantities smaller loads (*e.g.* type (b)) have a lower marginal cost while at higher quantities larger loads (*e.g.* type (e)) have lower cost.

Via Theorem 2.1 we determine the optimal menu of contracts the aggregator offers to the heterogeneous population of loads (Table 2.2). The optimal menu of

[8] If p_r is not exogenous, the aggregator's interactions with the reserve market on one hand and the demand population on the other hand become coupled. In this case these interactions must be studied simultaneously.

contracts can be interpreted also as a nonlinear pricing that the aggregator offers to loads (Fig. 2.1).

The optimal choice and the resulting payoff for each type of load are summarized in Table 2.3. We note that, unlike one-dimensional contracts, a type with a higher quantity does not necessarily get a higher payoff.

type	$L_{i(kWh)}$	α_i^0 (¢)	α_i^1 $(\frac{¢}{kWh})$	α_i^2 $(\frac{¢}{kWh^2})$	f_i
(a)	0.5	0.1	5	10	0.1
(b)	1	0.1	4	10	0.3
(c)	1.5	0.6	8	5	0.2
(d)	2	0.6	5	8	0.3
(e)	2.5	1.2	6	5	0.1

Table 2.1: Different types of loads

Quantity $q(\cdot)$ (kWh)	0.38	0.64	0.82	1.10	1.40
Payment $t(\cdot)$ (¢)	3847	7569	10498	15450	20991

Table 2.2: Options menu offered by the aggregator

Type	Quantity	Payment	Cost	Profit
(a)	0.382	3847	3469	378
(b)	0.643	7569	6897	762
(c)	1.100	15450	15450	0
(d)	0.8185	10498	10052	446
(e)	1.400	20991	19400	1591

Table 2.3: Optimal contract for each and the resulting outcome

2.5 Forward Contracts with Random Allocation

In some instances of the problem considered in this chapter, the buyer has a reserve resource [130] or wants to supply deferrable loads [21] that gives him the

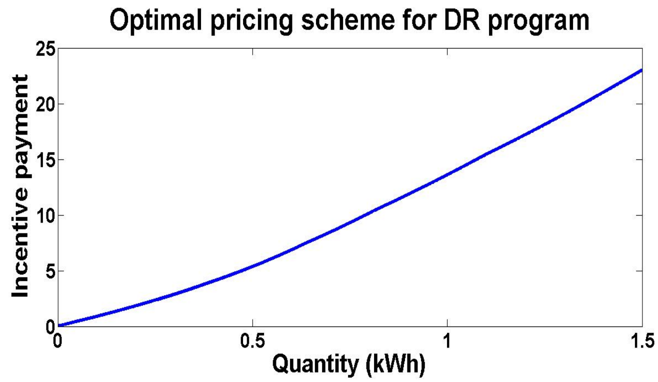


Figure 2.1: The optimal pricing scheme for DR program

flexibility to accept a random allocation $q(\mathbf{x}, W)$ that depends on the uncertainty W, and compensate the randomness in the allocation utilizing the existing flexibility. In this section, we formulate and analyze a forward contract design problem with random allocation. We assume that the realization of the random variable W is common knowledge between the buyer and the seller.

2.5.1 Problem Formulation

Let $e(\mathbf{x})$ denote the forward scheduled quantity (deterministic) by the buyer and $q(\mathbf{x}, w)$ denote the random delivered quantity by the seller with type $\mathbf{x}$. Let $C_R(e(\mathbf{x}) - q(\mathbf{x}, w))$ denote the cost incurred by the buyer to compensate the real-time deviation $e(\mathbf{x}) - q(\mathbf{x}, w)$ from the forward schedule $e(\mathbf{x})$. Then, for a given set of contract menus $(q(\mathbf{x}, w), t(\mathbf{x}, w))$, the buyer's optimal schedule $e(\mathbf{x})$ for the seller's type $\mathbf{x}$ is defined by

$$e(\mathbf{x}) = arg \max_{\hat{e}} \mathbb{E}_W\{\mathcal{V}(\hat{e}) - t(\mathbf{x}, W) - C_R(\hat{e} - q(\mathbf{x}, W))\}, \tag{2.27}$$

and the buyer's expected utility is given by

$$\mathbb{E}_{W,\mathbf{x}}\{\mathcal{V}(e(\mathbf{x})) - t(\mathbf{x}, W) - C_R(e(\mathbf{x}) - q(\mathbf{x}, W))\}\,. \tag{2.28}$$

The buyer wants to design a mechanism $(q(\mathbf{x}, w), t(\mathbf{x}, w))$ so as to maximize his expected utility given by (2.28), subject to the voluntary participation of the seller. Formally, the contract design problem with random allocation for the buyer, called **(Q1)**, can be stated as follows:

$$\underset{\{q(\cdot,\cdot),t(\cdot,\cdot)\}}{\text{maximize}} \mathbb{E}_{W,\mathbf{x}}\{\mathcal{V}(e(\mathbf{x}))-t(\mathbf{x},W)-C_R(e(\mathbf{x})-q(\mathbf{x},W))\} \tag{2.29}$$

subject to

$$\mathbb{E}_W\{t(\mathbf{x}, W) - C(q(\mathbf{x}, W), \mathbf{x}, W)\} \geq 0, \forall \mathbf{x} \in \mathcal{X}. \tag{2.30}$$

2.5.2 Analysis & Results

We show, via Theorem 2.2 below, that the optimal forward contract with random allocation is a menu of pricing schemes, one for each type of the seller.

Theorem 2.2. *The optimal forward contract with random allocation for problem **(Q1)** is a menu of pricing schemes given by*

$$e(\mathbf{x}) = \tilde{q}(\mathbf{x}) \tag{2.31}$$

$$q(\mathbf{x}, w) = \tilde{q}(\mathbf{x}) - q_R(\mathbf{x}, w), \tag{2.32}$$

$$t(\mathbf{x}, w) = \tilde{t}(\mathbf{x}) - C_R(q_R(\mathbf{x}, w)), \tag{2.33}$$

where $\left\{\tilde{q}(x), \tilde{t}(x)\right\}$ *denotes the optimal solution to the optimization problem*

$$\underset{\left\{\tilde{q}(\cdot),\tilde{t}(\cdot)\right\}}{\text{maximize}} \mathbb{E}_{W,\mathbf{x}} \left\{\mathcal{V}(\tilde{q}) - \tilde{t}\right\} \tag{2.34}$$

subject to

$$\mathbb{E}_{W,\mathbf{x}} \left\{\tilde{t}(\mathbf{x}) - \tilde{C}(\tilde{q}(\mathbf{x}), \mathbf{x}, W)\right\} \geq 0, \tag{2.35}$$

$$\tilde{C}(q, \mathbf{x}, w) := \min_{l} \{C(l, \mathbf{x}, w) + C_R(\tilde{q} - l)\}, \tag{2.36}$$

and

$$q_R(\mathbf{x}, w){:=}arg\min_{l}\left\{C(\tilde{q}(\mathbf{x}){-}l, \mathbf{x}, w){+}C_R(l)\right\}. \tag{2.37}$$

Proof. Consider the following contract design problem where the seller's cost function is defined as

$$\tilde{C}(\tilde{q}, \mathbf{x}, w) = \min_{l}\left\{C(l, \mathbf{x}, w) + C_R(\tilde{q} - l)\right\},$$

where $C(\cdot, \cdot, \cdot)$ is the seller's cost function in **(Q1)**, and the buyer's utility is defined as

$$\mathbb{E}_{W,\mathbf{x}}\left\{\mathcal{V}(\tilde{q}) - \tilde{t}(\tilde{q})\right\}. \tag{2.38}$$

The optimal contract design problem for the defined environment above, called **(Q2)**, can be stated as

$$\underset{\{\tilde{q},\tilde{t}\}}{\text{maximize}}\mathbb{E}_{\mathbf{x},W}\left\{\mathcal{V}(\tilde{q}) - \tilde{t}\right\} \tag{2.39}$$

subject to

$$\text{IC: } \mathbf{x}{=}arg\max_{\mathbf{x}'}\mathbb{E}_W\Big\{\tilde{t}(\mathbf{x}'){-}\tilde{C}(\tilde{q}(\mathbf{x}'), \mathbf{x}, W)\Big\}, \forall\mathbf{x}{\in}\chi \tag{2.40}$$

$$\text{interim VP:} \, \mathbb{E}_W\Big\{\tilde{t}(\mathbf{x}){-}\tilde{C}(\tilde{q}(\mathbf{x}), \mathbf{x}, W)\Big\}{\geq}0, \forall\mathbf{x}{\in}\chi \tag{2.41}$$

where $\tilde{q}$ and $\tilde{t}$ denote the quantity and payment function for the defined problem above. By construction, problem **(Q2)** is the same as problem **(P2)**. Let $\left\{\tilde{q}(\mathbf{x}), \tilde{t}(\mathbf{x})\right\}$ denote the optimal contract for problem **(Q2)** obtained via Theorem 2.1. Note that through the cost function $\tilde{C}(\tilde{q}(\mathbf{x}), \mathbf{x}, w)$, defined by (2.36), we absorb the optimal schedule choice $e(\mathbf{x})$, given by (2.27), and internalize the compensation cost $C_R(e(\mathbf{x}) - q(\mathbf{x}, W))$ for the random deviation $e(\mathbf{x}) - q(\mathbf{x}, W)$ in problem **(Q1)** into the seller's cost function. Therefore, the optimal scheduled quantity $e(\mathbf{x})$ for problem **(Q1)** is equal to the optimal function $\tilde{q}(\mathbf{x})$ for problem **(Q2)**, *i.e.* $e(\mathbf{x}) = \tilde{q}(\mathbf{x})$. Consequently, one can reconstruct the optimal contract $\{q(\mathbf{x}, w), t(\mathbf{x}, w)\}$ for problem **(Q1)** using the optimal contract $\left\{\tilde{q}(\mathbf{x}), \tilde{t}(\mathbf{x})\right\}$ for the equivalent problem **(Q2)** as

$$\begin{aligned} q(\mathbf{x}, w) &= \tilde{q}(\mathbf{x}) - q_R(\mathbf{x}, w), \\ t(\mathbf{x}, w) &= \tilde{t}(\mathbf{x}) - C_R(q_R(\mathbf{x}, w)), \end{aligned}$$

where

$$q_R(\mathbf{x}, w) := arg \min_l \left\{ C(\tilde{q}(\mathbf{x}) - l, \mathbf{x}, w) + C_R(l) \right\},$$

denotes the random reserve quantity required to compensate the random allocation $q(\mathbf{x}, w)$. □

Theorem 2.2 has the following interpretation. The buyer offers different pricing schemes (quantity-payment curves), and each type of the seller chooses one based on her private information and expectation about W. Then, in real time as W is realized, based on the realization w, one point from the chosen pricing scheme is selected and the payment t and the energy delivery q are determined.

2.5.3 Imperfect Commitment and Ex-post Voluntary Participation

The voluntary participation constraint imposed in problem (Q1) is interim. That is, the expected profit with respect to W must be non-negative for each type of the seller. Up until now (problem **(P1)** and **(Q1)**) we have assumed that once the seller agrees to sign the contract (such an agreement takes place before the realization of random variable W) she is fully committed to following the agreement, even if the realized profit is negative (due to some realization w)[9]. Therefore, it would be desirable to modify the contract in order to ensure a positive payoff for the seller for every realization of W and full commitment without any outside enforcement. To ensure that the seller's realized profit is non-negative for every realization w,

[9]Since the seller's reserved utility is zero by not participating (outside option), we can always think of the seller walking away from the agreement for these negative profit realizations and not following the mechanism rules.

we impose an *ex-post voluntary participation* constraint and replace the interim VP constraint (2.30) by

$$\text{Ex-post VP: } t(\mathbf{x}) - C(q, \mathbf{x}, w) \geq 0, \forall w,\ \forall \mathbf{x} \in \chi. \tag{2.42}$$

To satisfy the ex-post voluntary participation constraint, we modify the payment function of the mechanism given by Theorem 2.2 as follows:

$$\begin{aligned} \check{t}(\mathbf{x}, w) = \mathbb{E}_W \{t(\mathbf{x}, W)\} - \mathbb{E}_W \{C(q(\mathbf{x}, W), \mathbf{x}, W)\} \\ + C(q(\mathbf{x}, w), \mathbf{x}, w). \end{aligned} \tag{2.43}$$

We have $\mathbb{E}_W\{\check{t}(\mathbf{x}, W)\} = \mathbb{E}_W \{t(\mathbf{x}, W)\}$, and therefore, the seller always chooses the same quantity q under the modified payment function $\check{t}$ as under the original payment function t given by (2.33). Furthermore, we have $\check{t}(\mathbf{x}, w) - C(q(\mathbf{x}, w), \mathbf{x}, w) = \mathbb{E}_W \{t(\mathbf{x}, W)\} - \mathbb{E}_W \{C(q(\mathbf{x}, W), \mathbf{x}, W)\} \geq 0$ for all $w, \mathbf{x}$, where the last inequality is true since $\{q, t\}$ satisfies the interim VP constraint (2.30). Therefore, under the modified payment function $\check{t}$, $\{q, \check{t}\}$ satisfies the expost VP constraint (2.42).

2.6 Example - Forward Bilateral Trade

Consider a forward bilateral trade between a buyer and a seller with wind generation. The buyer has an (almost inelastic) energy demand curve given by Fig. 2.2.

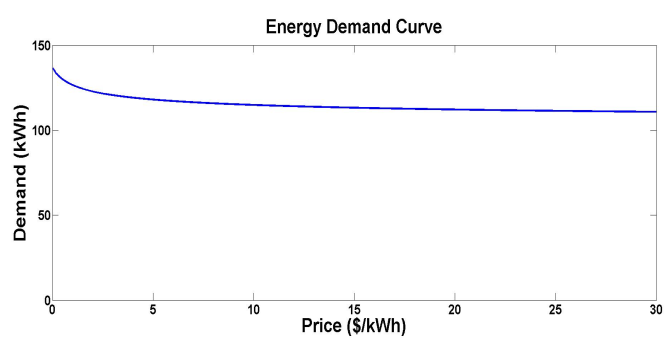


Figure 2.2: The buyer's demand curve

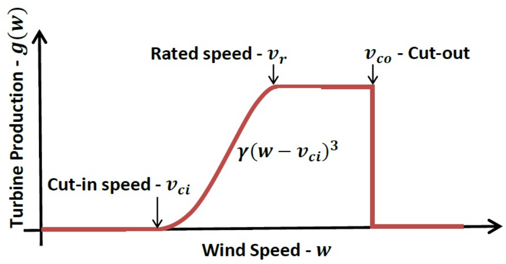


Figure 2.3: The wind turbine generation curve $g(w, v_{ci}, v_r, v_{co}, \gamma)$

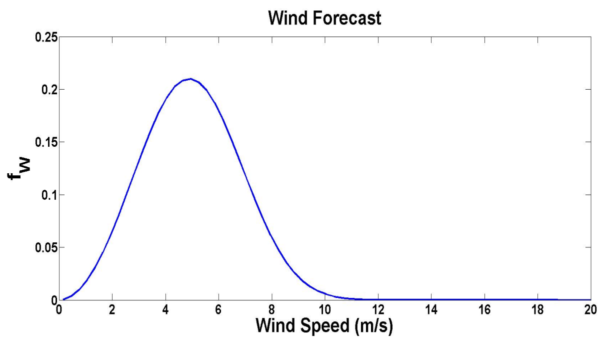


Figure 2.4: The wind forecast F_W

The seller has a wind farm and (possibly) a reserve generator/storage that can be used to compensate for wind generation intermittency. The seller's wind generation is given by $g(w, v_{ci}, v_r, v_{co}, \gamma)$ as in Fig. 2.3, where w denotes the wind speed and $(v_{ci}, v_r, v_{co}, \gamma)$ denotes the specification of the wind turbine. The wind speed is random and the wind forecast f_W is given by Fig. 2.4, which is a Weibull distribution with shape parameter $k = 3$ and average wind speed of $5m/s$. We assume that the wind forecast f_W as well as the wind realization w are common knowledge between the buyer and the seller. The wind generation has a marginal operational cost θ_w. The seller (possibly) has a reserve generator/storage with capacity r and a marginal cost θ_r that can be utilized if needed. The seller has a zero-production cost c_0 which accounts for her capital cost and the start-up cost of her facilities. Therefore, the seller's private information is as $\mathbf{x} = (c_0, \theta_w, \theta_r, v_{ci}, v_r, v_{co}, \gamma, r)$.

We assume that the buyer has a reserve generator/deferrable load that can be utilized to compensate the real-time random energy delivery by the seller. We assume that deviation from the scheduled energy has an increasing marginal cost for the buyer given by $b_0 + b_1 q$.

We consider four types for the seller as in Table 2.4, and set $b_0 = 1.4\frac{\$}{kwh}$ and $b_1 = 0.05\frac{\$}{kWh^2}$.

type	c_0	θ_w	θ_c	v_{ci}	v_r	v_{co}	γ	r	f_i
a	90	0.20	1.2	0.4	13	23	1	60	0.1
b	60	0.25	1.4	0.8	17	25	1.25	30	0.3
c	40	0.10	1.0	0.1	15	20	1.5	10	0.2
d	20	0.15	-	1.0	17	28	1.7	0	0.3

Table 2.4: Different types of the seller

The optimal forward contract menu for the buyer is given by Fig. 2.5. Since the energy demand considered in this example is almost inelastic, the scheduled quantity $e(\mathbf{x})$, and therefore, the quantity-demand curves are also close to each others.[10] Table 2.5 summarizes the optimal energy schedule $e(\mathbf{x})$, and the expected utility $U(\mathbf{x})$ for different types of the seller.

type	a	b	c	d
$e(\mathbf{x})$	122.3	118.5	120.4	116.5
$U(\mathbf{x})$	84.47	35.10	101.82	0

Table 2.5: The outcomes of the optimal contract menus

The energy $q(\mathbf{x}, w)$ delivered to the buyer, the payment $t(\mathbf{x}, w)$ made to the seller, and the seller's utility $u(\mathbf{x}, w)$ in terms of wind w are given by Figures 2.6, 2.7, and 2.8, respectively.

For low realizations of wind speed, the delivered energy is low and the seller may even incur some penalty for very low energy delivery (Fig. 2.5). For higher realization of wind speed, the energy delivery increases, and therefore, the payment

[10]For a completely inelastic energy demand, we have $e(\mathbf{x})$ fixed and independent of the seller's type $\mathbf{x}$. Therefore, all the quantity-demand curves coincide and are equal to the quantity-cost curve for the worst type.

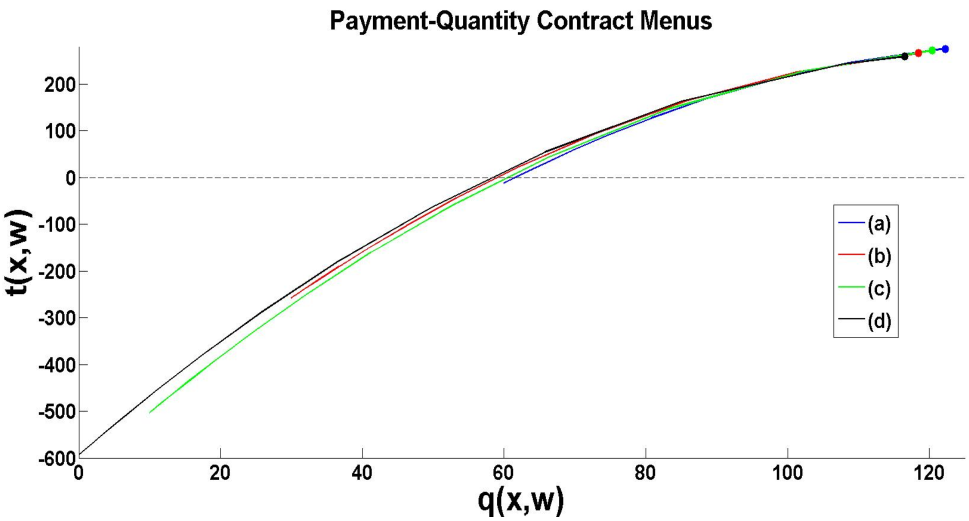


Figure 2.5: The optimal forward contract menus

and the realized utility increase. However, for very high realization of wind speed that surpasses the cut-off speed v_{co} (see Figure 2.3), the energy delivery, and consequently the payment and the realized utility, drop.

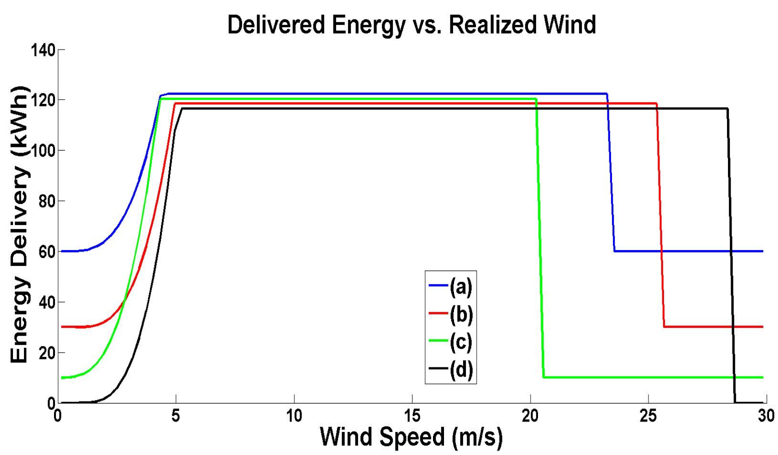


Figure 2.6: Energy deliver $q(\mathbf{x}, w)$ in terms of wind w

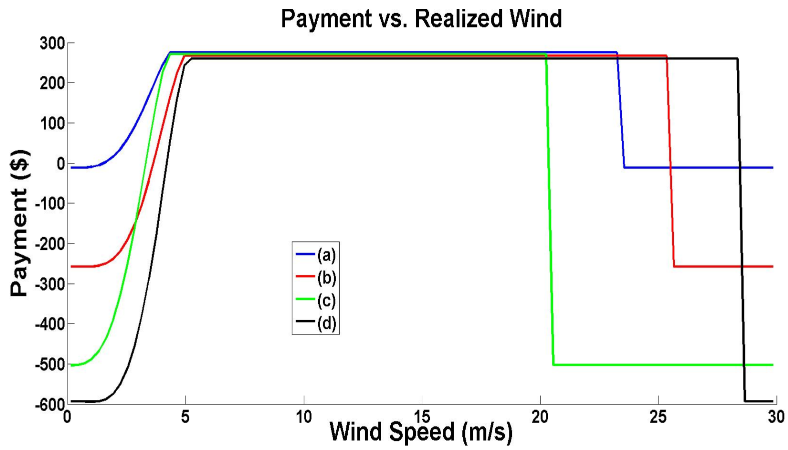


Figure 2.7: Payment $t(\mathbf{x}, w)$ in terms of wind w

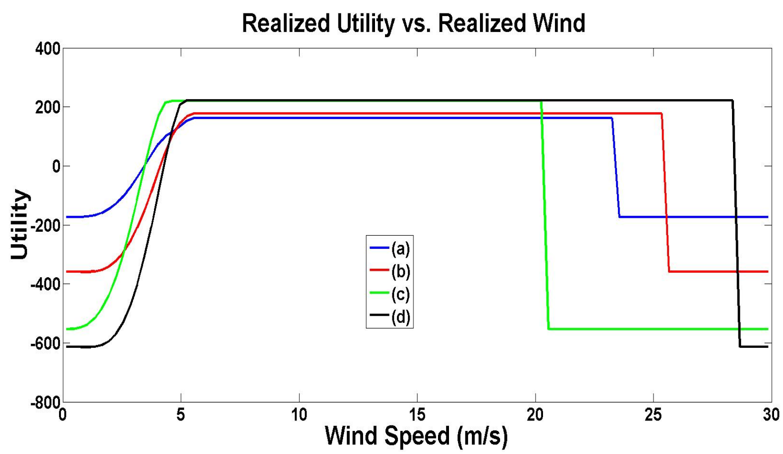


Figure 2.8: The realized seller's utility $u(\mathbf{x}, w)$ in terms of wind w

2.7 Discussion

For the problem on energy/service procurement formulated in this chapter the optimal mechanism is a menu of contracts/nonlinear pricing schemes. The nonlinearity is due to three factors. First, the buyer's utility function $\mathcal{V}(q)$ is not linear in the quantity q. Second, for each type of the seller, the cost function is a nonlinear function of the quantity . Third, the seller has private information about her technology and cost (seller's type).

The buyer has to pay information rent (monetary incentive) to the seller to incentivize her to reveal her true type. Therefore, the payment the buyer makes to the seller includes the cost of provision the seller incurs plus the information rent, which varies with the seller's type; the better the seller's type, the higher is the information rent.

The optimal forward contracts discovered in this chapter can be implemented as follows: the buyer offers the seller a menu of contracts (nonlinear pricing schemes); the seller chooses one of these contracts based on her type.

The optimal forward contracts induce some incentives for investment in infrastructure and technology development. From Lemma 2.1, the seller with the higher type has a higher utility. Therefore, there is an incentive for the seller to improve her technology and decrease her cost of generation.

It is well-known that in the presence of private information and strategic behavior, in general, there exists no mechanism/contract that is (1) individually rational, (2) incentive compatible, and (3) efficient (Pareto-optimal) [105]. In the optimal forward contract given by Theorems 2.1 and 2.2 the allocation for the seller's different types is not ex-post efficient (Pareto-optimal) except for the seller's worst type who gets zero utility.

In this chapter, we formulated the contract design problem in a principal-agent setup. Therefore, the result can be applied to the contract design problem for a setup with one buyer (principal) and a heterogeneous population of sellers (agents), if the buyer has a linear utility function, as in the example of Section 2.4, or if the share of each individual agent is small and their effect on the market is negligible. However, if one considers a setup with nonlinear utility for the buyer or market power for each individual agent, the associated problem for such setup with multiple agents becomes equivalent to the design of optimal multi-unit auctions in economics. It is known that there exist no closed form solution to the general problem of optimal multi-unit auctions, and their solutions can only be computed numerically or with approximation [66].

2.8 Conclusion

We investigated the problem of optimal forward contract design under uncertainty and multi-dimensional private information. The consideration of multi-dimensional private information and general utility/cost functions enables us to capture many applications in electricity markets as well as other disciplines. We assumed that the buyer and/or the seller has uncertainty in their utility/cost function which is realized after the time of contract signing. We considered froward contracts with random allocation that depends on the real-time realization of the uncertainty. We characterized the optimal forward contract under uncertainty as a menu of contracts. We addressed the problem of commitment (ex-post voluntary participation), and risk sharing in the presence of uncertainty. We demonstrated our results by two examples; an optimal contract design for a demand response program, and an optimal forward bilateral trade between a buyer and a seller with wind energy generation.

Chapter 3

Dynamic Market Mechanisms for Wind Energy

3.1 Introduction

3.1.1 Background and Motivation

Wind generation is intermittent and uncertain. An energy producer with wind energy (seller) has neither complete control over his generation nor does he have an exact prediction of his generation in advance. The information about wind realization arrives dynamically over time and an accurate prediction is only available within a few (5-15) minutes of the generation time [56]. The stochastic and dynamic nature of wind energy makes the integration of wind generation into grids a challenging task.

The common practice for the integration of wind energy is to incorporate it into the existing two-settlement market architecture for conventional generators along with *extra-market treatments* such as *feed-in tariffs*, *investment tax credit*, and *production tax credit*. The two-settlement market architecture consists of forward markets (*e.g.* day-ahead market) and real-time markets, where the outcome of forward markets is fed to real-time markets.

For energy markets with low share of wind energy, like the U.S., it is possible to include wind energy in real-time markets, and treat it as negative load [19]; we call this approach *real-time mechanism*. One advantage of incorporating wind energy into real-time markets is that the allocation for wind generation is decided when all the information about wind generation is available. Moreover, a seller does not face any penalty risk as he commits to a certain level of generation only if he can produce it. However, in energy markets with high share of wind energy, due to reliability concerns, inclusion of wind energy in real-time markets as negative load is not possible.

For high shares of wind energy, the system operator needs to have information about wind generation in advance, and to incorporate wind energy as an active generation into its forward planning for power flow. Thus, wind energy is included in forward markets [62]; we call this approach *forward mechanism*.

In forward markets, knowledge about wind generation in real time is imperfect. Nevertheless, a seller needs to commit to certain levels of generation in advance even without knowing the exact amount he will be able to produce in real time. Therefore, the energy allocation decision in forward markets is determined only based on the incomplete information available at the time of forward markets, and all the new information that arrives after forward markets are closed is not incorporated into the energy allocation decision. In some countries, like the U.K., a seller is exposed to penalty risk if his real-time generation is different from his commitment in forward markets. We note that this is not an issue for a conventional generator as he has perfect knowledge of his real-time generation in advance.

The limitations of the real-time and forward market mechanisms discussed above motivate our work to study alternative market mechanisms for the integration of wind energy into grids. The U.S. department of energy encourages the development of "rules for market evolution that enable system flexibility rationale", *i.e.* market mechanisms that give a seller flexibility in generation, and provide opportunities for the demand side to actively respond to changes in market over time [30]. It is desired that such mechanisms provide truthful (probabilistic) information about the seller's generation in real-time, and assign commitment to the seller in advance [126]. To achieve the above features, we need to study market mechanisms in a dynamic setting that accounts for the dynamic and intermittent nature of wind generation as well as the strategic behavior of the seller. The forward and real-time mechanisms discussed above are static mechanisms in the strategic sense. That is, for each market mechanism, the sellers and buyers make simultaneous decisions only once, and their one-shot decisions determine the energy allocations and payments at that market; the outcome of the market is then assumed to be fixed and is fed as an exogenous parameter to the next market in time.

In this chapter, we consider a simple two-step model that captures the dynamic and intermittent nature of wind generation. We propose a dynamic mechanism that provides a coupling between the forward and real-time mechanisms, and, unlike the (static) real-time and forward mechanisms, allows for flexible generation of wind energy, incorporates all the information that arrives over time, and provides forward commitment of the seller.

To demonstrate the main ideas, we first consider a strategic setting with one buyer and one seller with wind generation. The buyer and the seller trade energy through a mechanism determined by a mechanism designer. The seller's cost depends on his private technology and the wind condition which he learns dynamically over time. Since the seller is strategic and profit maximizer, he must be incentivized to reveal his private information about his cost function. We formally define such incentive payment, and utilize it in the formulation and solution of the mechanism design problems that we consider in this chapter. We determine such incentive payments for different market mechanisms. We then characterize the set of feasible outcomes under each market mechanism. We show that the dynamic mechanism outperforms the real-time and forward mechanisms for a general objective of the designer.

After we establish our results for a setting with one seller, we consider a setting with many sellers. We discuss how the problem of market design with many sellers is similar to that with one seller, and argue that our results generalize to this setting.

Specifically, we formulate and study three different mechanism design problems. Two of these problems capture the real-time and forward mechanisms. For the third problem, we propose a new dynamic market mechanism that dynamically couples the outcome of the real-time and forward markets. In the dynamic market mechanism, the seller is required to sequentially reveal his private information to the designer as it becomes available, and accordingly refines his commitment for energy generation over time.

We show that the set of constraints that the designer faces due to the seller's strategic behavior and private information is less restrictive under the dynamic mechanism than under the forward and real-time mechanisms. Consequently, we show that the proposed dynamic market mechanism outperforms the real-time and forward mechanisms. We further consider two variants of the dynamic mechanism; one guarantees no penalty risk to the seller, and in the other the designer monitors the wind speed. By analyzing the outcome of these variants of the dynamic mechanism, we characterize the effect of penalty risk exposure and wind monitoring on the performance of the dynamic mechanism.

3.1.2 Related work

Most of the literature on market design for wind energy assumes a static information structure and has mainly taken the forward mechanism approach. The works of [20, 139] study the problem of optimal bidding in a forward market with an exogenous price and penalty rate. The works of [72, 92] investigate the problem of mechanism design for wind aggregation among many wind producers that jointly participate in a forward market. The work in [115] studies the problem of auction design for a forward market that determines the penalty rate endogenously.

The concept of (flexible) contracts with risk in electricity market has been proposed in [113, 21, 139, 122]. The authors of [139] propose and investigate risky contracts for wind aggregation where there is no private information. The work in [113] studies the problem of efficient pricing of interruptible energy services. The authors of [21] look at the problem of optimal pricing for deadline-differentiated deferrable loads. The authors of [122] study the problem of forward risky contracts when the seller's private information is multi-dimensional and the wind is monitored.

The various mechanism design problems formulated in this chapter belong to the *screening* literature in economics. We formulate static and dynamic mechanism design problems. Our approach to the static mechanism design problems is similar to the one in [25]. Our approach to the dynamic mechanism design problems is inspired by the ones in [31, 38, 63]. We provide a unified approach for all the mechanism design problems that enables us to demonstrate the advantage of dynamic mechanisms over static mechanisms for the wind energy market.

3.1.3 Contribution

We propose a dynamic model that enables us to provide a comparison among various market mechanisms for wind energy. We propose and analyze a dynamic market mechanism that couples the outcomes of the real-time and forward mechanisms. We show that the proposed dynamic mechanism outperforms the real-time and forward mechanisms for a general objective of the market designer. The proposed dynamic mechanism reveals to the system designer the information required for planning in advance, incorporates the new information that arrives over time, and provides flexibility for the intermittent wind energy generation. We further study the effect of providing penalty insurance to the seller and monitoring the wind condition. We show that the performance of the dynamic mechanism with no penalty risk is in general inferior to the dynamic mechanism with penalty risk. Moreover, with wind monitoring, the outcome of the dynamic mechanism improves, as the seller cannot manipulate the outcome of the mechanism by misrepresenting his wind condition.

3.1.4 Organization

We present our model in Section 3.2. In Section 3.3, we discuss the market mechanism design problems as well as the seller's strategic behavior and private information. In Section 3.4, we propose the dynamic market mechanism, and formulate three mechanism design problems accordingly. We analyze the formulated mechanism design problems and compare them in Section 3.5. In Section 3.6, we provide some additional remarks and consider variants of the proposed dynamic mechanism. We illustrate our results with an example in Section 3.7. We discuss how our results generalize to settings with many sellers in Section 3.8, and conclude in Section 3.9. All the technical proofs can be found in Appendix B.

3.2 Model

Consider a buyer and a diversified seller with wind energy generation who trade energy through a mechanism determined by a designer. We refer to the seller as "he" and the designer and the buyer as "she". The buyer gets utility $\mathcal{V}(\hat{q})$ from receiving energy $\hat{q}$; $\mathcal{V}(\hat{q})$ is increasing and strictly concave in $\hat{q}$. The seller has production cost $C(\hat{q};\theta)$ parametrized by his type $\theta \in \{\underline{\theta}, \overline{\theta}\}$.[1] The seller's type depends on his technology τ (*i.e.* technology of his wind turbines and the operational status of them, the size of his wind farm and its location) and the wind speed ω. We assume that at $T=1$ (*e.g.* day ahead), *ex-ante* the seller knows privately his technology τ for wind generation, which takes values from one of M possible technologies $\mathcal{T} := \{\tau_1, \tau_2, \cdots, \tau_M\}$, $\tau_1 < \tau_2 < ... < \tau_M$, with probability $p_1, p_2, \cdots, p_M$, $\sum_{i=1}^{M} p_i = 1$, respectively. At $T = 2$ (*e.g.* real-time), the seller receives additional information $\omega \in [\underline{\omega}, \overline{\omega}]$ about the wind condition and can refine his private information about his cost function at $T=2$ as $\theta = \Theta(\tau, \omega)$. The probability distribution of wind ω is independent of the technology τ and is denoted by $G(\omega)$. We assume that $C(\hat{q};\theta)$ is increasing in θ, and $\Theta(\tau,\omega)$ is decreasing in ω and τ. Define $C_\theta(\hat{q};\theta) := \frac{\partial C(\hat{q};\theta)}{\partial \theta}$, $c(\hat{q};\theta) := \frac{\partial C(\hat{q};\theta)}{\partial \hat{q}}$, $c_\theta(\hat{q};\theta) := \frac{\partial c(\hat{q};\theta)}{\partial \theta}$, and $\Theta_\omega(\tau,\omega) := \frac{\partial \Theta(\tau,\omega)}{\partial \omega}$.

Assumption 3.1. *The production cost $C(\hat{q};\theta)$ is increasing and convex in $\hat{q}$. The marginal cost $c(\hat{q};\theta)$ is increasing in θ, i.e. $C_\theta(\hat{q},\theta) \geq 0$ and $c_\theta(\hat{q},\theta) \geq 0$. The seller's type $\Theta(\tau,\omega)$ is decreasing in τ, i.e. $\Theta(\tau_i,\omega) \leq \Theta(\tau_j,\omega)$ for $i > j$, $\forall \omega$, and strictly decreasing in ω, i.e. $\Theta_\omega(\tau,\omega) < 0$.*

Let $F_i(\theta)$ denote the resulting conditional probability distribution of θ given τ_i. Then, Assumption 3.1 states that a seller with technology τ_i expects to have a lower production cost than the one with technology τ_j for $i > j$, *i.e.* $F_{\tau_j}(\theta)$ *first order stochastically dominates* $F_{\tau_i}(\theta)$.

We also make the following technical assumption.

[1] We assume that $C(\hat{q};\theta)$ captures the seller's operational cost, capital cost, and an exogenous opportunity cost associated with wind generation participation in his outside option. Moreover, we assume that the seller has a diversified energy portfolio, so he does not face a strict capacity constraint for generation as he can produce energy from other resources that are more expensive. Therefore, $C(\hat{q};\theta)$ is non-zero but decreasing in the realization of wind speed.

Assumption 3.2. *The distribution $F_\tau(\theta)$ has non-shifting support, i.e. f_τ is strictly positive on the interval $\left[\underline{\theta}, \overline{\theta}\right]$ for $\tau \in \mathcal{T}$.*

Assumption 3.2 implies that the range of achievable values of θ is the same for all technologies τ_i, $i=1,...,M$. However, the same realization of the wind ω results in different values of θ for different technologies. Thus, the probability distribution of θ, given by $F_{\tau_i}(\theta)$, is different for different technologies (*cf.* Assumption 3.1).

3.3 Mechanism Design and Strategic Behavior

Consider an arbitrary mechanism that determines the agreement between the buyer and the seller. Let $t(\tau,\omega)$ denote the payment the buyer ends up paying to the seller with technology τ and wind speed ω, and $q(\tau,\omega)$ denote the amount of energy the seller delivers according to the mechanism. The social welfare $\mathcal{S}$, the seller's revenue $\mathcal{R}$, and the buyer's utility $\mathcal{U}$ can be written as,

$$\mathcal{S} := \mathbb{E}_{\tau,\omega}\{\mathcal{V}(q(\tau,\omega)) - C(q(\tau,\omega); \Theta(\tau,\omega))\}, \tag{3.1}$$

$$\mathcal{R} := \mathbb{E}_{\tau,\omega}\{t(\tau,\omega) - C(q(\tau,\omega); \Theta(\tau,\omega))\}, \tag{3.2}$$

$$\mathcal{U} := \mathbb{E}_{\tau,\omega}\{\mathcal{V}(q(\tau,\omega)) - t(\tau,\omega)\} = \mathcal{S} - \mathcal{R}. \tag{3.3}$$

The social welfare $\mathcal{S}$ only depends on q and is independent of t. Thus, by the first order optimality condition, an allocation q^{e*} is socially efficient (maximizing $\mathcal{S}$) if and only if at q^{e*} the marginal utility $v(\hat{q}) := \frac{\partial \mathcal{V}(\hat{q})}{\partial \hat{q}}$ is equal to the marginal cost $c(\hat{q}; \theta)$, *i.e.*, for all τ, ω,

$$v(q^{e*}(\tau,\omega)) = c(q^{e*}(\tau,\omega); \Theta(\tau,\omega)). \tag{3.4}$$

If the seller is not strategic or does not have any private information, the mechanism designer can set $q(\tau,\omega)$ equal to $q^{e*}(\tau,\omega)$, to maximize the social welfare $\mathcal{S}$. Then, the designer can set payment $t(\tau,\omega)$ so as to achieve any arbitrary surplus distribution between the buyer and the seller.

However, the strategic seller does not simply reveal his production cost function, which is his private information. Therefore, the socially efficient allocation (3.4) is not sinply implementable by the mechanism designer. The seller has incentives to manipulate the outcome (misrepresent his cost function, by misreporting τ and ω) in order to gain a higher revenue. Thus, the mechanism's output would differ from the efficient allocation q^{e*}.

Due to the seller's strategic behavior, any mechanism must (i) incentivize the seller to truthfully reveal his private information, and (ii) leave a non-negative revenue to the seller so that he voluntarily participates in the mechanism.

Following the literature on regulation and market design [68], we assume that the mechanism designer has the following general objective,

$$\mathcal{W} := \mathcal{U} + \alpha\mathcal{R}, \tag{3.5}$$

where $\alpha \in [0,1]$. When $\alpha = 1$, the designer wants to maximize the social welfare $\mathcal{S}$. When $\alpha = 0$, the designer seeks to maximize the utility of the buyer (demand side). For $\alpha \in (0,1)$, the designer's objective is to maximize a weighted sum of the buyer's utility $\mathcal{U}$ and the seller's revenue $\mathcal{R}$. We assume that the designer knows the buyer's utility function $\mathcal{V}(\hat{q})$.[2]

We invoke the revelation principle for multistage games [87], and, without loss of generality, restrict attention to direct revelation mechanisms that are incentive compatible. In these mechanisms, the designer determines a mechanism for the payment and the allocation $\{t(\tau,\omega), q(\tau,\omega), \tau \in \mathcal{T}, \omega \in [\underline{\omega}, \overline{\omega}]\}$ based on the seller's technology τ and the wind speed ω, and asks the seller to report his private information about τ and ω over time. The designer determines $\{t(\tau,\omega), q(\tau,\omega), \tau \in \mathcal{T}, \omega \in [\underline{\omega}, \overline{\omega}]\}$ so as to ensure the truthful report of the seller; this is called *incentive compatibility* (IC). A mechanism $\{t(\tau,\omega), q(\tau,\omega), \tau \in \mathcal{T}, \omega \in [\underline{\omega}, \overline{\omega}]\}$ is incentive compatible if it is always optimal for the seller to report truthfully his private information. The seller voluntarily participates in the mechanism if he earns a positive (expected) revenue from the agreement; this

[2] In practice, even when the designer's objective is to maximize the social welfare, we have $\alpha < 1$ due to cost/loss associated with financial transaction between the buyer and the seller (see [68] for more discussion). We note that the case where $\alpha > 1$ is of no interest, as it implies that simply the money flow from the buyer to the seller increases the designer's objective.

is known as *individual rationality* (IR).

Define $\mathcal{R}_{\tau,\omega} := t(\tau,\omega) - C(q(\tau,\omega);\Theta(\tau,\omega))$ as the seller's revenue with technology τ and wind ω, and $\mathcal{R}_\tau := \mathbb{E}_\tau\{\mathcal{R}_{\tau,\omega}\}$ as the seller's expected revenue with technology τ. Then, $\mathcal{R} = \mathbb{E}_\tau\{\mathcal{R}_\tau\} = \mathbb{E}_{\tau,\omega}\{\mathcal{R}_{\tau,\omega}\}$. In the next section, we show that the IC and IR constraints can be written in terms of $\mathcal{R}_{\tau,\omega}$, and $\mathcal{R}_\tau$. We show that any mechanism design problem can be formulated as a constrained functional optimization problem, where we determine the optimal allocation and the seller's revenue that maximize $\mathcal{W}$ subject to the IC and IR constraints.

3.4 Market Mechanisms

We consider different structures and timings of mechanisms, and formulate the resulting mechanism design problems accordingly. We consider a forward mechanism that takes place at $T = 1$. We also consider a real-time mechanism that takes place at $T = 2$. Moreover, we propose a dynamic mechanism that takes place at $T = 1$ and $T = 2$. Therefore, we formulate three mechanism design problems for the model of Section 3.2: (A) real-time mechanism, (B) forward mechanism, (C) dynamic mechanism. By comparing the solutions to these mechanism design problems, we demonstrate the advantage of the dynamic mechanism over the static forward and real-time mechanisms.

We show that when the objective of the designer is to maximize $\mathcal{W}$, given by (3.5), each of the mechanism design problems mentioned above can be formulated as a functional optimization problem with different sets of constraints. In Section 3.5, we determine the restrictions that each of these sets of constraints implies on the outcome of the mechanism design problems. Thus, we demonstrate how each of the three market structures impact the market outcome.

3.4.1 Forward Mechanism

In the forward mechanism, the designer specifies the mechanism that determines the agreement between the buyer and the seller at $T = 1$; the information about wind speed ω that becomes available at $T = 2$ is ignored. This mechanism is similar to the current day-ahead integration in Europe, and the proposed firm contracts in the literature.

Since information ω is not available at $T=1$, the allocation function $q(\tau,\omega)$ and the payment function $t(\tau,\omega)$ are independent of ω. Therefore, we drop the dependence on ω and denote the allocation and payment functions for the forward mechanism by $q(\tau)$ and $t(\tau)$, respectively. The optimal forward mechanism is then given by the solution to the following optimization problem:

$$\max_{q(\cdot),t(\cdot)} \mathcal{W}$$

subject to

$$IC\colon \mathcal{R}_\tau \geq t(\hat{\tau}) - \mathbb{E}_\omega\{C(q(\hat{\tau});\Theta(\tau,\omega))\} \quad \forall \tau, \hat{\tau}, \tag{3.6}$$

$$IR\colon \mathcal{R}_\tau \geq 0 \quad \forall \tau. \tag{3.7}$$

We note that (3.7) only ensures a positive expected revenue for the seller with technology τ, and, therefore, exposes him to the risk of penalty (*i.e.* $\mathcal{R}_{\tau,\omega} < 0$) for low realizations of ω.[3]

3.4.2 Real-Time Mechanism

In the real-time mechanism, the designer specifies the mechanism that determines the agreement between the buyer and the seller at $T=2$, after the information about wind speed ω is available to the seller. We assume that the wind speed ω is not monitored by the designer. This mechanism is similar to the current real-time market integration in the U.S. The allocation and payment functions $\{q(\tau,\omega), t(\tau,\omega)\}$ depend on (τ,ω), and the seller reports/reveals his technology τ and his private knowledge about ω simultaneously at $T=2$. The optimal real-time mechanism is then given by

[3]For simplicity, we do not model forward mechanisms with explicit penalty rate for shortfalls here. In Section 3.6, we discuss how our results continue to hold when we consider forward market mechanisms with explicit penalty for shortfalls.

the solution to the following optimization problem:

$$\max_{q(\cdot,\cdot),t(\cdot,\cdot)} \mathcal{W}$$

subject to

$$IC\colon \mathcal{R}_{\tau,\omega} \geq t(\hat{\tau},\hat{\omega}) - C(q(\hat{\tau},\hat{\omega});\Theta(\tau,\omega)) \quad \forall \tau,\omega,\hat{\tau},\hat{\omega}, \tag{3.8}$$

$$IR\colon \mathcal{R}_{\tau,\omega} \geq 0 \quad \forall \tau,\omega. \tag{3.9}$$

Equation (3.8) ensures that the seller's revenue by reporting the true value of (τ,ω) is higher than the one with any misreport $(\hat{\tau},\hat{\omega})$. Equation (3.9) guarantees a positive revenue for the seller for all wind realizations, thus, ensuring no penalty risk for him.

3.4.3 Dynamic Mechanism

In the dynamic mechanism, the designer specifies the mechanism that determines the agreement between the buyer and the seller at $T=1$. However, unlike the forward mechanism, the dynamic mechanism determines an agreement that is contingent on the information about the wind speed ω that becomes available at $T=2$. Moreover, unlike the real-time mechanism, the seller reports τ and ω sequentially; at $T=1$, he reports τ, then at $T=2$, he reports ω. We assume that the wind speed ω is not monitored by the designer. Therefore, the optimal dynamic mechanism is given by the solution to the following optimization problem:

$$\max_{q(\cdot,\cdot),t(\cdot,\cdot)} \mathcal{W}$$

subject to

$$IC_1\colon \mathcal{R}_{\tau} \geq \mathbb{E}_{\omega}\{t(\hat{\tau},\sigma(\omega)) - C(q(\hat{\tau},\sigma(\omega));\Theta(\tau,\omega))\} \ \forall \tau,\hat{\tau},\sigma(\cdot), \tag{3.10}$$

$$IC_2\colon \mathcal{R}_{\tau,\omega} \geq t(\tau,\hat{\omega}) - C(q(\tau,\hat{\omega});\Theta(\tau,\omega)) \quad \forall \tau,\omega,\hat{\omega}, \tag{3.11}$$

$$IR\colon \mathcal{R}_{\tau} \geq 0 \quad \forall \tau. \tag{3.12}$$

The above dynamic mechanism design problem involves two IC constraints (3.10) and (3.11). By (3.10), the designer ensures the seller's true report of τ, even when he can potentially coordinate his misreport $\hat{\tau}$ at $T=1$ with an arbitrary (mis)report strategy $\sigma(\omega)$ of ω at $T=2$. The designer ensures the seller's truthful report of ω at $T=2$ by (3.11), assuming that the seller already reported the true τ at time $T=1$.

3.5 Comparison of Market Mechanisms

In all three mechanism design problems formulated in Section 3.4, the designer wants to maximize $\mathcal{W} = \mathcal{U} + \alpha\mathcal{R} = \mathcal{S} + (\alpha - 1)\mathcal{R}$. However, in each problem the designer faces a different set of constraints due to the seller's strategic behavior and private information about his cost function, as well as the specific mechanism structure and rules. In this section, we analyze these sets of constraints so as to compare the performance of the three mechanism design problems presented in Section 3.4.

We proceed as follows. We consider the set of constraints that the designer needs to satisfy for each market mechanism: constraints (3.6)and (3.7) for the forward mechanism, constraints (3.8) and (3.9) for the real-time mechanism, and constraints (3.10)-(3.12) for the dynamic mechanism. We investigate the implications of each of these sets of constraints under the corresponding market structure. We provide a reduced form of these constraints in Theorems 3.1 and 3.2. Using the results of Theorems 3.1 and 3.2, we show that the set of constraints in the dynamic mechanism is less restrictive than the set of constraints in the forward and real-time mechanisms. Therefore, we show that the outcome of the dynamic market mechanism outperforms those of the forward and real-time mechanisms (Theorem 3.3).

We start our analysis by determining the condition that $\mathcal{R}_{\tau,\omega}$ must satisfy so as to ensure the seller's truthful report about ω in the real-time and dynamic mechanisms.

Theorem 3.1. *The real-time and dynamic mechanisms are incentive compatible for ω (constraints (3.8) and (3.11), respectively), if the allocation function $q(\tau,\omega)$ is increasing in ω, and the seller's revenue $\mathcal{R}_{\tau,\omega}$ satisfies*

$$\frac{\partial \mathcal{R}_{\tau,\omega}}{\partial \omega} = C_\theta(q(\tau,\omega); \Theta(\tau,\omega))\Theta_\omega(\tau,\omega) \geq 0, \tag{3.13}$$

where $C_\theta(q;\theta) := \frac{\partial C(q;\theta)}{\partial \theta}$ and $\Theta_\omega(\tau,\omega) := \frac{\partial \Theta(\tau,\omega)}{\partial \omega}$. The inequality (3.13) is strict if

$q(\tau, \omega) > 0$.

Theorem 3.1 states that, under the assumption that the seller reports truthfully his technology τ, his revenue $\mathcal{R}_{\tau,\omega}$ must be increasing in ω, as in (3.13), so that he is incentivized to report ω truthfully.

We next provide conditions on $\mathcal{R}_\tau = \mathbb{E}_\omega\{\mathcal{R}_{\tau,\omega}\}$ so as to ensure that the seller with technology τ_i reports truthfully in the forward, real-time, and dynamic mechanisms (constraints (3.6), (3.8), and (3.10) resp.). Moreover, we determine the conditions that $\mathcal{R}_\tau$ must satisfy so as to ensure the seller's voluntary participation in the forward, real-time, and dynamic mechanisms (constraints (3.7), (3.9), and (3.12) resp.).

For that matter, we define,

$$\mathcal{R}^T(\tau_j,\tau_i;q) := \int (C(q(\tau_j,\omega);\Theta(\tau_j,\omega)) - C(q(\tau_j,\omega);\Theta(\tau_i,\omega)))dG(\omega),$$

$$\mathcal{R}^W(\tau_j,\tau_i;q) := \iint_\omega^{\sigma^*(\tau_j;\tau_i,\omega)} (C_\theta(q(\tau_j,\omega);\Theta(\tau_j,\hat{\omega})) - C_\theta(q(\tau_j,\hat{\omega});\Theta(\tau_j,\hat{\omega})))$$
$$\Theta_\omega(\tau_j,\hat{\omega})d\hat{\omega}dG(\omega),$$

where $\sigma^*(\tau_j;\tau_i,\omega)$ is uniquely defined by $\Theta(\tau_j,\omega) = \Theta(\tau_i, \sigma^*(\tau_j;\tau_i,\omega))$[4]. That is, $\sigma^*(\tau_j;\tau_i,\omega)$ denotes the wind speed that the seller with technology τ_j requires so as to have a generation cost identical to that of the seller with technology τ_i and wind speed ω. Below, we elaborate on the role of $\mathcal{R}^T$ and $\mathcal{R}^W$.

Consider a seller with technology τ_i and wind speed ω. The seller can misreport his technology (say, by declaring τ_j), or his wind speed (say, by declaring $\hat{\omega}$), or both (by declaring $\tau_j,\hat{\omega}$). In Theorem 3.2, below, we prove the following result. Under the assumption that the seller reports truthfully his wind speed ω, the payment that incentivizes him to report truthfully τ_i instead of misreporting τ_j is given by $\mathcal{R}^T(\tau_j,\tau_i;q)$. The seller may also misreport his wind speed ω (say, by declaring $\hat{\omega}$) after misreporting his technology τ_i as τ_j. In this case, $\mathcal{R}^W(\tau_j,\tau_i;q)$ represents the additional expected payment that incentivizes the seller not to misreport his technology τ_i as

[4] This is true because of Assumption 3.2.

τ_j even when he can misreport his true wind speed. Consequently, $\mathcal{R}^T(\tau_j,\tau_i;q) + \mathcal{R}^W(\tau_j,\tau_i;q)$ represents the incentive payment that the designer must provide to the seller with technology τ_i so as to incentivize him not to misreport his technology as τ_j.

We also define,

$$\mathcal{R}^P(\tau_i;q) := \int [1-G(\omega)]\, C_\theta(q(\tau_i,\omega);\Theta(\tau_i,\omega))\Theta_\omega(\tau_i,\omega)d\omega,$$

which results from integrating (3.13), followed by an expectation with respect to ω. We show below that $\mathcal{R}^P(\tau_1;q)$ determines the minimum incentive payment that the seller with technology τ_1 (the worse technology) must receive in order to voluntarily participate in the real-time mechanism. The precise statement of these results is as follows.

Theorem 3.2. *In the mechanism design problems formulated in Section 3.4, the incentive compatibility and individual rationality constraints can be reduced to the following conditions.*

a) For the forward mechanism,

$$q(\tau,\omega) \text{ is independent of } \omega, \tag{3.14}$$

$$\mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_{i-1}} = \mathcal{R}^T(\tau_i,\tau_{i-1};q) \geq 0 \quad \forall i \in \{2,..,M\}, \tag{3.15}$$

$$\mathcal{R}_{\tau_1} = 0. \tag{3.16}$$

b) For the real-time mechanism,

$$q(\tau,\omega) \text{ is only dependent on } \Theta(\tau,\omega) \text{ and increasing in } \omega, \tag{3.17}$$

$$\frac{\partial \mathcal{R}_{\tau,\omega}}{\partial \omega} = C_\theta(q(\tau,\omega);\Theta(\tau,\omega))\Theta_\omega(\tau,\omega) \geq 0, \tag{3.18}$$

$$\mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_{i-1}} = \mathcal{R}^T(\tau_i,\tau_{i-1};q) + \mathcal{R}^W(\tau_i,\tau_{i-1};q) \geq 0 \quad \forall i \in \{2,\ldots,M\}, \tag{3.19}$$

$$\mathcal{R}_{\tau_1} = \mathcal{R}^P(\tau_1;q). \tag{3.20}$$

c) For the dynamic mechanism,

$$q(\tau,\omega) \text{ is increasing in } \omega, \tag{3.21}$$

$$\frac{\partial \mathcal{R}_{\tau,\omega}}{\partial \omega} = C_\theta(q(\tau,\omega);\Theta(\tau,\omega))\Theta_\omega(\tau,\omega) \geq 0, \tag{3.22}$$

$$\mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_j} \geq \mathcal{R}^T(\tau_j,\tau_i;q) + \mathcal{R}^W(\tau_j,\tau_i;q) \geq 0, \quad \forall i,j \in \{1,\ldots,M\}, i > j, \tag{3.23}$$

$$\mathcal{R}_{\tau_1} = 0. \tag{3.24}$$

Moreover, $R^T(\tau_j,\tau_i;q) \geq 0$ (strict if $q(\tau_j,\omega) \neq 0$ for some ω) and $R^W(\tau_j,\tau_i;q) \geq 0$ (strict if $q(\tau,\omega)$ is dependent on ω).

We now comment on the results of Theorem 3.2. The results of parts (a-c) reduce the set of constraints (3.6,3.7) for the forward mechanism, constraints (3.8,3.9) for the real-time mechanism, and constraints (3.10-3.12) for the dynamic mechanism to those given by (3.14-3.16), (3.17-3.20), and (3.21-3.24), respectively.

Part (a) of Theorem 3.2 follows from the fact that the forward mechanism takes place at $T=1$ when information about ω in not available. Therefore, the allocation function is independent of ω, and, thus, $\mathcal{R}^W(\tau_i,\tau_j;q)=0$ by its definition. Therefore, the designer only needs to provide the payment $\mathcal{R}^T(\tau_i,\tau_j;q)$ so as to incentivize the seller to report truthfully his technology τ_i instead of misreporting τ_j. Part (a) of Theorem 3.2 states further that when the seller's technology is τ_i, the payment $\mathcal{R}^T(\tau_{i-1},\tau_i;q)$ is enough to incentivize the seller not to misreport τ_i as τ_{i-1} or as any other technology τ_j.

In part (b) of Theorem 3.2, constraint (3.17) states that the allocation function $q(\tau,\omega)$ must be only a function of $\Theta(\tau,\omega)$, as the seller reports both of τ and ω simultaneously at $T=2$. Therefore, the designer cannot differentiate between different pairs of (τ,ω) that correspond to the same cost function $C(q;\Theta(\tau,\omega))$. Moreover, (3.17) requires the allocation function to be increasing in ω, which, along with constraint (3.18), ensures the seller's truth telling about ω (Theorem 3.1). Constraint (3.19) states that when the seller's technology is τ_i, the payment $\mathcal{R}^T(\tau_{i-1},\tau_i;q)+\mathcal{R}^W(\tau_{i-1},\tau_i;q)$ is enough to incentivize the seller not to misreport τ_i as τ_{i-1} or as any other technology τ_j. We note that, unlike the forward mechanism where the seller with technology τ_1 receives no incentive payment, in the real-time mechanism the seller with technology τ_1 receives a positive expected incentive payment $\mathcal{R}^P(\tau_1;q)$, given by (3.20), that ensures truth telling about ω.

In part (c) of Theorem 3.2, constraints (3.21) and (3.22) ensure the seller's truth telling about ω by Theorem 3.1. Equation (3.23) determines the incentive payment that the designer needs to provide to the seller with technology τ_i, so that he does not misreport his technology as τ_j, and does not misreport his wind speed ω along with τ_j. We note that, unlike the real-time market, the seller with technology τ_1 does not receive a positive expected incentive payment (see (3.24)).

From the designer's point of view, the dynamic mechanism has the following advantages over the forward and real-time mechanisms: (i) In contrast to the forward mechanism, the dynamic mechanism incorporates the information about the wind speed ω that becomes available at $T=2$ into the allocation and payment functions. As a result, the set of allocation and payment functions available to the designer in the dynamic mechanism is richer than the ones available in the forward mechanism. (ii) In the real-time mechanism, the seller reports τ and ω simultaneously. Therefore, he can perfectly coordinate his reports about τ and ω. In the dynamic mechanism, the seller reports τ and ω sequentially over time, thus, he cannot perfectly coordinate his reports about τ and ω. As a result, in the dynamic mechanism, the designer can distinguish among different pairs (τ,ω) that result in the same cost function $C(q;\Theta(\tau,\omega))$; this is not the case in the real-time mechanism (see (3.17) and (3.21)). Furthermore, in the dynamic mechanism the designer faces a less restrictive set of constraints on the seller's revenue $\mathcal{R}$ than the ones in the real-time mechanism (see (3.20) and (3.24)).

Using the result of Theorem 3.2, we can determine the adequate incentive payment to the seller that is associated with an allocation function $q(\tau,\omega)$ for the forward, real-time, and dynamic mechanisms. Consequently, we can omit the payment function $t(\tau,\omega)$ and optimize over the allocation function $q(\tau,\omega)$ to determine the optimal forward, real-time, and dynamic mechanisms.[5] Under a set of regularity conditions, the resulting functional optimization problems, which are in terms of the allocation function $q(\tau,\omega)$, can be solved point-wise separately for every pair (τ,ω) as a value optimization problem. The closed form solutions of the optimal forward, real-time, and dynamic mechanisms can be found in Appendix B.

Using the result of Theorem 3.2, we can compare the sets of constraints that the designer faces under the forward, real-time, and forward mechanisms. We prove below that for any arbitrary designer's objective of the form (3.5), the outcome is superior under the optimal dynamic mechanism than under the optimal forward or optimal real-time mechanisms.

Theorem 3.3. *The designer's objective under the optimal dynamic mechanism is higher than her objectives under the optimal forward and the real-time mechanisms, i.e.* $\mathcal{W}^{dynamic} > \mathcal{W}^{forward}$ *and* $\mathcal{W}^{dynamic} > \mathcal{W}^{real-time}$.

The result of Theorem 3.3 demonstrates the advantage of the dynamic mechanism over the forward and real-time mechanisms. We note that Theorem 3.3 does not provide a comparison between the forward and real-time mechanisms. That is because the designer faces different sets of constraints in the forward and the real-time mechanisms. On one hand, the forward mechanism ignores ω in its allocation function, while the real-time mechanism incorporates ω. On the other hand, the incentive payments are higher in the real-time mechanism than in the forward mechanism (see

[5]For dynamic mechanisms, the set of constraints (3.23) are in the form of inequality constraints, and it is not possible to determine a priori which of these inequality constraints are binding at the optimal solution (see [12] for more discussion). Therefore, we need to make assumptions about which subset of these inequality constraints are binding, and omit the payment function using the assumed binding conditions. We then need to verify that the set of assumed binding constraints are in fact binding at the computed optimal dynamic mechanism.

(3.20) and (3.16)) since the seller can perfectly coordinate his simultaneous reports about τ and ω in the real-time mechanism. The impact of these constraints on the performance of the forward and the real-time mechanisms depend on the specific parameters of the model, and thus, there is no generic ordering of the designer's objective under the forward and real-time mechanisms.

3.6 Additional Remarks

In this section, we examine the following variations of the problems formulated in Section 3.4, and analyzed in Section 3.5. First, we consider a forward mechanism with explicit penalty rate for shortfalls, and compare it with the dynamic mechanism proposed in Section 3.4. Second, we consider a dynamic mechanism that guarantees no loss for the seller for every realization of wind speed ω. We show that the performance of this mechanism is superior to that of the real-time mechanism and inferior to that of the dynamic mechanism proposed in Section 3.4. Third, we consider a dynamic mechanism with monitoring, where the designer monitors the wind speed ω. We show that the dynamic mechanism with monitoring outperforms the dynamic mechanism proposed in Section 3.4, where the designer does not monitor the wind speed ω. Moreover, in the dynamic mechanism with monitoring, the designer can guarantee no loss for the seller without compromising the performance of the mechanism.

3.6.1 Forward Mechanism with Penalty Rate

The forward mechanism formulated in Section 3.4 ignores the wind speed ω, and determines the allocation $q(\tau)$ and payment function $t(\tau)$ only as a function of τ. Here, we consider a variation of the forward mechanism where the seller is not bound to produce exactly $q(\tau)$ at $T = 2$. However, if the seller's generation at $T = 2$ falls short of his commitment $q(\tau)$, he faces a penalty charge rate $\lambda(\tau)$ by the designer for each unit of generation that he falls short of producing. The penalty rate $\lambda(\tau)$ is agreed at $T = 1$ based on the seller's report about his technology. We refer to this variation of the forward mechanism as *forward mechanism with penalty rate*. The work of [115] studies the design of such a mechanism for wind procurement, and the works of [20, 139] consider settings where λ is exogenously fixed and does not depend on τ.

We note that the forward mechanism with penalty rate does not fully ignore ω at $T = 2$, as it allows the seller to change his generation at $T = 2$ based on the realized wind condition ω. However, we argue below that such incorporation of ω into the forward mechanism with penalty rate is not efficient. Therefore, the dynamic mechanism outperforms the forward mechanism with penalty rate.

A forward mechanism with penalty rate that is incentive compatible and individually rational can be characterized by allocation and payment functions $\{q(\tau), t(\tau), \tau \in \mathcal{T}\}$, along with associated penalty rates $\{\lambda(\tau), \tau \in \mathcal{T}\}$. Let $e(\tau,\omega)$, $e(\tau,\omega) \leq q(\tau)$, denote the amount of energy that the seller with technology τ and wind speed ω actually produces at $T = 2$ (given the described forward mechanism with penalty rate). The seller with technology τ and wind ω chooses $e(\tau,\omega)$ so as to maximize his revenue as,

$$e(\tau,\omega) := \underset{0\leq \hat{e}\leq q(\tau)}{\arg\max} \left\{ t(\tau) - C(\hat{e};\Theta(\tau,\omega)) - \lambda(\tau)(q(\tau) - \hat{e}) \right\}.$$

Using $e(\tau,\omega)$, we can define a dynamic mechanism $\{\tilde{q}(\tau,\omega), \tilde{t}(\tau,\omega), \tau \in \mathcal{T}, \omega \in [\underline{\omega}, \overline{\omega}]\}$ that is equivalent to the forward mechanism with penalty rate described above. Define,

$$\tilde{q}(\tau,\omega) := e(\tau,\omega), \tag{3.25}$$

$$\tilde{t}(\tau,\omega) := t(\tau) - \lambda(\tau)(q(\tau) - e(\tau,\omega)). \tag{3.26}$$

The dynamic mechanism $\{\tilde{q}(\tau,\omega), \tilde{t}(\tau,\omega), \tau \in \mathcal{T}, \omega \in [\underline{\omega}, \overline{\omega}]\}$ defined above is incentive compatible and individually rational since it induces the same allocation and payment function for the seller with technology τ and wind ω as the forward mechanism with penalty rate.

In the equivalent dynamic mechanism $\{\tilde{q}(\tau,\omega), \tilde{t}(\tau,\omega), \tau \in \mathcal{T}, \omega \in [\underline{\omega}, \overline{\omega}]\}$ constructed above, the payment function $\tilde{t}(\tau,\omega)$ is linear in $\tilde{q}(\tau,\omega)$ (see (3.26)). However, in the optimal dynamic mechanism, the payment $t(\tau,\omega)$ is in general, nonlinear in $q(\tau,\omega)$

(see the example in Section 3.7). Therefore, even though the forward mechanism with penalty rate allows the seller to modify his generation according to the realized wind speed ω, such incorporation of ω is not as efficient as in the dynamic mechanism.

3.6.2 Dynamic Mechanism with no Penalty

The dynamic mechanism formulated in Section 3.4 promises a positive expected revenue to the seller for every technology τ (see the IR constraint (3.12)). However, once the seller signs the agreement at $T=1$, he is committed to following the terms of the agreement. Similar to the forward mechanism, it is possible that in the dynamic mechanism, the realized revenue of the seller becomes negative for very low realizations of wind speed, *i.e.* $\mathcal{R}_{\tau,\omega} < 0$ for small values of ω (note that the IR constraint (3.12) only guarantees $\mathcal{R}_{\tau} = \mathbb{E}\{\mathcal{R}_{\tau,\omega}\} \geq 0$).

In this section, we consider a variation of the dynamic mechanism that guarantees no loss for the seller for every realization of ω, *i.e.* $\mathcal{R}_{\tau,\omega} \geq 0$. We refer to this mechanism as *the dynamic mechanism with no penalty.* The optimal dynamic mechanism with no penalty is given by the solution to the following optimization problem:

$$\max_{q(\cdot,\cdot),t(\cdot,\cdot)} \mathcal{W}$$

subject to

$$IC_1: \mathcal{R}_{\tau} \geq \mathbb{E}_{\omega}\{t(\hat{\tau},\sigma(\omega)) - C(q(\hat{\tau},\sigma(\omega));\Theta(\tau,\omega))\} \ \forall \tau,\hat{\tau},\sigma(\cdot), \tag{3.27}$$

$$IC_2: \mathcal{R}_{\tau,\omega} \geq t(\tau,\hat{\omega}) - C(q(\tau,\hat{\omega});\Theta(\tau,\omega)) \quad \forall \tau,\omega,\hat{\omega}, \tag{3.28}$$

$$IR: \mathcal{R}_{\tau,\omega} \geq 0 \quad \forall \tau,\omega. \tag{3.29}$$

The above optimization problem is similar to the optimization problem formulated for the optimal dynamic mechanism in Section 3.4. The only difference is that we replace the *ex-ante* IR constraint $\mathcal{R}_{\tau} \geq 0$, given by (3.12), with the *ex-post* IR constraint $\mathcal{R}_{\tau,\omega} \geq 0$, given by (3.29).

It is clear that the IR constraint $\mathcal{R}_{\tau,\omega} \geq 0$ is more restrictive than the *ex-ante* IR constraint $\mathcal{R}_{\tau} \geq 0$. Therefore, the performance of the optimal dynamic mechanism with no penalty is inferior to that of the optimal dynamic mechanism. Nevertheless, we show below that the optimal dynamic mechanism with no penalty outperforms the optimal real-time mechanism.

Theorem 3.4. *(i) The set of IC and IR constraints for dynamic mechanism with no penalty, given by (3.27-3.29), can be reduced to the following conditions,*

$$q(\tau,\omega) \text{ is increasing in } \omega, \tag{3.30}$$

$$\frac{\partial \mathcal{R}_{\tau,\omega}}{\partial \omega} = C_\theta(q(\tau,\omega);\Theta(\tau,\omega))\Theta_\omega(\tau,\omega) \geq 0, \tag{3.31}$$

$$\mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_j} \geq \mathcal{R}^T(\tau_j,\tau_i;q) + \mathcal{R}^W(\tau_j,\tau_i;q) \geq 0, \quad \forall i,j \in \{1,\ldots,M\}, i>j, \tag{3.32}$$

$$\mathcal{R}_{\tau_1} \geq \mathcal{R}^P(\tau_1;q). \tag{3.33}$$

(ii) The designer's objective, given by (3.5), under the optimal dynamic mechanism with no penalty is higher than her objective under the optimal real-time mechanism and lower than her objective under the optimal dynamic mechanism, i.e. $\mathcal{W}^{dynamic} > \mathcal{W}^{dynamic\ no\ penalty} > \mathcal{W}^{real\text{-}time}$.

Part (i) of Theorem 3.4 states the analogue result of part (c) of Theorem 3.2 for dynamic mechanisms with no penalty. We note that the set of constraints (3.31) and (3.32) for the dynamic mechanisms with no penalty is the same as the set of constraints (3.22) and (3.23) for dynamic mechanisms in part (c) of Theorem 3.2. This is because the IC constraints (3.27) and (3.28) for dynamic mechanisms with no penalty are the same as the IC constraints (3.10) and (3.11) for dynamic mechanisms. However, constraint (3.33) for the dynamic mechanism with no penalty is different from constraint (3.24) for the dynamic mechanism. In the dynamic mechanism with no penalty we impose the ex-post IR constraint (3.29) instead of the less restrictive interim IR constraint (3.12) imposed for the dynamic mechanism. As a result, the designer cannot reduce the incentive payment she pays to the seller by exposing him to penalty risk for low realizations of wind speed ω. The ex-post constraint requires that $\mathcal{R}_{\tau_1,\underline{\omega}} \geq 0$, where τ_1 and $\underline{\omega}$ denote the worst technology and lowest wind speed realization for the seller, respectively. Using constraint (3.31), which is implied by the IC constraint about ω at $T=2$, the expected revenue $\mathbb{E}_{\omega}\{\mathcal{R}_{\tau_1,\omega}\}$ is strictly positive for the seller with technology τ_1, and must be greater than or equal to $\mathcal{R}^P(\tau_1;q)$ (see (3.33)).

Part (ii) of Theorem 3.4 follows directly from the result of part (i). First, the set of constraints (3.31-3.33) for the dynamic mechanism with no penalty is more restrictive than the set of constraints (3.22-3.24) for the dynamic mechanism (of Section 3.4). Therefore, the performance of the optimal dynamic mechanism with no penalty is inferior to that of the optimal dynamic mechanism.

Second, comparing the set of constraints (3.31)-(3.33) for the dynamic mechanism with no penalty with the set of constraints (3.18)-(3.20) for the real-time mechanism (in particular (3.19) and (3.32)), it is clear that the designer's objective under the optimal dynamic mechanism with no penalty is higher than her objective under the optimal real-time mechanism. This is because under the real-time mechanism, the designer cannot distinguish between different pairs of (τ,ω) corresponding to the same $\Theta(\tau,\omega)$ (see (3.17)), whereas, in the dynamic mechanism with no penalty, distinction among all pairs of (τ,ω) is possible as the seller reports τ and ω sequentially.

We note that Theorem 3.4 does not provide a comparison between the forward mechanism and the dynamic mechanisms with no penalty. This is because, on one hand, the forward mechanism ignores ω in its allocation function. On the other hand, in the dynamic mechanism with no penalty we impose the no penalty risk constraint, *i.e.* ex-post IR (3.29). The impact of these two constraints on the performance of the forward mechanism and the dynamic mechanism with no penalty depends on the specific parameters of the model, and thus, there is no generic ordering among them in terms of the designer's objective.

3.6.3 Dynamic Mechanism with Wind Monitoring

In the model of Section 3.2, we use ω to denote the wind speed information that becomes available only to the seller at $T=2$. In this section, we consider a scenario where the designer can also monitor the realization of wind speed ω.

In the following, we formulate the dynamic mechanism design problem under this assumption. We refer to this mechanism as *the dynamic mechanism with monitoring.* Assuming that the designer monitors ω, the seller is only required to reveal his private technology τ at $T=1$. Therefore, the optimal dynamic mechanism with monitoring is given by the solution to the following optimization problem:

$$\max_{q(\cdot,\cdot),t(\cdot,\cdot)} \mathcal{W}$$

$$\begin{aligned}
&\text{subject to}\\
&\quad IC\colon \mathcal{R}_\tau \geq \mathbb{E}_\omega\{t(\hat{\tau},\omega) - C(q(\hat{\tau},\omega);\Theta(\tau,\omega))\}\ \ \forall \tau,\hat{\tau}, &(3.34)\\
&\quad IR\colon \mathcal{R}_\tau \geq 0 \quad \forall \tau. &(3.35)
\end{aligned}$$

The above optimization problem is similar to the optimization problem for the optimal dynamic mechanism in Section 3.4. However, there is no IC constraint for ω in the dynamic mechanism with monitoring, as the designer monitors ω.

Below, we show that wind monitoring provides two advantages to the designer. First, it improves the outcome of the dynamic mechanism. Second, it enables the designer to render the dynamic mechanism with monitoring free of any penalty risk for the seller simply by modifying the payment function for different realizations of ω.

Theorem 3.5. *(i) The set of IC and IR constraints for the dynamic mechanism with monitoring, given by (3.34,3.35) can be reduced to the following conditions,*

$$\mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_j} = \mathcal{R}^T(\tau_j, \tau_i; q) \geq 0,\ \ \forall i,j \in \{1,\ldots,M\}, i>j, \tag{3.36}$$

$$\mathcal{R}_{\tau_1} = 0. \tag{3.37}$$

(ii) The designer's objective under the optimal dynamic mechanism with monitoring is higher than her objective under the optimal dynamic mechanism, the optimal dynamic mechanism with no penalty risk, and the optimal real-time and forward mechanisms. i.e. $\mathcal{W}^{\textit{dynamic with monitoring}} > \mathcal{W}^{\textit{dynamic}} > \mathcal{W}^{\textit{dynamic no penalty}} > \mathcal{W}^{\textit{real-time}}$*, and* $\mathcal{W}^{\textit{dynamic with monitoring}} > \mathcal{W}^{\textit{dynamic}} > \mathcal{W}^{\textit{forward}}$.

	Wind monitoring	No wind monitoring
Penalty risk	$\mathcal{R}^T(q)$	$\mathcal{R}^T(q)+\mathcal{R}^W(q)$
No penalty risk	$\mathcal{R}^T(q)$	$\mathcal{R}^T(q)+\mathcal{R}^W(q)+\mathcal{R}^P(q)$

Table 3.1: Incentive payment to the seller

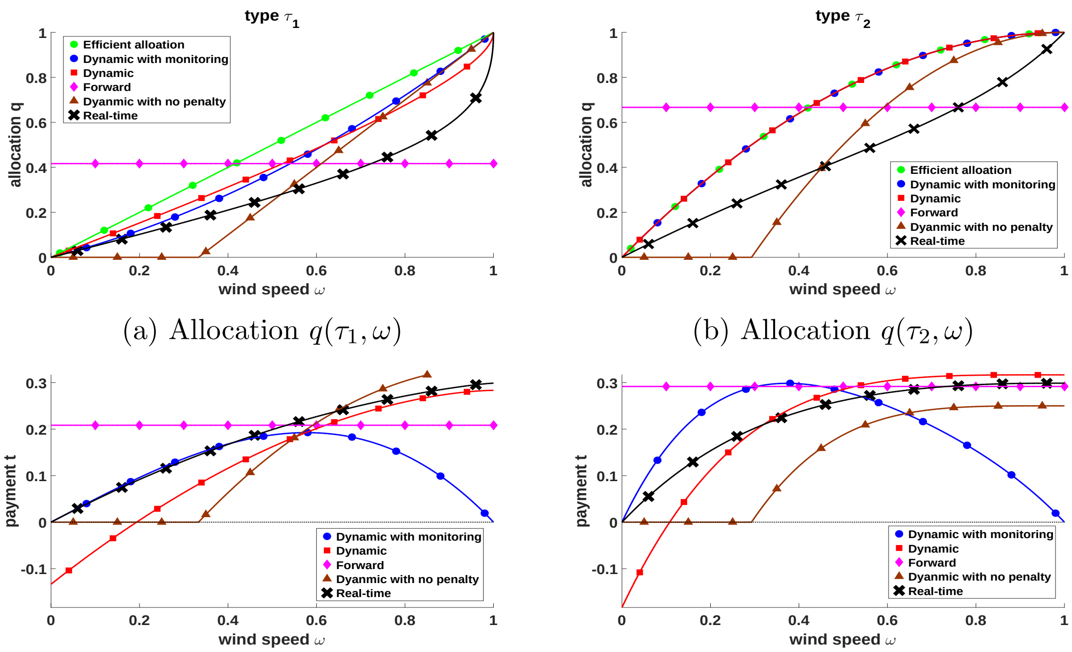
(iii) In the dynamic mechanism with wind monitoring, the designer can guarantee no penalty risk for the seller, without changing the mechanism outcome in terms of her objective $\mathcal{W}$, the buyer's utility $\mathcal{U}$, and the seller's revenue $\mathcal{R}$.

We summarize the results of Theorems 3.2-3.5 on incentive payments to the seller under different market mechanisms in Table 3.1. For a given allocation function $q(\tau,\omega)$, define $\mathcal{R}^T(q):=\sum_{i=2}^M p_i \sum_{j=2}^i \mathcal{R}^T(\tau_j,\tau_{j-1};q)$, $\mathcal{R}^W(q):=\sum_{i=2}^M \sum_{j=2}^i p_i \mathcal{R}^T(\tau_j,\tau_{j-1};q)$, and $\mathcal{R}^P(q):=\mathcal{R}^P(\tau_1;q)$. Then, $\mathcal{R}^T(q)$ denotes the incentive payment that the designer must pay to the seller so that he reveals truthfully his private technology τ. If the designer wants to incorporate ω in energy allocation and does not monitor ω, then she needs to pay the additional incentive payment $\mathcal{R}^W(q)$ so that the seller reveals ω truthfully. Without monitoring ω, if the designer wants to guarantee no penalty risk for the seller, then she must pay an additional incentive payment $\mathcal{R}^P(q)$ to the seller.

3.7 Example

We consider an environment where the designer's objective is $\mathcal{W}=\mathcal{U}+0.5\mathcal{R}$ (*i.e.* $\alpha=0.5$). That is, the designer assigns more weight on the welfare of the demand than on the seller's revenue. The buyer has utility function $\mathcal{V}(\hat{q})=\hat{q}-\frac{1}{2}\hat{q}^2$. The seller has linear cost function $C(\hat{q};\theta)=\theta\hat{q}$. The seller has two possible technologies; technology τ_1 (inferior technology) with marginal cost $\Theta(\tau_1,\omega)=1-\omega$, and technology τ_2 (superior technology) with marginal cost $\Theta(\tau_2,\omega)=(1-\omega)^2$. We assume that the designer believes that both technologies are equally likely and each has probability 0.5. The wind speed ω is uniformly distributed in $[0,1]$.

Using the results of Theorems 3.2-3.5, we compute the optimal forward, real-time, and dynamic mechanisms as well as the optimal dynamic mechanisms with no penalty



(a) Allocation $q(\tau_1, \omega)$

(b) Allocation $q(\tau_2, \omega)$

(c) Payment $t(\tau_1, \omega)$

(d) Payment $t(\tau_2, \omega)$

Figure 3.1: Example - optimal mechanisms

and with monitoring. (see Appendix B for the closed form solution of the general case). Figure 3.1 depicts the allocation and payment functions for each of the optimal mechanisms. The outcome of these different mechanisms is summarized in Table 3.2. Consistent with the results of Theorems 3.3-3.5, we find that $\mathcal{W}^{\text{dynamic with monitoring}} > \mathcal{W}^{\text{dynamic}} > \mathcal{W}^{\text{dynamic no penalty}} > \mathcal{W}^{\text{real-time}}$, and $\mathcal{W}^{\text{dynamic with monitoring}} > \mathcal{W}^{\text{dynamic}} > \mathcal{U}^{\text{forward}}$. We note that, as discussed in Section 3.6.B, there exists no general ordering between the dynamic mechanism with no penalty and the forward mechanism.

As we argued earlier, the seller's strategic behavior affects the efficiency of a mechanism and distorts its outcome from the efficient allocation (see Fig. 3.1). We note that for τ_2, the allocation functions in the dynamic mechanism and dynamic mechanism with monitoring are the same as the efficient allocation. Moreover, for τ_1, the distortion of the allocation function from the efficient allocation is less under the dynamic mechanism than under the real-time and forward mechanisms. These observations further illustrate the advantage of the dynamic mechanism over the

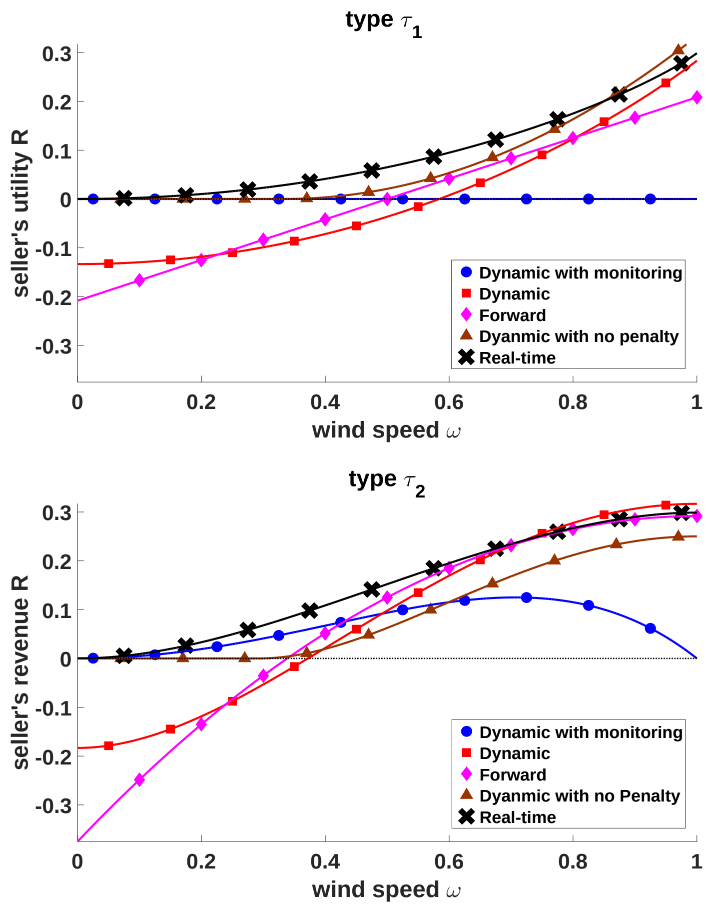


Figure 3.2: Example - the seller's revenue

forward and real-time mechanisms.

Next, we consider the seller's revenue $\mathcal{R}$. We note that the seller's revenue is the highest under the optimal real-time mechanism. This is because (i) the seller is not subject to any penalty risk under the real-time mechanism (as opposed to the forward and dynamic mechanisms); (ii) the seller reports τ and ω simultaneously, therefore, he has more power in manipulating his reports (about τ and ω) to the designer than in the dynamic mechanism, where he reports τ and ω sequentially. We also note that the seller's revenue under the optimal dynamic mechanism is higher than that of the forward mechanism because the forward mechanism ignores wind speed

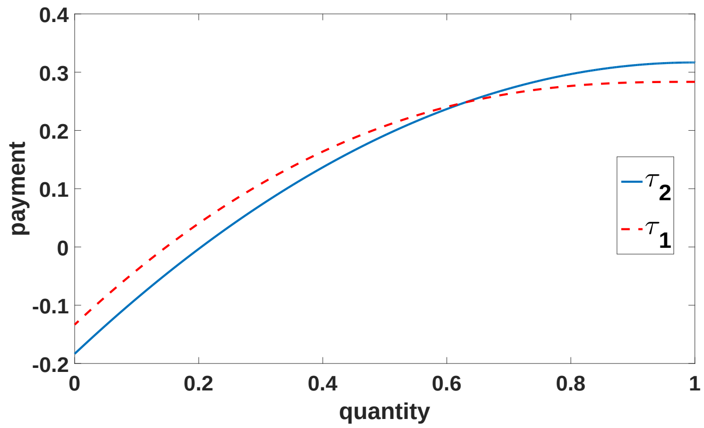


Figure 3.3: Example - different payment vs. quantity options in the dynamic mechanism

ω whereas wind speed ω is incorporated in the dynamic mechanism, and, therefore, the seller receives additional incentive (payment) to report ω truthfully. Next, we note that the seller's revenue under the optimal dynamic mechanism is lower than of the optimal dynamic mechanism with no penalty, and higher than of the optimal dynamic mechanism with monitoring. This is because the payment to the seller is lower in the optimal dynamic mechanism with monitoring and higher in the optimal dynamic mechanism with no penalty (see Table 3.1).

As pointed out above, the dynamic and forward mechanisms expose the seller to penalty risk (negative revenue) for low realizations of the wind speed ω. However, in the dynamic mechanism with monitoring, the dynamic mechanism with no penalty, and the real-time mechanism, the seller always receives a non-negative revenue for all realizations of ω. Figure 3.2 shows the seller's revenue for all the mechanisms and all realizations of ω.

Next, we consider the buyer's utility $\mathcal{U}$. The buyer's utility under the dynamic mechanism is higher than that of the forward and real-time mechanisms. This is because the dynamic mechanism's efficiency is higher than that of the optimal forward and real-time mechanisms, and this is also reflected in the buyer's utility. Moreover, for a given allocation function, the incentive payment made by the buyer to the seller

Mechanisms	Buyer's Utility $\mathcal{U}$	Seller's Revenue $\mathcal{R}$	Designer's Objective $\mathcal{W}$
Efficient allocation	—	—	0.2167
Forward	0.1372	0.0347	0.1545
Real-time	0.0693	0.1210	0.1298
Dynamic	0.1729	0.0417	0.1938
Dynamic with no penalty	0.1077	0.0842	0.1498
Dynamic with monitoring	0.1813	0.0333	0.1979

Table 3.2: Example - the buyer's utility $\mathcal{U}$, the information rent $\mathcal{R}$, and the designer's objective $\mathcal{W}$

under the optimal dynamic mechanism is less than the incentive payments under the optimal forward and real-time mechanisms. The buyer's utility under the optimal dynamic mechanism is higher than that of the optimal dynamic mechanism with no penalty, and lower than that of the optimal dynamic mechanism with monitoring. This is because the payment made by the buyer to the seller is lower in the optimal dynamic mechanism with monitoring and higher in the optimal dynamic mechanism with no penalty (see Table 3.1).

We note that in the forward and real-time mechanisms, which are static mechanisms, the payment the seller receives for producing a certain amount of energy $\hat{q}$ is independent of his type. However, this is not the case in the dynamic mechanism (see Fig. 3.3). In the dynamic mechanism the payment and allocation depend on the seller's reports at $T = 1$ and $T = 2$. Therefore, depending on the seller's reports at $T = 1$ and $T = 2$, the designer can provide different payments to the seller for producing the same amount of energy. In Fig. 3.3, one can see that the seller with the superior technology τ_2 receives a higher payment than the seller with the inferior technology τ_1 for high quantities of energy produced. On the contrary, for low quantities of energy produced the seller with inferior technology τ_1 receives a higher payment than the seller with superior technology τ_2. The difference in payments to different types of seller for the same quantity of energy produced, allows the designer to differentiate among different types of seller, thus, to increase the efficiency of the mechanism.

3.8 Extension to Many Sellers

To demonstrate the main ideas, namely the advantage of the dynamic mechanism over the forward and real-time mechanisms, we considered a setting with only one seller in the model of Section 3.2. However, in a general electricity market there exist many sellers competing with one another. In this section, we discuss how our results on the advantage of the dynamic mechanism over the forward and real-time mechanisms also hold for environments with many sellers. We provide only the sketch of the proof following steps similar to those presented for the model with one seller.

Consider a model similar to that of Section 3.2, with N sellers. Seller n, $n \in \{1,...,N\}$, has generation cost $C(\hat{q}^n;\Theta(\tau^n,\omega^n))$, where $\hat{q}^n$ denotes the amount of energy he produces, and τ^n and ω^n denotes seller n's technology and wind speed, respectively. We assume that τ^n takes values in $\{\tau_1^n,...,\tau_M^n\}$ with probability $(p_1^n...,p_M^n)$. The probability distribution of $\omega^n \in [\underline{\omega},\overline{\omega}]$ is independent of τ_n. The wind speeds $\omega^1,...,\omega^N$ may be correlated as the sellers can be located in geographically close locations. We assume that joint distribution of $(\omega^1,...,\omega^N)$ (resp. $(\tau^1,...,\tau^N)$), is commonly known to all sellers as well as the designer. Furthermore, Assumptions 3.1 and 3.2 of Section 3.2, hold for every seller n, $n \in \{1,...,N\}$.

Similar to the approach presented in Section 3.3, we invoke the revelation principle and restrict attention to direct revelation mechanisms that are incentive compatible and individually rational. Let τ^{-n} (resp. ω^{-n}) denote the set of all sellers' technologies (resp. wind speeds) except seller n's technology τ^n (resp. wind speed ω^n). In the model with many sellers, seller n's allocation and payment are functions of seller n's reports about τ^n and ω^n as well as all the other sellers' reports about τ^{-n} and ω^{-n}. Let $q^n(\tau^n,\omega^n,\tau^{-n},\omega^{-n})$, and $t^n(\tau^n,\omega^n,\tau^{-n},\omega^{-n})$ denote the allocation and payment functions for seller n, respectively.

We assume that the designer does not reveal the reports of seller n to other sellers before finalizing the payments and allocations of all sellers. Define,

$$\bar{C}^n(\hat{\tau}^n,\hat{\omega}^n;\tau^n,\omega^n):=\mathbb{E}_{\tau^{-n},\omega^{-n}}\Big\{C(q^n(\hat{\tau}^n,\hat{\omega}^n,\tau^{-n},\omega^{-n});\Theta(\tau^n,\omega^n))\\ \Big|\tau^n,\omega^n\Big\}, \tag{3.38}$$

$$\bar{t}^n(\hat{\tau}^n,\hat{\omega}^n):=\mathbb{E}_{\tau^{-n},\omega^{-n}}\{t^n(\hat{\tau}^n,\hat{\omega}^n,\tau^{-n},\omega^{-n})\big|\tau^n,\omega^n\}, \tag{3.39}$$

as seller n's conditional expected cost and payment, respectively, when he has technology τ^n and wind ω^n, and he reports $\hat{\tau}^n$ and $\hat{\omega}^n$. We note that, the expectations above are written using the fact that at equilibrium, seller i believes that other sellers report their private information truthfully.

Similar to Section 3.3, let

$$\mathcal{R}^n_{\tau^n,\omega^n}:=\mathbb{E}_{\tau^{-n},\omega^{-n}}\Big\{t^n(\tau^n,\omega^n,\tau^{-n},\omega^{-n})\\ -C\Big(q^n(\tau^n,\omega^n,\tau^{-n},\omega^{-n});\Theta(\tau^n,\omega^n)\Big)\Big\},$$

denote the expected revenue of seller n with technology τ^n and wind ω^n. Also, define,

$$\mathcal{R}^n_{\tau^n}:=\mathbb{E}_{\omega^n}\{\mathcal{R}^n_{\tau^n,\omega^n}\},$$
$$\mathcal{R}^n:=\mathbb{E}_{\tau_n}\{\mathcal{R}^n_{\tau^n}\mathcal{R}^n_{\tau^n}\}.$$

Moreover, let,

$$\mathcal{U}_{\text{many}}:=\mathbb{E}_{\{\tau^n,\omega^n\}_{n=1}^N}\left\{\mathcal{V}\left(\sum_{n=1}^N q^n(\tau^n,\omega^n,\tau^{-n},\omega^{-n})\right)\right.\\ \left.-\sum_{n=1}^N t^n(\tau^n,\omega^n,\tau^{-n},\omega^{-n})\right\},$$
$$\mathcal{W}_{\text{many}}:=\mathcal{U}_{\text{many}}+\alpha\sum_{n=1}^N\mathcal{R}^n,$$

denote the buyer's expected utility and the designer's objective, respectively.

We discuss how the results of Theorems 3.1-3.3 (on the advantage of the dynamic mechanism over the forward and real-time mechanisms) continue to hold in the model with many sellers described above.

The key idea is the following. We write each mechanism design problem with many sellers in a form similar to the corresponding mechanism design problem with one seller (formulated in Section 3.4). We write these mechanism design problems using the conditional expected cost and payment functions, defined by (3.38) and (3.39), respectively, instead of the cost and payment functions that appear in the problems with a single seller. We show that in each mechanism design problem and for each seller, the designer faces a set of constraints that are similar to those that arise in the case with a single seller. Therefore, the arguments used in the proofs of Theorems 3.1-3.3 can be directly extended to environments with many sellers.

We proceed with the formulation of the mechanism design problems with many sellers, and discuss how the results of Theorems 3.1-3.3 continue to hold in the model with many sellers.

Forward mechanism: In the forward mechanism, the allocation function $q^n(\hat{\tau}^n, \hat{\omega}^n, \tau^{-n}, \hat{\omega}^{-n})$ and payment function $t^n(\hat{\tau}^n, \hat{\omega}^n, \tau^{-n}, \hat{\omega}^{-n})$ for seller n are independent of the reported wind speeds $\hat{\omega}^n$ and $\hat{\omega}^{-n}$. Therefore, we drop the dependence on $(\hat{\omega}^n, \hat{\omega}^{-n})$, and denote the conditional expected cost and payment functions by $\bar{C}^n(\hat{\tau}^n; \tau^n, \omega^n)$ and $\bar{t}^n(\hat{\tau}^n)$, respectively. The optimal forward mechanism with many sellers is given by the solution to the following optimization problem:

$$
\begin{aligned}
&\max_{\{q^n(.),t^n(.)\}_{n=1}^N} \mathcal{W}_{\text{many}} \\
&\text{subject to} \\
&IC^n\colon \mathcal{R}^n_{\tau^n} \geq \bar{t}^n(\hat{\tau^n}) - \mathbb{E}_\omega\{\bar{C}^n(\hat{\tau}; \tau, \omega^n)\} \quad \forall \tau^n, \hat{\tau}^n, n && (3.40) \\
&IR^n\colon \mathcal{R}^n_{\tau^n} \geq 0 \quad \forall \tau^n, n. && (3.41)
\end{aligned}
$$

We note that in the above optimization problem, the designer faces a set of constraints, given by (3.40) and (3.41), for every seller n. These constraints are similar to those given by (3.6) and (3.7) for the forward mechanism with a single seller. Therefore, we can show that seller's expected revenue $\mathcal{R}^n_{\tau^n}$ for every seller i must satisfy a set of constraints similar to those given by (3.15) and (3.16) in part (a) of Theorem 3.2 in terms of conditional expected cost $\bar{C}^n(\hat{\tau}^n;\tau^n,\omega^n)$.

Real-time mechanism: The real-time mechanism design problem with many sellers can be formulated as follows.

$$\max_{\{q^n(.),t^n(.)\}_{n=1}^N} \mathcal{W}_{\text{many}}$$

subject to

$$IC^n\colon \mathcal{R}^n_{\tau^n,\omega^n} \geq \bar{t}^n(\hat{\tau}^n,\hat{\omega}^n) - \bar{C}^n(\hat{\tau}^n,\hat{\omega}^n;\tau^n,\omega^n) \quad \forall \tau^n,\omega^n,\hat{\tau}^n,\hat{\omega}^n,n, \tag{3.42}$$

$$IR^n\colon \mathcal{R}^n_{\tau^n,\omega^n} \geq 0 \quad \forall \tau^n,\omega^n,n. \tag{3.43}$$

Note that the set of constraints (3.42) and (3.43), for every seller n, is similar to the set of constraints (3.8) and (3.9) in the real-time mechanism with a single seller. Therefore, we can show that expected revenue $\mathcal{R}^n_{\tau^n,\omega^n}$ must satisfy a set of constraints similar to those given by Theorem 3.1. Consequently, we can show that the set of constraints that the designer faces for every seller n can be reduced to a set of constraints that are similar to those given by (3.18)-(3.20) in part (b) of Theorem 3.2 in terms of expected seller's revenue $\mathcal{R}^n_{\tau^n,\omega^n}$ and $\mathcal{R}^n_{\tau^n}$, and the conditional expected cost $\bar{C}^n(\hat{\tau}^n,\hat{\omega}^n;\tau^n,\omega^n)$. We can also show that $q(\tau^n,\omega^n,\tau^{-n},\omega^{-n})$ must only be a function of $\Theta^n(\tau^n,\omega^n)$ and $\Theta^m(\tau^m,\omega^m)$, $m \neq n$, since all sellers report simultaneously about their own technologies and wind speeds (see (3.17)). Using an argument similar to the one in Theorem 3.2 for seller n, we can also show that $\mathbb{E}_{\tau^{-n},\omega^{-n}}\{q^n(\tau^n,\omega^n,\tau^{-n},\omega^{-n})\}$ must be increasing in ω^n.

Dynamic mechanism: The optimal dynamic mechanism with many sellers is given by the solution to the following optimization problem:

$$\max_{\{q^n(.),t^n(.)\}_{n=1}^N} \mathcal{W}_{\text{many}}$$

$$\begin{aligned}
&\text{subject to}\\
&IC_1^n\!:\mathcal{R}^n_{\tau^n}\geq\mathbb{E}_{\omega^n}\{\bar{t}^n(\hat{\tau}^n,\sigma^n(\omega^n))-\bar{C}^n(\hat{\tau}^n,\sigma^n(\omega^n);\tau^n,\omega^n)\}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\forall\tau^n,\hat{\tau}^n,\sigma^n(\cdot),n, && (3.44)\\
&IC_2^n\!:\mathcal{R}^n_{\tau^n,\omega^n}\geq\bar{t}^n(\tau^n,\hat{\omega}^n)-\bar{C}^n(\tau^n,\hat{\omega}^n;\tau^n,\omega^n)\ \ \forall\tau^n,\omega^n,\hat{\omega}^n,n, && (3.45)\\
&IR^n:\mathcal{R}^n_{\tau^n}\geq 0\quad\forall\tau,n. && (3.46)
\end{aligned}$$

As we mentioned above, we assume that the designer does not reveal other sellers' reports about τ^{-n} at $T=1$. Therefore, seller n's IC constraint about ω at $T=2$ is given by (3.45), where seller n assumes that other sellers report truthfully τ^{-n} and ω^{-n}. We note that the set of constraints (3.44)-(3.46), for every seller n, is similar to the set of constraints (3.10)-(3.12) in the dynamic mechanism with a single seller. Therefore, we can show that the expected seller's revenue $\mathcal{R}^n_{\tau^n,\omega^n}$ must satisfy a set of constraints similar to those given by Theorem 3.1. Furthermore, we can show that the expected seller's revenue $\mathcal{R}^n_{\tau^n,\omega^n}$ and $\mathcal{R}^n_{\tau^n}$ for every seller n must satisfy a set of constraints that are similar to the ones given by (3.22)-(3.24) in part (c) of Theorem 3.2. Using an argument similar to Theorem 3.2 for seller n, we can also show that $\mathbb{E}_{\tau^{-n},\omega^{-n}}\{q^n(\tau^n,\omega^n,\tau^{-n},\omega^{-n})\}$ must be increasing in ω^n as in (3.21) in part (c) of Theorem 3.2.

The above arguments show that all the mechanism design problems with many sellers considered in this section, have sets of constraints for each seller that are similar to those that arise in the corresponding problems with a single seller. Thus, we can show, by arguments similar to those given in the proof of Theorem 3.3, that the set of constraints for the dynamic mechanism with many sellers is less restrictive than the set of constraints for the forward and real-time mechanisms with many sellers. Therefore, we can establish that the dynamic mechanism with many sellers outperforms the forward and real-time mechanism with many sellers.

Remark 3.1. *The optimal forward and real-time mechanism with many sellers are* *standard static multi-unit auctions (see [66]). However, the optimal dynamic mechanism with many sellers is a form of handicap auction, first introduced in [38].*

In a standard auction, seller n bids his generation cost function simultaneously and the auctioneer determines the outcomes based on all sellers' bids.

In a handicap auction, at $T=1$, seller n bids his current information about his generation cost (i.e. τ^n), so as to pick a menu of payment-quantity curve, from which he can select his generation at $T=2$. Then, at $T=2$, seller n observes the realized wind speed ω^n, and competes with other sellers for generation based on the payment-quantity curve he won at $T=1$. Sellers with better technologies anticipate to have higher generations in real-time. Therefore, in a handicap auction, at $T=1$, they bid for payment-quantity curves that give them high marginal prices for high generation quantities (which are more likely for them), but low marginal prices for low generation quantities (which are less likely for them). On the other hand, sellers with worse technologies bid for payment-quantity curves that give them high marginal prices for low generation quantities (which are more likely for them), but low marginal prices for high generation quantities (which are less likely for them); see Fig. 3.3 in the example of Section 3.7.

3.9 Conclusion

We considered a dynamic model for market design for wind energy, and studied the problem of market mechanism design for wind energy. We proposed dynamic market mechanisms that couple the outcome of the real-time and forward mechanisms. We showed that the dynamic market mechanisms proposed in this chapter outperform the real-time and forward mechanisms with respect to the integration of wind energy into the grid. Dynamic mechanisms dynamically incorporate the new information that arrives over time, and require the wind generator to sequentially reveal his private information to the market designer, and to refine his generation commitment accordingly. We investigated the effects of wind monitoring and penalty risk exposure on the market outcome. Generalization of the these results to mechanisms with multiple wind generators is our main direction for future work.

Chapter 4

Informational Incentives in Congestion Games

4.1 Introduction

4.1.1 Background and Motivation

In a transportation network, the condition of every link varies over time due to changes in weather conditions, accidents, traffic jams, etc. Traditionally, drivers receive public information about these changes at every route through various infrastructures, *e.g.* regional traffic updates via radio broadcasts, and/or variable (dynamic) message signs on road sides displaying specific information about the onward routes [11, 29, 36]. In recent years, the advent of GPS-enabled routing devices and navigation applications (*e.g.* Waze and Google maps) has enabled drivers to receive private, real-time data about the transportation network's condition for their own intended origin-destination [10]. The development of these technologies creates new opportunities to reduce congestion in the network, and improve its overall performance (as measured by various metrics including *social welfare*).

Several studies have investigated the effects of information provision to drivers on the social welfare of the transportation network [4, 14, 33, 73, 124, 1, 74, 69]. These studies have shown that the effect of information provision on social welfare is ambiguous, and in general, is not necessarily socially beneficial. For exogenously fixed information provision structures, these works have identified instances where the provision of information can in fact increase congestion in parts of the network, leading to a decrease in social welfare. In addition to the above-mentioned theoretical and experimental works, there are empirical evidences that identify negative impacts of information provision on the network's congestion [52, 94, 103, 45, 26]. Therefore, it is important to investigate how to design appropriate information provision mechanisms, in a manner that is socially beneficial and leads to a reduction in the overall congestion in the transportation network.

In this chapter, we study the problem of designing a socially optimal information disclosure mechanism. We consider a congestion game [83, 104] in a parallel two-link network. We consider an information provider (principal) who wants to disclose information about the condition of the network to a fixed population of drivers (agents). We assume that the condition of one route/link, called safe route, is constant and known to everyone, while the condition of the other route, called risky route, is random and only known to the principal. The principal wants to design an information disclosure mechanism so as to maximize the social welfare.

We study the problem of designing an optimal information provision mechanism in two cases: (1) when the principal can only provide information that is publicly available to all drivers, and (2) when the principal can provide private information to each driver individually.

We first consider a static setting where the drivers do not learn from their past experiences. We determine a condition under which the principal can achieve the maximum social welfare using an optimal information provision mechanism. That is, the principal can utilize her *superior information* about the network to provide *informational incentives* so as to align the drivers' objectives with the overall social welfare. Next, we consider a dynamic two-stage setting where the drivers learn from their experience at $t = 1$, and the risky routes' condition evolves according to an uncontrolled Markov chain. We consider three scenarios for the drivers' learning at $t = 1$: (i) the drivers only learn from the information they receive directly from the principal at $t = 1$, (ii) in addition to the information they receive directly from the principal at $t = 1$, the drivers who take the risky route learn perfectly its condition at $t = 1$, and (iii) in addition to the information he receives directly from the principal at $t = 1$, each driver perfectly observes the number of cars/drivers on the route he takes at $t = 1$. Using numerical simulations, we show that in scenario (i) the principal can achieve the same outcome as in the static setting in which there is no learning. However, in scenarios (ii) and (iii) the performance of an optimal dynamic information provision mechanism decreases due to the drivers' learning only. In particular, we identify instances in scenario (ii) where it is optimal for principal to reveal perfectly risky route's condition at $t = 2$ so that the drivers do not have an incentive to experiment and learn the risky route's condition at $t = 1$. Moreover, in scenario (iii) we identify instances where it is optimal for the designer to not implement different routing outcome, *i.e.* reveal her information about risky route's condition, so as to increase her information superiority at $t = 2$.

4.1.2 Related literature

The problems investigated in [78, 4, 14, 33, 73, 69, 124, 1, 74] are the most closely related to our problem. The authors of [4, 14] consider a *bottleneck* model [128] with stochastic capacities, where each route is modeled as a queue with a first-come, first-served policy, with the service rate determined by the route's condition. In [4], the authors consider a scenario where each driver decides the time and the route/queue he wants to join, considering the behavior of the other drivers. Through numerical simulations, they show that when drivers receive a low quality/highly noisy signal about the routes' conditions, social welfare decreases. In [14], the authors consider a scenario where each driver does not take into account the other drivers' response to the information provided by the principal. They show that social welfare may decrease when the drivers receive accurate information about the routes, due to the drivers' *overreaction* and/or higher *congestion concentration* in parts of the network.

The authors of [33, 73] consider a parallel two-link network that is similar to our model. In [33], the authors assume that only one of the routes has a random condition with two possible values. They show that social welfare may decrease when the drivers receive public information about the routes' conditions irrespective of whether they are risk-neutral or risk-averse. The authors of [73] assume that the condition of both routes are random. They show that social welfare can decrease when the drivers receive public perfect information about the realizations of the routes' conditions.

The authors of [124, 74, 1] consider a model where a subset of drivers (informed drivers) has access to more accurate information about the condition of every route than the remaining drivers. In [124], the authors assume that the drivers that do not receive the more accurate information prefer to use high-capacity routes (*i.e.* highways) rather than low-capacity routes. By numerical simulation, they show that as the number of informed drivers increases, the congestion in low-capacity routes (*i.e.* urban areas) increases. The authors of [74] consider a model similar to ours, in which the condition of one of the routes is random. They show that when the number of informed drivers is low, the expected utilities of both groups of drivers are higher compared to the case where all drivers are uninformed. However, as the number of informed drivers increases, the social welfare decreases, even compared to the case where all drivers are uninformed. The authors of [1] study a model where the informed drivers become aware of the existence of additional routes in a network. They show that the provision of information can create a *Braess' Paradox* phenomenon, and thus, can reduce not only the social welfare, but also the utility of the informed drivers.

In contrast to [78, 4, 14, 33, 73, 69, 124, 1, 74], which analyze the performance of *fixed* information provision mechanisms, in this chapter we investigate the problem of designing an optimal information provision mechanism; the information provision mechanisms analyzed in [78, 4, 14, 33, 73, 69, 124, 1, 74] are within the set of feasible mechanisms the principal can choose from when maximizing the social welfare.

Our work is also related to the literature on improving efficiency in resource allocation problems with externalities (*i.e.* congestion games). It is known that the equilibrium outcome in congestion games is not socially optimal [106]. Several approaches have been proposed in the literature to address this inefficiency. One approach is to utilize monetary mechanisms in order to align the agents' objectives with the social welfare (see [99] and references therein). A second approach, which is applicable when the principal has control over a fraction of agents, is for the principal to choose routes for this fraction so as to influence the behavior of selfish agents. This can lead to improvement in the social welfare (see [65, 112] and references therein). We propose an alternative approach to improving the efficiency in congestion game by utilizing informational incentives when the principal has an informational advantage over the agents. Specifically, our approach can prove promising in transportation networks where the application of tolls (pricing) is limited and agents are typically selfish.

Within the economics literature, the problems studied in this chapter belong to the class of information design problems (see [17] and references therein). Our approach to the public information mechanism design problem (Section 4.4) is similar to the ones in [35, 59]. Our approach to the private information mechanism design problems (Section 4.5 and 4.6) is similar to the ones in [15, 16]. The work in [64] is closely related to our work. The authors of [64] consider a model with two possible actions where the payoff of one of the actions is not known, even to the principal. The principal (*e.g.* the Waze application) faces a group of short-lived agents that arrive sequentially over time. She wants to design an information mechanism that provides information about the agents' past experience to the incoming agents over time. Our work is different from [64] as (i) in contrast to the single-agent decision problem considered in [64], our model assumes that the principal faces a population of agents that create negative externalities on one other at each time step, and (ii) in the dynamic setting, agents are long-lived and learn from their past (private) experience, while in [64], the agents are short-lived.

The dynamic two-stage problems studied in this chapter are also related to literature on strategic experimentation in economics [22, 61, 60]. The authors of [47] study a monetary mechanism design in a principal-agent relationship. Our problem is different from that of [47] since we study an information disclosure mechanism design instead of a monetary mechanism design. The authors of [48, 18] study the problem of information disclosure mechanism design in an innovation contest. Our problem is different from those of [48, 18] since in contrast to their models where agents do not observe each others' actions and payoffs unless the principal discloses information and an agent's payoff is independent of other agents' action, in our model agents' actions create negative externality and influence agents' payoffs, and each agent may have an indirect observation about other agents' actions and payoffs at $t = 1$.

The information design problems that we consider in this chapter are also related to the problem of designing real-time communication systems [77, 123, 131, 136]. However, in contrast to these studies, where the receivers are cooperative and have the same objective as the transmitter, in our problem the drivers are strategic and have objectives that are different from that of the principal. The authors of [2] consider a problem of real-time communication with a strategic transmitter and receiver, Gaussian source, and quadratic estimation cost. They follow an approach that is similar to that of [59] and our approach for the design of public information disclosure mechanism in Section 4.4. However, our problem is different from that of [2] since in our model there exist many agents, and each agent's utility depends on his action and the routes' conditions as well as other agents' actions. Moreover, in this chapter, we study the problem of private information disclosure mechanism design that is not present in the work of [2].

4.1.3 Contribution

We determine optimal public and private information provision mechanisms that maximize the social welfare in a transportation network. Our results propose a solution to the concern raised in [78, 4, 14, 33, 73, 69, 124, 1, 74] about the potential negative impact of information provision on congestion in transportation networks. We show that the principal can utilize his superior information about the condition of the network, and provide informational incentives to the drivers so as to improve the social welfare. When the principal can disclose information to every driver privately, we show that the principal can benefit from providing coordinated routing recommendations to the drivers. We identify a condition under which the principal can achieve the efficient routing outcome in a static setting. Moreover, we consider a dynamic setting with two-time steps under three scenarios, each capturing a possible piece of information that the drivers can learn from it in dynamic setting. Using numerical simulations, we discuss the effect of each piece of information on the solution to the optimal dynamic information mechanism and its qualitative properties.

4.1.4 Organization

The rest of this chapter is organized as follows. In Section 4.2, we present our model in a static setting. In Section 4.3, we consider two naive information mechanisms and compare their outcomes with the socially efficient outcome. We study the problem of designing an optimal public information mechanism in Section 4.4. In Section 4.5, we study the problem of designing an optimal private information mechanism. We consider the design of optimal dynamic information mechanisms in a two-step setting in Section 4.6, and investigate the effect of different types of drivers' observations on the performance and qualitative properties of an optimal dynamic mechanism through numerical simulations. We conclude in Section 4.7. All proofs appear in Appendix C.

4.2 Model

Consider a two-link network managed by a principal who wants to maximize social welfare (Figure 4.1). There is a unit mass of agents traveling from the origin O to the destination D. There are two routes/links that agents can take. The top route, denoted route s (*i.e.* safe route) has condition $a > 0$ that is known to all agents and the principal. The bottom route, denoted route r (*i.e.* risky route), has a condition $\theta \in \Theta := \{\theta^1, \ldots, \theta^M\}$, $\theta^1 < \theta^2 < \ldots < \theta^M$, that is not known to the agents and is only known to the principal. It is common knowledge among the agents and the principal that θ takes values in $\{\theta^1, \ldots, \theta^M\}$ with probability $\{p_{\theta^1}, p_{\theta^2}, ..., p_{\theta^M}\}$, respectively. Let x^s and x^r (where $x^s + x^r = 1$), denote the mass of agents that choose route s and r, respectively. The utility of each agent depends on the condition of the route that he chooses to travel as well as on the congestion (negative externality) that he observes along his route. Given x^s and x^r, let $C^s(x^s)$ and $C^r(x^r)$ denote the congestion cost at route s and route r, respectively. The functions $C^s(\cdot)$ and $C^r(\cdot)$ are strictly increasing, with $C^s(0) = 0$ and $C^r(0) = 0$. For the ease of exposition, we assume that $C^s(x^s) = x^s$ and $C^r(x^r) = x^r$. Throughout this chapter, we discuss how our results extend to general congestion functions.

We assume that the utility of an agent taking route s (resp. r) is given by $a - C^s(x^s) = a - x^s$ (resp. $\theta - C^r(x^r) = \theta - x^r$); that is, the effect of a route's condition on an agent's utility is separable from the effect of the congestion cost. Therefore, the expected social welfare W is given by

$$W := \mathbb{E}\left\{x^s(a - C^s(x^s)) + x^r(\theta - C^r(x^r))\right\}. \tag{4.1}$$

We make the following assumption about the possible values of θ.

Assumption 4.1. *The risky route's types θ are such that $\theta^M - C^r(1) \leq a$ and $a - C^s(1) \leq \theta^1$.*

Assumption 4.1 ensures that for every realization θ of route r's condition, there will be a positive mass of agents taking either route.

The principal wants to design an information disclosure mechanism that provides information about the condition θ of route r to the agents, so as to maximize the expected social welfare W. We consider two classes of information disclosure mechanisms by the principal: (i) public information disclosure mechanisms, where the principal sends a public signal about θ which is observed by all agents (Section 4.4), (ii) private information disclosure mechanisms, where the principal sends a private signal about θ to each agent, and this signal is only observed by that agent (Section 4.5).

Before proceeding with the study of optimal public and private information disclosure mechanisms which maximize social welfare, we present two naive information disclosure mechanisms in Section 4.3. By exploring the agents' routing decisions under these two naive information disclosure mechanisms, along with the socially efficient routing decisions, we will elaborate on the main insights underlying some of the results appearing in the rest of the chapter.

4.3 Naive Mechanisms

We study two naive information disclosure mechanisms that the principal can employ to disclose information about the condition θ of route r, namely, the no in-

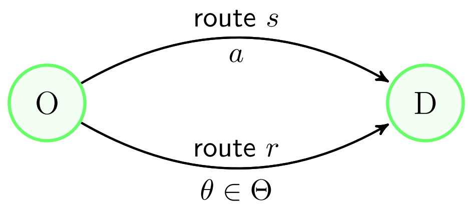


Figure 4.1: The two-link network

formation disclosure and full information disclosure mechanisms. We then present the socially optimal outcome and compare it to the outcomes of the naive mechanisms. Let $\mu := \mathbb{E}\{\theta\} = \sum_{\theta\in\Theta} p_\theta \theta$ denote the expected condition of route r. Define $\Delta := a - \mu$ and $\Delta_\theta := a - \theta$ as the expected and realized difference between the conditions of routes s and r. Let $\sigma^2 = \mathbb{E}\left\{(\theta - \mu)^2\right\}$ denote the variance of route r's condition. In the sequel, we characterize the traffic outcome under different information that the agents may receive as a function of μ, Δ, and Δ_θ.

4.3.1 No Information Disclosure

Consider an information disclosure mechanism where the principal discloses no information about θ to the agents. In this case, the expected utility from route r, given by $\mu - x^r$, must be equal to the utility $a - x^s$ from route s; this is because otherwise, some agents would switch from the route with lower utility to the one with higher utility[1]. Therefore, the traffic at routes s and r are given by

$$x^{s,\text{no info}} = \frac{1}{2} + \frac{1}{2}\Delta, \tag{4.2}$$

$$x^{r,\text{no info}} = \frac{1}{2} - \frac{1}{2}\Delta. \tag{4.3}$$

That is, the difference between the traffic on routes s and r depends on the expected difference between the routes' conditions, given by Δ.

Consequently, the expected social welfare $W^{\text{no info}}$ under the no information dis-

[1]Note that by Assumption 4.1, both routes are non-empty for any realization of the condition of the risky route.

closure mechanism is given by

$$W^{\text{no info}} = \frac{a+\mu-1}{2}, \tag{4.4}$$

where $\frac{a+\mu-1}{2}$ denotes the expected utility of an agent taking either of the routes.

4.3.2 Full Information Disclosure

Consider an information disclosure mechanism where the principal reveals perfectly the condition θ of route r to all agents. In this case, agents choose their route knowing θ. By an argument similar to the one given above for the no information disclosure mechanism, the utility from taking either of the routes must be equal. Therefore, the traffic at routes s and r are given by

$$x^{s,\text{full info}}(\theta) = \frac{1}{2} + \frac{1}{2}\Delta_\theta, \tag{4.5}$$

$$x^{r,\text{full info}}(\theta) = \frac{1}{2} - \frac{1}{2}\Delta_\theta. \tag{4.6}$$

In this case, the traffic difference between routes s and r depends on the *realized* difference Δ_θ between the routes' conditions, as opposed to the *expected* difference Δ that determines the outcome under the no information disclosure mechanism.

Using (4.5) and (4.6), we can obtain the expected social welfare $W^{\text{full info}}$ under the full information disclosure mechanism as

$$W^{\text{full info}} = \mathbb{E}\{\frac{a+\theta-2}{2}\} = \frac{a+\mu-1}{2}. \tag{4.7}$$

Remark 4.1. *We note that the expected social welfare $W^{\text{full info}}$ under the full information disclosure mechanism and $W^{\text{no info}}$ under the no information disclosure mechanism are the same in the model of Section 4.2 with linear congestion costs. This is because under the full information disclosure mechanism the social welfare is linear in θ. As we discuss in Remark 4.2 below, for congestion functions $C^s(x^s)$ and $C^r(x^r)$ that are nonlinear in x^s and x^r, respectively, the social welfares under the full information and no information disclosure mechanisms are not identical in general.*

4.3.3 Socially Efficient Outcome

When each agent chooses his route, under either the no information or full information disclosure mechanisms, he does not take into account the congestion (*i.e.* negative externality) that his decision creates on the other agents. Therefore, the social welfare under the no information and full information disclosure mechanisms are different from the one under the socially efficient outcome. The socially efficient routing outcome is given by

$$x^{s,\text{eff}}(\theta) = \frac{1}{2} + \frac{1}{4}\Delta_\theta, \tag{4.8}$$

$$x^{r,\text{eff}}(\theta) = \frac{1}{2} - \frac{1}{4}\Delta_\theta, \tag{4.9}$$

and the corresponding expected social welfare is given by

$$W^{\text{eff}} = \frac{a+\mu-1}{2} + \frac{\Delta^2}{8} + \frac{\sigma^2}{8}. \tag{4.10}$$

We observe that the difference in the traffic of routes s and r is doubled under the full information mechanism (see (4.8) and (4.9)), where agents make routing decisions selfishly, compared to the socially efficient routing. This is an instance of the *tragedy of commons*, where each agent maximizes his own utility and does not take into account the congestion cost he imposes on the other agents on his route. Therefore, it may not be optimal for the principal to perfectly reveal his information about θ to the agents, as in the full information disclosure mechanism.

We now compare the optimal social welfare W^{eff} with the social welfare $W^{\text{no info}}$ under the no information disclosure mechanism. Under the no information disclosure mechanism, agents do not know θ and make their routing decisions only based on their ex-ante belief about θ (see (4.2) and (4.3)). Therefore, the social welfare under the no information disclosure mechanism, given by (4.4) is lower than the efficient social welfare because (i) the agents make their routing decisions selfishly, and (ii) the agents make their routing decisions without any knowledge about the realization of θ. The terms $\frac{\Delta^2}{8}$ and $\frac{\sigma^2}{8}$ in (4.10) capture the social welfare loss due to factors (i) and (ii) above, respectively.

In order to reduce the social welfare loss due to the agents' lack of information about θ, the principal may want to disclose information about the realization of θ to the agents. As discussed earlier, disclosing the realization of θ perfectly does not improve social welfare (see (4.4) and (4.7)). Therefore, the principal must utilize her superior information about route r's condition to strategically disclose information to the agents and influence their routing decision so as to improve the expected social welfare. This can be interpreted as providing *informational incentives* to the agents that align their objectives with that of the principal.

In the sequel, we explore various information disclosure mechanisms that the principal can employ to improve the expected social welfare. In Section 4.4, we explore public information disclosure mechanisms, where the principal reveals information about the realization of θ which is publicly observed by all agents. In Section 4.5, we explore private information disclosure mechanisms, where the principal reveals information to each agent individually through a private communication channel.

4.4 Public Information Disclosure

In this section, we consider mechanisms through which the principal reveals public information about the realization of θ to all agents. For instance, the principal can post traffic information on public road signs, or broadcast traffic updates through radio stations. Let $\mathcal{M}$ denote the set of all messages through which the principal can reveal information about the realization of θ. For instance, $\mathcal{M}$ can be the set of possible commute times on route r, or the number of congestion-causing accidents that have happened on route r. Given a message space $\mathcal{M}$, a public information disclosure mechanism can be fully described by $\psi : \Theta \to \Delta(\mathcal{M})$. For every realization of θ, ψ determines a probability distribution over the set of messages $\mathcal{M}$ that the principal sends. We note that the no information and full information disclosure mechanisms presented in Section 4.3 can be described as special instances of public information disclosure mechanisms by setting $\mathcal{M} = \emptyset$, and $\mathcal{M} = \Theta$ along with $\psi(\theta) = \theta$, respectively.

Given a public information disclosure mechanism $(\mathcal{M}, \psi)$, the agents update their belief about route r's condition θ after receiving a public message $m \in \mathcal{M}$ as,

$$\mathbb{P}\{\theta = \hat{\theta} | m\} = \frac{p_{\hat{\theta}} \psi(\hat{\theta})(m)}{\sum_{\tilde{\theta} \in \Theta} p_{\tilde{\theta}} \psi(\tilde{\theta})(m)}. \tag{4.11}$$

Using an argument similar to the one given in Section 4.3.1, for every message realization $m \in \mathcal{M}$, the traffic at routes s and r are given by

$$x^{s,\text{public}}(m) = \frac{1}{2} + \frac{1}{2}\Delta_m, \tag{4.12}$$

$$x^{r,\text{public}}(m) = \frac{1}{2} - \frac{1}{2}\Delta_m, \tag{4.13}$$

where $\Delta_m := a - \mathbb{E}\{\theta | m\}$.

The principal's objective is to design a message space $\mathcal{M}$ along with a public information disclosure mechanism ψ so as to maximize the expected social welfare W. Formally,

$$\max_{\mathcal{M}, \psi} \quad W$$

$$\text{subject to (4.12) and (4.13).}$$

Even though the principal can influence the agents' routing decisions for different realizations of θ by employing various public information disclosure mechanisms $(\mathcal{M}, \psi)$, we prove below that the expected social welfare W is independent of $(\mathcal{M}, \psi)$ for the model of Section 4.2.

Theorem 4.1. *For every public information disclosure mechanisms $(\mathcal{M}, \psi)$, the expected social welfare W is given by $\frac{a+\mu-1}{2}$.*

The result of Theorem 4.1 states that the principal cannot benefit from employing a public information disclosure mechanism. We would like to note that for the model of Section 4.2 (i) the congestion functions $C^s(x^s)$ and $C^r(x^r)$ are linear in x^s and x^r, and (ii) the effect of route r's condition θ on the utility of an agent taking route r is linearly separable from the congestion cost $C^r(x^r)$. Because of features (i) and (ii), conditioned on the realization of message m, the expected social welfare is a linear function of Δ_m; this leads to the result of Theorem 4.1. In a model where either feature (i) or (ii) is absent, the result of Theorem 4.1 does not hold.

Remark 4.2. *Consider a model where the congestion costs $C^s(x^s)$ and $C^r(x^r)$ are nonlinear functions of x^s and x^r, respectively. Define a function $G : [a-\theta^M, a-\theta^1] \to [0,1]$ as $G(\delta) := \{x : \ C^r(1-x) - C^s(x) = \delta\}$. Note that since $C^s(x^s)$ and $C^r(x^r)$ are strictly increasing in x^s and x^r, respectively, function $G(\delta)$ is well-defined. When $C^s(x^s) = x^s$ and $C^r(x^r) = x^r$, we have $G(\delta) = \frac{1}{2} + \frac{\delta}{2}$. Under a public information disclosure mechanism $(\mathcal{M}, \psi)$, conditioned on the realization of message m, the traffics at routes s and r are given by*

$$x^{s,public}(m) = G(\Delta_m), \tag{4.14}$$

$$x^{r,public}(m) = 1 - G(\Delta_m). \tag{4.15}$$

We can verify that if the function $C^s(G(\delta))$ is convex (resp. concave) in δ, the optimal public information disclosure mechanism is the no information (resp. full information) mechanism.[2] *In particular, if $C^s(x^s)$ and $C^r(x^r)$ are convex and concave (resp. concave and convex) in x^s and x^r, respectively, the function $C^s(G(\delta))$ is convex (resp. concave) in δ; thus, the optimal public information disclosure mechanism is the no information (resp. full information) mechanism. However, if the function $C^s(G(\delta))$ is neither convex nor concave in δ, there may exist instances of a set Θ of possible values for θ, along with a probability distribution over Θ, such that the optimal public information disclosure mechanisms is a public partial information disclosure mechanism.*

4.5 Private Information Disclosure

In this section, we study various private information disclosure mechanisms that the principal can use to reveal information about the realization of θ to the agents

[2] The result directly follows from an application of Jensen's inequality since $W = \mathbb{E}\{(a - G(\cdot)\}$.

so as to improve the expected social welfare. For instance, the principal can provide individualized and private information to every agent through GPS-enabled devices such as routing suggestions in smart phone applications. Under a private information disclosure mechanism, the principal sends a private signal that is based on the realization of θ to every agent. Similar to a public information disclosure mechanism, the principal needs to determine (i) a set of messages (*i.e.* language) that he wants to use, and (ii) a mapping that determines for every realization of θ the probability according to which every signal is sent.

One class of private information disclosure mechanisms is the set of mechanisms where the principal sends to every agent a private and individualized routing recommendation (*i.e.* which route to take) based on the realization of θ. We refer to this subset of private information disclosure mechanisms as recommendation policies. We note that since the agents are strategic, they do not necessarily follow the principal's recommendation unless it is a best response for them. Using the revelation principle argument for information design problems (see [15]), we can restrict attention, without loss of generality, to the set of recommendation policies where it is a best response for every agent to follow the recommendation he receives.

To avoid measure theoretic difficulties, we first assume that the principal sends $N > 0$ different recommendations to N groups of agents that have equal masses of $\frac{1}{N}$. We then consider the asymptotic case where $N \to \infty$.

Let $\sigma^N : \Theta \to \Delta(\{s, r\}^N)$ denote the recommendation policy that the principal employs for a given N. With some abuse of notation, let $\sigma^N(m^N|\theta)$ denote the probability that the principal sends routing recommendation $m^N := (m_1^N, ..., m_N^N) \in \{s, r\}^N$ to the N groups of agents, given that the state realization is $\theta \in \Theta$. Given a recommendation policy σ^N, each agent must be willing to take the recommended route given his information about route r's condition θ. This is captured by the following *obedience* condition for each agent belonging to group n, for $1 \leq n \leq N$,

(i) if $m_n^N = s$

$$\frac{1}{\sum_{\theta\in\Theta} p_\theta \sigma^N((s, m_{-n}^N)|\theta)} \sum_{\theta\in\Theta, m_{-n}^N\in\{s,r\}^{N-1}} p_\theta \sigma^N((s, m_{-n}^N)|\theta)(a - \frac{1}{N}\sum_{1\le i\le N} \mathbf{1}_{\{m_i^N=s\}})$$
$$\geq$$
$$\frac{1}{\sum_{\theta\in\Theta} p_\theta \sigma^N((s, m_{-n}^N)|\theta)} \sum_{\theta\in\Theta, m_{-n}^N\in\{s,r\}^{N-1}} p_\theta \sigma^N((s, m_{-n}^N)|\theta)(\theta - \frac{1}{N}\sum_{1\le i\le N} \mathbf{1}_{\{m_i^N=r\}}), \tag{4.16}$$

(ii) if $m_n^N = r$

$$\frac{1}{\sum_{\theta\in\Theta} p_\theta \sigma^N((r, m_{-n}^N)|\theta)} \sum_{\theta\in\Theta, m_{-n}^N\in\{s,r\}^{N-1}} p_\theta \sigma^N((r, m_{-n}^N)|\theta)(\theta - \frac{1}{N}\sum_{1\le i\le N} \mathbf{1}_{\{m_i^N=r\}})$$
$$\geq$$
$$\frac{1}{\sum_{\theta\in\Theta} p_\theta \sigma^N((r, m_{-n}^N)|\theta)} \sum_{\theta\in\Theta, m_{-n}^N\in\{s,r\}^{N-1}} p_\theta \sigma^N((r, m_{-n}^N)|\theta)(a - \frac{1}{N}\sum_{1\le i\le N} \mathbf{1}_{\{m_i^N=s\}}). \tag{4.17}$$

The above obedience constraints are the analogue of the *incentive compatibility* constraints in mechanism design problems, and can be interpreted similarly as follows. The left hand side of condition (4.16) (resp. (4.17)) expresses the expected utility of an agent in group n, $1 \le n \le N$, if he follows the recommendation to take route s (resp. r) given his ex-post belief about θ after he receives the recommendation, assuming that the other agents are following their recommendations. The right hand side of condition (4.16) (resp. (4.17)) expresses the expected utility of an agent in group n, if he deviates from his recommendation and takes route r (resp. s) instead of s (resp. r) given his ex-post belief about θ. The obedience constraint (4.16) therefore requires that it is a best response for every agent to follow the recommendation, given his ex-post belief about θ, assuming that other agents follow their routing recommendations. We note that unlike standard mechanism design problems, there is no individual rationality constraint, since an agent can simply ignore the recommendation and choose any route he wishes.

Let $x^{s,N}(\theta) \in \{\frac{1}{N}, \frac{2}{N}, \ldots, \frac{N}{N}\}$ denote the mass of agents that take route s when the state is θ under σ^N. Note that the set of obedience constraints (4.16) and (4.17) are linear in $\sigma^N(\cdot|\cdot)$ and identical for all N groups of agents. Therefore, by symmetry, we can restrict attention to the set of recommendation policies for the principal where she selects $N \cdot x^{s,N}(\theta)$ groups randomly, recommends to them to take route s, and recommends to the agents in the remaining groups to take route r.

Therefore, for $N \to \infty$ the set of recommendation policies for the principal can be characterized by $y(\theta) \in [0,1]$, where $y(\theta)$ denotes the mass of agents receiving the recommendation to take route s, *i.e.* $x^s(\theta) = y(\theta)$ and $x^r(\theta) = 1 - y(\theta)$. When the state is θ, the principal recommends route s (resp. r) to every agent with probability $y(\theta)$ (resp. $1 - y(\theta)$) independent of her recommendation to other agents.

Under the information policy σ, let $U^\sigma(s,\theta) := a - y(\theta)$ and $U^\sigma(r,\theta) := \theta - (1 - y(\theta))$ denote an agent's utility from taking routes s and r, respectively, when route r's condition is θ. The set of obedience constraints (4.16) and (4.17) for each agent can be then written as

$$\sum_{\theta \in \Theta} p_\theta y(\theta) U^\sigma(s,\theta) \geq \sum_{\theta \in \Theta} p_\theta y(\theta) U^\sigma(r,\theta), \tag{4.18}$$

$$\sum_{\theta \in \Theta} p_\theta (1 - y(\theta)) U^\sigma(r,\theta) \geq \sum_{\theta \in \Theta} p_\theta (1 - y(\theta)) U^\sigma(s,\theta). \tag{4.19}$$

Therefore, the problem that the principal faces is to determine a recommendation policy that maximizes the expected social welfare subject to the obedience constraints above; this optimization problem is given by[3]

$$\max_{\{y(\theta), \theta \in \Theta\}} W$$
$$\text{subject to (4.18) and (4.19).}$$

4.5.1 Implementable Outcomes

To determine an optimal recommendation policy, we first specify the set of feasible routing outcomes/recommendation policies that satisfy the obedience constraints (4.18) and (4.19).

Lemma 4.1. *A routing outcome* $\{x^s(\theta), x^r(\theta), x^s(\theta) + x^r(\theta) = 1,\ \theta \in \Theta\}$ *is implementable if and only if*

$$\mathbb{E}\left\{x^s(\theta)\left[\left(\frac{1}{2} + \frac{\Delta_\theta}{2}\right) - x^s(\theta)\right]\right\} \geq 0, \tag{4.20}$$

$$\mathbb{E}\left\{x^r(\theta)\left[\left(\frac{1}{2} + \frac{\Delta_\theta}{2}\right) - x^r(\theta)\right]\right\} \geq 0. \tag{4.21}$$

We note that the outcomes under the no information and full information disclosure policies, given by (4.2)-(4.3) and (4.5)-(4.6), respectively, satisfy conditions (4.20) and (4.21) with equality. That is, they are the corner points of the set of implementable outcomes. The set of implementable outcomes is depicted in Figure 4.2 for an example with $|\Theta| = 2$.

4.5.2 Incentivizing the Socially Efficient Routing

Using the result of Lemma 4.1, we can determine the necessary and sufficient condition to implement the efficient allocation $\{x^{s,\text{eff}}(\theta), x^{r,\text{eff}}(\theta), \theta \in \Theta\}$ through the recommendation policy below.

[3]We note that we can restrict attention, without loss of optimality, to policies where $y(\theta)$ is deterministic. This is because the set of obedience constraints only depends on the expected value of $y(\theta)$. Moreover, the principal's objective is a concave function of $y(\theta)$ (see (4.1)). Thus, by the Jensen's inequality, an optimal recommendation policy is a recommendation policy where $y(\theta)$ is deterministic for every $\theta \in \Theta$.

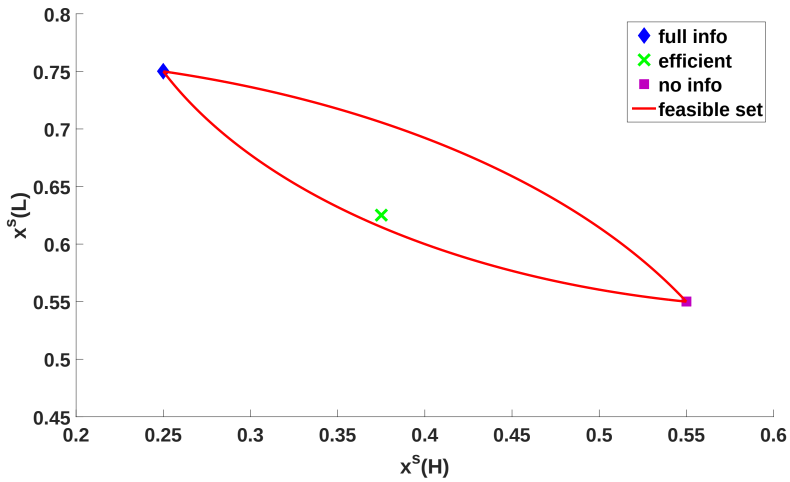


Figure 4.2: The set of implementable outcomes for $a = 2$, $\Theta = \{L, H\}$, $L = 1.5$, $H = 2.5$, $p_L = 0.6$, and $p_H = 0.4$.

Theorem 4.2. *The efficient routing policy x^{eff} is implementable through an information disclosure policy if and only if*

$$\sigma^2 \geq 2|\Delta| - \Delta^2. \tag{4.22}$$

We note that $|\Delta| = \frac{|a-\mu|}{m} \leq 1$ by Assumption 4.1; thus, $2|\Delta| - \Delta^2 \geq 0$. For ex-ante symmetric routes (*i.e.* $\mu = a$), we have $\Delta = 0$, and the efficient outcome is always implementable for any distribution of θ. However, if the two routes are ex-ante asymmetric (*i.e.* $\mu \neq a$), to incentivize the efficient policy, the variance of θ must be greater than the threshold (4.22), which depends on the expected difference between the routes. We further elaborate on this issue below.

As we discussed above, we can view the routing recommendation by the principal to the agents as an informational incentive that she provides so as to influence the routing decision of each agent. When the routes are symmetric, *i.e.* $\Delta = 0$, under the no information disclosure policy, each agent (at equilibrium) is indifferent between taking either of the routes; see (4.2) and (4.3). Therefore, the principal can persuade (*i.e.* recommend to) an agent to take a specific route even when she does not have significant information superiority over him (*i.e.* σ^2 is small). However, when the routes are asymmetric, *i.e.* $\Delta \neq 0$, under the no information disclosure policy, each agent has a strict preference over route s (resp. r) if $\Delta > 0$ (resp. $\Delta < 0$). Thus, the principal needs a strictly positive incentive to persuade an agent to take the route that is not aligned with his original preference. This implies that the information the principal holds must be valuable enough to enable her to offer adequate informational incentives to persuade an agent to follow her recommendation. Condition (4.22) captures the value of the principal's information about θ in terms of σ^2.

Figure 4.3 depicts the maximum expected social welfare the principal can achieve for different combinations of σ^2 and Δ by utilizing a recommendation policy in an example with $|\Theta| = 2$. We note that for pairs (σ, Δ) that satisfy condition (4.22) of Theorem 4.2, the principal can implement the socially efficient outcomes. However, when this condition is violated, the performance of the best outcome decreases.

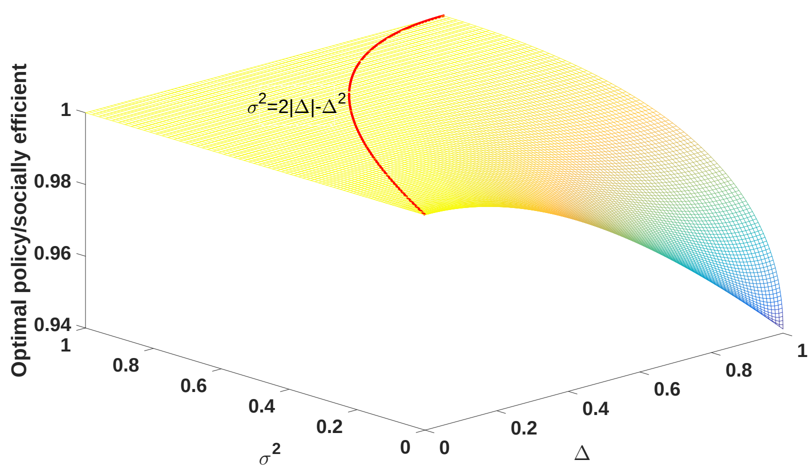


Figure 4.3: The best implementable outcomes with respect to the socially efficient outcome for $a = 2$, $\Theta = \{L, H\}$, and $p_L = p_H = 0.5$.

Remark 4.3. *A result similar to that of Theorem 4.2 can be obtained for general congestion functions $C^r(x^r)$ and $C^s(x^s)$, where the condition that is the analogue of (4.22) depends on higher order moments of θ.*

4.6 Dynamic Setting

In this section, we study a dynamic setting with time horizon $T = 2$, *i.e* $t \in \{1, 2\}$, where route r's condition θ_t, $t = 1, 2$, has uncontrolled Markovian dynamics with transition probability $\mathbf{P} \in \mathbb{R}^{|\Theta| \times |\Theta|}$. We assume that $\mathbf{P}[p_{\theta^1}, \ldots, p_{\theta^M}]^T = [p_{\theta^1}, \ldots, p_{\theta^M}]^T$, that is, the marginal probability distribution of θ_2 is the same that of θ_1.

We consider a situation where the same group of agents commute from the origin to the destination every day. Therefore, agents at $t = 2$ may have learnt new information from their observations at $t = 1$. We study the problem of designing an optimal dynamic private information disclosure policy by the principal. The investigation of this two-step dynamic mechanism provides some insight into how the results for a static setting change in a dynamic setting where agents can learn from their past experience. We note that by the result of Theorem 4.1, the study of dynamic public information disclosure mechanisms in a dynamic setting does not introduce any issue in addition to those present in the study of static mechanisms within the context of the model of Section 4.2.

We consider three scenarios depending on the agents' observations at $t = 1$ as follows: (i) agents do not make any environmental observations (*i.e.,* neither the condition of the risky route nor the traffic (*i.e.* mass of agents/cars) on routes s and r, (ii) agents who take route r observe only its condition θ_1, and (iii) each agent observes only the traffic on the route he takes at $t = 1$. In a real world situation, the agents can have noisy observations of θ_1 as well as a noisy observation of the number of cars traveling the route. Therefore, the study of the three scenarios described above will allow us to understand the effect of each type of learning (piece of information) on the solution of the dynamic problem and uncover its qualitative properties.

In all of these scenarios, we assume that the principal's routing recommendation policy at $t = 2$ does not depend on the agent's decisions at $t = 1$. We make this assumption for the following reasons. (1) If the principal wants to incorporate the agents' past decisions into her routing recommendation policy, she needs to monitor every agent's location over time; this may not be feasible due to technological limitations and/or privacy concerns. (2) If the principal can incorporate the agents' past decisions into her routing recommendation policy, then her optimal strategy would be to not disclose any further information to every agent that does not follow her routing recommendation (*i.e.* punish him). On one hand, such a punishment scheme may not be desirable in practical settings. On the other hand, if such a punishment scheme is permitted, then the principal can incentivize any desired routing behavior in a dynamic setting with long enough horizon if the agents are sufficiently patient and θ_t does not have deterministic dynamics (*i.e.* the principal has information superiority over the agents for all times). Therefore, in the sequel we restrict attention to the class of dynamic recommendation policies where the principal does not observe/incorporate the agents' decisions at $t = 1$ when designing her policy.

As a result, the set of recommendation policies for the principal can be characterized by $\sigma := \{y_1(\theta_1), y_2^s(\theta_2, \theta_1), y_2^r(\theta_2, \theta_1),\ \forall \theta_1, \theta_2 \in \Theta\}$. The principal's routing policy at $t = 1$ is given by $y_1(\theta)$ (resp. $1 - y_1(\theta)$), which denotes the probability that route s (resp. r) is recommended when route r's condition is θ_1. For $t = 2$, $y_2^s(\theta_2, \theta_1)$ (resp. $y_2^r(\theta_2, \theta_1)$) denotes the probability that route s is recommended to agents who took route s (resp. r) at $t = 1$, when route r's condition at $t = 1$ and $t = 2$ are θ_1 and θ_2, respectively. Similarly, $1 - y_2^s(\theta_2, \theta_1)$ (resp. $1 - y_2^r(\theta_2, \theta_1)$) denotes the probability that route r is recommended to agents who took route s (resp. r) at $t = 1$, when route r's condition at $t = 1$ and $t = 2$ are θ_1 and θ_2, respectively.

4.6.1 Case (i): No Environmental Observations

Consider a situation where the agents do not make any observations about the road condition θ_1 and/or the number of agents at the route they took at $t = 1$. Therefore, the agents infer information about θ_1 only based on the recommendation that they receive at $t = 1$.

Let

$$U_2^\sigma(s,\theta_2,\theta_1) := a - y_1(\theta_1)y_2^s(\theta_2,\theta_1) - (1-y_1(\theta_1))y_2^r(\theta_2,\theta_1),$$
$$U_2^\sigma(r,\theta_2,\theta_1) := \theta_2 - y_1(\theta_1)(1-y_2^s(\theta_2,\theta_1)) - (1-y_1(\theta_1))(1-y_2^r(\theta_2,\theta_1)),$$

denote the utility of routes s and r at $t=2$ given that all agents follow their recommendations. Moreover, define,

$$\mathbb{P}^\sigma\{\theta_2,\theta_1,s,s\} = y_2^s(\theta_2,\theta_1)y_1(\theta_1)\mathbf{P}(\theta_2,\theta_1)p_{\theta_1},$$
$$\mathbb{P}^\sigma\{\theta_2,\theta_1,s,r\} = y_2^r(\theta_2,\theta_1)(1-y_1(\theta_1))\mathbf{P}(\theta_2,\theta_1)p_{\theta_1},$$
$$\mathbb{P}^\sigma\{\theta_2,\theta_1,r,s\} = (1-y_2^s(\theta_2,\theta_1))y_1(\theta_1)\mathbf{P}(\theta_2,\theta_1)p_{\theta_1},$$
$$\mathbb{P}^\sigma\{\theta_2,\theta_1,r,r\} = (1-y_2^r(\theta_2,\theta_1))(1-y_1(\theta_1))\mathbf{P}(\theta_2,\theta_1)p_{\theta_1}.$$

Then the set of obedience constraints for $t=2$ are as follows:

(a) Recommendation s at $t=2$ and s at $t=1$:

$$\sum_{\theta_1,\theta_2\in\Theta}\mathbb{P}^\sigma\{\theta_2,\theta_1,s,s\}U_2(s,\theta_2,\theta_1) \geq \sum_{\theta_1,\theta_2\in\Theta}\mathbb{P}^\sigma\{\theta_2,\theta_1,s,s\}U_2(r,\theta_2,\theta_1). \tag{4.23}$$

(b) Recommendation s at $t=2$ and r at $t=1$:

$$\sum_{\theta_1,\theta_2\in\Theta}\mathbb{P}^\sigma\{\theta_2,\theta_1,s,r\}U_2(s,\theta_2,\theta_1) \geq \sum_{\theta_1,\theta_2\in\Theta}\mathbb{P}^\sigma\{\theta_2,\theta_1,s,r\}U_2(r,\theta_2,\theta_1) \tag{4.24}$$

(c) Recommendation r at $t=2$ and s at $t=1$:

$$\sum_{\theta_1,\theta_2\in\Theta}\mathbb{P}^\sigma\{\theta_2,\theta_1,r,s\}U_2(r,\theta_2,\theta_1) \geq \sum_{\theta_1,\theta_2\in\Theta}\mathbb{P}^\sigma\{\theta_2,\theta_1,r,s\}U_2(s,\theta_2,\theta_1) \tag{4.25}$$

(d) Recommendation r at $t=2$ and r at $t=1$:

$$\sum_{\theta_1,\theta_2\in\Theta}\mathbb{P}^\sigma\{\theta_2,\theta_1,r,r\}U_2(s,\theta_2,\theta_1) \geq \sum_{\theta_1,\theta_2\in\Theta}\mathbb{P}^\sigma\{\theta_2,\theta_1,r,r\}U_2(r,\theta_2,\theta_1) \tag{4.26}$$

We note that since the agents' observations at $t = 1$ include only the routing recommendation they receive at that time, the agents act myopically at $t = 1$, as they cannot learn additional information by themselves. Thus, the set of obedience constraints at $t = 1$ are the same as those of the static problem; these constraints are given by

$$\sum_{\theta_1 \in \Theta} p_\theta y_1(\theta_1) U_1^\sigma(s, \theta_1) \geq \sum_{\theta_1 \in \Theta} p_\theta y_1(\theta_1) U_1^\sigma(r, \theta_1), \tag{4.27}$$

$$\sum_{\theta_1 \in \Theta} p_\theta (1 - y_1(\theta_1)) U_1^\sigma(r, \theta_1) \geq \sum_{\theta_! \in \Theta} p_\theta (1 - y_1(\theta_1)) U_1^\sigma(s, \theta_1), \tag{4.28}$$

where $U_1^\sigma(s, \theta_1) = a - y_1(\theta_1)$ and $U_1^\sigma(r, \theta_1) = \theta_1 - (1 - y_1(\theta_1))$.

The expected social welfares at $t = 1$ and $t = 2$ are given by

$$W_1 := \sum_{\theta_1 \in \Theta} p_{\theta_1} \left[y(\theta_1) U_1^\sigma(s, \theta_1) + (1 - y(\theta_1)) U_1^\sigma(r, \theta_1) \right], \tag{4.29}$$

$$\begin{aligned} W_2 := \sum_{\theta_1, \theta_2 \in \Theta} & p_{\theta_1} \mathbf{P}(\theta_2, \theta_1) \Big[\big[y(\theta_1) y_2^s(\theta_2, \theta_1) + (1 - y(\theta_1)) y^r(\theta_2, \theta_1) \big] U_1^\sigma(s, \theta_2, \theta_1) \\ & + \big[y(\theta_1)(1 - y_2^s(\theta_2, \theta_1)) + (1 - y(\theta_1))(1 - y_2^r(\theta_2, \theta_1)) \big] U_2^\sigma(r, \theta_2, \theta_1) \Big] \end{aligned} \tag{4.30}$$

Therefore, the optimal routing recommendation policy by the principal when the agents do not have any environmental observations is given by the solution to the following optimization problem

$$\begin{aligned} &\max_\sigma \quad W_1 + W_2 \\ &\text{subject to } (4.23) - (4.28). \end{aligned}$$

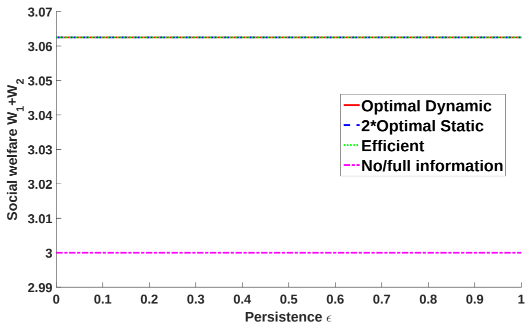
In this chapter, we do not provide a closed solution to the above maximization problem for a general transition matrix $\mathbf{P}$. Nevertheless, we consider two special cases below: (a) when θ_2 is identically distributed and independent of θ_1 (i.e. no correlation), and (b) when θ_2 is identical to θ_1 (*i.e.* perfect correlation). We argue below that that for special cases (a) and (b) the performance of an optimal dynamic recommendation policy per time step is equal to that of optimal static recommendation policy.

For case (a), it is easy to verify that the repetition of the optimal static recommendation policy is an optimal dynamic recommendation policy. For case (b), at $t = 1$, the optimal static routing policy is implementable since the obedience constraints at $t = 1$ are identical to those in the static problem, At $t = 2$, consider the recommendation policy $y^s(\theta_2) = 1$ and $y^r(\theta_2) = 0$. That is, at $t = 2$, the principal recommends to every agent to take the same route he took at $t = 1$. If an agent is willing to follow his recommendation at $t = 1$, then he is also willing to take the exact route for the next day, since he does not learn any new information after $t = 1$, and route r's condition remains the same. Therefore, $y^s(\theta_2) = 1$ and $y^r(\theta_2) = 0$ is implementable at $t = 2$. It is easy to verify that the performance of the dynamic recommendation policy described above is identical to that of the optimal static recommendation policy.

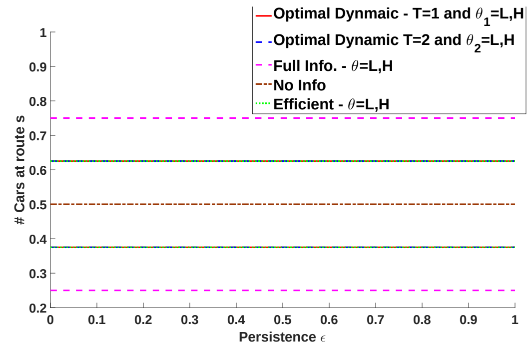
Given a general transition matrix $\mathbf{P}$, every agent forms an updated belief about θ_1 after receiving his recommendation at $t = 1$. As the correlation between θ_1 and θ_2 increases, the information that the agent learns at $t = 1$ becomes more valuable to him, and consequently, the principal's information superiority decreases. The argument given above for case (b) states that even when $\theta_1 = \theta_2$, the principal can achieve the same expected social welfare at $t = 2$ as in the static setting. Therefore, we conjecture that the (partial) results we proved above for the special cases (a) and (b), hold in general for any correlation between 0 and 1.

Conjecture 1. *For every transition matrix $\mathbf{P}$, the performance of an optimal dynamic recommendation policy per time step is equal to that of an optimal static recommendation policy.*

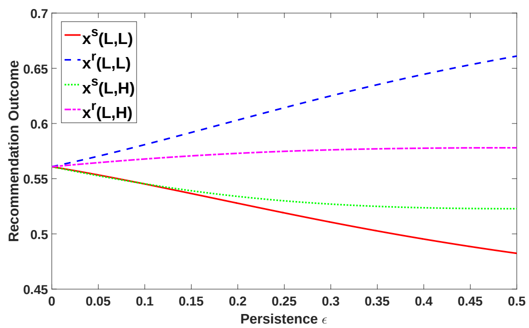
We examine the above conjecture through a numerical simulation below. Consider a setting where $\Theta = \{L, H\}$ with $p_L = p_H = 0.5$. Assume that the transition matrix



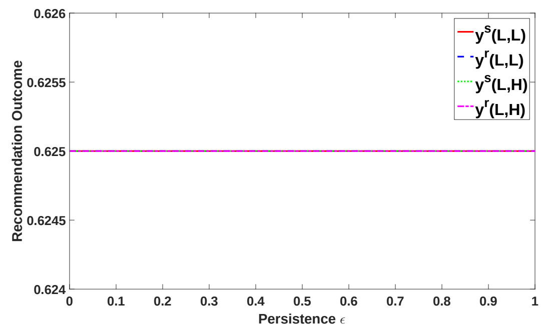
(a) The social welfare



(b) The number of agents at routes at $t = 1, 2$ for different route r's condition



(c) The recommendation outcome at $t = 2$ for $\theta_1 = L$



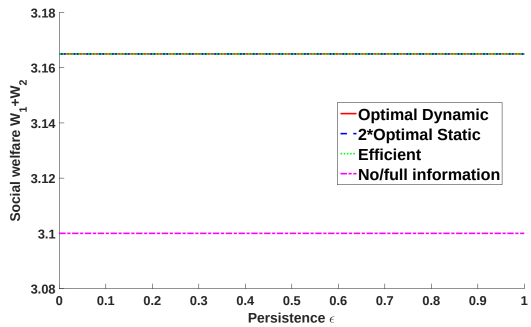
(d) The recommendation outcome at $t = 2$ for $\theta_1 = H$

Figure 4.4: Case (i) - the optimal dynamic recommendation policy for different values of persistence ϵ, for $\theta \in \{L, H\}$, $L = 1.5$, $H = 2.5$, $p_L = p_H = 0.5$ ($\Delta = 0$, $\sigma^2 = 0.25$)

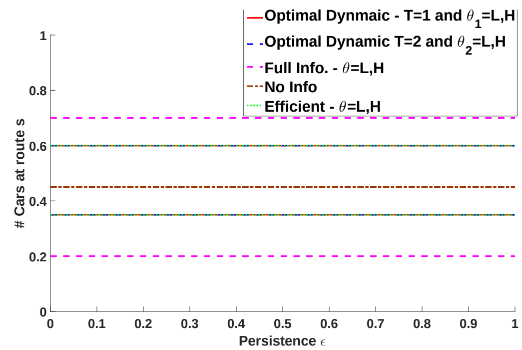
$\mathbf{P}$ is given by

$$\mathbf{P} := \begin{pmatrix} p_L + \frac{\epsilon}{2} & p_L - \frac{\epsilon}{2} \\ p_H - \frac{\epsilon}{2} & p_H + \frac{\epsilon}{2} \end{pmatrix}, \tag{4.31}$$

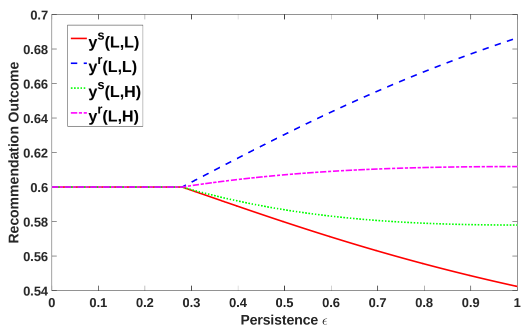
where $\epsilon \in [0, 1]$ denotes the persistence (i.e. correlation) of route r's condition from $t = 1$ to $t = 2$. Figures 4.4-4.6 depict the optimal dynamic recommendation policies vs. different values of ϵ for three pairs of σ^2 and Δ. As seen in Fig. 4.4a-4.6a, in all three examples the performance of the dynamic recommendation policy per time step is the same as in the optimal static recommendation policy for different values of ϵ. Moreover, the total number of cars at each route at $t = 1, 2$ are identical under the optimal dynamic recommendation policy; see Figures 4.4b-4.6b.



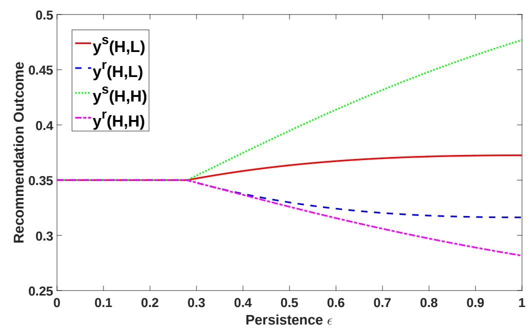
(a) The social welfare



(b) The number of agents at routes at $t = 1, 2$ for different route r's condition



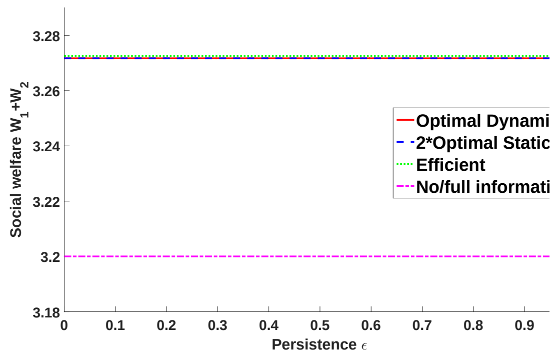
(c) The recommendation outcome at $t = 2$ for $\theta_1 = L$



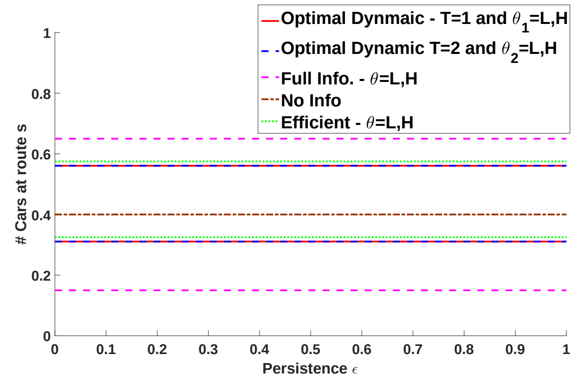
(d) The recommendation outcome at $t = 2$ for $\theta_1 = H$

Figure 4.5: Case (i) - the optimal dynamic recommendation policy for different values of persistence ϵ, for $\theta \in \{L, H\}$, $L = 1.6$, $H = 2.6$, $p_L = p_H = 0.5$ ($\Delta = 0.1$, $\sigma^2 = 0.25$)

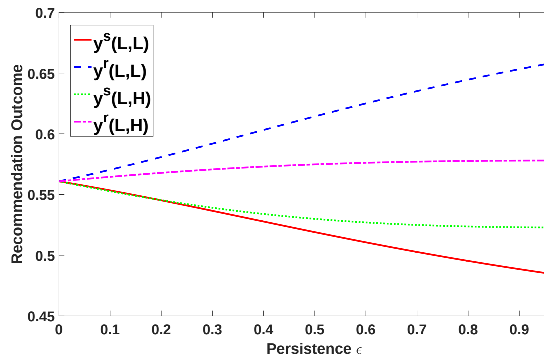
In the first two examples (see Figures 4.4 and 4.5), the pair (σ^2, Δ) satisfies condition (4.22) of Theorem 4.2; thus, the performance of the optimal dynamic and static recommendation policies are the same as the efficient social welfare. However, in the third example (see Figure 4.6), where the pair (σ^2, Δ) does not satisfy condition (4.22), the performance of the optimal dynamic and static recommendation policies are inferior to that of the social welfare maximizing policy.



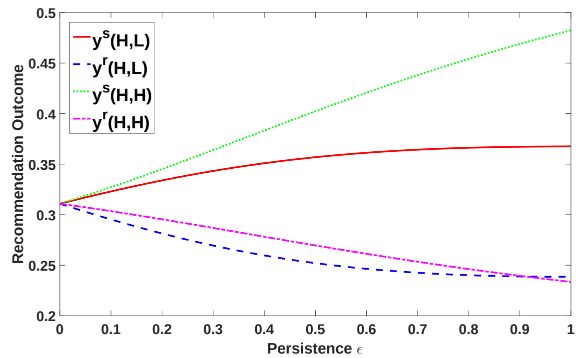
(a) The social welfare



(b) The number of agents at routes at $t = 1, 2$ for different route r's condition



(c) The recommendation outcome at $t = 2$ for $\theta_1 = L$



(d) The recommendation outcome at $t = 2$ for $\theta_1 = H$

Figure 4.6: Case (i) - the optimal dynamic recommendation policy for different values of persistence ϵ, for $\theta \in \{L, H\}$, $L = 1.7$, $H = 2.7$, $p_L = p_H = 0.5$ ($\Delta = 0.2$, $\sigma^2 = 0.25$)

In Section 4.5.2, we argued that σ^2 represents a measure of the principal's power in terms of the informational incentives she can provide to the agents. Moreover, we argued that Δ indicates the agents' ex-ante preference towards one of the routes. Accordingly, we interpreted the condition of Theorem 4.2 as requiring that the principal's informational power to be greater than the agents' ex-ante preference towards one of the routes. A similar interpretation can be given here by comparing the recommendation outcomes for different pairs of (σ^2, Δ). As seen in Figures 4.4c and

4.4d, when the agents do not have an ex-ante preference towards either of the routes (*i.e.* $\Delta = 0$), the optimal recommendation policy prescribes the same routing suggestion for all groups of agents at $t = 2$ irrespective of what they have learnt at $t = 1$. As the agents develop an ex-ante preference towards one of the routes (see Figure 4.5 where $\Delta = 0.1$), the optimal recommendation policy prescribes the same routing suggestion for low values of ϵ; but as ϵ increases, the optimal recommendation policy prescribes different routing suggestions depending on what every agent has learnt at $t = 1$. When agents have a high ex-ante preference towards one of the routes (see Figure 4.5 where $\Delta = 0.2$), the optimal recommendation policy prescribes different routing suggestions for every value of $\epsilon \neq 0$ depending on what every agent has learnt at $t = 1$.

4.6.2 Case (ii): Observing θ

Consider a situation where agents who take route r at $t = 1$ observe θ_1 perfectly. Therefore, at $t = 2$ agents have heterogeneous/asymmetric information about θ_2 depending on which route they took at $t = 1$.

As a result, the set of obedience constraints at $t = 2$ is as follows:

(a) Recommendation s at $t = 2$ and s at $t = 1$:

$$\sum_{\theta_1, \theta_2 \in \Theta} \mathbb{P}^{\sigma}\{\theta_2, \theta_1, s, s\} U_2(s, \theta_2, \theta_1) \geq \sum_{\theta_1, \theta_2 \in \Theta} \mathbb{P}^{\sigma}\{\theta_2, \theta_1, s, s\} U_2(r, \theta_2, \theta_1). \tag{4.32}$$

(b) Recommendation s at $t = 2$ and r at $t = 1$: for every $\theta_1 \in \Theta$

$$\sum_{\theta_2 \in \Theta} \mathbb{P}^{\sigma}\{\theta_2, s, r\} U_2(s, \theta_2, \theta_1) \geq \sum_{\theta_2 \in \Theta} \mathbb{P}^{\sigma}\{\theta_2, s, r\} U_2(r, \theta_2, \theta_1) \tag{4.33}$$

(c) Recommendation r at $t = 2$ and s at $t = 1$:

$$\sum_{\theta_1, \theta_2 \in \Theta} \mathbb{P}^{\sigma}\{\theta_2, \theta_1, r, s\} U_2(r, \theta_2, \theta_1) \geq \sum_{\theta_1, \theta_2 \in \Theta} \mathbb{P}^{\sigma}\{\theta_2, \theta_1, r, s\} U_2(s, \theta_2, \theta_1) \tag{4.34}$$

(d) Recommendation r at $t = 2$ and r at $t = 1$: for every $\theta_1 \in \Theta$

$$\sum_{\theta_2 \in \Theta} \mathbb{P}^\sigma\{\theta_2, \theta_1, r, r\} U_2(r, \theta_2, \theta_1) \geq \sum_{\theta_2 \in \Theta} \mathbb{P}^\sigma\{\theta_2, \theta_1, r, r\} U_2(s, \theta_2, \theta_1) \tag{4.35}$$

We note that the obedience constraints (4.32) and (4.34) are the same as (4.23) and (4.25). This is because if an agent takes route s at $t = 1$, his only new information at $t = 2$ is the routing recommendation that he receives at $t = 1$, and the situation at $t = 2$ is similar to the scenario with no environmental observations in Section 4.6.1. However, if an agent takes route r at $t = 1$, he observes θ_1 perfectly. Therefore, for every possible value of θ_1 at $t = 1$, there exists a corresponding obedience constraint at $t = 2$ expressed (4.33) and (4.35).

Since agents can learn/observe θ_1 at $t = 1$ by taking route r, the agents' incentives to follow the principal's routing recommendation are different from those in a static setting. An agent may want to deviate from the recommendation to take route s and instead take route r so as to observe θ_1; he can then utilize this observation to his benefit, as he now has better information about route r's condition at $t = 2$. In other words, an agent can coordinate his routing decision at $t = 2$ with his routing decision at $t = 1$. Therefore, the set of obedience constraints at $t = 1$ must consider all possible future plans that an agent can utilize at $t = 2$ after his deviation at $t = 1$; this set of obedience constraints can be described as follows:

(a) Recommendation s at $t = 1$: for every $\lambda : \{s, r\} \times \Theta \to \{s, r\}$

$$\sum_{\theta_1 \in \Theta} p_{\theta_1} y(\theta_1) \Bigg[U_1^\sigma(s, \theta_1) + \sum_{\theta_2 \in \Theta} \mathbf{P}(\theta_2, \theta_1) \Big(y_2^s(\theta_2, \theta_1) U_2^\sigma(s, \theta_2, \theta_1). \\ + (1 - y_2^s(\theta_2, \theta_1)) U_2^\sigma(r, \theta_2, \theta_1) \Big) \Bigg] \\ \geq \\ \sum_{\theta_1 \in \Theta} p_{\theta_1} y(\theta_1) \Bigg[U_1^\sigma(r, \theta_1) + \sum_{\theta_2 \in \Theta} \mathbf{P}(\theta_2, \theta_1) \Big(y_2^s(\theta_2, \theta_1) U_2^\sigma(\lambda(s, \theta_1), \theta_2, \theta_1) \\ + (1 - y_2^t(\theta_2, \theta_1)) U_2^\sigma(\lambda(r, \theta_1), \theta_2, \theta_1) \Big) \Bigg], \tag{4.36}$$

(b) Recommendation r at $t = 1$: for every $\lambda : \{s, r\} \to \{s, r\}$

$$\sum_{\theta_1 \in \Theta} p_{\theta_1} (1 - y(\theta_1)) \Bigg[U_1^\sigma(r, \theta_1) + \sum_{\theta_2 \in \Theta} \mathbf{P}(\theta_2, \theta_1) \big(y_2^r(\theta_2, \theta_1) U_2^\sigma(s, \theta_2, \theta_1). \\ + (1 - y_2^r(\theta_2, \theta_1)) U_2^\sigma(r, \theta_2, \theta_1) \Big) \Bigg] \\ \geq \\ \sum_{\theta_1 \in \Theta} p_{\theta_1} (1 - y(\theta_1)) \Bigg[U_1^\sigma(s, \theta_1) + \sum_{\theta_2 \in \Theta} \mathbf{P}(\theta_2, \theta_1) \Big(y_2^r(\theta_2, \theta_1) U_2^\sigma(\lambda(s), \theta_2, \theta_1) \\ + (1 - y_2^r(\theta_2, \theta_1)) U_2^\sigma(\lambda(r), \theta_2, \theta_1) \Big) \Bigg]. \tag{4.37}$$

In the obedience constraints above, an agent's plan at $t = 2$ after his deviation at $t = 1$ is denoted by λ. If an agent deviates from a recommendation to take route s at $t = 1$ and takes route r instead, his plan λ at $t = 2$ depends on his observation of θ_1 as well as his routing recommendation at $t = 2$ (see (4.36)). If an agent deviates from a recommendation to take route r at $t = 1$ and instead takes route s, his plan λ

at $t = 2$ depends only his routing recommendation at $t = 2$ since he does not observe θ_1 at $t = 1$ (see (4.37)).

Consequently, the optimal routing recommendation policy by the principal, when agents taking route r at $t = 1$ observe θ_1, is given by the solution to the following optimization problem

$$\max_{\sigma} \quad W_1 + W_2$$
$$\text{subject to } (4.32) - (4.37).$$

The above optimization problem has $6+2|\Theta|+2^{2|\Theta|}$ number of constraints, which grows exponentially in $|\Theta|$, making it difficult to provide a closed form solution to the above problem in general. Therefore, we investigate the properties of an optimal dynamic recommendation policy, when agents can observe θ_1, through numerical simulations below.

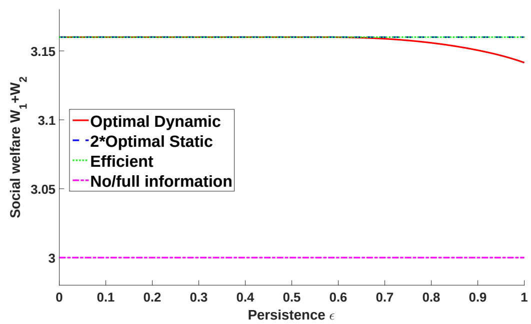
Consider a setting similar to the one in Section 4.6.1, where $\Theta = \{L, H\}$ with $p_L = p_H = 0.5$ and the transition matrix $\mathbf{P}$ is given by (4.38). Figures 4.7-4.8 depict the optimal dynamic recommendation policies vs. different values of ϵ for two pairs of σ^2 and Δ. As seen in Figures 4.7a-4.8a, in both examples the performance of the dynamic recommendation policy is decreasing in ϵ.

In the first example (see Figure 4.7)), the pair (σ^2, Δ) satisfies condition (4.22) of Theorem 4.2. Therefore, for low values of ϵ, where the information the agents learn at $t = 1$ does not reduce the principal's information superiority to the point where condition (4.22) is not satisfied at $t = 2$, the principal can implement the efficient routing policy (see Figure 4.7a)). However, as ϵ increases, the principal cannot implement the efficient routing at $t = 1, 2$. As a result, the optimal recommendation policy is different from the efficient routing policy for higher values of ϵ (see Figure 4.7b)). Moreover, for higher values of ϵ the optimal recommendation policy at $t = 2$ depends on the route an agent took at $t = 1$ (see Figures 4.7c and 4.7d). We note that for $\epsilon = 1$, $y^r(L, L) = 1$ and $y^r(H, H) = 0$. This is because, when $\epsilon = 1$, an agent who takes route r at $t = 1$ perfectly knows θ_2 since $\theta_2 = \theta_1$. Therefore, an agent who takes route r (*i.e.* observes θ_1) at $t = 1$ always chooses the route with the better condition at $t = 2$.

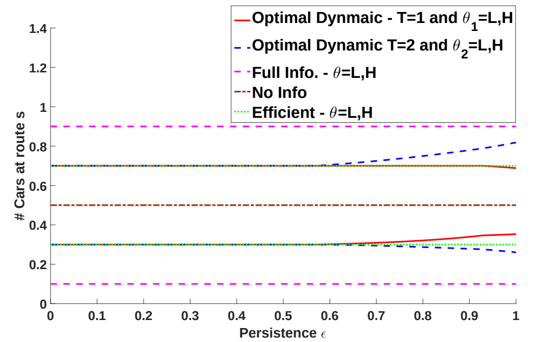
In the second example (see Figure 4.8), the pair (σ^2, Δ) does not satisfy condition (4.22) of Theorem 4.2. Therefore, the performance of the optimal dynamic recommendation policy is *strictly* decreasing in ϵ for all values of ϵ (see Figure 4.8a). Moreover, the optimal recommendation policy at $t = 1, 2$ is always different from the efficient routing policy (see Figure 4.8b). As seen in Figures 4.8c and 4.8d, for all $\epsilon \neq 0$, the optimal recommendation policy at $t = 2$ depends on which route an agent took at $t = 1$. We note that when $\epsilon = 1$, the optimal routing recommendation policy at $t = 2$ results in the same traffic on the routes as the one under the full information disclosure mechanisms (see Figure 4.8b). That is, for very high values of ϵ, where the agents have considerable incentive to experiment at $t = 1$ by taking route r, the principal promises to perfectly disclose his information at $t = 2$ so that the agents are willing to follow her recommendation at $t = 1$. As a consequence, in contrast to Figures 4.7c and 4.7d, in Figures 4.8c and 4.8d we have $y^r(L, L) \neq 1$ and $y^r(H, H) \neq 0$ for $\epsilon = 1$. This is because under the optimal recommendation policy the utility from taking either route is the same when $\epsilon = 1$, and thus, an agent is indifferent between them even though he knows θ_2 perfectly.

4.6.3 Case (iii): Observing the Traffic

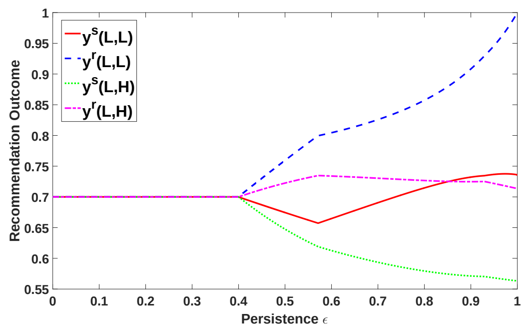
Consider a situation where each agent observes the traffic on the route he has taken at $t = 1$, , given by $y_1(\theta_1)$ or $1 - y_1(\theta_1)$. Since we assume that there exists a unit mass of agents, which is common knowledge among all agents, each agent can determine the traffic at $t = 1$ on both routes, and thus, all agents have identical information at $t = 2$. Consequently, the set of obedience constraints for $t = 1, 2$ are similar to that of the static problem and are given by (4.18) and (4.19). Therefore, when the agents observe the traffic at $t = 1$, the problem of designing an optimal dynamic recommendation policy does not introduce any conceptual issue in addition to those present in the study of an optimal static recommendation policy. In the following, we argue that as the correlation between θ_1 and θ_2 increases the performance of the dynamic optimal recommendation policy decreases. Moreover, we



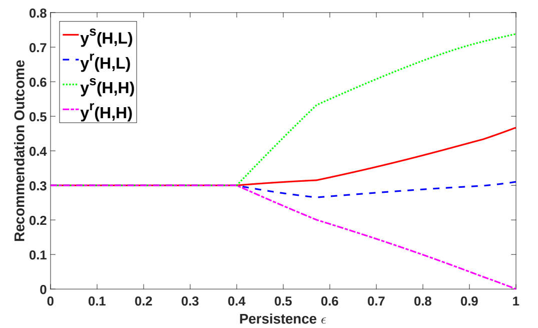
(a) The social welfare



(b) The number of agents at routes at $t = 1, 2$ for different route r's condition



(c) The recommendation outcome at $t = 2$ for $\theta_1 = L$

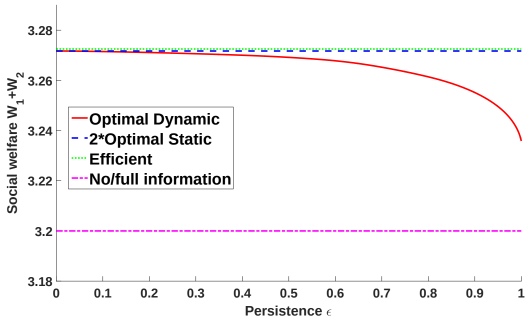


(d) The recommendation outcome at $t = 2$ for $\theta_1 = H$

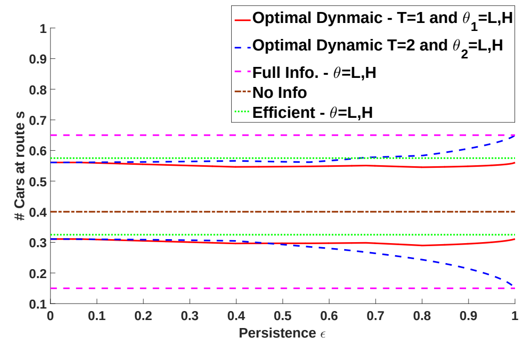
Figure 4.7: Case (ii) - the optimal dynamic recommendation policy for different values of persistence ϵ, for $\theta \in \{L, H\}$, $L = 1.2$, $H = 2.8$, $p_L = p_H = 0.5$ ($\Delta = 0$, $\sigma^2 = 0.64$)

show that when the correlation between θ_1 and θ_2 is high the designer's optimal recommendation policy at $t = 1$ is a partial information disclosure mechanism.[4]

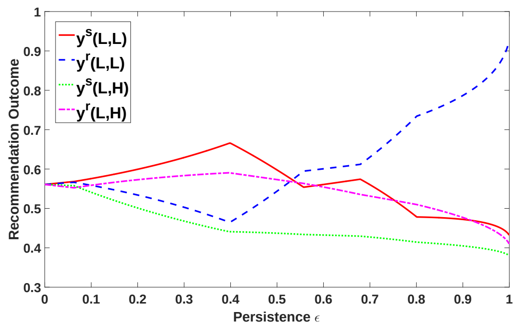
[4]For the sake of discussion we restrict attention to deterministic policies at $t = 1$. We can show that this restriction is without loss of optimality. For $t = 2$, fix the recommendation policy at $t = 1$. Then by an argument similar to the one given in the static setting in Section 4.5, we can restrict without loss of optimality, to deterministic recommendation policies. For $t = 1$, fix the recommendation policy at $t = 2$, which is a deterministic recommendation policy. Then the problem of designing the optimal recommendation policy for $t = 1$ can be written as a linear program in terms of the probabilities of different routing policies for every realization of θ_1 (see [16] and [129]). It is known that in a linear program, the optimal solutions are the corner points. Therefore, under the optimal recommendation policy the probability of each routing policy is either 0 or 1. Therefore, we can restrict attention to the set of deterministic policies without loss of optimality.



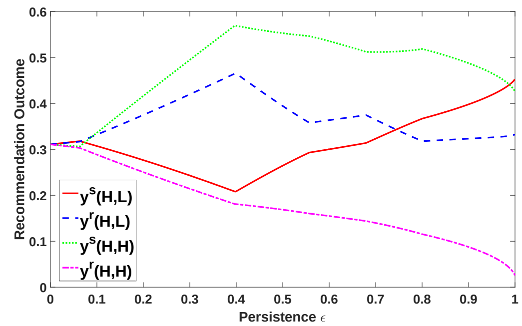
(a) The social welfare



(b) The number of agents at routes at $t = 1, 2$ for different route r's condition



(c) The recommendation outcome at $t = 2$ for $\theta_1 = L$



(d) The recommendation outcome at $t = 2$ for $\theta_1 = H$

Figure 4.8: Case (iii) - the optimal dynamic recommendation policy for different values of persistence ϵ, for $\theta \in \{L, H\}$, $L = 1.7$, $H = 2.7$, $p_L = p_H = 0.5$ ($\Delta = 0.2$, $\sigma^2 = 0.25$)

We consider two classes of recommendation policies for the designer: (a) *dynamic separating recommendation policies* in which for every $\hat{\theta}_1 \neq \tilde{\theta}_1$, $\hat{\theta}, \tilde{\theta} \in \Theta$, we have $y_1(\hat{\theta}_1) \neq y_1(\tilde{\theta}_1)$, and (b) *dynamic pooling recommendation policies* in which there exists $\hat{\theta}_1 \neq \tilde{\theta}_1$, $\hat{\theta}, \tilde{\theta} \in \Theta$, we have $y_1(\hat{\theta}_1) = y_1(\tilde{\theta}_1)$.

First, consider a case where the principal employs a separating recommendation policy. As argued above, the agents can infer perfectly θ_1 at $t = 1$. Therefore, as ϵ increases the principal's information superiority is reduced at $t = 2$, and thus, the performance of the dynamic recommendation policy decreases at $t = 2$. Therefore, the overall performance of a dynamic separating recommendation policy is decreasing in ϵ. In particular, when $\theta_1 = \theta_2$ (*i.e.* perfect correlation), the agents learn θ_2 perfectly and the performance of the recommendation policy at $t = 2$ is the same as that of the full information disclosure mechanism.

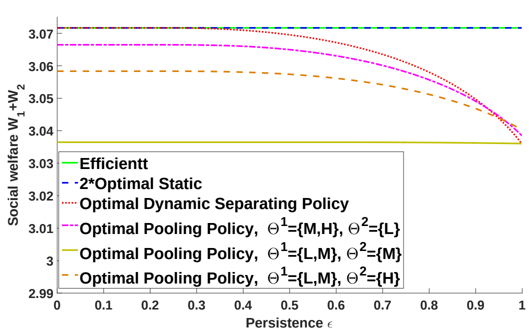
Next, consider a case where the principal employs a dynamic pooling recommendation policy. When the correlation between θ_1 and θ_2 is high, the principal may prefer to use a dynamic pooling recommendation policy so as to not reveal θ_1 perfectly at $t = 1$, and consequently, increase his information superiority at $t = 2$ as compared to the outcome under a dynamic separating recommendation policy. Given a dynamic pooling recommendation policy, let $\{\Theta^1, \Theta^2, \ldots \Theta^m\}$ denote a partition of Θ such that for every $\hat{\theta}_1, \tilde{\theta}_1 \in \Theta$, we have $y_1(\hat{\theta}_1) = y_2(\tilde{\theta}_1)$ if and only if $\hat{\theta}_1, \tilde{\theta}_1 \in \Theta^i$ for some $1 \leq i \leq M$. That is, if Θ^i, $1 \leq i \leq m$, is a singleton, the principal implements a distinct routing outcome at $t = 1$ and the drivers perfectly learn the realization of θ_1 at $t = 1$; if Θ^i, $1 \leq i \leq m$, is not a singleton, the principal implements the same routing outcome for all realization in Θ^i and the drivers only learn that $\theta_1 \in \Theta^i$ at $t = 1$. Similar to the outcome under a dynamic separating recommendation policy, under a dynamic pooling recommendation policy the principal's information superiority decreases at $t = 2$ as the correlation between θ_1 and θ_2 increases since the drivers learn the partition to which θ_1 belongs. Therefore, the overall performance of the principal's optimal dynamic recommendation policy decreases as the correlation between θ_1 and θ_2 increases irrespective of the exact form of the optimal dynamic recommendation policy.

It is clear that the performance of an optimal dynamic pooling recommendation policy at $t = 1$ is inferior to that of an optimal dynamic separating recommendation policy. However,the performance of an optimal dynamic pooling recommendation policy at $t = 2$ is higher than that of an optimal dynamic separating recommendation policy since the principal has a higher information superiority under an optimal dynamic pooling recommendation policy rather than the one under an optimal dynamic separating recommendation policy. Using a numerical simulation, we show below that, when the correlation between θ_1 and θ_2 is high, there are instances where the principal's optimal recommendation policy is a dynamic pooling recommendation policy.

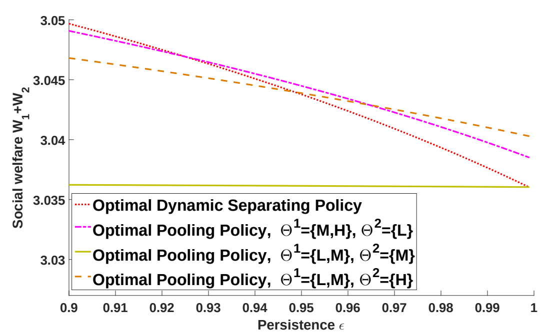
Consider a setting where $\Theta = \{L, M, H\}$ with $p_L = p_M = p_H = \frac{1}{3}$. Assume that the transition matrix $\mathbf{P}$ is given by

$$\mathbf{P} := \begin{pmatrix} p_L + \frac{2\epsilon}{3} & p_M - \frac{\epsilon}{3} & p_H - \frac{\epsilon}{3} \\ p_L - \frac{\epsilon}{3} & p_M + \frac{2\epsilon}{3} & p_H - \frac{\epsilon}{3} \\ p_L - \frac{\epsilon}{3} & p_M - \frac{\epsilon}{3} & p_H + \frac{2\epsilon}{3} \end{pmatrix}, \tag{4.38}$$

where $\epsilon \in [0, 1]$ denotes the persistence (i.e. correlation) of route r's condition over time. Figures 4.9 and 4.10 depict the optimal dynamic recommendation policy for two pairs of σ^2 and Δ. In the first example (see Figure 4.9), the parameters σ^2 and Δ satisfy the condition (4.22) of Theorem 4.2, while in the second example (see Figure 4.10) they do not satisfy it. As seen in Figures 4.9a and 4.9a, the performance of an optimal dynamic pooling recommendation policy and an optimal dynamic separating recommendation policy are decreasing in persistence ϵ. Figures 4.9b and 4.9b depict the optimal dynamic recommendation policies for high values of ϵ. As we discussed above, when the correlation between θ_1 and θ_2 is high, the principal prefers to employ a dynamic pooling recommendation policy.

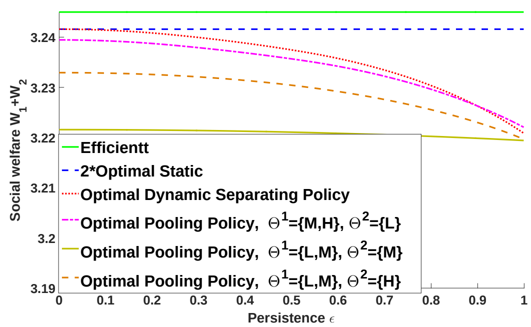


(a) The social welfare under the optimal dynamic recommendation policy for different values of persistence ϵ

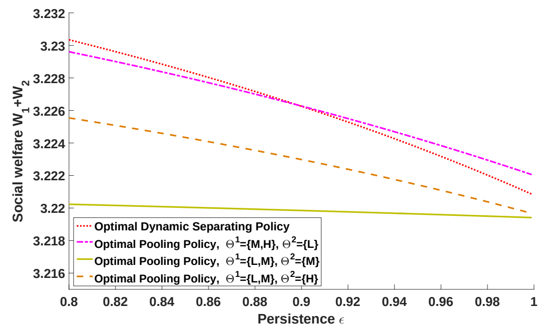


(b) The optimal recommendation policy for high values of ϵ

Figure 4.9: Case (iii) - The optimal dynamic recommendation policy for $\{L, M, H\} = \{1.3, 2.1, 2.6\}$ ($\Delta = 0$, $\sigma^2 = 0.2867$)



(a) The social welfare under the optimal dynamic recommendation policy for different values of persistence ϵ



(b) The optimal recommendation policy for high values of ϵ

Figure 4.10: Case(iii) - The optimal dynamic recommendation policy for different values of persistence ϵ, for $\{L, M, H\} = \{1.7, 2.3, 2.6\}$ ($\Delta = 0.2$, $\sigma^2 = 0.0867$)

4.7 Conclusion

We investigated the problem of information disclosure mechanisms design in transportation networks. We showed that the principal can improve the social welfare by strategically disclosing information to the drivers, and coordinating the routing recommendations she provides to them. We characterized a condition under which the principal can implement the efficient routing outcome by utilizing her superior information to provide informational incentives to the drivers. We also investigated a two-time step dynamic setting where the drivers learn from their experience at $t = 1$. We characterized different pieces of information from which the drivers can learn, and examined the effect of each of them using numerical simulations. For future research, we will investigate the dynamic setting more extensively and consider the extension of our results for nonlinear congestion cost functions.

Chapter 5

Stochastic Dynamic Games with Asymmetric Information: A Common Information Approach

5.1 Introduction

5.1.1 Background and Motivation

Stochastic dynamic games with asymmetric information have been used to model many situations arising in engineering, economic, and socio-technological network applications. In these applications many decision makers/agents interact with each other as well as with a dynamic system. They make private imperfect observations over time, and influence the evolution of the dynamic system through their actions.

In this chapter we study a general class of dynamic games where the underlying system has Markovian dynamics. Given the agents' actions at every time, the system state at the next time is a stochastic function of the current system state. The instantaneous utility of each agent depends on the agents' joint actions as well as the system state. At every time, each agent makes a private noisy observation that depends on the current system state and the agents' past actions. Therefore, at every time agents have asymmetric and imperfect information about the history of the game. Moreover, at every time the information that an agent possesses about the history of the game depends on the other agents' past actions and strategies; this phenomenon is known as *signaling* in the control theory literature. Therefore, the agents' decisions and information are coupled and interdependent over time in these games because (i) an agent's utility depends on the other agents' actions, (ii) the evolution of the system state depends on the agents' actions, (iii) the agents have imperfect and asymmetric information about the history of the game, and (iv) at every time an agent's information depends on the agents' (including his own) past actions and strategies.

There are two main challenges in the study of dynamic games with asymmetric information. First, because of the coupling and interdependence among the agents' decisions and information over time, we need to determine the agents' strategies simultaneously for all times. Second, as the agents acquire more information over time, the domains of their strategies grow.

In this chapter, we propose a general approach for the study of the class of games described above and address the challenges stated above. We provide a set of conditions sufficient to characterize an information state for every agent, where private and common information are compressed over time in a mutually consistent manner among the agents. Based on this information state, we propose the notion of Common Information Based Pefect Bayesian Equilibrium (CIB-PBE) that characterizes a set of outcomes for dynamic games. We provide a sequential decomposition of the game over time based on the notion of CIB-PBE, and formulate a dynamic program that enables us to compute the set of CIB-PBEs via backward induction. We characterize specific instances of dynamic games where we can determine a set of information states for the agents that have time-invariant domain. We determine conditions that guarantee the existence of CIB-PBEs. We show that the proposed approach to dynamic games can also be used to study dynamic teams with asymmetric information, thus, we provide a framework for the study of a broad class of dynamic multi-agent decision problems with asymmetric information.

To present clearly the key ideas and results appearing in this chapter we attempt to connect and compare them with existing key ideas and results in stochastic control, dynamic teams, and dynamic games with symmetric information. In the following we briefly discuss some of these existing results. We then provide a quick overview of our approach and results and compare them with related literature.

The centralized stochastic control problem presents the simplest form of a decision making problem with only one agent where the main two challenges described above are present. Partially Observed Markov Decision Process (POMDP) provides a general model to describe a centralized control problem. In a POMDP, an agent acquires imperfect observations about the system state over time; thus, he has a growing domain for his strategies over time. Moreover, the information that the agent acquires at any time is affected by his previous actions since his past actions influence the evolution of the dynamic system. To address these two challenges the notion of *information state* is introduced [67]. An information state in POMDP can be defined as the agent's belief about the current system state conditioned on his information history. The definition of information state provides an approach to compress the agent's information in a way that is sufficient for decision making purposes. Assuming that the agent has perfect recall, it is shown that his conditional belief about the system state (*i.e.* information state) is independent of his strategies over time; this result is known as *policy-independence belief* property [67, Lemma 6.5.10]. As a result, the problem of finding an optimal policy for the agent can be sequentially decomposed over time so that the complexity of the agent's decision problem does not grow over time.

Our main objective in this chapter is to present an approach to compress the agents' information and to provide a decomposition of dynamic games with asymmetric information similar to the one described for POMDPs above. Therefore, we highlight the three main properties that underly the definition of an information state in POMDPs as follows (see [76, 137]): (1) the information state can be updated recursively, that is, at any time t the information state at time t can be written as a function of the information state at $t-1$ and the new information that becomes available at t, (2) the agent's belief about the information state at the next time conditioned on the current information state and action is independent of his information history, and (3) at any time t and for any arbitrary action the agent's expected instantaneous utility conditioned on the information state is independent of his information history. In this chapter, we provide a generalization of these properties to decision problems with many strategic agents, and accordingly, present a general approach to study dynamic games with asymmetric information. Before, we present our approach and results, we discuss the existing results that generalize the approach and results for POMDPs to dynamic games with symmetric information and dynamic teams below.[1]

The authors of [81] study dynamic games with symmetric information and propose an approach to compress the agents' information over time, and sequentially decompose the game over time. In a strategic setting, each agent has its own objective, thus, he chooses his strategy individually and privately so as to maximize his own utility. The authors in [81] show that when agents have symmetric information, the belief about the system state conditioned on the agents' information satisfies properties (1)-(3) described above; thus, it can be defined as an information state for all agents. Similar to the policy-independence belief property in centralized stochastic control, they show that this information state is independent of the agents' private strategy choices. They consider a class of strategies for the agents, called *Markov strategies*, that utilizes this information state, and show that this class of strategies are closed under the agents' best response mapping. Consequently, they introduce the notion of Markov Perfect Equilibrium that characterizes a subset of Subgame Perfect Equilibria (SPE) for dynamic games with symmetric information. The notion of MPE characterizes a class of equilibria where the agents' strategies have time-invariant domain, and they can be computed sequentially via backward induction.

The results of [67, 81] show that when agents have symmetric information in dynamic teams and games, the conditional belief about the system state defines an information state that can be used to compress the agents' information and to sequentially decompose the problem over time. However, this approach is not directly applicable when the agents have asymmetric information over time. When the agents have asymmetric information, they need to form beliefs about the other agents' information, beliefs about the agents' belief about the agents' information, and so on.

[1] We note that dynamic teams with symmetric information do not introduce additional difficulties compared to POMDPs. This is because in these problems all the agents share the objective and have identical information at all times. Therefore, they can be treated as centralized control problems with many actions (one for each agent).

On one hand, these hierarchies of beliefs depend on the agents' strategies over time. On the other hand, the agents' optimal strategies depend on the agents' beliefs over time. Thus, there is an interdependence between the agents' strategies and beliefs overtime. As a result, the results for dynamic games and teams with symmetric information do not directly apply to dynamic games and teams with asymmetric information.

The authors of [90, 91] study dynamic team problems with non-classical information structure [2], where at every time agents possess common and private information. They propose an approach to construct a dynamic team problem with symmetric information that is equivalent to the original problem as follows. For every agent, they consider a *fictitious agent* who has access to the agent's common information but not his private information. At every time t, each fictitious agent chooses a function/prescription that determines the corresponding agent's action at t for every possible realization of the agent's private information. They show that the dynamic team problem among the fictitious agents is equivalent to the original dynamic team problem. However, the problem with the fictitious agents is a dynamic team problem with symmetric information. Therefore, it can be solved using the existing results for centralized stochastic control problems. In the problem with the fictitious agents, the information state is given by the Common Information Based (CIB) belief about the agents' private information and the system state. Thus, in the original problem the information state for each agent is given by the described-above CIB-belief and his own private information. Using this information state, the authors of [90, 91] provide a sequential decomposition of dynamic teams with asymmetric information, and provide an approach to compress the agents' common information over time. We would like to point out that the proposed information state in [90, 91] requires the agents to keep track of their private information (or his private memory that is a predetermined function of his private information if they do not have perfect recall) over time, and does not provide an approach on how to effectively compress it.

The approach proposed in [90, 91] for dynamic teams does not apply for dynamic

[2] A dynamic team has a non-classical information structure when each agent's information depends on other agents' strategies (see [135])

games with asymmetric information. This is because in games, an agent's strategy is his private information. Therefore, in the equivalent dynamic games among the fictitious agents, the decisions of each fictitious agent, which prescribes an action for every realization of the associated agent's private information, is his own private information; that is, the resulting game among fictitious agents is a dynamic game with asymmetric information (hidden actions).

In this chapter, we propose a general approach for the study of dynamic games with asymmetric information. We present a set of conditions sufficient to characterize information states where the agents common and private information are compressed in a mutually consistent manner. Based on these information states, we define a class of equilibria called Common Information Baesd Perfect Bayesian Equilibria (CIB-PBE) that generalize the notion of MPE to dynamic games with asymmetric information. We provide a sequential decomposition of the game that enables us to compute the set of CIB-PBEs by backward induction. Moreover, we show that our results also apply to dynamic teams with asymmetric information, thus, generalize the results of [90, 91] by providing an approach to effectively compress the agents' private information in a mutually consistent manner. As a result, our results provide an universal approach for the study of dynamic decision problems with many agents in strategic and cooperative settings.

5.1.2 Related Literature

Dynamic games with asymmetric information have been investigated extensively in literature in the context of repeated games; see [138, 42, 5, 79] and the references therein. The key feature of these games is the absence of a dynamic system. Moreover, the works on repeated games study primarily their asymptotic properties when the horizon is infinite and agents are sufficiently patient (*i.e.* the discount factor is close one). In repeated games, agents play a stage (static) game repeatedly over time. As a result, in the absence of strategic interactions with other agents, the decision making problem that each agent faces is very simple. The main objective of this strand of literature is to explore situations where agents can form self-enforcing punishment/reward mechanisms so as to create additional equilibria that improve upon the payoffs that agents can get by simply playing an equilibrium of the stage game over time. Recent works (see [55, 37, 111]) adopt approaches similar to those used in repeated games to study infinite horizon dynamic games with asymmetric information when there is an underlying dynamic Markovian system. Under certain conditions on the system dynamics and the agents' information structure, the authors of [55, 37, 111] characterize a set of asymptotic equilibria when the agents are sufficiently patient.

The problem that we study in this chapter is different from the ones in [138, 42, 5, 79, 55, 37, 111] in two aspects. First, we consider a class of stochastic dynamic games where the underlying dynamic system has a general Markovian dynamics and information structure, and we do not restrict our attention to asymptotic behaviors when the horizon is infinite and/or the agents are sufficiently patient. Second, we study situations where the decision problem that each agent faces, in the absence of strategic interactions with other agents, is a POMDP, which is a complex problem to solve by itself. Therefore, reaching (and computing) a set of equilibrium strategies, which take into account the strategic interactions among the agents, is a very challenging task. As a result, it is not very plausible for the agents to seek reaching equilibria that are generated by the formation of self-enforcing punishment/reward mechanisms similar to those used in infinitely repeated games (see Section 5.7 for more discussion). We believe that our results provide new insight into the behavior of strategic agents in complex and dynamic environments, and complement the existing results in the repeated games literature with simple and (mostly) static environments.

The works in [101, 27, 44, 70] consider dynamic zero-sum games with asymmetric information. The authors of [27, 101] study zero-sum games with Markovian dynamics and lack of information on one side (*i.e.* one informed player and one uninformed player). The authors of [44, 70] study zero-sum games with Markovian dynamics with lack of information on both sides. The problem that we study in this chapter is different from the ones in [101, 27, 44, 70] in three aspects. First, we study a general class of dynamic games that include dynamic zero-sum games with asymmetric information as a special case. Second, we consider a general Markovian dynamics for the underlying system whereas the authors of [27, 101, 44, 70] consider a specific Markovian dynamics where each agent observes perfectly a local state that evolves independently of the other local states conditioned on the agents' observable actions. Third, we consider a general information structure for each agent that allows us to capture scenarios with unobservable actions and imperfect observations that are not captured in [27, 101, 44, 70].

The problems investigated in [89, 46, 97, 98, 127, 108] are the most closely related to our problem. The authors of [89, 46] study a class of dynamic games where the agents' common information based belief is independent of their strategies; that is, there is no signaling among them. This property allows them to apply ideas from the common information approach developed in [90, 91], and define an equivalent dynamic game with symmetric information among the fictitious agents. Consequently, they characterize a class of equilibria for dynamic games called *Common Information Based Markov Perfect Equilibrium*. Our results are different from those in [89, 46] in two aspects. First, we consider a general class of dynamic games where the agents' CIB beliefs are strategy-dependent, thus, signaling is present. Second, the proposed approach in [89, 46] requires the agents to keep track of all of their private information over time. We propose an approach to effectively compress the agents' private information, and consequently, reduce the number of variables on which the agents need to form CIB belief.

The authors of [97, 98, 127, 108] study a class of dynamic games with asymmetric information where signaling occurs. When the horizon in finite, the authors of [97, 98] introduce the notion of Common Information Based Perfect Bayesian Equilibrium, and provide a sequential decomposition of the game over time. The authors of [127, 108] extend the results of [97, 98] to finite horizon Linear-Quadratic-Gaussian (LQG) dynamic games and infinite horizon dynamic games, respectively. The class of dynamic games studied in [97, 98, 127, 108] satisfies the following assumptions: (i) agents' actions are observable (ii) each agent has a perfect observation of his own local states/type (iii) conditioned on the agents' actions, the evolution of the local states are independent.

In this chapter we relax assumptions (i)-(iii) of [97, 98, 127, 108], and study a general class of dynamic games with asymmetric information, hidden actions, imperfect observations, and controlled and coupled dynamics. As a result, each agent needs to form a belief about the other agents' past actions and private (imperfect) observations. Moreover, in contrast to [97, 98, 127, 108], an agent's, say agent i's, belief about the system state and the other agents' private information is his own private information and is different from the CIB belief. In this chapter, we extend the methodology developed in [97, 98] for dynamic games, and generalize the notion of CIB-PBE. Furthermore, we propose an approach to effectively compress the agents' private information and obtain the results of [97, 98, 127, 108] as special cases.

5.1.3 Contribution

We develop a general methodology for the study and analysis of dynamic games with asymmetric information, where the information structure is non-classical. We propose an approach to characterize a set of information states that effectively compress the agents' private and common information in a mutually consistent manner. As a result, we characterize a subclass of Perfect Bayesian Equilibria for dynamic games with asymmetric information, called CIB-PBE, and provide a sequential decomposition of these dynamic games over time. This decomposition provides a backward induction algorithm to determine the set of CIB-PBEs. We characterize special instances of dynamic games where we can identify a set of information states with time-invariant domain. We provide conditions that guarantee the existence of CIB-PBEs in dynamic games with asymmetric information. We show that the methodology developed in this chapter generalizes the existing results for dynamic teams with non-classical information structure. The information state characterized in this chapter provides a sufficient statistic for decision making purposes in strategic and non-strategic settings. Therefore, we provide a universal approach to decision making problems with strategic and non-strategic agents; our approach can be applied to study dynamic games among teams of agents.

5.1.4 Organization

The rest of the chapter is organized as follows. In Section 5.2, we describe our model and formulate the dynamic game problem. In Section 5.3, we discuss the main issues that arise in the study of dynamic games with asymmetric information. We provide the formal definition of Perfect Bayesian Equilibrium in Section 5.4. In Section 5.5, we propose an approach to compressing the agents' common and private information and define an information state for each agent. Accordingly, we propose the notion of CIB assessment and CIB-PBE for dynamic games. In Section 5.6, we present our main results and provide a sequential decomposition of dynamic games. We compare the notion of CIB-PBE with other equilibrium concepts appropriate for dynamic games with asymmetric information, and extend our results to dynamic teams in Section 5.7. In Section 5.8, we discuss the role of assumptions we make in the model of Section 5.2, and provide the extension of our results by relaxing them under certain conditions. In Section 5.9, we determine conditions that guarantee the existence of CIB-PBE. We conclude in Section 5.10. The proofs of all the theorems and lemmas appear in Appendix D.

5.2 Model

1) System dynamics: There are N strategic agents who live in a dynamic Markovian world over horizon $\mathcal{T} := \{1, 2, ..., T\}$, $T < \infty$. Let $X_t \in \mathcal{X}_t$ denote the state of the world at $t \in \mathcal{T}$. At time t, each agent, indexed by $i \in \mathcal{N} := \{1, 2, ..., N\}$, chooses an action $a_t^i \in \mathcal{A}_t^i$, where $\mathcal{A}_t^i$ denotes the set of available actions to him at t. Given the collective action profile $A_t := (A_t^1, ..., A_t^N)$, the state of the world evolves according to the following stochastic dynamic equation,

$$X_{t+1} = f_t(X_t, A_t, W_t^x), \tag{5.1}$$

where $W_{1:T-1}^x$ is a sequence of independent random variables. The initial state X_1 is a random variable that has a probability distribution $\eta \in \Delta(\mathcal{X}_1)$ with full support.

At every time $t \in \mathcal{T}$, before taking an action, agent i receives a noisy private observation $Y_t^i \in \mathcal{Y}_t^i$ of the current state of the world X_t and the action profile A_{t-1}, given by

$$Y_t^i = O_t^i(X_t, A_{t-1}, W_t^i), \tag{5.2}$$

where $W_{1:T}^i$, $i \in \mathcal{N}$, are sequences of independent random variables. Moreover, at every $t \in \mathcal{T}$, all agents receive a common observation $Z_t \in \mathcal{Z}_t$ of the current state of the world X_t and the action profile A_{t-1}, given by

$$Z_t = O_t^c(X_t, A_{t-1}, W_t^c), \tag{5.3}$$

where $W_{1:T}^c$, is a sequence of independent random variables. We note that the agents' actions A_{t-1} is commonly observable at t if $A_{t-1} \subseteq Z_t$. We assume that the random variables X_1, $W_{1:T-1}^x$, $W_{1:T}^c$, and $W_{1:T}^i$, $i \in \mathcal{N}$ are mutually independent.

2) Information structure: Let $H_t \in \mathcal{H}_t$ denote the aggregate information of all agents at time t. Assuming that agents have perfect recall, we have $H_t = \{Z_{1:t}, Y_{1:t}^{1:N}, A_{1:t-1}^{1:N}\}$, *i.e.* H_t denotes the set of all agents' past observations and actions. The set of all possible realizations of the agents' aggregate information is given by $\mathcal{H}_t := \prod_{\tau \leq t} \mathcal{Z}_\tau \times \prod_{i \in \mathcal{N}} \prod_{\tau \leq t} \mathcal{Y}_\tau^i \times \prod_{i \in \mathcal{N}} \prod_{\tau < t} \mathcal{A}_\tau^i$.

At time $t \in \mathcal{T}$, the aggregate information H_t is not fully known to all agents, and each agent may have asymmetric information about H_t. Let $C_t := \{Z_{1:t}\} \in \mathcal{C}_t$ denote the agents' common information about H_t and $P_t^i := \{Y_{1:t}^i, A_{1:t-1}^i\} \backslash C_t \in \mathcal{P}_t^i$ denote agent i's private information about H_t, where $\mathcal{P}_t^i$ and $\mathcal{C}_t$ denote the set of all possible realizations of agent i' private and common information at time t, respectively. In Section 5.2.1, we consider and discuss several instances of information structures that can be captured as special cases of our general model.

3) Strategies and Utilities: Let $H_t^i := \{C_t, P_t^i\} \in \mathcal{H}_t^i$ denote the information available to agent i at t, where $\mathcal{H}_t^i$ denote the set of all possible realizations of agent i's information at t. Agent i's *behavioral strategy* g_t^i, $t \in \mathcal{T}$, is defined as a sequence of mappings $g_t^i : \mathcal{H}_t^i \to \Delta(\mathcal{A}_t^i)$, $t \in \mathcal{T}$, that determine agent i's action A_t^i for every realization $h_t^i \in \mathcal{H}_t^i$ of the history at $t \in \mathcal{T}$.

Agent i's instantaneous utility at t depends on the state of the world X_t and the collective action profile A_t, and is given by $u_t^i(X_t, A_t)$. Agent i chooses his behavioral strategy $g_{1:T}^i$ so as to maximize his total (expected) utility over horizon $\mathcal{T}$, given by,

$$U^i(X_{1:T}, A_{1:T}) = \sum_{t \in \mathcal{T}} u_t^i(X_t, A_t). \tag{5.4}$$

To avoid measure-theoretic technical difficulties and for clarity and convenience of exposition, we assume that all the random variables take values in finite sets.

Assumption 5.1. *(Finite game) The sets $\mathcal{X}_t$, $\mathcal{Z}_t$, $\mathcal{Y}_t^i$, $\mathcal{A}_t^i$, $i \in \mathcal{N}$, $t \in \mathcal{T}$, are finite.*

Moreover, we assume that given any sequence of actions $a_{1:t-1}$ up to time $t-1$, every realization $x_t \in \mathcal{X}_t$ for system state at t has a strictly positive probability of realization.

Assumption 5.2. *(Strictly positive transition matrix) For all $t \in \mathcal{T}$, $x_t \in \mathcal{X}_t$ and $a_{1:t-1} \in \mathcal{A}_{1:t-1}$, we have $\mathbb{P}\{x_t|a_{1:t-1}\} > 0$.*

Furthermore, we assume that for any sequence of actions $\{a_{1:T}\}$, all realizations of private observations $\{y_{1:T}^{1:N}\}$ have a positive probability. That is, no agent can infer perfectly another agent's action based only on his private observations.

Assumption 5.3. *(Imperfect private monitoring) For all $t \in \mathcal{T}$, $y_{1:t} \in \mathcal{Y}_{1:t}$, and $a_{1:t-1} \in \mathcal{A}_{1:t-1}$, we have $\mathbb{P}\{y_{1:t}|a_{1:t-1}\} > 0$.*

We discuss the role of Assumptions 5.1-5.3 in Section 5.8, where we determine conditions under which we can relax these assumptions, and obtain results similar to those of Sections 5.5, and 5.6.

5.2.1 Special Cases

We discuss several instances of dynamic games with asymmetric information that can be described as special cases of general model described above.

1) Nested information structure: Consider a two-player game with one informed player and one uninformed player. At every time $t \in \mathcal{T}$, the informed player makes a private perfect observation of the state X_t, *i.e.* $Y_t^1 = X_t$. The uninformed player does not have any observation of the state X_t. Both the informed and uninformed players observe each others' actions, *i.e.* $Z_t = \{A_{t-1}\}$. Therefore, we have $P_t^1 = \{X_{1:t}\}$, $P_t^2 = \emptyset$, and $C_t = \{A_{1:t-1}^1, A_{1:t-1}^2\}$ for all $t \in \mathcal{T}$. The above nested information structure corresponds to dynamic games with asymmetric information considered in [101, 102, 71], where in [102, 71] the underlying state X_t is static.

2) Independent dynamics with observable actions: Consider an N-player game where the state $X_t := (X_t^0, X_t^1, X_t^2, ..., X_t^N)$ has N components. The agents' actions A_t are observable by all agents, *i.e.* $A_{t-1} \subset Z_t$ for all $t \in \mathcal{T}$. At every time $t \in \mathcal{T}$, agent i makes a perfect observation of its local state X_t^i as well as a global state X_t^0. Moreover, at time t all agents make a common imperfect observation of state X_t^i given by $Z_t^i = O_t^c(X_t^i, A_{t-1}, W_t^{c,i})$, $i \in \mathcal{N}$. Conditioned on the agents' collective action A_t, each X_t^i evolves independently over time as $X_{t+1}^i = f_t(X_t^i, A_{t-1}, W_t^{x,i})$ for all $i \in \mathcal{N}$ and $t \in \mathcal{T}$, where $W_t^{x,i}$, $i \in \mathcal{N}$, $t \in \mathcal{T}$ are mutually independent. Therefore, we have $P_t^i = \{X_{1:t}^i\}$ and $C_t = \{X_{1:t}^0, Z_{1:t}^{1:N}, A_{1:t-1}\}$. The above environment includes the dynamic game with asymmetric information considered in [98, 97].

3) Delayed sharing information structure: Consider a N-player game with observable actions where agents observe each others' observations with d-step delay. That is, $P_t^i = \{Y_{t-d+1:t}^i\}$ and $C_t = \{Y_{1:t-d}, A_{1:t-1}\}$. We note that in the model we assume that the agents' common observation Z_t at t is only a function of X_t and and A_{t-1}. Therefore, to describe the game with delayed sharing information structure within the context of our model we need to augment our state space to include the agents' last d observations as part of the augmented state. Define $\tilde{X}_t := \{X_t, M_t^1, M_t^2, ..., M_t^d\}$ as the augmented system state where $M_t^i := \{A_{t-i}, Y_{t-i}\} \in \mathcal{A}_{t-i} \times \mathcal{Y}_{t-i}$, $i \in \mathcal{N}$; that is, M_t^i serves as a temporal memory for the agents' observation Y_{t-i} at $t-i$. Then, we have $\tilde{X}_{t+1} = \{X_{t+1}, M_{t+1}^1, M_{t+1}^2, ..., M_{t+1}^d\} = \{f_t(X_t, A_t, W_t^x), (Y_t), M_t^1, ..., M_t^{d-1}\}$ and $Z_t = \{M_t^d\} = \{Y_{t-d}\}$.

The above environment captures a connection between the symmetric information structure and asymmetric information structure. The information asymmetry among the agents increases as d increases.

4) Perfectly controlled dynamics with hidden actions: Consider a N-player game where the state $X_t := (X_t^1, X_t^2, ..., X_t^N)$ has N components. Agent i, $i \in \mathcal{N}$, perfectly controls X_t^i, *i.e.* $X_{t+1}^i = A_t^i$. Agent i's actions A_t^i, $t \in \mathcal{T}$, is not observable by all other agents $-i$. Every agent i, $i \in \mathcal{N}$, makes a noisy private observation $Y_i^t(X_t, W_t^i)$ of the system state at $t \in \mathcal{T}$. Therefore, we have $P_t^i := \{A_{1:t}, Y_{1:t}^i\}$, $C_t = \emptyset$.

5.3 Equilibrium Solution Concept

In this section we discuss the notion of an equilibrium solution concept for dynamic games with asymmetric information. We argue that an equilibrium solution concept must consist of a pair of a strategy profile and a belief system (to be defined below). We provide a comparison between approaches to dynamic games with asymmetric information and dynamic teams with non-classical (asymmetric) information structure, and discuss the importance of off-equilibrium path beliefs in dynamic games.[3]

In a dynamic game, as described in Section 5.2, agents have private information about the evolution of the game, and they do not observe the complete history of the game given by $\{H_t, X_t\}$, $t \in \mathcal{T}$. Therefore, at every time $t \in \mathcal{T}$, each agent, say agent $i \in \mathcal{N}$, needs to form (i) an appraisal about the current state of the system X_t and the other agents' information H_t^{-i} (appraisal about the history), and (ii) an appraisal about how other agents will play in the future so as to evaluate the performance of his strategy choices (appraisal about the future). Given the other agents' strategies g^{-i}, agent i can utilize his own information H_t^i at $t \in \mathcal{T}$, along with (i) other agents' past strategies $g_{1:t-1}^{-i}$ and (ii) other agents' future strategies $g_{t:T}^{-i}$ to form these appraisals about the history and future of the game, respectively.[4]

[3] We refer the interested reader to the papers by Battigalli [13], Myerson and Remy [88], and Watson [132]

[4] In dynamic teams, agents share the same objective, and thus, coordinate their strategies so as to maximize their shared objective. This implies that in dynamic teams the agents' strategies $g_{1:T}^{1:N}$ are common knowledge among them. Therefore, agent $i \in \mathcal{N}$ can form appraisals about the system's history and its future by using his private information H_t^i along with commonly known strategies g^{-i}. As a result, the outcome of dynamic team problems can be fully characterized by the agents' strategy profile g.

In dynamic games, each agent has his own objective and chooses his strategy g^i so as to maximize his objective. Thus, in contrast to dynamic teams, in dynamic games strategy g^i is agent i's private information and not known to other agents. Therefore, in dynamic games, each agent needs to form a prediction about the other agents' strategies. We denote this prediction by $g^{*1:N}_{1:T}$ to distinguish it from the strategy profile $g^{1:N}_{1:T}$ that is actually being played by the agents. Following Nash's idea, we assume that agents share a common prediction g^* about the actual strategy g. We would like to emphasize that the prediction g^* does not necessarily coincide with the actual strategy g. As we point out later, one requirement of an equilibrium of a game is that for every agent $i \in \mathcal{N}$, the prediction g^{*i} must be an optimal strategy for him given the other agents prediction strategy g^{*-i}.

Since an agent's actual strategy, say agent i, is his own private information, it is possible that agent i's actual strategy g^i is different from the prediction g^{*i}. Below we discuss the implication of an agent's deviation from the prediction strategy profile g^*. For that matter, we first consider an agent who may want to deviate from g^*, and then we consider an agent who faces such a deviation and his response.

In dynamic games, when agent $i \in \mathcal{N}$ chooses his strategy g^i, he needs to know how other agents will play for any choice of g^i which can be different from the prediction g^{*i}. Therefore, the prediction g^* has to be defined at all the information sets of every agent, those that have positive probability under g^* as well as those that have zero probability under g^*. Using the prediction g^*, any agent, say agent i, can form an appraisal about the future of the game for any strategy choice g^i, and evaluate the performance of g^i.[5]

By the same rationale, when agent i chooses g^i he needs to determine his strategy for all of his information sets, even those that have zero probability under g^{*-i}. This is because it is possible that some agent $j \in \mathcal{N}$ may deviate from g^{*j} and play a strategy g^j that is different from the prediction g^{*j}. Agent i must foresee these possible deviations by other agents and determine his response to these deviations.

[5]We note that this is not an issue in dynamic teams. In dynamic teams, agents coordinate in advance their choice of strategy profile g, and no agent has an incentive to (privately) deviate from it. Hence, the agents' strategy profile g is only needed to be defined on information sets of positive probability under g.

To determine his optimal strategy g^i at any information set, agent i needs to first form an appraisal about the history of the game at t as well as an appraisal about the future of the game using the strategy prediction g^{*-i}. For an information set h_t^i that is *compatible* with the prediction g^{*-i} given his strategy g^i at $t \in \mathcal{T}$ (*i.e.* h_t^i has positive probability of being realized under g^*), agent i can use Bayes' rule to derive the appraisal about the history of the game at t. However, for an information set h_t^i that has zero probability under the prediction g^{*-i} given g^i, agent i cannot anymore rely on the prediction g^* and use Bayes' rule to form his appraisal about the history of the game at t. The realization of history h_t^i tells agent i that his original prediction $g_{1:t-1}^{*-i}$ is not (completely) correct, thus, he needs to revise his original prediction $g_{1:t-1}^{*-i}$ and to form a revised appraisal about the history of the game up to t. Therefore, agent i must determine how to form/revise his appraisal about the history of the game for every realization $h_t^i \in \mathcal{H}_t^i$, $t \in \mathcal{T}$, that has zero probability under g^{*-i}. We note that upon reaching an information set of measure zero, agent i only revises his prediction $g_{1:t-1}^{*-i}$ about other agents' past strategies, but does not change his prediction $g_{t:T}^{*-i}$ about their future strategies. This is because we assume that at equilibrium, the prediction $g_{t:T}^{*-i}$ specifies a set of strategies for other agents that are optimal in the continuation game that takes place after the realization of the information set of zero probability under $g_{1:t-1}^*$.[6]

We now describe how we can formalize the above issues that we need to consider in the study of dynamic games with asymmetric information. Following the game theory literature [43], agents' appraisals about the history and future of the game can be captured by an *assessment* that all agents commonly hold about the game. We

[6]We note that in dynamic teams, agents only need to determine their optimal strategy g for information sets that have positive probability of realizations under g. As a result, a collective choice of strategy is optimal at every information set with positive probability if and only if it maximizes the (expected) discounted utility of the team from $t = 1$ up to T. However, in dynamic games agents need to determine their strategies for all information sets irrespective of whether they have zero or positive probability of realization under g^*. Therefore, if a choice of strategy g^i maximizes agent i's (expected) discounted utility from $t = 1$ to $T = 1$, it does not imply that it is also optimal at all information sets that have zero probability of realization under $\{g^{*-i}, g^i\}$. Consequently, unlike team problems, in dynamic games a choice of agent i's strategy must be optimal for all continuation games that follow after a realization of an information set h_t^i irrespective of whether it has zero or positive probability of realization.

define an *assessment* as a pair of mappings (g^*, μ), where $g^* := \{g_t^{*i}, i \in \mathcal{N}, t \in \mathcal{T}\}$, $g_t^{*i} : \mathcal{H}_t^i \to \Delta(\mathcal{A}_t^i)$ denotes a *prediction* about agent i's strategy at t, and $\mu := \{\mu_t^i, i \in \mathcal{N}, t \in \mathcal{T}\}$, $\mu_t^i : \mathcal{H}_t^i \to \Delta(\mathcal{X}_t \times \mathcal{H}_t^{-i})$, denotes agent i's *belief* about the system state X_t and agents $-i$'s information H_t^{-i} given his information H_t^i. The collection of mappings $\mu := \{\mu_t^i, i \in \mathcal{N}, t \in \mathcal{T}\}$ is called a *belief system.* For every $i \in \mathcal{N}$, $t \in \mathcal{T}$, and $h_t^i \in \mathcal{H}_t^i$, $\mu_t^i(h_t^i)$ denotes agent i's belief about the history $\{X_t, H_t^{-i}\}$ of the game, and $g_{t:T}^{*-i}$ denotes agent i's prediction about all other agents' continuation strategy from t onward. We note that $\mu_t^i(h_t^i)$ determines agent i's appraisal about the history of the game when h_t^i has either positive or zero probability under g^*. Therefore, using an assessment (g^*, μ) each agent can fully construct at any $t \in \mathcal{T}$ appraisals about the history and future of the game.

Using the definition of an assessment, we can extend the idea of Nash equilibrium to dynamic games with asymmetric information. An equilibrium of the dynamic game is defined as a common assessment (g^*, μ) among the agents that satisfies the following conditions under the assumption that the agents are rational. (1) Agent $i \in \mathcal{N}$ chooses his strategy $g_{1:T}^i$ so as to maximize his total expected utility (5.4) in all continuation games given the assessment (g^*, μ) about the game. Therefore, the prediction $g_{1:T}^{*i}$ that other agents hold about agent i's strategy must be a maximizer of his total expected utility under the assessment (g^*, μ). (2) For all $t \in \mathcal{T}$, agent i's, $i \in \mathcal{N}$, belief $\mu_t^i(h_t^i)$ at information set $h_t^i \in \mathcal{H}_t^i$ that has positive probability under g^*, must be equal to the conditional probability distribution of $\{X_t, H_t^{-i}\}$ given the realization h_t^i via Bayes' rule when agents $-i$ play according to $g_{1:t}^{*-i}$. When h_t^i has zero probability under the assessment g^*, the belief $\mu_t^i(h_t^i)$ cannot be determined via Bayes' rule and must be revised. The revised belief must satisfy a certain set of *"reasonable"* conditions so as to be compatible with agent i's rationality. Various sets of conditions have been proposed in the literature (see [43, 96]) to capture the notion of "reasonable" beliefs that are compatible with the agents' rationality. Different sets of conditions for off-equilibrium beliefs $\mu_t^i(h_t^i)$ result in the different equilibrium concepts that are proposed for dynamic games with asymmetric information.

In this chapter, we consider Perfect Bayesian Equilibrium (PBE) as the equilibrium solution concept. In the next section we provide the formal definition of

PBE.

5.4 Perfect Bayesian Equilibrium

Perfect Bayesian Equilibrium (PBE) is a solution concept that has been widely used in the economic literature for the study of dynamic games with asymmetric information. The formal definition of PBE for dynamic games in extensive form can be found in [96]. In this chapter, we use a state space representation for dynamic games instead of an extensive game form representation, therefore, we need to adapt the definition of PBE to this representation. We define a PBE as an assessment (g^*, μ) that satisfies the *sequential rationality* and *consistency* conditions. The sequential rationality condition requires that for all $i \in \mathcal{N}$, the prediction g^{*i} is optimal for agent i given the assessment (g^*, μ). The consistency condition requires that for all $i \in \mathcal{N}$, $t \in \mathcal{T}$, and $h_t^i \in \mathcal{H}_t^i$, agent i's belief $\mu(h_t^i)$ must be compatible with prediction g^*. We formally define these conditions below.

Let $\mathbb{P}_{\mu_t^i}^{(g_{t:T}^{*-i}, g_{t:T}^{*i})}\{.|h_t^i\}$ denote the probability measure induced by the stochastic process that starts at time t with initial condition $\{X_t, P_t^{-i}, p_t^i, c_t\}$, $h_t^i = \{c_t, p_t^i\}$, where $\{X_t, P_t^{-i}\}$ is distributed according to probability distribution $\mu_t^i(h_t^i)$, assuming that agents i and $-i$ take actions according to strategies $g_{t:T}^{*i}$ and $g_{t:T}^{*-i}$, respectively. In the sequel, to save some notation, we write $\mathbb{P}_{\mu}^{g^*}\{.\}$ instead of $\mathbb{P}_{\mu_t^i}^{(g_{t:T}^{*-i}, g_{t:T}^{*i})}\{.\}$ whenever there is no confusion.

Definition 5.1 (Sequential rationality). *We say that an assessment (g^*, μ) is sequentially rational if $\forall i \in \mathcal{N}$, $t \in \mathcal{T}$, and $h_t^i \in \mathcal{H}_t^i$, the strategy prediction $g_{t:T}^{*i}$ is a solution to*

$$\sup_{g_{t:T}^i} \mathbb{E}_{\mu_t^i}^{(g_{t:T}^{*-i}, g_{t:T}^{i})} \left\{ \sum_{\tau=t}^{T} u_t^i(X_t, A_t) | h_t^i \right\} \tag{5.5}$$

The sequential rationality condition (5.5) requires that, given the assessment (g^*, μ), the prediction strategy g_t^{i*} for agent i is an optimal strategy for him for all continuation games after history realization $h_t^i \in H^i$, irrespective of whether

h_t^i has positive or zero probability under (g^*, μ). That is, the common prediction g^{*i} about agent i's strategy must be an optimal strategy choice for him since it is common knowledge that he is a rational agent. We note that the sequential rationality condition defined above is more restrictive than the optimality condition for Bayesian Nash Equilibrium (BNE) which only requires (5.5) to hold at $t = 1$. By the sequential rationality condition, we require the optimality of prediction g^* even along off-equilibrium paths, and thus, we rule out the possibility of *non-credible threats*. For example, an agent can threaten to play an action that is suboptimal for himself upon a realization of a history that has zero probability under g^*. Such a non-credible threat is not ruled out by considering only the optimality at $t = 1$ (see [43] for more discussion).

The sequential rationality condition results in a set of constraints that the strategy prediction g^* must satisfy given a belief system μ. As we argued in Section 5.3, the belief system μ must be also compatible with the strategy prediction g^*. The following *consistency* condition captures such compatibility between the belief system μ and the prediction g^*.

Definition 5.2 (Consistency). *We say that an assessment (g^*, μ) is consistent if*

i) For all $i \in \mathcal{N}$, $t \in \mathcal{T}\backslash\{1\}$, $h_{t-1}^i \in \mathcal{H}_{t-1}^i$, and $h_t^i \in \mathcal{H}_t^i$ such that $\mathbb{P}_\mu^{g^*}\{h_t^i | h_{t-1}^i\} > 0$, the belief $\mu_t^i(h_t^i)$ must satisfy Bayes' rule, i.e.

$$\mu_t^i(h_t^i)(x_{1:t}, p_t^{-i}) = \frac{\mathbb{P}^{g^*}\{h_t^i, x_t, p_t^{-i} | h_{t-1}^i\}}{\mathbb{P}^{g^*}\{h_t^i | h_{t-1}^i\}}. \tag{5.6}$$

ii) For all $i \in \mathcal{N}$, $t \in \mathcal{T}\backslash\{1\}$, $h_{t-1}^i \in \mathcal{H}_{t-1}^i$, and $h_t^i \in \mathcal{H}_t^i$ such that $\mathbb{P}_\mu^{g^*}\{h_t^i | h_{t-1}^i\} = 0$, we have

$$\mu_t^i(h_t^i)(x_{1:t}, p_t^{-i}) > 0$$

only if there exists an open loop strategy $(A_{1:t-1}^{-i} = \hat{a}_{1:t-1}^{-i}, A_{1:t-1}^i = a_{1:t-1}^i)$ such *that*

$$\mathbb{P}^{(A^{-i}_{1:t-1}=\hat{a}^{-i}_{1:t-1},A^{i}_{1:t-1}=a^{i}_{1:t-1})}\{x_t, p^{-i}_t\} > 0. \tag{5.7}$$

The above consistency condition places a restriction on the belief system μ so that it is compatible with the strategy prediction g^*. For information sets along equilibrium paths, *i.e.* $\mathbb{P}^{g^*}_{\mu^i_1}\{h^i_t\} > 0$, $\mu^i_t(h^i_t)$ must be updated according to (5.6) via Bayes' rule since the observations of agent i are consistent with the prediction g^*. For information sets along off-equilibrium paths, *i.e.* $\mathbb{P}^{g^*}_{\mu^i_1}\{h^i_t\} = 0$, agent i needs to revise his belief about the strategy of agents $-i$ as the realization of h^i_t indicates that some agent has deviated from prediction $g^{*-i}_{1:t}$. As pointed out before, the revised belief $\mu^i(h^i_t)$ must be "reasonable". Definition 5.2 provides a set of such "reasonable" conditions captured by (5.6) and (5.7) that we discuss further below.

First, consider an information set h^i_t along an off-equilibrium path such that $\mathbb{P}^{g^*}_{\mu^i_{t-1}}\{h^i_t|h^i_{t-1}\} > 0$. That is, conditioned on reaching information set h^i_{t-1} at $t-1$, h^i_t has a positive probability of realization under the prediction strategy g^*. Since $\mathbb{P}^{g^*}_{\mu^i_1}\{h^i_t\} = \mathbb{P}^{g^*}_{\mu^i_{t-1}}\{h^i_t|h^i_{t-1}\}\mathbb{P}^{g^*}_{\mu^i_1}\{h^i_{t-1}\}$ and $\mathbb{P}^{g^*}_{\mu^i_1}\{h^i_t\} = 0$, we have $\mathbb{P}^{g^*}_{\mu^i_1}\{h^i_{t-1}\} = 0$. Therefore, h^i_{t-1} is also an information set along an off-equilibrium path, and $\mu^i(h^i_{t-1})$ is a revised belief that agent i holds at $t-1$. Note that if the assessment (g^*, μ) satisfies the sequential rationality condition, g^* is a best response for all agents in all continuation games that follow the realization of every information set of positive or zero probability. Moreover, since $\mathbb{P}^{g^*_{t-1}}_{\mu^i_{t-1}}\{h^i_t|h^i_{t-1}\} > 0$, the realization of h^i_t conditioned on reaching h^i_{t-1} is consistent with the strategy prediction g^*_{t-1}. Therefore, agent i does not have any reason to further revise his belief about agents $-i$'s strategy beyond the revision that results in $\mu^i(h^i_{t-1})$. As a result, agent i determines his belief $\mu^i(h^i_t)$ by utilizing his belief $\mu^i(h^i_{t-1})$ at $t-1$ and updating it via Bayes' rule assuming that agents $-i$' play according to the prediction g^{*-i}_{t-1} (see (5.6) in part (i)).

Next, consider an information set h^i_t along an off-equilibrium path such that $\mathbb{P}^{g^*}_{\mu^i_t}\{h^i_t|h^i_{t-1}\} = 0$. That is, conditioned on reaching information set h^i_{t-1} at $t-1$, h^i_t has a zero probability of realization under the prediction g^*. In this case, the realization of h^i_t indicates that agents $-i$ have deviated from prediction $g^{*-i}_{1:t-1}$, and this deviation has not been detected by agent i before. Therefore, agent i needs to form a new belief on agents $-i$'s private information P_t^{-i} and the state X_t by revising $\mu(h_t^i)$. Part (ii) of the consistency condition concerns such belief revisions and requires that the support of agent i's revised belief $\mu(h_t^i)$ includes only the states and private information that are feasible under the system and information dynamics (5.1-5.2), that is, they are reachable under some open-loop control strategy $(A_{1:t-1}^{-i} = \hat{a}_{1:t-1}^{-i}, A_{1:t-1}^{i} = a_{1:t-1}^{i})$. We note that since we are using a state representation of the dynamic game, we need to impose such a requirement, whereas in the equivalent extensive form representation of the game such a requirement is satisfied by the construction of the game-tree.

Remark 5.1. *Under Assumptions 5.2 and 5.3, we have* $\mathbb{P}_{\mu_1}^{(A_{1:t-1}=\hat{a}_{1:t-1})}\{x_{1:t}, p_t^{-i}\} > 0$ *for all* $(A_{1:t-1} = \hat{a}_{1:t-1})$*. Therefore part (ii) of the consistency conditions is trivially satisfied. In the rest of the chapter, we ignore part (ii) and only consider part (i) of the consistency condition. In Section 5.8, we discuss the case where we relax Assumptions 5.2 and 5.3.*

We can now provide the formal definition of PBE for the dynamic game of Section 5.2.

Definition 5.3. *An assessment* (g^*, μ) *is called a PBE if it satisfies the sequential rationality and consistency conditions.*

The definition of Perfect Bayesian equilibrium provides a general formalization of outcomes that are *rationalizable* (*i.e.* consistent with agents' rationality) under some strategy profile and belief system. However, in the following we argue that there are computational and philosophical reasons that motivate us to define a sub class of PBEs that provide a *simpler* and more *tractable* approach to characterizing the outcomes of dynamic games with asymmetric information.

There are two major challenges in computing a PBE (g^*, μ). First, there is an inter-temporal coupling between the agents' strategy prediction g^* and belief system μ. According to the consistency requirement, the belief system μ has to satisfy a set of conditions given a strategy prediction g^*. On the other hand, by sequential rationality, a strategy prediction g^* must satisfy a set of optimality condition given belief system μ. Therefore, there is a circular dependency between a prediction strategy g^* and a belief system μ over time. For instance, by sequential rationality, agent i's strategy g_t^{i*} at time t depends on the agents' future strategies $g_{t:T}^*$ and on the agents' past strategies $g_{1:t-1}^*$ indirectly through the consistency condition for μ_t^i. As a result, one needs to determine the strategy prediction g^* and belief system μ simultaneously for the whole time horizon so as to satisfy the sequential rationality and consistency conditions, and thus, cannot sequentially decompose the computation of PBE over time. Second, the agents' information h_t^i, $i \in \mathcal{N}$, has a growing domain over time. Hence, the agents' strategies have growing domains over time, and this feature further complicates the computation of PBEs of dynamic games with asymmetric information.

The definition of PBE requires agents to keep track of all the observations they acquire over time and form beliefs about the private information of all other agents. As we show in the next section, agents do not need to keep track of all of their past observations to reach an equilibrium. They can take into account fewer variables for decision making purposes and ignore part of their past observations that are not *relevant* to the continuation game at any time. As we argue in Section 5.7.2, the class of simpler strategies proposed in this chapter characterize a more plausible prediction about the outcome of the interaction among agents when the underlying system is highly dynamic and there exists considerable information asymmetry among the agents.

5.5 The Common Information Approach

We generalize the notion of Common Information Based PBE (CIB-PBE), first introduced in [98, 97], and characterize a class of PBEs that utilize strategy choices that are simpler than general behavioral strategies as they require agents to keep track of only a compressed version of their information over time. We proceed as follows. In Section 5.5.1 we provide sufficient conditions for the subset of private information an agent needs to keep track of it over time for decision making purposes. In Section 5.5.2, we introduce the common information based belief as a compressed version of the agents' common information that is sufficient for decision making purposes. Based on these compressions of the agents' private and common information, we introduce the notion of common information based assessments and common information based perfect Bayesian equilibrium in Sections 5.5.3 and 5.5.4, respectively.

5.5.1 Sufficient Private Information

The key ideas for compressing an agent's private information appear in Definitions 5.4 and 5.5 below. We first characterize the subset of an agent's private information that is necessary for the agent's decision making process, irrespective of what strategies other agents play over time.

Definition 5.4 (Private payoff-relevant information). *Let $P_t^{i,pr} = \bar{l}_t^i(P_t^i, C_t)$ denote a private signal that agent $i \in \mathcal{N}$ forms at $t \in \mathcal{T}$ based on his private information P_t^i and common information C_t. We say $P_t^{i,pr}$ is a private payoff-relevant information for agent i if for all open-loop strategies $(A_{1:T}^{1:N} = \hat{a}_{1:T}^{1:N})$, and for all $t \in \mathcal{T}$,*

(i) it can be updated recursively as $P_t^{i,pr} = \zeta_t(P_{t-1}^{i,pr}, H_t^i \backslash H_{t-1}^i)$,

(ii) it satisfies

$$\mathbb{P}^{(A_{1:T}^{1:N}=\hat{a}_{1:T}^{1:N})}\left\{P_{t+1}^{i,pr}\middle|P_t^i, C_t, a_t\right\} = \mathbb{P}^{(A_{1:T}^{1:N}=\hat{a}_{1:T}^{1:N})}\left\{P_{t+1}^{i,pr}\middle|P_t^{i,pr}, C_t, a_t\right\} \quad w.p.1 \quad (5.8)$$

(iii) for all realizations $\{c_t, p_t^i\} \in \mathcal{C}_t \times \mathcal{P}_t^i$ such that $\mathbb{P}^{(A_{1:T}^{1:N}=\hat{a}_{1:T}^{1:N})}\{c_t, p_t^i\} > 0$,

$$\mathbb{E}^{(A_{1:t-1}^{1:N}=\hat{a}_{1:t}^{1:N})}\left\{u_t^i(X_t, A_t)\middle|c_t, p_t^i, a_t\right\} = \mathbb{E}^{(A_{1:t-1}^{-i}=\hat{a}_{1:t}^{-i})}\left\{u_t^i(X_t, A_t)\middle|c_t, p_t^{i,pr}, a_t\right\}. \quad (5.9)$$

By assuming that all other agents play open loop strategies we remove the interdependence between agents $-i$'s strategy choices and agent i's information structure, thus, we eliminate *signaling* among the agents. Fixing the open-loop strategies of agents $-i$, agent i faces a centralized stochastic control problem. Definition 5.4 says that $P_t^{i,pr}$, $t \in \mathcal{T}$, is a private payoff-relevant information for agent i if (i) it can be recursively updated, (ii) $P_t^{i,pr}$ includes all information in P_t^i that is relevant to $P_{t+1}^{i,pr}$ and (iii) agent i's instantaneous conditional expected utility at any $t \in \mathcal{T}$ is only a function of C_t, $P_t^{i,pr}$, and his action A_t^i at t. These three conditions are similar to properties (1)-(3), described in Section 5.1, that define an information state for a centralized stochastic control problem [76], but they concern only agent i's private information P_t^i instead of the collection $H_t^i = \{C_t, P_t^i\}$ of his private and common information.[7] We would like to point out that conditions (i)-(iii) can have many solutions including the trivial solution $P_t^{i,pr} = P_t^i$. [8]

While the definition of private payoff-relevant information suggests a possible way to compress the information required for an agent's decision making process, it assumes that other agents play open-loop strategies and do not utilize the information they acquire in real-time for decision making purposes. However, open-loop strategies are not in general optimal for agents $-i$. As a result, to evaluate the performance of any strategy choice g^i agent i needs also to form a belief about the information that other agents utilize to make decisions.

Definition 5.5 (Sufficient private information)**.** *We say $S_t^i = l_t^i(P_t^i, C_t)$, $i \in \mathcal{N}$, $t \in \mathcal{T}$, is sufficient private information for agents if,*

(i) it can be updated recursively as $S_t^i = \phi_t(S_{t-1}^i, H_t^i \backslash H_{t-1}^i)$ for $t \in \mathcal{T} \backslash \{1\}$,

(ii) it satisfies

$$\mathbb{P}\left\{S_{t+1}, Z_{t+1} \middle| P_t, C_t, A_t\right\} = \mathbb{P}\left\{S_{t+1}, Z_{t+1} \middle| S_t, C_t, A_t\right\} \quad w.p.1, \tag{5.10}$$

(iii) for every strategy profile $\tilde{g}^* : \{\tilde{g}_t^{*i} : \mathcal{S}_t^i \times \mathcal{C}_t \to \Delta(\mathcal{A}_t^i), i \in \mathcal{N}, t \in \mathcal{T}\}$ and $a_t \in \mathcal{A}_t$,

[7] We note that if we interpret a centralized control problem as a special case of our model where $N = 1$, $H_t^1 = P_t$ and $C_t = \emptyset$ for all $t \in \mathcal{T}$, Definition 5.4 coincides with the definition of information state for the single agent decision problem.

[8] An interesting research direction is to determine whether a minimal private payoff-relevant information exists, and if so, characterize such a minimal payoff-relevant information. However, such a direction is beyond the scope of this chapter, and we leave this topic for future research.

$t \in \mathcal{T}$,

$$\mathbb{E}^{\tilde{g}^{*-i}_{1:t-1}}\left\{u^i_t(X_t, A_t)\Big|c_t, p^i_t, a_t\right\} = \mathbb{E}^{\tilde{g}^{*-i}_{1:t-1}}\left\{u^i_t(X_t, A_t)\Big|c_t, s^i_t, a_t\right\}, \tag{5.11}$$

for all realizations $\{c_t, p^i_t\} \in \mathcal{C}_t \times \mathcal{P}^i_t$.

(iv) for every strategy profile $\tilde{g}^* : \{\tilde{g}^{*i}_t : \mathcal{S}^i_t \times \mathcal{C}_t \to \Delta(\mathcal{A}^i_t), i \in \mathcal{N}, t \in \mathcal{T}\}$ *and* $a_t \in \mathcal{A}_t$, $i \in \mathcal{N}$, *and* $t \in \mathcal{T}$,

$$\mathbb{P}^{\tilde{g}^{*-i}_{1:t-1}, g^{*i}_{1:t-1}}\left\{S^{-i}_t\Big|P^i_t, C_t\right\} = \mathbb{P}^{\tilde{g}^{*-i}_{1:t-1}}\left\{S^{-i}_t\Big|S^i_t, C_t\right\} \quad w.p.1, \tag{5.12}$$

In general, the sufficient private information $S^{1:N}_t$ is more restrictive than that of private payoff relevant information $P^{1:N,pr}_t$. This is because $S^{1:N}_t$, $t \in \mathcal{T}$, needs to satisfy condition (iv) in addition to conditions (i)-(iii). Moreover, condition (5.10) requires that the belief about Z_{t+1} conditioned on $\{S_t, C_t, A_t\}$ must be independent of $\{H_t, A_t\}$. Furthermore, in contrast to condition (5.9) that assumes that agents $-i$ play open-loop strategies, condition (5.11) must be satisfied when agents play closed-loop strategies. We note the definition of $S_{1:T}$ provides an interdependence among agents' sufficient private information $S^{1:N}_{1:T}$ through condition (iii) and (iv). Specifically, by condition (iv) agent i's sufficient private information S^i_t must be rich enough so that he can form beliefs about agents $-i$'s sufficient private information S^{-i}_t. Note that in (5.11) and (5.12) the conditional probability distributions do depend on the strategy prediction g^*. As we pointed out in Section 5.4, the agents' actual strategy profile g may be different from the prediction g^*. We will discuss the robustness of sufficient private information to possible unilateral deviations of agents from g^* in Section 5.6. We would like to point out that conditions (i)-(iv) of Definition 5.5 can have many solutions including the trivial solution $S^i_t = P^i_t$.[9]

Definition 5.5 provides sufficient conditions under which agents can compress their private information in a "mutually consistent' manner. Below, we discuss a

[9]We do not discuss the possibility of finding a minimal set of sufficient private information in this chapter, and leave it for future research as such investigation is beyond the scope of this chapter.

few special instances of the general model 5.2, and identify the sufficient private information $S^{1:N}_{1:T}$.

Special Cases:

1) Nested information structure: The uninformed agent (agent 2) has no private information, $P^2_t = \emptyset$. Thus, $S^2_t = \emptyset$. For the informed agent (agent 1) consider $P^{1,pr}_t = X_t$. Consequently, we can set $S^1_t = X_t$. Note that $P^2_t = \emptyset$, thus, the uninformed agent's belief about P^1_t is the same as common belief.

2) Independent dynamics with observable actions: Consider $P^{i,pr}_t = X^i_t$. Note that X^j_t, $j = 1, 2, ..., N$ have independent dynamics given the collective action A_t that is commonly observable by all agents. Therefore, agent i's belief about X_j, $j \neq i$, is the same as common belief, $P^g\{X^j_t | P^i_t, C_t\} = \mathbb{P}^g\{X^j_t | C_t\}$. Consequently, we can set $S^i_t = X^i_t$.

3) Delayed sharing information structure: Consider $P^{i,pr}_t = \{Y^i_{t-d+1:t}\}$. Since we do not assume any specific structure for system dynamics and agents' observations, agent i's complete private information P^i_t is payoff-relevant for him. Therefore, we set $S_t = P^i_t$.

4) Perfectly controlled dynamics with hidden actions: Since agent i, $i \in \mathcal{N}$, perfectly controls X^i_t over time $t \in \mathcal{T}$, we set $S^i_t = \{A^i_{t-1}, Y^i_t\}$.

5.5.2 Common Information based Belief

Based on the characterization of sufficient private information, we present a statistic (compressed version) of the common information C_t that agents need to keep track of over time for decision making purposes.

Consider the sufficient private information $S^{1:N}_t$, $t \in \mathcal{T}$. Define $\mathcal{S}^i_t$ to be the set of all possible realizations of S^i_t, and $\mathcal{S}_t := \prod_{i=1}^N \mathcal{S}^i_t$. Let $\gamma_t : \mathcal{C}_t \to \Delta(\mathcal{X}_t \times \mathcal{S}_t)$ denote a mapping that determines a conditional probability distribution over the system state X_t and the agents' sufficient private information S_t given the common information C_t at time t. We call the collection of mappings $\gamma := \{\gamma_t, t \in \mathcal{T}\}$ a *common information based belief system* (CIB belief system). Note that γ_t is only a function of the common information C_t, and thus, it is computable by all agents. Let $\Pi_t^\gamma := \gamma_t(C_t)$ denote the (random) common information based belief that agents hold under belief system γ at t. We can interpret Π_t^γ as the common belief that each agent holds about the system state X_t and all the agents' (including himself) sufficient private information S_t at time t. In the rest of the chapter, we write Π_t and drop the superscript γ whenever such a simplification in notation is clear.

We show below that using the sufficient private information S_t along with the CIB belief Π_t, we can form a common information based assessment about the game. We prove that such a common information based assessment is rich enough to capture a subset of PBE.

5.5.3 Common Information based Assessment

As we discussed in Section 5.4, to form a prediction about the game we need to determine an assessment about the game that is sequentially rational and consistent. In the following we present an assessment that is based on the common information based belief Π_t and the sufficient private information S_t^i, $i \in \mathcal{N}$.

Consider a class of strategies that utilize the information given by (Π_t, S_t^i) for agent $i \in \mathcal{N}$ at time t. We call the mapping $\sigma_t^i : \Delta(X_t \times \mathcal{S}_t) \times \mathcal{S}_t^i \rightarrow \Delta(\mathcal{A}_t^i)$ a *Common Information Based (CIB) strategy* for agent i at time t. A CIB strategy σ^i determines a probability distribution for agent i's action A_t^i at time t given his information (Π_t, S_t^i). A CIB strategy is a behavioral strategy where agents only use the common belief $\Pi_t = \gamma_t(C_t)$ (instead of the common information C_t), and the sufficient private information $S_t^i(P_t^i, C_t)$ (instead of complete private information P_t^i). A collection of CIB strategies $\{\sigma_{1:T}^1, ..., \sigma_{1:T}^N\}$ is called a *CIB strategy profile* σ. The set of CIB strategies is a subset of behavioral strategies, defined in Section 5.2, as we can define,

$$g_t^{(\sigma,\gamma),i}(h_t^i) := \sigma_t^i(\pi_t^\gamma, s_t^i).$$

In Section 5.4, we defined a consistency condition between strategy prediction g^* and a belief system μ. Below, we provide an analogous consistency condition between a CIB strategy prediction σ^* and a CIB belief system γ.

Definition 5.6. *A pair (σ^*, γ) of a CIB strategy prediction profile σ^* and belief system γ satisfies the consistency condition if*

(i) for all $t \in \mathcal{T} \backslash \{1\}$[10], $z_t \in \mathcal{Z}_t$, $\pi_{t-1} = \gamma_{t-1}(c_{t-1})$, and $\pi_t = \gamma_t(\{c_{t-1}, z_t\})$ such that $\mathbb{P}^{\sigma_t^*}_{\pi_{t-1}}\{z_t\} > 0$, π_t must satisfy Bayes' rule, i.e.,

$$\pi_t(x_t, s_t) = \frac{\mathbb{P}^{\sigma_t^*}_{\pi_{t-1}}\{x_t, s_t, z_t\}}{\mathbb{P}^{\sigma_t^*}_{\pi_{t-1}}\{z_t\}}, \quad \forall x_t \in \mathcal{X}_t, \forall s_t \in \mathcal{S}_t, \tag{5.13}$$

(ii) for all $t \in \mathcal{T} \backslash \{1\}$, $c_{t-1} \in \mathcal{C}_{t-1}$, $\pi_{t-1} = \gamma_{t-1}(c_{t-1})$, $z_t \in \mathcal{Z}_t$, and $\pi_t = \gamma_t(\{c_{t-1}, z_t\})$ such that $\mathbb{P}^{\sigma_t^*}_{\pi_{t-1}}\{z_t\} = 0$, we have

$$\pi_t(x_t, s_t) > 0, \quad \forall x_t \in \mathcal{X}_t, \forall s_t \in \mathcal{S}_t,$$

only if there exists an open-loop strategy $(A_{1:t-1} = a_{1:t-1})$ such that $\mathbb{P}^{(A_{1:t-1}=a_{1:t-1})}_{\pi_1}$ $\{c_{t-1}, z_t\} > 0$, and

$$\mathbb{P}^{(A_{1:t-1}=a_{1:t-1})}_{\pi_1}\{x_t, s_t\} > 0, \tag{5.14}$$

(iii) for all $t \in \mathcal{T} \backslash \{1\}$, $c_{t-1} \in \mathcal{C}_{t-1}$, $\pi_{t-1} = \gamma_{t-1}(c_{t-1})$, $z_t \in \mathcal{Z}_t$, and $\pi_t = \gamma_t(\{c_{t-1}, z_t\})$ such that $\mathbb{P}^{\sigma_t^*}_{\pi_{t-1}}\{z_t\} = 0$, we have

$$\sum_{x_t \in \mathcal{X}_t} \pi_t(x_t, s_t) > 0, \quad \forall s_t \in \mathcal{S}_t$$

if there exists an open-loop strategy $(A_{1:t-1} = a_{1:t-1})$ such that $\mathbb{P}^{(A_{1:t-1}=a_{1:t-1})}_{\pi_1}$ $\{c_{t-1}, z_t\} > 0$, and

$$\sum_{x_t \in \mathcal{X}_t} \mathbb{P}^{(A_{1:t-1}=a_{1:t-1})}_{\pi_1}\{x_t, s_t\} > 0. \tag{5.15}$$

[10]For $t = 1$, Π_t is given by the conditional probability at $t = 1$ as $\Pi_1(x_1, s_1) := \frac{\mathbb{P}\{s_1|x_1, z_1\}}{\sum_{\hat{x}_1 \in \mathcal{X}_1} \mathbb{P}\{z_1|\hat{x}_1\}\eta(\hat{x}_1)}$.

Parts (i) and (ii) of Definition 5.6 follow from rationales similar to their analogues in Definition 5.2, and require a CIB belief system to satisfy a sets of constraints with respect to a CIB strategy profile that are similar to those for an assessment (g^*, μ). Definition 5.6 requires an additional condition described by part (iii). By (5.15), we require a CIB belief system γ consistent with the CIB strategy profile σ^* to assign a positive probability to every realization s_t of the agents' sufficient private information S_t that is "plausible" given the common information realization $c_t = \{c_{t-1}, z_t\}$; plausibility of s_t given c_t means that there exists an open-loop strategy profile $(A_{1:t-1} = a_{1:t-1})$ consistent with the realization c_t that leads to the realization of s_t with positive probability. Therefore, part (iii) ensures that there exists no incompatibility between the CIB belief Π_t and the agents' sufficient private information S_{t+1}. As we show later (Section 5.5.4), such a compatibility allows each agent to refine the CIB belief Π_t using his own private information S_t^i, and to form his private belief about the game.

Remark 5.2. *Similar to Remark 5.1, Assumptions 5.2 and 5.3 imply that (5.14) holds for all* $(A_{1:t-1} = a_{1:t-1})$ *such that* $\mathbb{P}^{(A_{1:t-1}=a_{1:t-1})}\{c_t\} > 0$*. Therefore, in the rest of the chapter, we ignore part (ii) of the consistency condition for CIB belief systems. Moreover, under Assumptions 5.2 and 5.3, condition (5.15) is always satisfied. Therefore, condition (iii) is equivalent to having* $\sum_{x_t \in \mathcal{X}_t} \pi_t(x_t, p_t) > 0$ *whenever* $\mathbb{P}^{\sigma^*}_{\pi_{t-1}}\{z_t\} = 0$*. In Section 5.8, we discuss the case where we relax Assumptions 5.2 and 5.3.*

Given a CIB strategy profile prediction σ^*, a consistent CIB belief must satisfy (5.13), which determines the CIB belief Π_t at t in terms of the CIB belief Π_{t-1} at $t-1$ and the new common information Z_t at t. We define a *CIB belief update rule* as a mapping $\psi_t : \Delta(\mathcal{X}_{t-1} \times \mathcal{S}_{t-1}) \times \mathcal{Z}_t \to \Delta(\mathcal{X}_t \times \mathcal{S}_t), t \in \mathcal{T}$ that determines recursively the common belief

$$\Pi_t^\psi := \psi_t(\Pi_{t-1}^\psi, Z_t), \tag{5.16}$$

as a function of new common observation Z_t at t and the common belief Π_{t-1}^ψ at

$t-1$.[11] The superscript ψ in Π_t^{ψ} indicates that the CIB belief Π_t^{ψ} is generated using the CIB update rule ψ. Let γ^{ψ} denote the common belief system that is equivalent to the CIB update rule ψ. We call a CIB belief update rule ψ consistent with a CIB strategy profile σ^* if the equivalent CIB belief system γ^{ψ} is consistent with σ^* (Definition 5.6).

Define a CIB assessment (σ^*, γ) as a pair of CIB strategy profile σ^* and a CIB belief system γ. Below, we show that a consistent CIB assessment (σ^*, γ) is equivalent to a consistent assessment (g^*, μ) as defined in Section 5.4 (Definition 5.2).

Lemma 5.1. *For any given CIB assessment (σ^*, γ), there exists an equivalent assessment (g^*, μ) of a behavioral strategy prediction g^* and belief system μ such that:*

i) the behavioral strategy g^* is defined by

$$g_t^{*i}(h_t^i) := \sigma_t^{*i}(\pi_t^{\gamma}, s_t^i); \tag{5.17}$$

ii) the belief system μ is consistent with g^* and satisfies

$$\mathbb{P}^{g^*}\left\{s_t^{-i}|h_t^i\right\} = \mathbb{P}\left\{s_t^{-i}|\pi_t, s_t^i\right\}, \tag{5.18}$$

for all $i \in \mathcal{N}$, $t \in \mathcal{T}$,$h_t^i \in \mathcal{H}_t^i$, and $s_t^{-i} \in \mathcal{S}_t^{-i}$.

Lemma 5.1 shows that the set of consistent CIB assessment (σ^*, γ) is equivalent to a subset of consistent assessments (g^*, μ). That is, using the CIB belief system γ and CIB strategy profile σ^*, agents can form a consistent assessment about the evolution of the game. Moreover, condition (5.18) implies that the CIB belief Π_t along with agent i's sufficient information S_t^i capture all the information in H_t^i that is relevant to agent i's belief about S_t^{-i}.

[11] Upon reaching an information set of measure zero (parts (ii) and (iii) of Definition 5.6), the revised CIB belief could be a function of $C_t = \{C_{t-1}, Z_t\}$, not only $\Pi_{t-1}(C_{t-1})$ and Z_t. Therefore, the set of CIB belief systems that are generated from CIB update rules is a subset of all consistent CIB belief systems given by Definition 5.6. However, we argue that upon reaching an information set of measure zero, it is more plausible to revise the CIB belief only as a function of relevant information $\Pi_{t-1}(C_{t-1})$ and Z_t; C_t is irrelevant given $\Pi_{t-1}(C_{t-1})$ and Z_t.

5.5.4 Common Information based PBE

Using the result of Lemma 5.1, we can define a class of PBE, called Common Information based PBE (CIB-PBE), as a set of equilibria for dynamic games with asymmetric information that can be expressed as CIB assessments.

Definition 5.7. *A CIB assessment* (σ^*, γ) *is called a CIB-PBE if* γ *is consistent with* σ^* *(Definition 5.6), and the equivalent consistent assessment* (g^*, μ)*, given by Lemma 5.1, is a PBE.*

In the following, we also call a consistent pair (σ^*, ψ) of a CIB strategy prediction profile σ^* and CIB belief update rule ψ a CIB-PBE if $(\sigma^*, \gamma^{\psi})$ is a CIB-PBE.

Throughout Section 5.6, we assume that agents play according to the strategy predictions g^* (or CIB strategy predictions σ^*). However, an agent's, say agent $i \in \mathcal{N}$'s, actual strategy g^i is his private information and could be different from g^* if such a deviation is profitable for him. The proposed class of CIB assessments imposes two restrictions on agents' strategies and beliefs compared to the general class of assessment presented in Section 5.4. First, it requires that each agent i, $i \in \mathcal{N}$, must play a CIB strategy σ^{*i} instead of a general behavioral strategy g^{*i}. Second, it requires that each agent i, $i \in \mathcal{N}$, must form a belief about the status of the game using only the CIB belief Π_t along with his sufficient private information S_t^i (instead of a general belief μ_t^i). A strategic agent $i \in \mathcal{N}$ does not restrict his choice of strategy to CIB strategies, and may deviate from σ^{*i} to a non-CIB strategy g^i if it is profitable to him. Moreover, a strategic agent i, $i \in \mathcal{N}$, does not limit himself to form belief about the current status of the game only based on Π_t and S_t^i, and may instead use a general belief μ^i if it enables him to improve his expected utility. In the next section, we address these strategic concerns, and show that no agent $i \in \mathcal{N}$ wants to deviate from (Π, σ^*) and play a non-CIB strategy g^i when all other agents are playing according to CIB assessment (Π, σ^*). This result allows us to focus on the class of CIB assessments, and develop a methodology to sequentially decompose the dynamic game over time.

5.6 Main Results

In this section, we show that the class of CIB assessments is rich enough to capture the agents' strategic interactions.We first show that when agents $-i$ play according to a CIB assessment (γ_t, σ^*), agent i, $i \in \mathcal{N}$, cannot mislead these agents by playing a strategy g^i different from σ^{*i}, thus, creating dual beliefs, one belief that is based on the CIB assessment (γ_t, σ^*) the functional form of which is known to all agents, and another belief that is based on his private strategy g^i that is only known to him (Theorem 5.1). Then, we show that given that agents $-i$ play CIB strategy σ^{*-i}, agent i's best response is a CIB strategy (Theorem 5.2).

We prove the result of Theorems 5.1 (resp. 5.2) for agent $i \in \mathcal{N}$ assuming that all other agents $-i$ are playing according to strategy prediction g^{*-i} (resp. σ^{*-1}). The same results hold for every continuation game that starts at any time $t \in \mathcal{T}$ along an off-equilibrium path; they can be proved by relabeling time t as time 1, and using the CIB belief $\pi_t = \gamma_t(c_t)$ and the corresponding $\mu_t^i(h_t^i)$, defined by Lemma 5.1, as the initial common belief for the continuation game.

Using the results of Theorems 5.1 and 5.2, we present a methodology to determine the set of CIB-PBEs of stochastic dynamic games with asymmetric information (Theorem 3). The proposed methodology leads to a sequential decomposition of stochastic dynamic games with asymmetric information. This decomposition gives rise to a dynamic program that can be utilized to compute CIB-PBEs via backward induction. We proceed to formally states these results. All the proofs can be found in the Appendix.

Theorem 5.1 (Policy-independence belief property)**.** *Consider a consistent assessment (g^*, μ). If agents $-i$ play according to strategy prediction g^{*-i}, then for every behavioral strategy g^i that agent i actually plays, we have*

$$\mathbb{P}^{g^*, g^i}\left\{x_t, h_t^{-i} \middle| h_t^i\right\} = \mathbb{P}^{g^{*-i}}\left\{x_t, h_t^{-i} \middle| h_t^i\right\}. \tag{5.19}$$

Theorem 5.1 provides a generalization of the standard *policy-independence belief* *property* in the centralized control literature [67] into multi-agent decision making problems. It states that under perfect recall agent i's belief is independent of his actual strategy g^i. Therefore, agent i cannot mislead agents $-i$ by deviating from the CIB strategy prediction g^{*i} to a behavioral strategy g^i so as to create dual beliefs (described above) that he can use to his advantage.

Using the result of Theorem 5.1, we show that agent $i \in \mathcal{N}$ does not gain by playing a non-CIB strategy $\tilde{g}^i$ when all other agents $-i$ are playing CIB strategies σ^{*-i}.

Theorem 5.2 (Closeness of CIB strategies)**.** *Consider a consistent CIB assessment (σ^*, γ). If every agent $j \in \mathcal{N}$, $j \neq i$ plays the CIB strategy σ^{*j}, then, there exists a CIB strategy σ^i for agent i that is a best response to σ^{*-i}.*

The results of Theorems 5.1 and 5.2 address the two restrictions (discussed above) imposed in CIB assessments on the agents' beliefs and strategies, respectively.

Based on these results, we restrict attention to CIB assessments, and provide a sequential decomposition of dynamic games with asymmetric information. A CIB-PBE is CIB assessment that is a fixed point under the best response map for all agents. Below, we formulate a dynamic program that enables us to compute CIB-PBEs of dynamic games with asymmetric information.

Consider a dynamic program over time horizon $\mathcal{T} \cup \{T+1\}$ with information state $\{\Pi_t, S_t\}$, $t \in \mathcal{T}$. Let $V_t := \{V_t^i : \Delta(\mathcal{X}_t \times \mathcal{S}_t) \times \mathcal{S}_t \rightarrow \mathbb{R}, i \in \mathcal{N}\}$ denote the value function that captures the continuation payoffs for all agents, for all realizations of the CIB belief Π_t and the agents' private sufficient information S_t, $t \in \mathcal{T}$. Set $V_{T+1}^i = 0$ for all $i \in \mathcal{N}$. For each stage $t \in \mathcal{T}$ of the dynamic program consider the following static game.

Stage game $\mathbf{G_t}(\pi_\mathbf{t}, \mathbf{V_{t+1}}, \psi_\mathbf{t+1})$: Given the value function V_{t+1} and CIB update rule ψ_{t+1}, we define the stage game $G_t(\pi_t, V_{t+1}, \psi_t)$ as a static game of asymmetric information among agents for every realization π_t. Each agent $i \in \mathcal{N}$ has private information S_t^i that is distributed according to π_t, which is common knowledge among the agents. Given a realization a_t of the agents' collective action profile and a realization s_t of the agents' sufficient private information, agent i's utility is given by

$$\bar{U}_t^i(a_t, s_t, \pi_t, V_{t+1}, \psi_{t+1}) := \mathbb{E}_{\pi_t}\left\{u_t^i(X_t, a_t) + V_{t+1}^i\left(\psi_{t+1}(\pi_t, Z_{t+1}), S_{t+1}\right)\middle|\pi_t, s_t, a_t\right\}. \tag{5.20}$$

BNE correspondence: We define the correspondence $BNE_t\left(V_{t+1}, \psi_{t+1}\right), t \in \mathcal{T}$, as the correspondence mapping that characterizes the set of BNEs of the stage game $\mathbf{G_t}(\pi_\mathbf{t}, \mathbf{V_{t+1}}, \psi_\mathbf{t+1})$ for every realization of π_t; this correspondence is given by

$$BNE_t\left(V_{t+1}, \psi_{t+1}\right) := \left\{\sigma_t^* : \forall \pi_t \in \Delta(\mathcal{X}_t \times \mathcal{S}_t), \sigma(\pi_t, \cdot) \text{ is a BNE of } G_t(\pi_t, V_{t+1}, \psi_{t+1})\right\}. \tag{5.21}$$

We say $\sigma^*(\pi_t, \cdot)$ is a BNE of the stage game $G_t(\pi_t, V_{t+1}, \psi_t)$ if for all agents $i \in \mathcal{N}$, and for all $s_t^i \in \mathcal{S}_t^i$,

$$\sigma^{*i}(\pi_t, s_t^i) \in \underset{\alpha \in \Delta(\mathcal{A}_t^i)}{\arg\max}\, \mathbb{E}_{\pi_t}\left\{\bar{U}_t^i((\alpha, \sigma^{*-i}(\pi_t, S_t^{-i})), S_t, \pi_t, V_{t+1}, \psi_{t+1})\middle|\pi_t, s_t^i\right\}. \tag{5.22}$$

Below, we provide a sequential decomposition of dynamic games with asymmetric information using the stage game and the BNE correspondence defined above.

Theorem 5.3 (Sequential decomposition)**.** *A pair (σ^*, ψ) of a CIB strategy profile σ^* and a CIB update rule ψ (equivalently, a CIB assessment $(\sigma^*, \gamma_t^\psi)$) is a CIB-PBE if (σ^*, ψ) solves the following dynamic program:*

$$V_{T+1}^i(\cdot) := 0 \qquad \forall\, i \in \mathcal{N}, \tag{5.23}$$

$$\tag{5.24}$$

for $t \in \mathcal{T}$,

$$\sigma_t^* \in BNE_t\left(V_{t+1}, \psi_{t+1}\right), \tag{5.25}$$

$$\psi_{t+1} \textit{ is consistent with } \sigma^*, \tag{5.26}$$

$$V_t^i(\pi_t, s_t) := \mathbb{E}^{\sigma_t^*}_{\pi_t,\psi_{t+1}}\left\{\bar{U}_t^i((\sigma^*(\pi_t, S_t)), S_t, \pi_t, V_{t+1}, \psi_{t+1})\middle|\pi_t, s_t^i\right\}, i \in \mathcal{N}. \tag{5.27}$$

5.7 Discussion

Using the notion of CIB-PBE proposed in Sections 5.5 and 5.6, we provide a sequential decomposition of dynamic games with asymmetric information over time. The set of CIB-PBEs characterizes a set of equilibria in which the agents compress their private and common information, thus, they play simpler strategies whose domains do not grow as rapidly as those of behavioral strategies. We also identified a special cases of the general model of Section 5.2, where the domain of CIB strategies is time-invariant. In the following, we elaborate further on the notion of CIB-PBE. In Section 5.7.1, we discuss the connections between the common information approach to dynamic games presented in this chapter and the existing results on the common information approach to dynamic teams. We show that the approach proposed in this chapter to compress the agents' private and common information extends to dynamic teams. In Section 5.7.2, we discuss the relation between CIB-PBEs and other equilibrium concepts for dynamic teams. We argue that the notion of CIB-PBE provides a more plausible and robust characterization of the outcomes in dynamic games where the information asymmetry among agents is high and the underlying system is highly dynamic.

5.7.1 Dynamic Games vs. Teams

As pointed out in Section 5.1, the approach proposed in this chapter is inspired by the common information approach proposed in [91, 90]. However, in dynamic games among strategic agents there are additional challenges that are not present in dynamic teams where agents cooperate. We discussed these challenges in Section 5.3, and presented an approach to address them in Sections 5.5 and 5.6. The approach proposed to compress the agents' private information (Definitions 5.4 and 5.5) can be also applied to dynamic team problems, thus, it provides a new result and insight on the common information approach to dynamic teams.

Consider a setting where each agent $i \in \mathcal{N}$ commits to play strategy g^i; that is, he is non-strategic and his strategy g^i is known to all the other agents. Moreover, assume that the agents have access to a *public randomization device*; that is at every time $t \in \mathcal{T}$ they observe a public random signal ω_t that is uniformly distributed on $[0, 1]$ and is independent across time. At each time $t \in \mathcal{T}$, all agents can condition their actions on the realization of ω_t. We can then establish the following result.

Theorem 5.4. *Assume that the agents are non-strategic and have access to a public randomization device. Then, for any strategy profile g there exists an equivalent CIB strategy profile σ that results in the same expected flow of utility, i.e.* $\mathbb{E}^g \left\{ \sum_{\tau=t}^{T} u_\tau^i(g_\tau^{1:N}(H_\tau^{1:N}), X_\tau) \right\} = \mathbb{E}^\sigma \left\{ \sum_{\tau=t}^{T} u_\tau^{1:N}(\sigma_\tau^i(\Pi_\tau, S_\tau^{1:N}), X_\tau) \right\}$, *for all $i \in \mathcal{N}$ and $t \in \mathcal{T}$.*

Consider the team problem that corresponds to the model of Section 5.2, where all agents have the same utility $u_t^i(\cdot, \cdot) = u_t^{\text{team}}(\cdot, \cdot)$. For this dynamic team problem, the result of Theorem 5.4 implies that we can restrict attention to the set of CIB strategies without loss of generality. The result of Theorem 5.4 extends the results of [91, 90] in two directions. First, it states that the restriction to the set of CIB strategies is without loss of generality, while the results of [91, 90] only states that this restriction is without loss of optimality. Second, the definition of CIB strategies first presented in [91, 90] requires the agents to use all of their private information P_t^i, $i \in \mathcal{N}$ (or all their private memory that is a predetermined function of their private information if they do not have perfect recall); the result of Theorem 5.4 holds for CIB strategies where the agents' private information is effectively compressed , thus, it generalizes the definition of CIB strategies. We would also like to note that to achieve a general expected flow of utility agents may need to utilize a public randomization device to construct correlated strategies. However, when our objective is to determine an optimal strategy profile, we can restrict attention to CIB strategies without a public randomization device; this is because in dynamic teams we can restrict attention to deterministic strategies without loss of optimality.

The results of Theorem 5.4 for dynamic teams along with the results in Section 5.6 for dynamic games, provide a general approach for the study of dynamic decision problems with many agents and asymmetric information. As a result, our results also extend to dynamic games among teams of agents and characterize a set of information states for each agent in a mutually consistent manner.

5.7.2 CIB-PBE vs. Other Notions of Equilibrium

In Section 5.6, we showed that CIB assessments proposed in this chapter are rich enough to capture a set of PBE. However, we would like to point out that the concept of CIB-PBE does not capture all PBEs of a dynamic game in general. We expand on the relation between the of CIB-PBE and PBE below. We argue that the set of CIB-PBEs are more plausible to arise as the information asymmetry among the agents increases and the underlying system is dynamic.

In Sections 5.5, we presented an approach to compress the agents' private and common information by providing conditions sufficient to characterize the information that is relevant for decision making purposes. Such information compression means that the agents do not incorporate into their decision making processes their observations that are irrelevant to the continuation game. As we showed in Section 5.7.1, this information compression is without loss of generality for dynamic team problems. However, this is not the case in dynamic games. In general, the set of CIB-PBEs of a dynamic game is a subset of all PBEs of that game. This is because in a dynamic game agents can incorporate their past irrelevant observations into their future decisions so as to create rewards (resp. punishments) that incentivize the agents to play (resp. not play) specific actions over time. By compressing the agents' private and common information in CIB assessments, we do not capture such punishment/reward schemes that are based on past irrelevant observations. Below, we present an example where there exists a PBE that cannot be captured as a CIB-PBE.

Consider a two-agent repeated game with $T = 2$ and a payoff matrix given in Table 5.1. At each stage, agent 1 chooses from $\{U, D\}$, and agent 2 chooses from $\{L, M, R\}$. We assume that agents observe each others' actions. Therefore, the agents have no private information, and the sufficient private information and CIB belief are trivial. The stage game has two equilibria in pure strategies given by (D, M) and (U, R). Using the results of Theorem 3, we can characterize four CIB-PBEs of the repeated game that correspond to the different combinations of the two equilibria of the stage game as follows: (DD, MM), (UU, RR), $DU, MR)$, and (UD, RM). However, there exists an another PBE of the repeated game that cannot be captured as a CIB-PBE. Consider the following equilibrium: *Play* (U, L) *at* $t = 1$. *If agent* 2 *plays* L *at* $t = 1$ *then play* (U, R)*; otherwise, play* (D, M) *at* $t = 2$. Note that the agent 1's decision at $t = 2$ depends on the agent 2's action at $t = 1$, which is a payoff-irrelevant information since the two stages of the game are independent.

	L	M	R
U	(8,3)	(0,2)	(2,10)
D	(0,1)	(1,2)	(0,0)

Table 5.1: Payoff matrix

We would like to point out that there are instances of dynamic games with asymmetric information, such as zero-sum dynamic games [109], where the equilibrium payoffs for the agents are unique. In these games it is not possible to incorporate payoff irrelevant information so as to construct additional equilibria where the agents' payoffs are different from the ones corresponding to CIB-PBEs; this is clearly the case for zero-sum games since the agents do not cooperate on creating punishment/reward schemes due to the zero-sum nature of the game.[12]

While it is true that in general, the set of PBEs of a dynamic game is larger than the set of CIB-PBEs of that game, in the remainder of this section, we provide three reasons on why in a highly dynamic environment with information asymmetry among agents, CIB-PBEs are more plausible to arise as an outcome of a game.

[12]See Section 5.9 for the proof of existence for CIB-PBEs in zero-sum games.

First, we argue that in the face of a highly dynamic environment, an agent with partial observations of the environment should not behave fundamentally different whether he interacts in a strategic or cooperative environment. From the single-agent decision making point of view (*i.e.* control theory), CIB strategies are the natural choice of an agent for decision making purposes (See Theorem 5.4). The notion of CIB-PBE proposed in this chapter for dynamic games along with the CIB approach to dynamic teams proposed in [91, 90] and extended in Section 5.7.1 provide a universal foundation for agents' behavior in a dynamic environment with information asymmetry among agents.

Second, we argue that in a highly dynamic environment with information asymmetry among the agents, the formation of punishment/reward schemes that utilize the agents' payoff-irrelevant information requires prior complex agreements among the agents; these complex agreements are sensitive to the parameters of the model and are not very plausible to arise in practice when the decision making problem for each agent is itself a complex task. We note that the set of PBEs that cannot be captured as CIB-PBEs are the ones that utilize payoff-irrelevant information to create punishment/reward schemes in the continuation game as in the example above. However, such punishment/reward schemes require the agents to form a common agreement among themselves on how to utilize such payoff-irrelevant information and how to implement such punishment/reward schemes. The formation of such a common agreement among the agents is more likely in games where the underlying system is not highly dynamic (as in repeated games [79]) and there is no much information asymmetry among agents. However, in a highly dynamic environment with information asymmetry among agents the formation of such common agreement becomes less likely for the following reasons. First, in those environments each agent's individual decision making process is described by a complex POMDP; thus, strategic agents are less likely to form a prior common agreement (that depend on the solution of the individual POMDPs) in addition to solving their individual POMDPs. Second, as the information asymmetry increases among agents, punishment/reward schemes that utilize payoff-irrelevant information require a complex agreement among the agents that is sensitive and not robust to changes in the assumptions on the information structure of the game. For instance, consider the example described above, but assume that agents observe imperfectly each others' actions at each stage (Assumption 5.3). Let $1-\epsilon$, $\epsilon \in (0,1)$, denote the probability that agents observe each others' actions perfectly, and ϵ denote the probability that their observation is different from the true action of the other agent. Then, the described non-CIB strategy profile above remains as a PBE of the game only if $\epsilon \leq \frac{1}{5}$. The author of [86] provides a general result on the robustness of above-mentioned punishment/reward schemes in repeated games; he shows that the set of equilibria that are robust to changes in information structure that affect only payoff-irrelevant signals does not include the set of equilibria that utilize punishment/reward schemes described above.

Third, the proposed notion of CIB-PBE can be viewed as a generalization of Markov Perfect Equilibrium [81] to dynamic games with asymmetric information. Therefore, a similar set of rationales that support the notion of MPE also applies to the notion of CIB-PBE as follows. First, the set of CIB assessments describe the simplest form of strategies capturing the agents' behavior that is consistent with the agents' rationality. Second, the class of CIB assessments captures the notion that "bygones are bygones", which also underlies the requirement of subgame perfection in equilibrium concepts for dynamic games. That is, the agents' strategies in two continuation games that only differ in the agents' information about payoff-irrelevant events must be identical. Third, the class of CIB assessments embodies the principle that "minor changes in the past should have minor effects". This implies that if there exists a small perturbation in the specifications of the game or the agents' past strategies that are irrelevant to the continuation game, the outcome of the continuation game should not change drastically. The two-step example above presents one such situation, where one equilibrium that is not CIB-PBE disappears suddenly as $\epsilon \to \frac{1}{5}$.

5.8 Extensions

In the model of Section 5.2, we presented a class of finite horizon dynamic games under Assumptions 5.1-5.3. Assumption 5.1 enables us to avoid measure-theoretic technical difficulties and to provide a simple but comprehensive presentation of the key ideas of the common information approach to dynamic games along with the notion of CIB-PBE. Assumptions 5.2 and 5.3 enable us to simplify the computation of beliefs along off-equilibrium paths. Below, we discuss Assumptions 5.1-5.3 and present extensions of our results. In Section 5.8.1, we extend our results to infinite horizon dynamic games with asymmetric information (*i.e.* $T = \infty$). In Section 5.8.2, we argue that when the agents' action are observable we can relax Assumptions 5.2 and 5.3, and obtain a result similar to that of Theorem 5.3. In Section 5.8.3, we discuss the assumption that is most crucial to the development of the common information approach to dynamic games proposed in this chapter.

5.8.1 Infinite Horizon Games

In the model of Section 5.2, we assume that the horizon T is finite. Below, we present a model similar to that of Section 5.2 with infinite horizon, *i.e.* $T = \infty$.

Infinite Horizon Dynamic Game: There are N strategic agents who live in a dynamic Markovian world over an infinite horizon. We consider a time-invariant model where the system state, actions, and observations spaces are finite and time-invariant, *i.e.* $\mathcal{X}_{\text{in}} = \mathcal{X}_t$, $\mathcal{A}_{\text{in}} = \mathcal{A}_t$, $\mathcal{Z}_{\text{in}} = \mathcal{Z}_t$, and $\mathcal{Y}_{\text{in}} = \mathcal{Y}_t$ for all $t \in \mathbb{N}$. Let $X_t \in \mathcal{X}_{\text{in}}$ denote the system state at $t \in \mathbb{N}$. Given the agents' actions A_t at t, the system state evolution is given by

$$X_{t+1} = f_{\text{in}}(X_t, A_t, W_t^x), \tag{5.28}$$

where $\{W_t^x, t \in \mathbb{N}\}$ is a sequence of independent and identically distributed random variables. The initial state X_1 is a random variable with probability distribution $\eta \in \Delta(\mathcal{X}_{\text{in}})$ with full support that is common knowledge among the agents.

At every time $t \in \mathbb{N}$, each agent $i \in \mathcal{N}$, receives a noisy observation Y_t^i given by

$$Y_t^i = O_{\text{in}}^i(X_t, A_{t-1}, W_t^i), \tag{5.29}$$

where $\{W_t^i, t \in \mathbb{N}, i \in \mathcal{N}\}$ is a sequence of independent and identically distributed random variables.

In addition, at every $t \in \mathcal{T}$ all agents receive a common observation $Z_t \in \mathcal{T}$ given by

$$Z_t = O^c_{\text{in}}(X_t, A_{t-1}, W^c_t), \tag{5.30}$$

where $\{W^c_t, t \in \mathbb{N}\}$ is a sequence of independent and identically distributed random variables; the sequences $\{W^x_t, t \in \mathbb{N}\}$, $\{W^c_t, t \in \mathbb{N}\}$, and $\{W^i_t, t \in \mathbb{N}, i \in \mathcal{N}\}$ and the initial state X_1 are mutually independent.

Similar to the model of Section 5.2, let P^i_t and C_t denote agent i's, $i \in \mathcal{N}$, private and common information at $t \in \mathcal{T}$, respectively. Agent i chooses his strategy so as to maximize his discounted (expected) utility given by

$$U^i_{\text{in}}(X, A) := \sum_{t\in\mathbb{N}} \delta^{t-1} u^i_{\text{in}}(X_t, A_t), \tag{5.31}$$

where δ denotes the discount factor. We assume that Assumptions 5.2 and 5.3 are satisfied.

We provide an extension of our results to infinite horizon dynamic games. For that matter, we need the following definition.

Definition 5.8 (Time-invariant sufficient private information)**.** *We say S^i_t, $i \in \mathcal{N}$, $t \in \mathbb{N}$ is a time-invariant sufficient private information if it is a sufficient private information and has a time-invariant domain denoted by $\mathcal{S}^i_{in}$, $i \in \mathcal{N}$.*

We note that for the special cases presented in Section 5.5 the characterized sufficient private information is time-invariant. Let Π_t denote the CIB belief about (X_t, S_t) at t. Consider a class of CIB strategies that are based on time-invariant sufficient private information. We call the mapping $\sigma^i_{\text{s}} : \Delta(\mathcal{X}_t \times \mathcal{S}_{\text{in}}) \times \mathcal{S}^i_{\text{in}} \to \Delta(\mathcal{A}^i_{\text{in}})$ a *stationary CIB strategy* for agent i if S^i_t, $i \in \mathcal{N}$, $t \in \mathcal{T}$, is a time-invariant sufficient private information. Similarly, we define a *stationary CIB update rule* as a time-invariant mapping $\psi_s : \Delta(\mathcal{X}_{\text{in}} \times \mathcal{S}_{\text{in}}) \times \mathcal{Z}_t$[13] $\to \Delta(\mathcal{X}_{\text{in}} \times \mathcal{S}_{\text{in}})$, that recursively determines

[13]Note that $\mathcal{Z}_t$ is time-invariant.

the CIB belief for all $t \in \mathcal{T}$.

Definition 5.9. *We say that a pair* (σ_s^*, ψ_s) *of a CIB strategy profile* σ_s^* *and a CIB update rule* ψ_{in} *is a stationary CIB-PBE if* (σ_s^*, ψ_s) *is a CIB-PBE,* σ_s^* *is a stationary CIB strategy profile, and* ψ_s *is a stationary update rule.*

Based on Definition 5.9, we provide a sequential decomposition of dynamic games with infinite horizon below.

Let $V_s := \{V_s^i : \Delta(\mathcal{X}_t \times \mathcal{S}_{\text{in}}) \times \mathcal{S}_{\text{in}}^i \to \mathbb{R}, i \in \mathcal{N}\}$ denote a stationary value function that captures the continuation payoff for all agents. Given a value function V_s and a stationary update rule ψ_s, for every realization π_t define *the stationary stage game* $G_s(\pi_t, V_s, \psi_s)$ as a static game of asymmetric information among agents, where the agents' utilities are given by

$$\bar{U}_s^i(a_t, s_t, \pi_t, V_s, \psi_s) := \mathbb{E}_{\pi_t}\left\{u_{\text{in}}^i(X_t, a_t) + V_s^i\left(\psi_s(\pi_t, Z_{t+1}), S_{t+1}\right) \middle| \pi_t, s_t, a_t\right\}, \quad (5.32)$$

for every realization of $s_t \in \mathcal{S}_{\text{in}}$ and $a_t \in \mathcal{A}_{\text{in}}$.

Similar to (5.21), define the correspondence $BNE_s\left(V_{t+1}, \psi_{t+1}\right)$ as the stationary correspondence that characterizes the set of BNEs of the stationary stage game $\mathbf{G_s}(\pi_\mathbf{t}, \mathbf{V_s}, \psi_\mathbf{s})$ for every realization of π_t; this correspondence is given by

$$BNE_s\left(V_s, \psi_s\right) := \left\{\sigma_s^* : \forall \pi_s \in \Delta(\mathcal{X}_t \times \mathcal{S}_t), \sigma^*(\pi_s, \cdot) \text{ is a BNE of } G_s(\pi_s, V_s, \psi_s)\right\}. \quad (5.33)$$

The following theorem extends the result of Theorem 5.3 to dynamic games with infinite horizon

Theorem 5.5. *Consider a infinite-horizon dynamic game with asymmetric information where there exists a time-invariant sufficient private information* $S_t \in \mathcal{S}_{in}$, $t \in \mathbb{N}$, *for all agents. Then, a pair of* (σ_s^*, ψ_s) *of a stationary CIB strategy profile* σ_s^* *and a stationary CIB update rule* ψ_s *is a stationary CIB-PBE if there exist a value* *function $V_s(\cdot,\cdot)$ that satisfies the following set of equations:*

$$\sigma_s^* \in BNE_s(V_s, \psi_s), \tag{5.34}$$

$$\psi_s \text{ is consistent with } \sigma_s^*, \tag{5.35}$$

$$V_s^i(\pi, s^i) = \mathbb{E}^{\sigma^*}_{\pi_t,\psi_s}\left\{\bar{U}_s^i((\sigma^*(\pi,S)), S, \pi, V_s, \psi_s)\Big|\pi, s^i\right\}, \forall i \in \mathcal{N}, s \in \mathcal{S}_{in}, \pi \in \Delta(\mathcal{X}_{in} \times \mathcal{S}_{in}). \tag{5.36}$$

5.8.2 Signaling-free Beliefs

In the model of Section 5.2, we assume that $\mathbb{P}\{x_t|a_{1:t-1}\} > 0$ (Assumption 5.2) and $\mathbb{P}\{y_{1:t}|a_{1:t}\} > 0$ (Assumption 5.3) for all $t \in \mathcal{T}$. In the following we argue that if the agents' actions are observable these assumptions can be relaxed and a result similar to that of Theorem 5.3 can be obtained.

As mentioned in Remark 5.2, condition (ii) of Definition 5.6 is trivially satisfied under Assumptions 5.2 and 5.3. However, when we relax Assumptions 5.2 and 5.3, we need to make sure that the solution of dynamic program described in Theorem 5.3 satisfies condition (ii) of Definition 5.6. To do so, we need to: (1) keep track of the set of system states $x_t \in \mathcal{X}_t$ and sufficient private information $s_t \in \mathcal{S}_t$ that are *feasible* under the common information c_t over time, *i.e.* they satisfy condition (ii) of Definition 5.6 (see (5.14)), and (2) assert that this feasible set can be recursively updated. These two conditions are required because in a dynamic program similar to that of Theorem 5.3 we need to ensure the consistency conditions based on update rule ψ_{t+1}, which utilizes the new common information Z_{t+1} at $t+1$; see (5.26). For that matter, we need to make the following assumption.

Assumption 5.4. *(Observable Actions) At every time $t \in \mathcal{T}$ the agents' actions A_t are commonly observable by all agents, i.e. $A_t \in Z_{t+1}$ for all $t \in \mathcal{T}\backslash\{T\}$.*[14]

[14]We would like to point out that in the absence of Assumption 5.4 the set of feasible system states cannot be updated recursively in general. For instance, consider a dynamic game with delay-sharing information structure, where the agents observe A_{t-d} at time t with d delay, $d > 1$. In this case, the set of feasible system states cannot be updated recursively since the agents' actions A_{t-d} observed at t affect the set of feasible system states at time $t-d+1 < t$.

With Assumption 5.4 replacing Assumptions 5.2 and 5.3, we propose an approach to keep track of feasible system states over time, and accordingly, present a result similar to that of Theorem 5.3.

For every $t \in \mathcal{T}$, let $\gamma_t^f : \mathcal{C}_t \to \Delta(\mathcal{X}_t \times \mathcal{S}_t)$, be such that $\gamma_t^f(c_t)$ assigns a positive probability to every system state $x_t \in \mathcal{X}_t$ and sufficient private information $s_t \in \mathcal{S}_t$ that are feasible under $c_t \in \mathcal{C}_t$. We note that under Assumption 5.4, the history of all actions $A_{1:t-1} \in C_t$. We call the collection of mappings γ_t^f, $t \in \mathcal{T}$ a *signaling-free CIB belief system.* Let $\Pi_t^f := \gamma_t^f(C_t)$ denote the (random) signaling-free CIB belief at t. Note that by using the realization π_t^f of the signaling-free CIB belief Π_t^f, we can rewrite condition (ii) of Definition 5.6 as

$$\pi_t(x_t, s_t) > 0 \text{ only if } \pi_t^f(x_t, s_t) > 0. \tag{5.37}$$

Under Assumption 5.4, Π_t^f can be updated recursively as a function of Π_{t-1}^f and the new common information Z_t at t. Define the *signaling-free CIB update rule* as a sequence of mappings $\psi^f := \{\psi_t^f : \Delta(\mathcal{X}_{t-1} \times \mathcal{S}_{t-1}) \times \mathcal{Z}_t \to \Delta(\mathcal{X}_t \times \mathcal{S}_t), t \in \mathcal{T}\backslash\{1\}\}$ that are given by

$$\begin{aligned}
\pi_1^f(x_1, s_1) &:= \pi_1(x_1, s_1), \\
\pi_t^f(x_t, s_t) &:= \psi_t^f(\pi_{t-1}, z_t) \\
&:= \frac{\sum_{x_{t-1}, s_{t-1}} \pi_{t-1}^f(x_{t-1}, s_{t-1}) \mathbb{P}\{x_t, s_t, z_t \backslash a_{t-1} | x_{t-1}, s_{t-1}, a_{t-1}\}}{\sum_{x_{t-1}, s_{t-1}} \pi_{t-1}^f(x_{t-1}, s_{t-1}) \mathbb{P}\{z_t \backslash a_{t-1} | x_{t-1}, s_{t-1}, a_{t-1}\}},
\end{aligned} \tag{5.38}$$

for all $x_t \in \mathcal{X}_t, s_t \in \mathcal{S}_t$, $z_t \in \mathcal{Z}_t$, $A_{t-1} \in \mathcal{A}_{t-1}$, and $t \in \mathcal{T}\backslash\{1\}$. We note that the signaling-free CIB update rule ψ^f does not depend on the agents' strategy prediction σ^* or actual strategy σ.

Using (5.38), we can write the consistency condition (Definition 5.6) between σ_{t-1}^* and ψ_t in terms of Π_{t-1} and ψ_t^f, $\forall t \in \mathcal{T}\backslash\{1\}$. Therefore, we can present a dynamic program similar to that of Theorem 5.3 by modifying the information state to be $\{\Pi_t, \Pi_t^f, S_t\}$. Accordingly, for all realizations of Π_t, Π_t^f, and S_t, $t \in \mathcal{T}$, we define the value function $\hat{V}_t := \{\hat{V}_t^i : \Delta(\mathcal{X}_t \times \mathcal{S}_t) \times \Delta(\mathcal{X}_t \times \mathcal{S}_t) \times \mathcal{S}_t \to \mathbb{R}, i \in \mathcal{N}\}$ as the continuation payoffs for all agents, and obtain the following result.

Theorem 5.6. *A pair (σ^*, ψ) of a CIB strategy profile σ^* and a CIB update rule ψ (equivalently, a CIB assessment $(\sigma^*, \gamma_t^{\psi})$) is a CIB-PBE if (σ^*, ψ) solves the following dynamic program:*

$$\hat{V}^i_{T+1}(\cdot) := 0 \qquad \forall\, i \in \mathcal{N}, \tag{5.39}$$

$$\tag{5.40}$$

for $t \in \mathcal{T}$,

$$\sigma_t^* \in BNE\left(G_t(\hat{V}_{t+1}, \psi_{t+1})\right), \tag{5.41}$$

$$\psi_{t+1} \textit{ is consistent with } \sigma^* \textit{ and } \psi^f_{t+1} \tag{5.42}$$

$$\hat{V}^i_t(\pi_t, \pi^f_t, s_t) := \mathbb{E}^{\sigma^*_t}_{\pi_t, \psi_{t+1}} \left\{ \bar{U}^i_t((\sigma^*(\pi_t, S_t)), S_t, \pi_t, \pi^f_t, \hat{V}_{t+1}, \psi_{t+1}) \middle| \pi_t, \pi^f_t, s^i_t \right\}, i \in \mathcal{N}. \tag{5.43}$$

We note that the dynamic program presented above is different from the one described in Theorem 5.3 in two aspects. First, the information state $\{\Pi_t, \Pi^f_t, S_t\}$ in Theorem 5.6 has an additional component, given by the signaling-free belief Π^f_t, compared to the information state $\{\Pi_t, S_t\}$ in Theorem 5.3. Second, the consistency condition (5.42) is in terms of σ^*_t, ψ_{t+1}, and ψ^f_{t+1} while (5.26) is in terms of σ^*_t and ψ_{t+1}

5.8.3 Common Observation of Deviations

In Sections 5.8.1 and 5.8.2, we demonstrated how our results can be extended to infinite horizon dynamic games and to instances where we relax Assumptions 5.2-5.3. In this section, we discuss the crucial assumption in our model that is necessary for our results to hold.

In the common information approach presented in Section 5.5, we utilize the CIB belief system γ to form a CIB assessment about the status of the game. As shown by

Lemma 5.1, every agent $i \in \mathcal{N}$ can use the CIB belief system γ along with the agents' CIB strategy σ^* and his sufficient private information S_t^i to form his private belief μ^i about the status of the game. The crucial requirement for the result of Lemma 5.1 to hold is that every realization $s_t^i \in \mathcal{S}_t^i$ of agent i's sufficient private information must be *compatible* with the CIB belief Π_t, *i.e.* $\sum_{x_t \in \mathcal{X}_t, s_t^{-i} \in \mathcal{S}_t^{-i}} \Pi_t(x_t, s_t^i, s_t^{-i}) > 0$. This requirement is satisfied by condition (iii) in Definition 5.6 for realizations z_t of the new common information at t that have zero probability under π_{t-1}, *i.e.* $\mathbb{P}^{\sigma^*}_{\pi_{t-1}}\{z_t\} = 0$. However, for realizations z_t that have positive probability under π_{t-1}, *i.e.* $\mathbb{P}^{\sigma^*}_{\pi_{t-1}}\{z_t\} > 0$, the compatibility condition between agent i's sufficient private information s_t^i and the common belief π_t^i is not satisfied in general. Therefore, we argue that the crucial requirement that underlies our results and guarantees the compatibility between the CIB belief Π_t and the sufficient private information S_t^i, $i \in \mathcal{N}$, at every $t \in \mathcal{T}$ can be summarized by the following assumption.[15]

Assumption 5.5. *For all* $t \in \mathcal{T}$*,* $a_{1:t} \in \mathcal{A}_{1:t}$*,* $i \in \mathcal{N}$*,* $p_t^i \in \mathcal{P}_t^i$ *and* $c_t \in \mathcal{C}_t$*,* $\mathbb{P}_{\pi_1}\{p_t^i, c_t | a_{1:t}\} = 0$ *only if* $\mathbb{P}_{\pi_1}\{c_t | a_{1:t}\} = 0$*.*

Assumption 5.5 implies that every deviation that can be detected by agent i at t must be also detectable by all agents at the same time t based on the common information C_t. We note that Assumption 5.5 is satisfied under Assumptions 5.2 and 5.3, or Assumption 5.4. We do not provide a formal proof for the sufficiency of Assumption 5.5, however, we provide an informal argument below.

Consider a dynamic game with unobservable actions where Assumption 5.5 is not satisfied, that is, there exist $i \in \mathcal{N}$, $t \in \mathcal{T}$, $p_t^i \in \mathcal{P}_t^i$, $c_t^i \in \mathcal{C}_t^i$, and $a_{1:t} \in \mathcal{A}_{1:t}$ such that $\mathbb{P}\{p_t^i, c_t | a_{1:t}\} = 0$ and $\mathbb{P}\{c_t | a_{1:t}\} > 0$. In this game, agent i can detect deviations from $a_{1:t}$ at t when he observes $\{p_t^i, c_t\}$. Upon detecting the deviation, agent i needs to revise his belief μ_t^i. However, if at time t no deviation can be detected based on the common information c_t, the CIB belief π_t is not revised. Therefore, agent i cannot rely anymore on the CIB belief π_t to form his private belief μ_t^i using the construction described by Lemma 5.1. Under Assumption 5.5, it is guaranteed that agent i wants

[15]As noted before, we maintain our assumptions that the system state, actions, and observations spaces are finite.

to revise his belief μ_t^i whenever the CIB belief Π_t is revised. Therefore, for all $t \in \mathcal{T}$, every realization of agent i's sufficient private information S_t^i is compatible with CIB belief Π_t. Consequently, the result of Lemma 5.1 holds, and we can utilize the CIB assessments to provide a decomposition of the dynamic game similar to that of Theorem 5.3.

5.9 Existence

As we discussed in Section 5.7, there exist PBEs that cannot be described as CIB-PBEs in general. Therefore, the standard results that guarantee the existence of a PBE for dynamic games with asymmetric information [96, Proposition. 249.1] cannot be used to guarantee the existence of a CIB-PBE in these games. In this Section, we discuss the existence of CIB-PBEs for dynamic games with asymmetric information. We provide conditions that are sufficient to guarantee the existence of CIB-PBEs (Lemmas 5.2 and 5.3). Using the result of Lemma 5.2, we prove the existence of CIB-PBEs for zero-sum dynamic games with asymmetric information (Theorem 5.7). Using the result of Lemma 5.3, we identify instances of non-zero-sum dynamic games with asymmetric information where we can guarantee the existence of CIB-PBEs.

Lemma 5.2. *The dynamic program given by (5.25)-(5.27) has at least one solution at stage t if the value function V_{t+1} is continuous in Π_{t+1}.*

We note that the condition of Lemma 5.2 is always satisfied for $t = T$ by definition of V_{T+1}; see (5.20) and (5.23). However, for $t < T$, it is not straightforward to prove the continuity of the value function V_t in π_t in general. Given V_{t+1} is continuous in π_{t+1}, the result of [85, Theorem 2] implies that the set of equilibrium payoffs for the state game at t is *upper hemicontinuous* in π_t. Therefore, if the stage game $\mathbf{G_t}(\pi_\mathbf{t}, \mathbf{V_{t+1}}, \psi_\mathbf{t+1})$ has a unique equilibrium payoff for every π_t, we can show that V_t is continuous in π_t for $t < T$. Using this approach, we prove the existence of CIB-PBEs for zero-sum games below, where the equilibrium payoff is unique.

Theorem 5.7. *For every dynamic zero-sum game with asymmetric information there exists a CIB-PBE that is a solution to the dynamic program given by (5.25)-(5.27).*

For dynamic non-zero-sum games, it is harder to establish that V_t is continuous in π_t for $t < T$ since the set of equilibrium payoffs is not a singleton in general. However, we conjecture that for every dynamic game with asymmetric information described in Section 5.2, at every stage of the corresponding dynamic program, it is possible to select a BNE for every realization of π_t so that the resulting V_t is continuous in π_t.

In addition to the results of Lemma 5.2 and Theorems 5.7, we provide another condition below that guarantees the existence of CIB-PBEs in some instances of dynamic games with asymmetric information.

Lemma 5.3. *A dynamic game with asymmetric information described in Section 5.2 has at least one CIB-PBE if there exits sufficient information $S_{1:t}^{1:N}$ such that the CIB update rule $\psi_{1:T}$ is independent of σ^*.*

The independence of CIB update rule from σ^* is a condition that is not satisfied for all dynamic games with asymmetric information. Nevertheless, we present below special instances where this condition is satisfied.

1) Nested information structure with one controller: Consider the nested information structure case described in Section 5.2. Assume that the evolution of the system state is controlled only by the uniformed player and is given by $X_{t+1} = f_t(X_t, A_t^2, W_t)$. For $S_t^1 = X_t$ and $S_t^2 = 0$, it is easy to check that $\mathbb{P}^{\sigma^*}\{\pi_{t+1}|\pi_t, a_t\} = \mathbb{P}\{\pi_{t+1}|\pi_t, a_t\}$ for all π_{t+1}, π_t ,a_t and $t \in \mathcal{T}$.

2) Independent dynamics with observable actions and no private valuation: Consider the model with independent dynamics and observable actions described in Section 5.2. Assume that agent i's, $i \in \mathcal{N}$, instantaneous utility is given by $u_t^i(A_t, X_t^{-i})$ (no private valuation); that is, agent i's utility at t does not depend on X_t^i. It is easy to verify that $S_t^i = \emptyset$ is sufficient private information for agent i. Hence, the condition of Lemma 5.3 is trivially satisfied.

3) Delayed sharing information structure with $\mathbf{d} = \mathbf{1}$ Consider the delayed sharing information structure described in Section 5.2 when delay $d = 1$ [7, 6]. Thus,

$P_t^i = \{Y_t^i\}$. Let $S_t^i = P_t^i = Y_t^i$. Then, it is easy to verify that the condition of Lemma 5.3 is satisfied.

4) Uncontrolled state process with hidden actions: Consider an N-player game with uncontrolled dynamics given by $X_{t+1} = f_t(X_t, W_t)$, $t \in \mathcal{T}$. At every time $t \in \mathcal{T}$, agent i, $i \in \mathcal{N}$, receives a noisy observation $Y_t^i = O_t^i(X_t^i, Z_t^i)$. The agents' actions are hidden. Thus, $P_t^i = \{Y_{1:t}^i, A_{1:t-1}^i\}$ and $C_t = \emptyset$. Hence, the condition of Lemma 5.3 is trivially satisfied. We note that in the case where a subset of the agents' observation reveals to all agents' with some delay, {i.e.} $C_t \subseteq \{Y_{1:t}\}$, the condition of Lemma 5.3 is also satisfied.

5.10 Conclusion

We proposed a general approach to study a dynamic game with asymmetric information with finite or infinite time horizon. We presented a set of conditions sufficient to characterize an information state for each agent that effectively compresses his common and private information in a mutually consistent manner. We showed that the above-mentioned information state provides a sufficient statistic for decision making purposes in strategic and non-strategic settings. We introduced the notion of Common Information based Perfect Bayesian Equilibrium that characterizes a set of outcomes for the dynamic game. We provided a sequential decomposition of the dynamic game over time, which leads to a dynamic program for the computation of the set of CIB-PBEs of the dynamic game. We determined conditions under which we can guarantee the existence of CIB-PBEs. Using these conditions, we proved the existence of CIB-PBE for dynamic zero-sum games and special instances of dynamic non-zero sum games.

For future research, we will investigate the problem of determining the minimal information state in a dynamic game with asymmetric information. In the examples presented in this chapter, we only characterized an information state that compresses an agent's private information by discarding a subset of his private information. Therefore, it will be interesting to identify instances of dynamic games with asymmetric information where we can find an information state that compresses an agent's information by applying a functional transformation on his private and common information. As another direction for future research, we will study the development of a computationally efficient algorithm to solve the dynamic program presented in this chapter and determine the set of CIB-PBEs for a general dynamic game with asymmetric information.

Chapter 6

Conclusion

6.1 A Brief Summary

In this dissertation, we investigated problems arising in the design and analysis of CPSs with strategic agents. We provided a general framework that captures a broad range of CPSs with strategic agents. In a CPS, strategic agents may have control over the decisions and/or possess information that is not available to the designer. Accordingly, we identified two classes of design problems. In the first class, the designer has control over decisions and resources but the strategic agents possess private information that is crucial for the efficient operation of the CPS. In the second class, the designer has superior information about the current status of the system but strategic agents have control over decisions and resources in the system. We identified specific design problems in power systems and transportation networks that can be formulated according to the above classification in Chapters 2-4.

Motivated by the increasing integration of renewable energy and flexible loads into power grids, we studied the design of electricity markets for renewable energy and flexible loads in Chapters 2 and 3 as instances of the first class of design problems described above. We proposed a stylized two-time step model that captures the dynamic and uncertain nature of the generation from renewable resources and the availability of flexible loads. We studied the design of forward bilateral contracts in Chapter 2 and showed that the optimal contract is a contingent contract that allows a renewable generator/flexible load to adjust its commitment according to the new information he receives over time. In Chapter 3, we studied the problem of designing a pooling market for wind energy. Assuming that wind generators receive private information about their generation capacity over time, we proposed a dynamic market mechanism that outperforms the existing sequence of static markets, *e.g.* the day-ahead market followed by the real-time market. We showed that the main advantage of the dynamic mechanism is due to the designer's ability to price discriminate sellers with different levels of uncertainty in their generation and to expose them to the risk of penalty charge. We characterized the benefit of wind monitoring on the overall performance of the market and showed that it vanishes as the number of possible generation technologies increases.

As an instance of the second class of design problems, we studied the problem of optimal information provision in a transportation network in Chapter 4. We investigated the design of public and private information disclosure mechanisms in a parallel two-link network where the designer can provide information to drivers about the condition of one of the links. We showed that the designer can improve the social welfare by strategically disclosing information to the drivers. In particular, we identified conditions under which the designer can achieve the socially efficient outcome. We also investigated the design of information disclosure mechanisms in a two-step dynamic setting where the condition of the network has an uncontrolled Markovian dynamics and the drivers can learn from their past experience. Using numerical simulations, we examined the effect of different pieces of information, from which the drivers can learn, on the designer's optimal information disclosure mechanisms and its performance.

We investigated the analysis of CPSs with strategic agents in Chapter 5. We consider a general dynamic game of asymmetric information with controlled Markovian dynamics where strategic agents make private observations and take actions over time. We presented a set of conditions sufficient to characterize an information state for each agent that effectively compresses the agent's private and common information. We showed that the characterized information state provides a sufficient statistics for decision making purposes in strategic and non-strategic settings. Consequently, the proposed methodology in this dissertation provides a universal approach to study dynamic games and teams including dynamic games among teams of agents. Using the above-mentioned information state, we introduced the notion of CIB-PBE which characterizes a subset of PBE of dynamic games. We provided a sequential decomposition of the dynamic game and formulated a dynamic program to determine the set of CIB-PBEs via backward induction.

6.2 Future Directions

The emergence of CPSs with strategic agents has created new challenges due to the decentralizations of decisions and information as well as the agents' autonomous selfish behavior. In this dissertation, we studied specific design problems that are motivated by applications in power systems and transportation networks and proposed a general approach to determine a set of outcomes in these systems. Below, we discuss s few directions which in our opinion provide valuable insights into the design and analysis of CPSs with strategic agents.

6.2.1 Games among Teams of Agents

Throughout this dissertation, we assumed that agents are selfish and have objectives that are different from each other and that of the designer. While this framework provides a general model to study CPSs with strategic and cooperative (non-strategic) agents, it does not consider explicitly situations where a group of agents have selfish objectives at a lower level but share a common objective at a higher level. For instance, in a networked system each agent/community might have a selfish objective at the microscopic level but all agents act as a team to protect the overall operation of the networked system against external attacks at the macroscopic level; *e.g.* in a power grid, every generator tries to maximize his own revenue, but at a network level all generators try to ensure the stability and reliability of the power grid and protect it against malicious and/or accidental disruptions.

A possible approach to study this class of problems is to investigate the problem at the microscopic level as a game among agents, and the problem at the macroscopic level as a game among teams of agents (*e.g.* attacker vs. defender). As we discussed in Chapter 5, the methodology proposed in this dissertation can be used to study dynamic games among teams of agents. We believe that investigations of this class of problems can provide valuable insights towards a better design of resilient CPSs that effectively respond to external disruptions.

6.2.2 Cyber-Physical-Social Systems and Bounded Rationality

Over the last few years, the emerging CPSs cease to be merely technological systems as they interact more than ever with human users and can be viewed as cyber-physical-social systems. In this dissertation, we attempted to study this new class of socio-technological systems assuming that the users in these systems act as rational Bayesian agents. This assumption enabled us to formulate stylized models and provide analytical results that help to better understand the key ideas in the design and analysis of CPSs with strategic agents. However, in many of these systems, the assumption that a normal human user acts as a rational Bayesian agent is inaccurate. For instance, a human user tends to develop habits over time or does not form his belief according to Bayesian inference methods [110, 107]. Therefore, it is important to develop models that take into account the human users' bounded rationality, validate the models with real-data, and investigate how existing results for rational Bayesian agents translate into these models.

6.2.3 Learning in CPS and Strategic Experimentation

The recent advances in data collection, storage, and processing technologies have created new opportunities for the application of data-driven techniques in the design and analysis of CPSs. However, many of these techniques assume a *passive* data generation process that is not affected by the design and the agents' strategic responses to it. In many of CPSs with strategic agents, the data available to the designer is generated by the agents' strategic behavior that incorporate the various effects of the data they generate on the overall characteristics of the CPS. For instance, there are empirical evidences about fake user-generated data in Waze App that attempt to trick the system into keeping the traffic out of certain residential areas [52, 94, 103] or hiding the police locations [45]. Therefore, it is important to consider the agents' strategic behavior when we use data-driven methods for the analysis and design of CPSs with strategic agents. Investigating agents' incentives in various environments where they strategically experiment and try to influence the outcome these methods [54, 22] would provide valuable insights into the design and improvement of data-driven CPSs that rely heavily on the user-generated data.

Appendices

Appendix A

Proofs of Chapter 2

Proof of lemma 2.1. The given mechanism (q, t) is incentive compatible, so we can rewrite $U(\mathbf{x})$ as

$$U(\mathbf{x}) = \max_{\mathbf{x}'} \mathbb{E}_W \left\{ t(\mathbf{x}') - C(q(\mathbf{x}'), \mathbf{x}, W) \right\} \tag{A.1}$$

By applying the envelope theorem [84] on (A.1), we get

$$\frac{\partial U}{\partial x_i} = - \left. \frac{\partial \mathbb{E}_W \left\{ C(q(\mathbf{x}'), \mathbf{x}, W) \right\}}{x_i} \right|_{\mathbf{x}'=\mathbf{x}}. \tag{A.2}$$

The above equation along with the assumption on the monotonicity of the marginal expected cost $c(q, \mathbf{x})$ with respect to $\mathbf{x}_i$ gives

$$\frac{\partial U}{\partial x_i} \leq 0, 1 \leq i \leq m \tag{A.3}$$

$$\frac{\partial U}{\partial x_i} \geq 0, m < i \leq n. \tag{A.4}$$

□

Proof of lemma 2.2. The proof is by contradiction. Assume that there exist $\mathbf{x}, \mathbf{x}' \in \chi$ such that $q(\mathbf{x}) = q(\mathbf{x}')$ but $t(\mathbf{x}') > t(\mathbf{x})$. Then a seller with type $\mathbf{x}$ is always better off by reporting $\mathbf{x}'$ instead of her true type $\mathbf{x}$, which contradicts the IC constraint. □

Proof of Lemma 2.3. Consider the buyer's objective (2.7). For any function $t(\cdot)$, we can determine from (2.14) the cumulative distribution function for q^*, called F_{q^*}. Consequently, we can rewrite the buyer's objective as

$$\begin{aligned} &\mathbb{E}_{q^*}\left[\mathcal{V}(q^*) - t(q^*)\right] = \int_0^\infty \left(\mathcal{V}(l) - t(l)\right) dF_{q^*}(l) \\ &= \left(F_{q^*}(l) - 1\right)\left(\mathcal{V}(l) - t(l)\right)\Big|_0^\infty \\ &+ \int_0^\infty \left(1 - F_{q^*}(l)\right) \frac{d\left(\mathcal{V}(l) - t(l)\right)}{dl} dl. \end{aligned} \tag{A.5}$$

We have

$$\left(F_{q^*}(l) - 1\right)\left(\mathcal{V}(l) - t(l)\right)\Big|_0^\infty = -t(0) \tag{A.6}$$

because $\mathcal{V}(0) = 0$ by assumption, and $\left(F_{q^*}(\infty) - 1\right) = 0$.

Because of (A.6), we can rewrite (A.5) as

$$\begin{aligned} \mathbb{E}_{q^*}\left[\mathcal{V}(q^*) - t(q^*)\right] = &\int_0^\infty P\left(q^* \geq l\right)\left(\mathcal{V}'(l) - p(l)\right) dl \\ &-t(0) \end{aligned} \tag{A.7}$$

where $\mathcal{V}'(l) = \frac{d\mathcal{V}(l)}{dl}$.

We can rewrite $P\left(q^* \geq l\right)$ as

$$P(q^* \geq l) = P[\mathbf{x} \in \chi | \arg\max_{\hat{l}} \mathbb{E}_W \left\{ t(\hat{l}) - C(\hat{l}, \mathbf{x}, W) \right\} \geq l]. \tag{A.8}$$

We implicitly assume that the seller's problem given by (2.14) is continuous and quasi-concave, so that from the first order optimality condition for (2.14) we obtain

$$p(q^*(\mathbf{x})) = \left.\frac{\partial \mathbb{E}_W \left\{C(l, \mathbf{x}, W)\right\}}{\partial l}\right|_{q^*(\mathbf{x})} = c(q^*(\mathbf{x}), \mathbf{x}). \tag{A.9}$$

Therefore, from the optimality of $q^*(\mathbf{x})$ and the quasi-concavity of (2.14), we must have $p(l) > c(l;\mathbf{x})$ and $p(l) < c(l;\mathbf{x})$ for $l < q*(\mathbf{x})$ and $l > q^*(\mathbf{x})$, respectively. That is, each type of the seller wishes to produce more than quantity l if and only if the marginal price $p(q)$ that she is paid at l is higher than the expected marginal cost of production $c(l, \mathbf{x})$ that she incurs at l. Consequently, combining (A.8) and (A.9) we obtain

$$P\left(q^* \geq l\right) = P\left[\mathbf{x} \in \chi | p(l) \geq c(l, \mathbf{x})\right]. \tag{A.10}$$

Substituting (A.10) in (A.7), we obtain the following alternative expression for the buyer's objective

$$\begin{aligned}\mathbb{E}_{q^*}\left[\mathcal{V}(q^*) - t(q^*)\right] = \int_0^\infty & P[\mathbf{x} \in \chi | p(l) \geq c(l, \mathbf{x})] \\ & (\mathcal{V}'(l) - p(l))\, dl - t(0).\end{aligned} \tag{A.11}$$

□

Proof of lemma 2.4. Let $\mathbf{x}, \mathbf{x}' \in \chi$, where $\mathbf{x}$ is a better type than $\mathbf{x}'$. From IC for seller's type $\mathbf{x}$ we have

$$\begin{gathered} t(q(\mathbf{x})) - \mathbb{E}_W\{C(q(\mathbf{x}), \mathbf{x}, W)\} \\ \geq \\ t(q(\mathbf{x}')) - \mathbb{E}_W\{C(q(\mathbf{x}'), \mathbf{x}, W)\} \end{gathered} \tag{A.12}$$

Similarly from IC for seller's type $\mathbf{x}'$ we have

$$\begin{gathered} t(q(\mathbf{x}')) - \mathbb{E}_W\left\{C(q(\mathbf{x}'), \mathbf{x}', W)\right\} \\ \geq \\ t(q(\mathbf{x})) - \mathbb{E}_W\left\{C(q(\mathbf{x}), \mathbf{x}', W)\right\} \end{gathered} \tag{A.13}$$

Subtracting (A.13) from (A.12), we get

$$\begin{aligned}\mathbb{E}_W\left\{C(q(\mathbf{x}),\mathbf{x}',W)\right\}-\mathbb{E}_W\left\{C(q(\mathbf{x}'),\mathbf{x}',W)\right\}\\ \geq \\ \mathbb{E}_W\left\{C(q(\mathbf{x}),\mathbf{x},W)\right\}-\mathbb{E}_W\left\{C(q(\mathbf{x}'),\mathbf{x},W)\right\}\end{aligned} \tag{A.14}$$

By assumption, $\frac{d\mathbb{E}_W\{C(l,\mathbf{x},W)\}}{dl}\leq\frac{d\mathbb{E}_W\{C(l,\mathbf{x}',W)\}}{dl}$ if $\mathbf{x}$ is a better type than $\mathbf{x}'$. Therefore, (A.14) holds if and only if

$$q(\mathbf{x})\geq q(\mathbf{x}'). \tag{A.15}$$

□

Proof of corollary 2.2. Because of corollary 2.1, the VP constraint implies

$$U(\underline{\mathbf{x}})=t(q(\underline{\mathbf{x}}))-\mathbb{E}_W\left[C(q^*(\underline{\mathbf{x}}),\underline{\mathbf{x}},W)\right]=0, \tag{A.16}$$

which is equivalent to

$$t(0)+\int_0^{q^*(\underline{\mathbf{x}})}p(l)dl=\mathbb{E}_W\left[C(q^*(\underline{\mathbf{x}}),\mathbf{x},W)\right]. \tag{A.17}$$

Furthermore, from Lemma 2.4 it follows that if the worst type wishes to produce more than $q^*(\underline{\mathbf{x}})$, then all types produce more than $q^*(\underline{\mathbf{x}})$. Therefore,

$$P\left[\mathbf{x}\in\chi|p(l)\geq c(l,\mathbf{x})\right]=1,\ \text{ for } l\leq q^*(\underline{\mathbf{x}}). \tag{A.18}$$

Using (A.18), we can rewrite the objective function of problem (P3) as,

$$\begin{aligned}&-\left(t(0)+\int_0^{q^*(\underline{\mathbf{x}})}p(l)dl\right)+\int_0^{q^*(\underline{\mathbf{x}})}\mathcal{V}'(l)dl\\ &+\int_{q^*(\underline{\mathbf{x}})}^{\infty}P\left[\mathbf{x}\in\chi|p(l)\geq c(l,\mathbf{x})\right]\left(\mathcal{V}'(l)-p(l)\right)dl.\end{aligned} \tag{A.19}$$

The term $t(0) + \int_0^{q^*(\underline{\mathbf{x}})} p(l)dl$ appears in both the objective (A.19) and the VP constraint (A.17). Therefore, without loss of optimality, we can assume $t(0) = C(0, \underline{\mathbf{x}})$, and set $t(q(\underline{\mathbf{x}})) = \mathbb{E}_W \left\{ C(q(\underline{\mathbf{x}}), \underline{\mathbf{x}}, W) \right\}$. $\square$

Appendix B

Proofs of Chapter 3

B.1 Proofs of the Main Results

The proofs of Theorems 3.1-3.5 are based on Lemmas B.1-B.7 that we state below. The proofs of Lemmas B.1-B.7 are given in Appendix B.2.

Let $g(\omega)$ denote the corresponding probability density function for ω, *i.e.* $g(\omega) := \frac{\partial G(\omega)}{\partial \omega}$.

First, we prove the following sufficient and necessary condition for the IC_2 constraint (reporting the true ω) for the dynamic mechanism.

Lemma B.1 (revenue equivalence)**.** *If the dynamic mechanism is incentive compatible, and the allocation rule $q(\tau, \omega)$ is continuous in ω, then, for all ω and $\acute{\omega}$,*

$$\mathcal{R}_{\tau,\omega} = \mathcal{R}_{\tau,\acute{\omega}} - \int_{\acute{\omega}}^{\omega} C_\theta(q(\tau, \hat{\omega}); \Theta(\tau, \hat{\omega}))\Theta_\omega(\tau, \hat{\omega})d\hat{\omega}, \tag{B.1}$$

*and $q(\tau, \omega)$ is increasing in ω. Moreover, if (*B.*1) holds and $q(\tau, \omega)$ is increasing in ω, then IC_2 is satisfied.*

We can now provide the proof for Theorem 3.1 using the result of Lemma B.1.

Proof of Theorem 3.1. We note that the IC constraint (3.8) for the real-time mechanisms implies the IC_2 constraint for the dynamic mechanism (given by (3.11)), by setting $\hat{\tau} = \tau$. By Lemma B.1, the IC_2 constraint for the dynamic mechanism is satisfied only if equation (3.13) holds.

□

Lemma B.1 characterizes the seller's revenue in the dynamic mechanism given that he tells the truth at $T = 1$ about his technology τ. To complete the characterization of the seller's optimal strategy at $T = 2$, we show, via Lemma B.2 below, that if the seller misreports his technology at $T = 1$ (off-equilibrium path), he later corrects his lie at $T = 2$.

Lemma B.2. *Consider the dynamic mechanism that satisfies the IC constraints for ω and τ. If a seller with technology τ misreports $\hat{\tau}$, $\hat{\tau} \neq \tau$, at $T = 1$, then, for every wind realization ω at $T = 2$, he corrects his lie by reporting $\hat{\omega} = \sigma^*(\hat{\tau}; \tau, \omega)$ such that,*

$$\Theta(\tau, \omega) = \Theta(\hat{\tau}, \hat{\omega}). \tag{B.2}$$

Remark: Note that by Assumption 3.2 on non-shifting support, for any $\tau, \hat{\tau}, \omega$, there exists a unique $\hat{\omega}$ (given the strict monotonicity of $\Theta(\tau, \omega)$ in ω) that satisfies equation (B.2), and $\sigma^*(\hat{\tau}; \tau, \omega)$ is well defined.

Using the results of Lemmas B.1 and B.2, we characterize below the seller's expected gain by misreporting his technology at $T = 1$ in the dynamic mechanism.

Lemma B.3. *For the dynamic mechanism that satisfies the set of IC constraints for ω and τ, the maximum utility of the seller with technology τ reporting $\hat{\tau}$ at $T = 1$ is given by,*

$$\begin{aligned} &\mathbb{E}_\omega\{t(\hat{\tau}, \sigma^*(\hat{\tau}; \tau, \omega)) - C(q(\hat{\tau}, \sigma^*(\hat{\tau}; \tau, \omega)); \Theta(\tau, \omega))\} \\ &= \mathcal{R}_{\hat{\tau}} - \int\int_{\omega}^{\sigma^*(\hat{\tau}; \tau, \omega)} C_\theta(q(\hat{\tau}, \hat{\omega}); \Theta(\hat{\tau}, \hat{\omega}))\Theta_\omega(\hat{\tau}, \hat{\omega}) d\hat{\omega} dG(\omega) \end{aligned} \tag{B.3}$$

The following result, for the dynamic mechanism, characterizes the payments that incentivize the seller to report truthfully his technology τ and wind speed ω.

Lemma B.4. *For the dynamic mechanism, the set of IC constraints for τ, given by (3.10), can be replaced by the following inequality constraints,*

$$\mathcal{R}_{\tau_i}-\mathcal{R}_{\tau_j}\geq\mathcal{R}^T(\tau_j,\tau_i;q)+\mathcal{R}^W(\tau_j,\tau_i;q)\quad\forall i,j\in\{2,..,M\},i>j.$$

To provide the proof for Theorems 3.2-3.3, we need the following results for the forward and the real-time mechanisms.

Lemma B.5. *For the forward mechanism, the set of IC constraints and IR constraints, given by (3.6) and (3.7), respectively, can be replaced by the following constraints,*

$$\begin{aligned}&q(\tau)\ \textit{only depends on}\ \tau\\&\mathcal{R}_{\tau_i}-\mathcal{R}_{\tau_{i-1}}=\mathcal{R}^T(\tau_{i-1},\tau_i;q)\quad\forall i\in\{2,..,M\},\\&\mathcal{R}_{\tau_1}\geq 0.\end{aligned}$$

Lemma B.6. *For the real-time mechanism, the set of IC and IR constraints given by (3.8) and (3.9), respectively, can be replaced by the following constraints,*

$$\begin{aligned}&q(\tau,\omega)\ \textit{only depends on}\ \Theta(\tau,\omega)\\&\mathcal{R}_{\tau_i}-\mathcal{R}_{\tau_{i-1}}=\mathcal{R}^T(\tau_{i-1},\tau_i;q)+\mathcal{R}^W(\tau_{i-1},\tau_i;q)\quad\forall i\in\{2,..,M-1\},\\&R_{\tau_1}\geq\mathcal{R}^P(\tau_1;q),\end{aligned}$$

and the seller's revenue satisfies (3.13).

Using the result of Lemmas B.4-B.6, we first provide the proof for Theorem 3.2.

Proof of Theorem 3.2. We first show that $R^T(\tau_j,\tau_i;q)\geq 0$ and $R^W(\tau_j,\tau_i;q)\geq 0$ for $i,j\in\{1,...,M\},i>j$ (strict if $q(\tau_j,\omega)\neq 0$ for some ω).

First, note that by Assumption 3.1 we have $\Theta(\tau_i, \omega) < \Theta(\tau_j, \omega)$ for $i > j$. Thus, $C(q(\tau_j, \omega); \Theta(\tau_j, \omega)) \geq C(q(\tau_j, \omega); \Theta(\tau_i, \omega))$ with strict inequality if $q(\tau_j, \omega) > 0$. Therefore,

$$\mathcal{R}^T(\tau_j,\tau_i;q) := \int (C(q(\tau_j,\omega);\Theta(\tau_j,\omega)) - C(q(\tau_j,\omega);\Theta(\tau_i,\omega)))dG(\omega) \geq 0,$$

with strict inequality if $q(\tau_j, \omega) > 0$ for a set of ω's with positive probability.

Second, we have $C_\theta(q(\tau_j, \omega); \Theta(\tau_j, \hat{\omega})) \leq C_\theta(q(\tau_j, \hat{\omega}); \Theta(\tau_j, \hat{\omega}))$ for $\omega \leq \hat{\omega}$ since $q(\tau; \hat{\omega})$ is increasing in $\hat{\omega}$ by Lemma B.1 and $C(q, \theta)$ is increasing in q by Assumption 3.1. Moreover, by the result of Lemma B.2, we have $\sigma^*(\tau_j; \tau_i, \omega) \geq \omega$ for $i > j$. Therefore, we have

$$\begin{aligned}\mathcal{R}^W(\tau_j,\tau_i;q) := \iint_\omega^{\sigma^*(\tau_j;\tau_i,\omega)} &(C_\theta(q(\tau_j,\omega);\Theta(\tau_j,\hat{\omega})) - C_\theta(q(\tau_j,\hat{\omega});\Theta(\tau_j,\hat{\omega}))) \\ &\Theta_\omega(\tau_j,\hat{\omega})d\hat{\omega}dG(\omega) \geq 0,\end{aligned}$$

since $\Theta_\omega(\tau_j,\hat{\omega}) < 0$; the inequality is strict if $q(\tau_j, \omega) > 0$ for a set of ω's with positive probability.

In the following, we provide the proof for each part of Theorem 3.2 separately.

a) The proof for part (a) directly follows from the result of Lemma B.5.

b) The proof for part (b) directly follows from the result of Lemma B.6.

c) By the result of Theorem 3.1, the set of IC constraints (3.11) can be replaced by (3.21,3.22). Moreover, by the result of Lemma B.4, we can reduce the set of IC constraints (3.10) to the set of inequality constraints (3.23).

   Furthermore, set of inequality constraints (3.23) implies that $\mathcal{R}_{\tau_1} \leq \mathcal{R}_{\tau_2} \leq ... \leq \mathcal{R}_{\tau_M}$ since $R^T(\tau_{i-1}, \tau_i; q) \geq 0$ and $R^W(\tau_{i-1}, \tau_i; q) \geq 0$. Thus, the set of IR constraints (3.12) can be reduced to $\mathcal{R}_{\tau_1} \geq 0$.

□

Next, we provide the proof for Theorem 3.3.

Proof of Theorem 3.3. The proof directly follows from the result of Theorem 3.2. We note that the objective functions in all the mechanism design problems are $\mathcal{W}$, and the problems differ only in the set of constraints they have to satisfy.

- The set of constraints for the real-time mechanism, given by part (b) of Theorem 3.2, is more restrictive than the set of constraints for the dynamic mechanism given by part (c) of Theorem 3.2. Therefore, the designer's objective $\mathcal{W}^{\text{dyanmic}}$ is higher than his objective $\mathcal{W}^{\text{real-time}}$.

- The set of constraints for the forward mechanism, given by part (a) of Theorem 3.2, is more restrictive than the set of constraints for the dynamic mechanisms, given by part (c) of Theorem 3.2. Therefore, the designer's objective $\mathcal{W}^{\text{dyanmic}}$ is higher than his objective $\mathcal{W}^{\text{forward}}$.

□

We now provide the proof of Theorem 3.4 on the dynamic mechanism with no penalty.

Proof of Theorem 3.4. **(i)** We note that the set of IC constraints (3.27,3.28) for the dynamic mechanism with no penalty is identical to the set of IC constraints (3.27,3.28) for dynamic mechanism. Therefore, constraints (3.30-3.32) directly follow from the result of part (c) of Theorem 3.2.

Next, we show that the set of IR constraints (3.29) is satisfied if and only if $\mathcal{R}_{\tau_i,\underline{\omega}} \geq \mathcal{R}^P(\tau_i; q)$ and the seller's revenue satisfies (3.13). First, by part (i) of Theorem 3.1, we have $\mathcal{R}_{\tau,\omega}$ is increasing in ω. Hence, the set of IR constraints (3.29) is satisfied if and only if $\mathcal{R}_{\tau_i,\underline{\omega}} \geq 0$, for all $i \in \{1, ..., M\}$. Second, using (3.13) along with

$\mathcal{R}_{\tau_i,\underline{\omega}} \geq 0$, $i \in \{1, ..., M\}$, we have,

$$\begin{aligned}
\mathcal{R}_{\tau_i} &\geq \int_{\underline{\omega}}^{\overline{\omega}} \int_{\underline{\omega}}^{\omega} C_\theta(q(\tau_i, \hat{\omega}); \Theta(\tau, \hat{\omega}))\Theta_\omega(\tau_i, \hat{\omega}) d\hat{\omega} dG(\omega) \\
&= \int_{\underline{\omega}}^{\overline{\omega}} \int_{\hat{\omega}}^{\overline{\omega}} C_\theta(q(\tau_i, \hat{\omega}); \Theta(\tau_i, \hat{\omega}))\Theta_\omega(\tau_i, \hat{\omega}) dG(\omega) d\hat{\omega} \\
&= \int_{\underline{\omega}}^{\overline{\omega}} [1 - G(\hat{\omega})]\, C_\theta(q(\tau_i, \hat{\omega}); \Theta(\tau_i, \hat{\omega}))\Theta_\omega(\tau_i, \hat{\omega}) d\hat{\omega} \\
&= \mathcal{R}^P(\tau_i; q).
\end{aligned}$$

Thus, the set of IR constraints (3.29) is satisfied if and only if (3.33) is satisfied.

(ii) We note that the objective functions in all the mechanism design problems are $\mathcal{W}$, and the problems differ only in the set of constraints they have to satisfy. The set of constraints for the real-time mechanism, given by part (b) of Theorem 3.2, is more restrictive than the set of constraints for the dynamic mechanisms with no penalty, given by part (i) above. Therefore, the designer's objective $\mathcal{W}^{\text{dyanmic no penalty}}$ is higher than his objective $\mathcal{W}^{\text{real-time}}$. Moreover, the set of constraints for the dynamic mechanism, given by part (c) of Theorem 3.2, is less restrictive than the set of constraints for the dynamic mechanisms with no penalty, given by part (i) above. Therefore, the designer's objective $\mathcal{W}^{\text{dyanmic no penalty}}$ is lower than his objective $\mathcal{W}^{\text{dynamic}}$. □

To provide the proof for Theorem 3.5, we need the following result for the dynamic mechanism with monitoring.

Lemma B.7. *For the dynamic mechanism with monitoring, the set of IC and IR constraints, given by (3.34) and (3.35), respectively, can be replaced by the following constraints,*

$$\begin{aligned}
&\mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_{i-1}} = \mathcal{R}^T(\tau_{i-1}, \tau_i; q) \quad \forall i \in \{2, .., M-1\}, \\
&R_{\tau_1} \geq 0.
\end{aligned}$$

Using the result of Lemma B.7, we provide the proof of Theorem 3.5 below.

Proof of Theorem 3.5. **(i)** The proof for part (i) directly follows from the result of Lemma B.7.

(ii) The objective functions in all the mechanism design problems are $\mathcal{W}$, and the problems differ only in the set of constraints they have to satisfy. The set of constraints for the dynamic mechanism, given by part (c) of Theorem 3.2, is more restrictive than the set of constraints for the dynamic mechanisms with monitoring, given by part (i) above. Therefore, the designer's objective $\mathcal{W}^{\text{dyanmic with monitoring}}$ is higher than $\mathcal{W}^{\text{dynamic}}$. Moreover, by the result of Theorems 3.3 and 3.4, we have $\mathcal{W}^{\text{dynamic}} \geq \mathcal{W}^{\text{dynamic no penalty}} \geq \mathcal{W}^{\text{real-time}}$ and $\mathcal{W}^{\text{dynamic}} \geq \mathcal{W}^{\text{forward}}$.

(iii) Define the following modified payment function:

$$\hat{t}(\tau,\omega) := \mathbb{E}_{\omega}\{t(\tau,\omega)\} + \left(C(q(\tau,\omega);\Theta(\tau,\omega)) - \mathbb{E}_{\omega}\{C(q(\tau,\omega);\Theta(\tau,\omega))\}\right).$$

We have $\mathbb{E}_{\omega}\{\hat{t}(\tau,\omega)\} = \mathbb{E}_{\omega}\{t(\tau,\omega)\}$. Thus, the seller's strategic report for τ at $T=1$ does not change. Consider a modified mechanism with the modified payment function $\hat{t}(\tau,\omega)$ and the original allocation function $q(\tau,\omega)$. This modified mechanism satisfies the set of IC constraint for τ and it satisfies the ex-post IR constraint, *i.e.*,

$$\mathcal{R}_{\tau,\omega} = \mathbb{E}_{\omega}\{t(\tau,\omega)\} - \mathbb{E}_{\omega}\left\{C(q(\tau,\omega);\Theta(\tau,\omega))\right\} = \mathcal{R}_{\tau} \geq 0.$$

We note that with the monitoring of ω, the modified mechanism keeps the same allocation function, and therefore, results in the same designer's objective $\mathcal{W}$, seller's revenue $\mathcal{R}$, and buyer's utility $\mathcal{U}$. $\square$

B.2 Proof of Lemmas

Proof of Lemma B.1. Assume that the IC for ω and $\breve{\omega}$ is satisfied. Then,

$$t(\tau,\omega) - C(q(\tau,\omega);\Theta(\tau,\omega)) \geq t(\tau,\breve{\omega}) - C(q(\tau,\breve{\omega});\Theta(\tau,\omega)),$$
$$t(\tau,\breve{\omega}) - C(q(\tau,\breve{\omega};\Theta(\tau,\breve{\omega})) \geq t(\tau,\omega) - C(q(\tau,\omega);\Theta(\tau,\breve{\omega})).$$

Therefore,

$$\begin{aligned} C(q(\tau,\breve{\omega});\Theta(\tau,\breve{\omega}))-C(q(\tau,\breve{\omega});\Theta(\tau,\omega)) &\leq \mathcal{R}_{\tau,\omega}-\mathcal{R}_{\tau,\breve{\omega}} \\ &\leq C(q(\tau,\omega);\Theta(\tau,\breve{\omega}))-C(q(\tau,\omega);\Theta(\tau,\omega)). \end{aligned} \tag{B.4}$$

Set $\breve{\omega} = \omega - \epsilon$ where $\epsilon \to 0$. We have $\frac{\partial \mathcal{R}_{\tau,\omega}}{\partial \omega} = C_\theta(q(\tau,\omega);\Theta(\tau,\omega))\Theta_\omega(\tau,\omega)$.

Moreover, (B.4) implies that $q(\tau,\breve{\omega}) < q(\tau,\omega)$ for $\breve{\omega} < \omega$, since by assumption $\Theta(\tau,\omega)$ is decreasing in ω and $C(\hat{q};\theta)$ is convex and increasing in $\hat{q}$.

To prove the converse, assume that (B.1) holds and $q(\tau,\omega)$ is increasing in ω. Then, for any $\tau,\omega,\breve{\omega}$, we have

$$\begin{aligned} &\mathcal{R}_{\tau,\omega}-[t(\tau,\breve{\omega}) - C(q(\tau,\breve{\omega});\Theta(\tau,\omega))]= \\ &[\mathcal{R}_{\tau,\omega} - \mathcal{R}_{\tau,\breve{\omega}}]-[C(q(\tau,\breve{\omega});\Theta(\tau,\breve{\omega}))-C(q(\tau,\breve{\omega});\Theta(\tau,\omega))] \\ &=\int_\omega^{\breve{\omega}} C_\theta(q(\tau,\hat{\omega});\Theta(\tau,\hat{\omega}))\Theta_\omega(\tau,\hat{\omega})d\hat{\omega} \\ &\quad -[C(q(\tau,\breve{\omega});\Theta(\tau,\breve{\omega}))-C(q(\tau,\breve{\omega});\Theta(\tau,\omega))] \geq 0, \end{aligned} \tag{B.5}$$

where the last last inequality is true since $C_\theta(q;\theta)$ is increasing in θ, $\Theta(\tau,\omega)$ is decreasing in ω, and $q(\tau,\omega)$ is increasing in ω. □

Proof of Lemma B.2. We first note that by Assumption 3.2 and the monotonicity of $\Theta(\hat{\tau},\hat{\omega})$ in ω, there exists a unique $\hat{\omega}$ such that $\Theta(\tau,\omega) = \Theta(\hat{\tau},\hat{\omega})$, *i.e.* $\sigma^*(\hat{\tau};\tau,\omega)$ is well defined.

Now, consider a seller with technology $\hat{\tau}$ and wind realization $\hat{\omega}$. Then, the IC_2

constraint requires

$$t(\hat{\tau},\hat{\omega})-C(q(\hat{\tau},\hat{\omega});\Theta(\hat{\tau},\hat{\omega}))\geq t(\hat{\tau},\omega')-C(q(\hat{\tau},\omega');\Theta(\hat{\tau},\hat{\omega}))\ \forall\omega'.$$

Replacing $\Theta(\hat{\tau},\hat{\omega})$ by $\Theta(\tau,\omega)$ in the above equation, we get

$$t(\hat{\tau},\hat{\omega})-C(q(\hat{\tau},\hat{\omega});\Theta(\tau,\omega))\geq t(\hat{\tau},\omega')-C(q(\hat{\tau},\omega');\Theta(\tau,\omega))\ \forall\omega'.$$

Now, consider a seller with technology τ that misreported $\hat{\tau}$ at $T=1$, and has a wind realization ω at $T=2$. The above inequality asserts that it is optimal for him to report $\hat{\omega}$ at $T=2$ so that $\Theta(\tau,\omega)=\Theta(\hat{\tau},\hat{\omega})$. □

Proof of Lemma B.3. We have,

$$\begin{aligned}&\mathbb{E}_\omega\{t(\hat{\tau},\sigma^*(\hat{\tau};\tau,\omega))-C(\hat{\tau},\sigma^*(\hat{\tau};\tau,\omega);\Theta(\tau,\omega))\}\\&=\int[t(\hat{\tau},\sigma^*(\hat{\tau};\tau,\omega))-C(q(\hat{\tau},\sigma^*(\hat{\tau};\tau,\omega));\Theta(\tau,\omega))]\,dG(\omega).\end{aligned}\tag{B.6}$$

Using Lemma B.2, we get

$$\begin{aligned}&\int[t(\hat{\tau},\sigma^*(\hat{\tau};\tau,\omega))-C(q(\hat{\tau},\sigma^*(\hat{\tau};\tau,\omega));\Theta(\tau,\omega))]\,dG(\omega)\\&=\int[t(\hat{\tau},\sigma^*(\hat{\tau};\tau,\omega))-C(q(\hat{\tau},\sigma^*(\hat{\tau};\tau,\omega));\Theta(\hat{\tau},\sigma^*(\hat{\tau};\tau,\omega)))]\,dG(\omega)\\&=\int\mathcal{R}_{\hat{\tau},\sigma^*(\hat{\tau};\tau,\omega)}dG(\omega).\end{aligned}\tag{B.7}$$

Then, by Lemma B.1,

$$\begin{aligned}&\int\mathcal{R}_{\hat{\tau},\sigma^*(\hat{\tau};\tau,\omega)}dG(\omega)=\\&\int\left[\mathcal{R}_{\hat{\tau},\omega}-\int_{\omega}^{\sigma^*(\hat{\tau};\tau,\omega)}C_\theta(q(\hat{\tau},\hat{\omega});\Theta(\hat{\tau},\hat{\omega}))\Theta_\omega(\hat{\tau},\hat{\omega})d\hat{\omega}\right]dG(\omega).\end{aligned}\tag{B.8}$$

Furthermore, $\mathcal{R}_{\hat{\tau}} = \int \mathcal{R}_{\hat{\tau},\omega} dG(\omega)$; thus, the RHS of (B.8) can be rewritten as,

$$\mathcal{R}_{\hat{\tau}} - \iint_{\omega}^{\sigma^*(\hat{\tau};\tau,\omega)} C_\theta(q(\hat{\tau},\hat{\omega});\Theta(\hat{\tau},\hat{\omega}))\Theta_\omega(\hat{\tau},\hat{\omega})d\hat{\omega}dG(\omega). \tag{B.9}$$

The assertion of Lemma B.3 follows from (B.6)-(B.9). □

Proof of Lemma B.4. We first prove that the set of IC constraints for τ given by (3.10) can be reduced to $\mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_j} \geq \mathcal{R}^T(\tau_j, \tau_i; q) + \mathcal{R}^W$ for all $i, j \in \{2, .., M\}$. Next, we show that the set of IC constraints for τ can be further reduced to $\mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_j} \geq \mathcal{R}^T(\tau_j, \tau_i; q) + \mathcal{R}^W(\tau_j, \tau_i; q)$ for only $i > j$, $i, j \in \{2, .., M\}$.

Using Lemma B.3, we can rewrite the set of IC_1 constraints , given by (3.10), as follows,

$$\mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_j} \geq - \iint_{\omega}^{\sigma^*(\tau_j;\tau_i,\omega)} C_\theta(q(\tau_j,\hat{\omega});\Theta(\tau_j,\hat{\omega}))\Theta_\omega(\tau_j,\hat{\omega})d\hat{\omega}dG(\omega). \tag{B.10}$$

Below, we prove that the RHS of (B.10) is equal to $\mathcal{R}^T(\tau_j,\tau_i;q) + \mathcal{R}^W(\tau_j,\tau_i;q)$. We have,

$$\begin{aligned}
\mathcal{R}^T(\tau_j,\tau_i;q) &= \int [C(q(\tau_j,\omega);\Theta(\tau_j,\omega)) - C(q(\tau_j,\omega);\Theta(\tau_i,\omega))]\, dG(\omega) \\
&= \iint_{\Theta(\tau_i,\omega)}^{\Theta(\tau_j,\omega)} C_\theta(q(\tau_j,\omega);\hat{\theta})d\hat{\theta}dG(\omega) \\
&= \iint_{\sigma^*(\tau_j;\tau_i,\omega)}^{\omega} C_\theta(q(\tau_j,\omega);\Theta(\tau_j,\hat{\omega}))\Theta_\omega(\tau_j,\hat{\omega})d\hat{\omega}dG(\omega) \\
&= - \iint_{\omega}^{\sigma^*(\tau_j;\tau_i,\omega)} C_\theta(q(\tau_j,\omega);\Theta(\tau_j,\hat{\omega}))\Theta_\omega(\tau_j,\hat{\omega})d\hat{\omega}dG(\omega),
\end{aligned} \tag{B.11}$$

where the third equality results from a change of variable from $\hat{\theta}$ to $\hat{\omega}$ as $\hat{\theta} := \Theta(\hat{\omega}, \tau_j)$. Note that $\Theta(\tau_i, \omega) = \Theta(\tau_j, \sigma^*(\tau_j; \tau_i, \omega))$ by Lemma B.2, thus, the new boundaries of integration with respect to $\hat{\omega}$ in (B.11) are given by $\sigma^*(\tau_j; \tau_i, \omega)$ and ω .

Adding $R^T(\tau_j,\tau_i;q)$, given by (B.11), to $R^W(\tau_j,\tau_i;q)$, we obtain,

$$\begin{aligned}&\mathcal{R}^T(\tau_j,\tau_i;q)+\mathcal{R}^W(\tau_j,\tau_i;q)=\\&-\iint_{\omega}^{\sigma^*(\tau_j;\tau_i,\omega)}C_\theta(q(\tau_j,\hat{\omega});\Theta(\tau_j,\hat{\omega}))\Theta_\omega(\tau_j,\hat{\omega})d\hat{\omega}dG(\omega),\end{aligned}\tag{B.12}$$

which is equal to RHS of (B.10).

Now we show that the set of IC constraints for τ can be further reduced to $\mathcal{R}_{\tau_i}-\mathcal{R}_{\tau_j}\geq\mathcal{R}^T(\tau_j,\tau_i;q)+\mathcal{R}^W(\tau_j,\tau_i;q)$ for only $i>j$, $i,j\in\{1,..,M\}$.

Using (B.12), we can write the set of inequality constraints (B.10) as,

$$\begin{aligned}&\mathcal{R}^T(\tau_j,\tau_i;q)+\mathcal{R}^W(\tau_j,\tau_i;q)\\&\qquad\leq\mathcal{R}_{\tau_i}-\mathcal{R}_{\tau_j}\leq\\&-\mathcal{R}^T(\tau_i,\tau_j;q)-\mathcal{R}^W(\tau_i,\tau_j;q)\quad\forall i,j\in\{1,...,M\},i>j,\end{aligned}\tag{B.13}$$

where the lower bound is from IC_1 constraint that ensures τ_i does not report τ_j, and the upper bound is from the IC_1 constraint that ensures τ_j does not report τ_i. We note that allocation $q(\tau,\omega)$ is implementable if and only if the set of constraints described by (B.13) has a feasible solution for $\mathcal{R}_{\tau_i}$, $i\in\{1,2,..,M\}$. Given an allocation rule $q(\tau,\omega)$, the set of constraints given by (B.13), has a feasible solutions only if for any arbitrary increasing sequence $(k_1,k_2),(k_2,k_3),\cdots,(k_{m-1},k_m)$, where $\tau_{k_{r-1}}<\tau_{k_r}$ for

$2 \leq r \leq m$,

$$\begin{aligned}
&\mathcal{R}^T(\tau_{k_1},\tau_{k_2};q)+\mathcal{R}^W(\tau_{k_1},\tau_{k_2};q)\\
&+\mathcal{R}^T(\tau_{k_2},\tau_{k_3};q)+\mathcal{R}^W(\tau_{k_2},\tau_{k_3};q)\\
&\ \vdots\\
&+\mathcal{R}^T(\tau_{k_{m-1}},\tau_{k_m};q)+\mathcal{R}^W(\tau_{k_{m-1}},\tau_{k_m};q)\\
&\quad \leq \sum_{r=2}^{m}\left[\mathcal{R}_{\tau_{k_r}}-\mathcal{R}_{\tau_{k_{r-1}}}\right]\\
&\quad = \mathcal{R}_{\tau_{k_m}}-\mathcal{R}_{\tau_{k_1}}\\
&\quad \leq -\mathcal{R}^T(\tau_{k_m},\tau_{k_1};q)-\mathcal{R}^W(\tau_{k_m},\tau_{k_1};q).
\end{aligned} \tag{B.14}$$

In the following, we show that for a given implementable allocation $q(\tau,\omega)$, the set of constraints (B.13) can be reduced to

$$\mathcal{R}_{\tau_i}-\mathcal{R}_{\tau_j} \geq \mathcal{R}^T(\tau_j,\tau_i;q)+\mathcal{R}^W(\tau_j,\tau_i;q) \quad \forall i,j\in\{1,...,M\}, i>j. \tag{B.15}$$

We note that $\mathcal{W}=\mathcal{S}-(1-\alpha)\mathcal{R}$, for $\alpha\in[0,1]$. Therefore, for a given allocation function $q(\tau,\omega)$, the designer wants to minimize $\mathcal{R}$. Let $\mathcal{R}^*_\tau$ denote the optimal seller's revenue that satisfies (B.15) for a given implementable allocation function $q(\tau,\omega)$. Construct a graph with M nodes, where there is an edge between node i and j if (B.15) is binding for i and j, $i>j$.

First, we note that the resulting graph must be connected. If not, then there exist at least two unconnected clusters of nodes. Consider a cluster that does not include node 1. Then one can reduce the value of $\mathcal{R}^*_\tau$ by $\epsilon>0$ for all the nodes in that cluster without violating any of the constraints (B.15), and improve the outcome of the mechanism by reducing the seller's revenue $\mathcal{R}$.

Now, assume that the optimal seller's revenue $\mathcal{R}^*_\tau$ that is determined by only considering the set of constraints (B.15), does not satisfy the set of constraints (B.13).

Then, there exists $i > j$ so that the following constraint is violated,

$$\mathcal{R}^*_{\tau_i} - \mathcal{R}^*_{\tau_j} \nleq -\mathcal{R}^T(\tau_i, \tau_j; q) - \mathcal{R}^W(\tau_i, \tau_j; q).$$

Let $C = \{(i, k_1), (k_1, k_2), \cdots, (k_m, j)\}$ denote a path between node i and node j in the connected graph constructed above. Then,

$$\begin{aligned}\mathcal{R}^*_{\tau_i} - \mathcal{R}^*_{\tau_j} &= \sum_{(\hat{k},k)\in C} \mathcal{R}^T(\tau_{\hat{k}}, \tau_k; q) + \mathcal{R}^W(\tau_{\hat{k}}, \tau_k; q) \\ &\nleq -\mathcal{R}^T(\tau_i, \tau_j; q) - \mathcal{R}^W(\tau_i, \tau_j; q), \qquad \text{(B.16)}\end{aligned}$$

which contradicts (B.14).

□

Proof of Lemma B.5. We can rewrite the set of IC constraints (3.6) as follows,

$$\begin{aligned}\mathcal{R}_{\tau_i} &\geq t(\tau_j) - \mathbb{E}_\omega\{C(q(\tau_j); \Theta(\tau_i, \omega))\} \\ &= \mathcal{R}_{\tau_j} + \Big[\mathbb{E}_\omega\{C(q(\tau_j); \Theta(\tau_j, \omega))\} - \mathbb{E}_\omega\{C(q(\tau_j); \Theta(\tau_i, \omega))\}\Big], \\ &= \mathcal{R}_{\tau_j} + \mathcal{R}^T(\tau_j, \tau_i; q) \quad \forall i, j \in \{1, ..., M\},\end{aligned}$$

where the last equality holds by definition. We proceed as follows. We first consider a relaxed version of the forward mechanism design problem. We determine the solution to the relaxed problem, and show it is also a feasible solution for the original forward mechanism design problem.

Consider the following relaxed version of the forward mechanism design problem,

$$
\begin{aligned}
&\max \mathcal{W} \\
&\text{subject to} \\
&\mathcal{R}_{\tau_i} \geq t(\tau_{i-1}) - \mathbb{E}_\omega\{C(q(\tau_{i-1}); \Theta(\tau_i, \omega))\} \\
&\quad = \mathcal{R}_{\tau_{i-1}} + \Big[\mathbb{E}_\omega\{C(q(\tau_{i-1}); \Theta(\tau_{i-1}, \omega))\} - \mathbb{E}_\omega\{C(q(\tau_{i-1}); \Theta(\tau_i, \omega))\}\Big], \\
&\qquad\qquad\qquad i \in \{2, ..., M\}, \qquad \text{(B.17)} \\
&\mathcal{R}_{\tau_i} \geq 0, i \in \{1, ..., M\}, \qquad \text{(B.18)}
\end{aligned}
$$

where we only include the set of IC constraints that ensures a seller with type τ_i does not report type τ_{i-1}, and the other IC constraints are omitted.

Note that $\mathbb{E}_\omega\{C(q(\tau_{i-1}); \Theta(\tau_{i-1}, \omega))\} \geq \mathbb{E}_\omega\{C(q(\tau_{i-1}); \Theta(\tau_i, \omega))\}$ since $\Theta(\tau_{i-1}, \omega) \geq \Theta(\tau_i, \omega)$ by Assumption 3.1. Thus, from (B.17) we have $\mathcal{R}_{\tau_i} \geq \mathcal{R}_{\tau_j}$ for $i > j$. The designer's objective is $\mathcal{W} = \mathcal{S} - (1 - \alpha)\mathcal{R}$. Therefore, for a given allocation function $q(\tau, \omega)$, the designer wants to minimize $\mathcal{R}_{\tau_i}$ $i \in \{1, 2, ..., M\}$. Therefore, at the optimal solution to the relaxed problem,

$$
\begin{aligned}
&\mathcal{R}_{\tau_i} = \mathcal{R}_{\tau_{i-1}} + \Big[\mathbb{E}_\omega\{C(q(\tau_{i-1}); \Theta(\tau_{i-1}, \omega))\} - \mathbb{E}_\omega\{C(q(\tau_{i-1}); \Theta(\tau_i, \omega))\}\Big], \\
&\mathcal{R}_{\tau_1} = 0.
\end{aligned}
$$

Therefore, we can write,

$$
\mathcal{R}_{\tau_i} = \sum_{j=1}^{i-1} \Big[\mathbb{E}_\omega\{C(q(\tau_j); \Theta(\tau_j, \omega))\} - \mathbb{E}_\omega\{C(q(\tau_j); \Theta(\tau_{j+1}, \omega))\}\Big]. \qquad \text{(B.19)}
$$

Substituting $\mathcal{R}_{\tau_i}$, $1 \le i \le M$, we get

$$
\begin{aligned}
&\mathcal{S} - (1-\alpha)\sum_{i=1}^{M} p_i \mathcal{R}_{\tau_i} \\
&\qquad = \sum_{i=1}^{n} p_i \left[\mathcal{V}(q(\tau_i)) - \mathbb{E}_\omega\{C(q(\tau_i);\Theta(\tau_i,\omega))\}\right] \\
&- (1-\alpha)\sum_{i=1}^{M} p_i \sum_{j=1}^{i-1}\Big[\mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_j,\omega))\} - \mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_{j+1},\omega))\}\Big] \\
&\qquad = \sum_{i=1}^{M} p_i \left[\mathcal{V}(q(\tau_i)) - \mathbb{E}_\omega\{C(q(\tau_i);\Theta(\tau_i,\omega))\}\right] - \\
&(1-\alpha)\sum_{j=1}^{M-1}\left(\sum_{i=j+1}^{M} p_i\right)\Big[\mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_j,\omega))\} - \mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_{j+1},\omega))\}\Big] \\
&\qquad = \sum_{i=1}^{M} p_i \Big(\mathcal{V}(q(\tau_i)) - \mathbb{E}_\omega\{C(q(\tau_i);\Theta(\tau_i,\omega))\} - \\
&(1-\alpha)\left(\sum_{j=i+1}^{M} \frac{p_j}{p_i}\right)\Big[\mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_j,\omega))\} - \mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_{j+1},\omega))\}\Big]\Big).
\end{aligned}
$$

By maximizing the above expression with respect to $q(\tau_i)$ and using the first-order condition, we find that the optimal $q(\tau_i)$ is determined by the following equation

$$
\begin{aligned}
v(q) =& \mathbb{E}_\omega\{c(q;\Theta(\tau_i,\omega))\} \\
&+ (1-\alpha)\sum_{j=i+1}^{M} \frac{p_i}{p_j}\Big[\mathbb{E}_\omega\{c(q;\Theta(\tau_i,\omega))\} - \mathbb{E}_\omega\{c(q;\Theta(\tau_{i+1},\omega))\}\Big].
\end{aligned} \tag{B.20}
$$

The above equation has a unique solution since the LHS is decreasing in q by the concavity of $\mathcal{V}(q)$, and the RHS is increasing in q by Assumption 3.1. Moreover, note that the RHS is decreasing in i by Assumption 3.1, therefore, $q(\tau_i) > q(\tau_j)$ for $i > j$.

Now, we show that the optimal solution to the relaxed problem satisfies the omitted constraints. Consider the IC constraint for type τ_i reporting τ_j. We need to show that the utility of a seller with technology τ_i is higher when he reports truthfully than his utility from misreporting τ_j. That is,

$$\begin{aligned}\mathcal{R}_{\tau_i} - [t(\tau_j) - \mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_i,\omega))\}] =& \mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_j} \\ &+ [\mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_i,\omega))\} - \mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_j,\omega))\}] \geq 0.\end{aligned}$$

For $i > j$, we can write,

$$\begin{aligned}&\mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_j} + [\mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_i,\omega))\} - \mathbb{E}_\omega\{C(q(\tau_j,\omega);\Theta(\tau_j,\omega))\}] \\ =& \sum_{k=j}^{i-1}\Big[\mathbb{E}_\omega\{C(q(\tau_k);\Theta(\tau_k,\omega))\} - \mathbb{E}_\omega\{C(q(\tau_k);\Theta(\tau_{k+1},\omega))\}\Big] \\ &+ [\mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_i,\omega))\} - \mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_j,\omega))\}] \\ =& \sum_{k=j}^{i-1}\Big[\mathbb{E}_\omega\{C(q(\tau_k);\Theta(\tau_k,\omega))\} - \mathbb{E}_\omega\{C(q(\tau_k);\Theta(\tau_{k+1},\omega))\}\Big] \\ &- \sum_{k=j}^{i-1}\Big[\mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_k,\omega))\} - \mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_{k+1},\omega))\}\Big] \geq 0,\end{aligned}$$

since $q(\tau_k) \geq q(\tau_j)$ and $C(\hat{q};\Theta(\tau_k,\omega)) - C(\hat{q};\Theta(\tau_{k+1},\omega))$ is increasing in $\hat{q}$ by Assumption 3.1.

Similarly for $i < j$, we have,

$$
\begin{aligned}
&\mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_j} + [\mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_i,\omega))\} - \mathbb{E}_\omega\{C(q(\tau_j,\omega);\Theta(\tau_j,\omega))\}] \\
&= -\sum_{k=i}^{j-1}\Big[\mathbb{E}_\omega\{C(q(\tau_k);\Theta(\tau_k,\omega))\} - \mathbb{E}_\omega\{C(q(\tau_k);\Theta(\tau_{k+1},\omega))\}\Big] \\
&\quad + \Big[\mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_i,\omega))\} - \mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_j,\omega))\}\Big] \\
&= -\sum_{k=i}^{j-1}\Big[\mathbb{E}_\omega\{C(q(\tau_k);\Theta(\tau_k,\omega))\} - \mathbb{E}_\omega\{C(q(\tau_k);\Theta(\tau_{k+1},\omega))\}\Big] \\
&\quad + \sum_{k=j}^{i-1}\Big[\mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_k,\omega))\} - \mathbb{E}_\omega\{C(q(\tau_j);\Theta(\tau_{k+1},\omega))\}\Big] \geq 0,
\end{aligned}
$$

where the last inequality is true since $q(\tau_k) \leq q(\tau_j)$ for $k \leq j$, and $C(\hat{q};\Theta(\tau_k,\omega)) - C(\hat{q};\Theta(\tau_{k+1},\omega))$ is increasing in $\hat{q}$ by Assumption 3.1.

□

Proof of Lemma B.6. First we note that in the real-time mechanism the seller reports τ and ω simultaneously, and his cost only depends $\Theta(\tau,\omega)$. Therefore, allocation function $q(\tau,\omega)$ and payment function $t(\tau,\omega)$ must only depend on $\Theta(\tau,\omega)$ rather than exact values of τ and ω. Therefore, the mechanism design problem can be written only in terms of θ, where the buyer designs $\{q(\theta), t(\theta)\}$, and asks the seller to report θ instead of (τ,ω). The reformulation of the real-time mechanism in terms of θ can be written as follows,

$$\max_{q(\cdot)} \mathcal{W} \tag{B.21}$$

subject to

$$IC : \mathcal{R}_\theta \geq t(\hat{\theta}) - C(q(\hat{\theta});\theta), \quad \forall\theta, \tag{B.22}$$

$$IR : \mathcal{R}_\theta \geq 0, \quad \forall\theta. \tag{B.23}$$

Claim B.1. *The sets of IC constraints (*B.*22) and IR constraints (*B.*23) are satisfied* *if and only if*

$$\frac{\partial \mathcal{R}_\theta}{\partial \theta} = -C_\theta(q(\theta);\theta), \tag{B.24}$$

$$\mathcal{R}_{\overline{\theta}} \geq 0, \tag{B.25}$$

and $q(\theta)$ is decreasing in θ.

We prove the result of Claim B.1 below. Assume (B.22) is true. Consider the following two IC constraints, where the first one requires that a seller with type θ does not gain by misreporting $\hat{\theta}$, and the second one requires that a seller with true type $\hat{\theta}$ does not gain by misreporting θ.

$$t(\hat{\theta}) - C(q(\hat{\theta});\hat{\theta}) \geq t(\theta) - C(q(\theta);\hat{\theta}),$$
$$t(\theta) - C(q(\theta);\theta) \geq t(\hat{\theta}) - C(q(\hat{\theta});\theta).$$

By subtracting the above two IC constraints, we obtain,

$$C(q(\hat{\theta});\hat{\theta}) - C(q(\hat{\theta});\theta) \leq \mathcal{R}_\theta - \mathcal{R}_{\hat{\theta}} \leq C(q(\theta);\hat{\theta}) - C(q(\theta);\theta). \tag{B.26}$$

Let $\hat{\theta} = \theta + \epsilon$ and take $\epsilon \to 0$. Then $\frac{\partial \mathcal{R}_\theta}{\partial \theta} = -C_\theta(q(\theta);\theta)$. Moreover, by (B.26),

$$C(q(\hat{\theta});\hat{\theta}) - C(q(\hat{\theta});\theta) \leq C(q(\theta);\hat{\theta}) - C(q(\theta);\theta),$$

which along with Assumption 3.1, implies that $q(\theta) > q(\hat{\theta})$ for $\theta < \hat{\theta}$.

Furthermore, by Assumption 3.1 $\frac{\partial \mathcal{R}_\theta}{\partial \theta} = -C_\theta(q(\theta);\theta) \leq 0$. Therefore, (B.23) is satisfied if and only if $\mathcal{R}_{\overline{\theta}} \geq 0$.

To prove the converse, assume that (B.24) and (B.25) hold and $q(\theta)$ is decreasing in θ. First, we show that the IC constraints (B.22) are satisfied. For any θ and $\hat{\theta}$, we have

$$\mathcal{R}_\theta - \Big[t(\hat\theta) - C(q(\hat\theta);\theta)\Big] = [\mathcal{R}_\theta - \mathcal{R}_{\hat\theta}] - \Big[C(q(\hat\theta);\hat\theta) - C(q(\hat\theta);\theta)\Big]$$
$$= \int_\theta^{\hat\theta} C_\theta(q(\breve\theta);\breve\theta)d\breve\theta - \Big[C(q(\hat\theta);\hat\theta) - C(q(\hat\theta);\theta))\Big] \geq 0,$$

where the second equality follows from (B.24) and the last inequality is true since $C_\theta(\hat q;\theta)$ is decreasing in θ by Assumption 3.1, and $q(\theta)$ is decreasing in θ.

Moreover, we have $\mathcal{R}_\theta = \mathcal{R}_{\overline\theta} - \int_\theta^{\overline\theta} C_\theta(q(\hat\theta);\hat\theta)d\hat\theta \geq 0$. This completes the proof of Claim 1.

We note that $q(\tau,\omega)$ only depends on $\Theta(\tau,\omega)$. Thus, the sets of constraints given by (B.24) and (3.13) are equivalent.

Next, we show that we can replace the set of IR constraints (3.9) by $\mathcal{R}_{\tau_1} = R^P(\tau_1;q)$. Equation (B.26) implies that $\mathcal{R}_\theta \geq \mathcal{R}_{\hat\theta}$ for $\theta < \hat\theta$, since $q(\theta)$ is decreasing in θ. That is, $\mathcal{R}_{\Theta(\tau_1,\underline\omega)} \leq \mathcal{R}_{\hat\theta}$ for all $\hat\theta$. Thus, the set of IR constraints (3.9) is satisfied if and only if $\mathcal{R}_{\tau_1,\underline\omega} \geq 0$. Therefore, by (B.24), we can write,

$$\begin{aligned}
\mathcal{R}_{\tau_1} &= \mathbb{E}_\omega\{\mathcal{R}_{\Theta(\tau_1,\omega)}\} + \mathcal{R}_{\tau_1,\underline\omega} \\
&\geq \mathbb{E}_\omega\{\mathcal{R}_{\Theta(\tau_1,\omega)}\} \\
&= \int_{\underline\omega}^{\omega}\int_{\underline\omega}^{\hat\omega} C_\theta(q(\tau_1,\hat\omega);\Theta(\tau_1,\hat\omega))\Theta_\omega(\tau_1,\hat\omega)d\hat\omega dG(\omega) \\
&= \int_{\underline\omega}^{\overline\omega}\int_{\hat\omega}^{\overline\omega} C_\theta(q(\tau_1,\hat\omega);\Theta(\tau_1,\hat\omega))\Theta_\omega(\tau_1,\hat\omega)dG(\omega)d\hat\omega \\
&= \int_{\underline\omega}^{\overline\omega} [1-G(\hat\omega)]\, C_\theta(q(\tau_1,\hat\omega);\Theta(\tau_1,\hat\omega))\Theta_\omega(\tau_1,\hat\omega)d\hat\omega \\
&= \mathcal{R}^P(\tau_1;q).
\end{aligned}$$

Finally, we show that $\mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_{i-1}} = \mathcal{R}^T(\tau_{i-1},\tau_i;q) + \mathcal{R}^W(\tau_{i-1},\tau_i;q)$. Using (B.24), we have

$$\begin{aligned}\mathcal{R}_{\tau_i}-\mathcal{R}_{\tau_{i-1}} &= \mathbb{E}_{\hat{\omega}}\{\mathcal{R}_{\tau_{i-1},\hat{\omega}}\}-\mathbb{E}_{\hat{\omega}}\{\mathcal{R}_{\tau_{i-1},\hat{\omega}}\}\\ &= \mathbb{E}_{\hat{\omega}}\{\mathcal{R}_{\Theta(\tau_i,\hat{\omega})}-\mathcal{R}_{\Theta(\tau_{i-1},\hat{\omega})}\}\\ &= -\iint_{\Theta(\tau_{i-1},\hat{\omega})}^{\Theta(\tau_i,\hat{\omega})} C_\theta(q(\breve{\theta});\breve{\theta})d\breve{\theta}dG(\hat{\omega}).\end{aligned}$$

By the change of variable $\breve{\theta}=\Theta(\tau_{i-1},\hat{\omega})$, we obtain,

$$\begin{aligned}&\iint_{\Theta(\tau_{i-1},\omega)}^{\Theta(\tau_i,\omega)} C_\theta(q(\breve{\theta});\breve{\theta})d\breve{\theta}dG(\omega)\\ &= -\iint_{\omega}^{\sigma^*(\tau_{i-1};\tau_i,\omega)} C_\theta(q(\tau_{i-1},\hat{\omega});\Theta(\tau_{i-1},\hat{\omega}))\Theta_\omega(\tau_{i-1},\hat{\omega})d\hat{\omega}dG(\omega)\\ &= \iint_{\sigma^*(\hat{\tau};\tau,\omega)}^{\omega} C_\theta(q(\hat{\tau},\hat{\omega});\Theta(\hat{\tau},\hat{\omega}))\Theta_\omega(\hat{\tau},\hat{\omega})d\hat{\omega}dG(\omega)\end{aligned}$$

which is the same as the RHS of (B.10). Thus, from the proof of Lemma B.4, we have $\mathcal{R}_{\tau_i}-\mathcal{R}_{\tau_{i-1}}=\mathcal{R}^T(\tau_{i-1},\tau_i;q)+\mathcal{R}^W(\tau_{i-1},\tau_i;q)$.

□

Proof of Lemma B.7. The IC constraint (3.34) for the dynamic mechanism with monitoring is given by,

$$\mathbb{E}_\omega\{\mathcal{R}_{\tau_i,\omega}\}\geq\mathbb{E}_\omega\{t(\tau_j,\omega)-C(q(\tau_j,\omega);\Theta(\tau_i,\omega))\}\ \ \forall\tau,\tau_j.$$

We can rewrite the above IC constraint as,

$$\begin{aligned}\mathcal{R}_{\tau_i}-\mathcal{R}_{\tau_j} &\geq \mathbb{E}_\omega\{t(\tau_j,\omega)-C(q(\tau_j,\omega);\Theta(\tau_i,\omega))\}-\mathcal{R}_{\tau_j}\\ &= \int\left[C(q(\tau_j,\omega);\Theta(\tau_j,\omega))-C(q(\tau_j,\omega);\Theta(\tau_i,\omega))\right]dG(\omega)\\ &= \mathcal{R}^T(\tau_j,\tau_i;q)\ \ \forall\tau_i,\tau_j.\end{aligned}$$

Therefore, we can replace the set of IC constraints (3.34) by $\mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_i} \geq R^T(\tau_j, \tau_i; q)$ for all $i, j \in [1, ..., M]$.

Using an argument similar to that of the proof of Lemma B.4, we can further reduce the set of IC constraints to $\mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_j} \geq \mathcal{R}^T(\tau_j, \tau_i; q)$ for only $i > j,\ i, j \in \{1, .., M\}$. $\square$

B.3 Closed Form Solutions

In this appendix, we provide closed form solutions for the mechanism design problems formulated in Sections 3.4 and 3.6.

B.3.1 The forward mechanism

Using the results of Lemmas B.5, the optimal allocation $q(\tau)$ is given by the unique solution to the following equation,

$$\begin{aligned} v(q) =& \mathbb{E}_\omega\{c(q;\Theta(\tau_i,\omega))\} \\ & + (1-\alpha)\sum_{j=i+1}^{i} \frac{p_j}{p_i}\Big[\mathbb{E}_\omega\{c(q;\Theta(\tau_i,\omega))\} - \mathbb{E}_\omega\{c(q;\Theta(\tau_{i+1},\omega))\}\Big]; \end{aligned}$$

the optimal payment function $t(\tau_i,\omega)$, $i \in \{1,2,...,M\}$ is given by,

$$\begin{aligned} t(\tau_i) =& \mathbb{E}_\omega\{C(q(\tau_i);\Theta(\tau_i,\omega))\} + \mathcal{R}_{\tau_i} \\ =& \mathbb{E}_\omega\{C(q(\tau_i);\Theta(\tau_i,\omega))\} \\ & + \sum_{j=1}^{i+1}\Big[\mathbb{E}_\omega\{C(q(\tau_j,\omega);\Theta(\tau_j,\omega))\} - \mathbb{E}_\omega\{C(q(\tau_j,\omega);\Theta(\tau_{j+1},\omega))\}\Big]. \end{aligned}$$

Proof. The proof follows directly from equations (B.19,B.20) (in the proof of Lemma B.5), which determine the optimal allocation function $q(\tau)$ and the seller's revenue $\mathcal{R}_\tau$. $\square$

B.3.2 The real-time mechanism

Using the result of Lemma B.6, the optimal allocation $q(\tau,\omega)$ is given by the unique solution to the following equation,[1]

$$v(q) = c(q;\Theta(\tau,\omega)) + (1-\alpha)\frac{F(\Theta(\tau,\omega))}{f(\Theta(\tau,\omega))}c_\theta(q;\Theta(\tau,\omega));$$

the optimal payment function $t(\tau,\omega)$ is given by,

$$\begin{aligned} t(\tau,\omega) &= C(q(\tau,\omega);\Theta(\tau,\omega)) + \mathcal{R}_{\tau,\omega} \\ &= C(q(\tau,\omega);\Theta(\tau,\omega)) + \int_{\Theta(\tau,\omega)}^{\overline{\theta}} C_\theta(q(\theta);\theta)d\theta. \end{aligned}$$

Proof. Using (B.24) in the proof of Lemma B.6, we can write,

$$\mathcal{R}_\theta = \int_\theta^{\overline{\theta}} C_\theta(q(\theta);\theta)d\theta,$$

by setting $\mathcal{R}_{\overline{\theta}} = \mathcal{R}_{\tau_1,\underline{\omega}} = 0$. Let $F(\theta) := \sum_{i=1}^M p_i F_i(\theta)$ and $f(\theta) := \sum_{i=1}^M p_i f_i(\theta)$. Then, we can write,

$$\begin{aligned} \mathcal{W} &= \mathcal{S} - (1-\alpha)\mathcal{R} \\ &= \int [\mathcal{V}(q(\theta)) - C(q(\theta);\theta)]\, dF(\theta) - (1-\alpha)\int\int_{\hat{\theta}}^{\overline{\theta}} C_\theta(q(\theta);\theta)d\theta dF(\hat{\theta}) \\ &= \int [\mathcal{V}(q(\theta)) - C(q(\theta);\theta)] dF(\theta) - (1-\alpha)\int\int_{\underline{\theta}}^{\theta} C_\theta(q(\theta);\theta)dF(\hat{\theta})d\theta \\ &= \int [\mathcal{V}(q(\theta)) - C(q(\theta);\theta)]\, dF(\theta) - (1-\alpha)\int C_\theta(q(\theta);\theta)F(\theta)d\theta \\ &= \int \left[\mathcal{V}(q(\theta)) - C(q(\theta);\theta) - (1-\alpha)\frac{F(\theta)}{f(\theta)}C_\theta(q(\theta);\theta)\right] dF(\theta). \end{aligned}$$

[1]We assume that $c(q;\theta) + (1-\alpha)\frac{F(\theta)}{f(\theta)}c_\theta(q;\theta)$ is increasing in q for all θ.

Hence, using Claim B.1 in the proof of Lemma B.6, we can rewrite the optimization problem (B.21) as follows,

$$\max_{q(\cdot)} \int \left[\mathcal{V}(q(\theta)) - C(q(\theta);\theta) - (1-\alpha)\frac{F(\theta)}{f(\theta)} C_\theta(q(\theta);\theta) \right] dF(\theta) \tag{B.27}$$

subject to (B.28)

$q(\theta)$ is decreasing in θ.

Consider a relaxed version of the above optimization problem by ignoring the monotonicity constraint on $q(\theta)$. The optimal solution $q(\theta)$ to the relaxed problem is determined by maximizing the integrand in (B.28) for each θ, and is given by unique solution to the following equation,

$$v(q) = c(q;\theta) + (1-\alpha)\frac{F(\theta)}{f(\theta)} c_\theta(q;\theta). \tag{B.29}$$

We note that if $c(q;\theta) + (1-\alpha)\frac{F(\theta)}{f(\theta)} c_\theta(q;\theta)$ is increasing in θ, then the optimal $q(\theta)$ determined above is decreasing in θ, and thus, automatically satisfies the ignored monotonicity condition on $q(\theta)$. □

B.3.3 The dynamic mechanism with monitoring

Using the results of Lemma B.7, the optimal allocation $q(\tau,\omega)$ is given by the unique solution to the following equation,

$$v(q) = c(q;\Theta(\tau_i,\omega)) + (1-\alpha)\sum_{j=i+1}^{M} \frac{p_j}{p_i} \Big[c(q;\Theta(\tau_i,\omega)) - c(q;\Theta(\tau_{i+1},\omega)) \Big];$$

the optimal payment function $t(\tau_i, \omega)$, $i \in \{1, 2, ..., M\}$ is given by,

$$\begin{aligned} t(\tau_i, \omega) &= C(q(\tau_i, \omega); \Theta(\tau_i, \omega)) + \mathcal{R}_{\tau_i, \omega} \\ &= C(q(\tau_i, \omega); \Theta(\tau_i, \omega)) \\ &\quad + \sum_{j=1}^{i+1} \Big[C(q(\tau_j, \omega); \Theta(\tau_j, \omega)) - C(q(\tau_j, \omega); \Theta(\tau_{j+1}, \omega)) \Big]. \end{aligned}$$

Proof. Using the result of Lemma B.7, the dynamic mechanism with monitoring is given by the solution to the following optimization problem,

$$\begin{aligned} &\max \ \ \mathcal{W} \\ &\text{subject to} \\ &\qquad \mathcal{R}_{\tau_i} - \mathcal{R}_{\tau_{i-1}} = R^T(\tau_{i-1}, \tau_i; q), i \in \{2, ..., M\}, \\ &\qquad \mathcal{R}_{\tau_1} \geq 0. \end{aligned}$$

Note that $\mathcal{W} = \mathcal{S} - (1 - \alpha)\mathcal{R}$. As a result, at the optimal solution of the relaxed problem we have,

$$\begin{aligned} &\mathcal{R}_{\tau_i} = \mathcal{R}_{\tau_{i-1}} + R^T(\tau_{i-1}, \tau_i; q), i \in \{2, ..., M\}, \\ &\mathcal{R}_{\tau_1} = 0. \end{aligned}$$

Substituting R_{τ_i}, $1 \le i \le M$, we obtain,

$$
\begin{aligned}
\mathcal{W} = & \mathcal{S} - (1-\alpha) \sum_{j=1}^{n} p_j \mathcal{R}_{\tau_j} \\
= & \mathcal{S} - (1-\alpha) \sum_{j=1}^{M} \left(p_j \sum_{i=1}^{j-1} R^T(\tau_i, \tau_{i+1}; q) \right) \\
= & \mathcal{S} - (1-\alpha) \sum_{i=1}^{M-1} \left(\sum_{j=i+1}^{M} p_j \right) R^T(\tau_i, \tau_{i+1}; q) \\
= & \sum_{i=1}^{M} p_i \int [\mathcal{V}(q(\tau_i,\omega)) - C(q(\tau_i,\omega); \Theta(\tau_i,\omega))] dG(\omega) - (1-\alpha) \Bigg[\\
& \sum_{i=1}^{M-1} \left(\sum_{j=i+1}^{M} p_j \right) \int [C(q(\tau_i,\omega); \Theta(\tau_i,\omega)) - C(q(\tau_i,\omega); \Theta(\tau_{i+1},\omega))] dG(\omega) \Bigg] \\
= & \sum_{i=1}^{M} p_i \int \Bigg[\mathcal{V}(q(\tau_i,\omega)) - C(q(\tau_i,\omega); \Theta(\tau_i,\omega)) - (1-\alpha) \Bigg[\\
& \sum_{j=i+1}^{M} \frac{p_j}{p_i} \Big[C(q(\tau_i,\omega); \Theta(\tau_i,\omega)) - C(q(\tau_i,\omega); \Theta(\tau_{i+1},\omega)) \Big] \Bigg] dG(\omega) \Bigg].
\end{aligned}
$$

By maximizing the integrand point-wise with respect to $q(\tau_i,\omega)$ and using the first-order condition, we find that the optimal value of $q(\tau_i,\omega)$ is determined by the following equation,

$$
v(q) = c(q; \Theta(\tau_i,\omega)) + (1-\alpha) \sum_{j=i+1}^{M} \frac{p_i}{p_j} \Big[c(q; \Theta(\tau_i,\omega)) - c(q; \Theta(\tau_{i+1},\omega)) \Big].
$$

The above equation has a unique solution since the LHS is decreasing in q by the concavity of $\mathcal{V}(q)$, and the RHS is increasing in q by Assumption 3.1.

$\square$

B.3.4 The dynamic mechanism

For the dynamic mechanism there exists no closed form solution for arbitrary M and parameters of the model, since the set of binding constraints from the inequality constraints given by (3.23) cannot be determined a priori, and depends on the allocation function $q(\tau,\omega)$ (see [12] for more discussion). However, for the case with two possible technologies, *i.e.* $M=2$, we provide the closed form solutions for the dynamic mechanism in the following.

If $c(\hat{q};\Theta(\tau_1,\omega))-\Theta_\omega(\tau_1,\omega)\frac{p_{\tau_2}}{p_{\tau_1}}\frac{G(\omega)-G(\sigma^*(\tau_2;\tau_1,\omega))}{g(\omega)}c_\theta(\hat{q};\Theta(\tau_1,\omega))$ is increasing in $\hat{q}$, for all ω, the optimal allocation $q(\tau,\omega)$ is given by the unique solution to the following equations,

$$
\begin{aligned}
v(q(\tau_2,\omega)) &= c(q(\tau_2,\omega);\Theta(\tau_2,\omega)) && \text{(B.30)}\\
v(q(\tau_1,\omega)) &= c(q(\tau_1,\omega);\Theta(\tau_1,\omega))-\Big[(1-\alpha)\frac{p_{\tau_2}}{p_{\tau_1}}\Theta_\omega(\tau_1,\omega)\\
&\qquad \frac{G(\omega)-G(\sigma^*(\tau_2;\tau_1,\omega))}{g(\omega)}c_\theta(q(\tau_1,\omega);\Theta(\tau_1,\omega))\Big]; && \text{(B.31)}
\end{aligned}
$$

the optimal payment function $t(\tau_i,\omega)$, is given by,

$$
\begin{aligned}
t(\tau_2,\omega) :=\,& C(q(\tau_2,\omega);\Theta(\tau_2,\omega))\\
&-\iint_{\sigma^*(\tau_2;\tau_1,\omega)}^{\omega} C_\theta(q(\tau_1,\omega);\Theta(\tau_1,\omega))\Theta_\omega(\tau_1,\omega)dG(\hat{\omega})d\omega\\
&-\int_{\underline{\omega}}^{\omega} C_\theta(q(\tau_2,\hat{\omega});\Theta(\tau_2,\hat{\omega}))\Theta_\omega(\tau_2,\hat{\omega})d\hat{\omega}\\
&+\int_{\underline{\omega}}^{\overline{\omega}}[1-G(\hat{\omega})]C_\theta(q(\tau_2,\hat{\omega});\Theta(\tau_2,\hat{\omega}))\Theta_\omega(\tau_2,\hat{\omega})d\hat{\omega}, && \text{(B.32)}\\
t(\tau_1,\omega) :=\,& C(q(\tau_1,\omega);\Theta(\tau_1,\omega))\\
&-\int_{\underline{\omega}}^{\omega} C_\theta(q(\tau_1,\hat{\omega});\Theta(\tau_1,\hat{\omega}))\Theta_\omega(\tau_1,\hat{\omega})d\hat{\omega}\\
&+\int_{\underline{\omega}}^{\overline{\omega}}[1-G(\hat{\omega})]C_\theta(q(\tau_1,\hat{\omega});\Theta(\tau_1,\hat{\omega}))\Theta_\omega(\tau_1,\hat{\omega})d\hat{\omega}. && \text{(B.33)}
\end{aligned}
$$

Proof. The designer's objective is $\mathcal{W} = \mathcal{S} - (1-\alpha)\mathcal{R}$. Therefore, given an allocation function $q(\tau,\omega)$, the designer wants to minimize the information rent $\mathcal{R}_\tau$ such that it satisfies the conditions of part (c) of Theorem 3.2. For $M = 2$, we have $\mathcal{R}_{\tau_1} = 0$ and $\mathcal{R}_{\tau_2} = \mathcal{R}^T(\tau_1,\tau_2;q) + \mathcal{R}^W(\tau_1,\tau_2;q)$. Using the results of Lemma B.1 and B.4, we can rewrite the dynamic mechanism design problem as,

$$\max_{q(\cdot,\cdot),t(\cdot,\cdot)} p_{\tau_2}\mathcal{S}_{\tau_2} - (1-\alpha)p_{\tau_2}\mathcal{R}_{\tau_2} + p_{\tau_1}\mathcal{S}_{\tau_1}$$

subject to

$$\mathcal{R}_{\tau_2} = \mathcal{R}^T(\tau_1,\tau_2;q) + \mathcal{R}^W(\tau_1,\tau_2;q),$$

$$\frac{\partial \mathcal{R}_{\tau_2,\omega}}{\partial \omega} = C_\theta(q(\tau_2,\omega);\Theta(\tau_2,\omega))\Theta_\omega(\tau_2,\omega), \tag{B.34}$$

$$\frac{\partial \mathcal{R}_{\tau_1,\omega}}{\partial \omega} = C_\theta(q(\tau_1,\omega);\Theta(\tau_1,\omega))\Theta_\omega(\tau_1,\omega), \tag{B.35}$$

$$q(\tau_2,\omega) \text{ and } q(\tau_1,\omega) \text{ are increasing in } \omega. \tag{B.36}$$

Consider the following relaxation of the above problem, where the last three constraints are omitted.

$$\max_{q(\cdot,\cdot),t(\cdot,\cdot)} p_{\tau_2}\mathcal{S}_{\tau_2} - (1-\alpha)p_{\tau_2}\mathcal{R}_{\tau_2} + p_{\tau_1}\mathcal{S}_{\tau_1} \tag{B.37}$$

subject to

$$\mathcal{R}_{\tau_2} = \mathcal{R}^T(\tau_1,\tau_2;q) + \mathcal{R}^W(\tau_1,\tau_2;q). \tag{B.38}$$

Below, we determine the solution to the above relaxed problem and show that its optimal solution also solves the original dynamic mechanism design problems. Using

(B.12) and (B.38), we obtain,

$$\begin{aligned}
\mathcal{R}_{\tau_2} &= -\iint_{\hat{\omega}}^{\sigma^*(\tau_1;\tau_2,\hat{\omega})} C_\theta(q(\tau_1,\omega);\Theta(\tau_1,\omega))\Theta_\omega(\tau_1,\omega)d\omega dG(\hat{\omega}) \\
&= -\iint_{\sigma^*(\tau_2;\tau_1,\omega)}^{\omega} C_\theta(q(\tau_1,\omega);\Theta(\tau_1,\omega))\Theta_\omega(\tau_1,\omega)dG(\hat{\omega})d\omega \\
&= -\int [G(\omega)-G(\sigma^*(\tau_2;\tau_1,\omega))]\, C_\theta(q(\tau_1,\omega);\Theta(\tau_1,\omega))\Theta_\omega(\tau_1,\omega)d\omega \\
&= -\int [G(\omega)-G(\sigma^*(\tau_2;\tau_1,\omega))]\, C_\theta(q(\tau_1,\omega);\Theta(\tau_1,\omega))\Theta_\omega(\tau_1,\omega)\frac{dG(\omega)}{g(\omega)},
\end{aligned}$$

where the second equality results from changing the order of integration and $\sigma^*(\tau_1;\tau_2,\sigma^*(\tau_2;\tau_1,\omega))=\omega$. Therefore, we can write the objective function (B.37) as,

$$\begin{aligned}
&p_{\tau_2}\mathcal{S}_{\tau_2}-(1-\alpha)p_{\tau_2}\mathcal{R}_{\tau_2}+p_{\tau_1}\mathcal{S}_{\tau_1}= \\
&p_{\tau_2}\int [\mathcal{V}(q(\tau_2,\omega))-C(q(\tau_2,\omega);\Theta(\tau_2,\omega)]\, dG(\omega)+ \\
&(1-\alpha)p_{\tau_2}\int [G(\omega)-G(\sigma^*(\tau_2;\tau_1,\omega))]\, C_\theta(q(\tau_1,\omega);\Theta(\tau_1,\omega))\Theta_\omega(\tau_1,\omega)\frac{dG(\omega)}{g(\omega)} \\
&+p_{\tau_1}\int [\mathcal{V}(q(\tau_1,\omega))-C(q(\tau_1,\omega);\Theta(\tau_1,\omega)]\, dG(\omega) \\
&= p_{\tau_2}\int [\mathcal{V}(q(\tau_2,\omega))-C(q(\tau_2,\omega);\Theta(\tau_2,\omega)]\, dG(\omega) \\
&+p_{\tau_1}\int \Bigg[\mathcal{V}(q(\tau_1,\omega))-C(q(\tau_1,\omega);\Theta(\tau_1,\omega)) \\
&\qquad +(1-\alpha)\frac{p_{\tau_2}}{p_{\tau_1}}\frac{G(\omega)-G(\sigma^*(\tau_2;\tau_1,\omega))}{g(\omega)}C_\theta(q(\tau_1,\omega);\Theta(\tau_1,\omega))\Theta_\omega(\tau_1,\omega)\Bigg] \\
&\qquad\qquad dG(\omega).
\end{aligned}$$

By maximizing the integrands point-wise with respect to $q(\tau_i,\omega)$ and using the first order condition, $i=1,2$, we obtain equations (B.30) and (B.31). We note that the allocation function $q(\tau_1,\omega)$, given by (B.31), is increasing in ω since by assumption $c(q;\Theta(\tau_1,\omega))-\Theta_\omega(\tau_1,\omega)\frac{p_{\tau_2}}{p_{\tau_1}}\frac{G(\omega)-G(\sigma^*(\tau_2;\tau_1,\omega))}{g(\omega)}C_\theta(q;\Theta(\tau_1,\omega))$ is increasing in q. Moreover, the allocation function $q(\tau_2,\omega)$, given by (B.30), is increasing in ω by Assumption

3.1. Therefore, the omitted constraint (B.36) in the original optimization problem is satisfied automatically.

Below, we construct payment function $t(\tau_i,\omega)$, $i=1,2$, such that the omitted constraints (B.34) and (B.35) are satisfied, $\mathbb{E}_\omega\{t(\tau_2,\omega)-C(q(\tau_2,\omega);\Theta(\tau_2,\omega))\} = \mathcal{R}^T(\tau_2,\tau_1;q)+\mathcal{R}^W(\tau_2,\tau_1;q)$, and $\mathbb{E}_\omega\{t(\tau_1,\omega)-C(q(\tau_1,\omega);\Theta(\tau_1,\omega))\}=0$. Define,

$$\begin{aligned} t(\tau_2,\omega):=&C(q(\tau_2,\omega);\Theta(\tau_2,\omega)) + \mathcal{R}^T(\tau_2,\tau_1;q)+\mathcal{R}^W(\tau_2,\tau_1;q)\\ &-\int_{\underline{\omega}}^{\omega} C_\theta(q(\tau_2,\hat\omega);\Theta(\tau_2,\hat\omega))\Theta_\omega(\tau_2,\hat\omega)d\hat\omega\\ &+\int_{\underline{\omega}}^{\overline{\omega}}[1-G(\hat\omega)]C_\theta(q(\tau_2,\hat\omega);\Theta(\tau_2,\hat\omega))\Theta_\omega(\tau_2,\hat\omega)d\hat\omega, \end{aligned} \tag{B.39}$$

$$\begin{aligned} t(\tau_1,\omega):=&C(q(\tau_1,\omega);\Theta(\tau_1,\omega))\\ &-\int_{\underline{\omega}}^{\omega} C_\theta(q(\tau_1,\hat\omega);\Theta(\tau_1,\hat\omega))\Theta_\omega(\tau_1,\hat\omega)d\hat\omega\\ &+\int_{\underline{\omega}}^{\overline{\omega}}[1-G(\hat\omega)]C_\theta(q(\tau_1,\hat\omega);\Theta(\tau_1,\hat\omega))\Theta_\omega(\tau_1,\hat\omega)d\hat\omega. \end{aligned} \tag{B.40}$$

By the above definition, we have $\frac{\partial \mathcal{R}_{\tau_i,\omega}}{\partial\omega} = C_\theta(q(\tau_i,\omega);\Theta(\tau_i,\omega))\Theta_\omega(\tau_i,\omega)$ for $i=1,2$. Thus, the omitted constraints (B.34) and (B.35) are satisfied. Moreover,

$$\begin{aligned} \mathbb{E}_\omega\{t(\tau_1,\omega)-C(q(\tau_1,\omega);\Theta(\tau_1,\omega))\}=&\\ &-\iint_{\underline{\omega}}^{\omega} C_\theta(q(\tau_2,\hat\omega);\Theta(\tau_2,\hat\omega))\Theta_\omega(\tau_2,\hat\omega)d\hat\omega dG(\omega)\\ &-\int_{\underline{\omega}}^{\overline{\omega}}[1-G(\omega)]C_\theta(q(\tau_2,\hat\omega);\Theta(\tau_2,\hat\omega))\Theta_\omega(\tau_2,\hat\omega)d\hat\omega\\ &=\\ &-\iint_{\hat\omega}^{\overline{\omega}} C_\theta(q(\tau_2,\hat\omega);\Theta(\tau_2,\hat\omega))\Theta_\omega(\tau_2,\hat\omega)dG(\omega)d\hat\omega\\ &-\int_{\underline{\omega}}^{\overline{\omega}}[1-G(\omega)]C_\theta(q(\tau_2,\hat\omega);\Theta(\tau_2,\hat\omega))\Theta_\omega(\tau_2,\hat\omega)d\hat\omega\\ &=0, \end{aligned}$$

where the second equality results from changing the order of integration. Similarly,

$$\mathbb{E}_\omega\{t(\tau_2,\omega)-C(q(\tau_2,\omega);\Theta(\tau_2,\omega))\} = \mathcal{R}^T(\tau_2,\tau_1;q)+\mathcal{R}^W(\tau_2,\tau_1;q).$$

Substituting $\mathcal{R}^T(\tau_2,\tau_1;q)+\mathcal{R}^W(\tau_2,\tau_1;q)$ in (B.39) using (B.12), we obtain (B.32) and (B.33). □

B.3.5 The dynamic mechanism with no penalty

Similar to the dynamic mechanism, for the dynamic mechanisms with no penalty there exists no closed form solution for arbitrary M and parameters of the model, since the set of binding constraints from the inequality constraints, given by (3.32), cannot be determined a priori, and it depends on allocation function $q(\tau,\omega)$. Moreover, for the dynamic mechanisms with no penalty, we face additional difficulties by imposing ex-post individual rationality, which results in the additional set of constraints (3.33) on the information rent. As a result, unlike the dynamic mechanism, we cannot determine a priori the set of binding constraints from the ones given by (3.32) and (3.33) even for $M = 2$.

For $M = 2$, the dynamic mechanisms with no penalty is given by the solution to the following optimization problem.

$$\begin{aligned}\max_{q(\cdot,\cdot),t(\cdot,\cdot)} p_{\tau_2}\int[\mathcal{V}(q(\tau_2,\omega))-C(q(\tau_2,\omega);\Theta(\tau_2,\omega))]dG(\omega)-(1-\alpha)p_{\tau_2}\mathcal{R}_{\tau_2}\\ +p_{\tau_1}\int[\mathcal{V}(q(\tau_1,\omega))-C(q(\tau_1,\omega);\Theta(\tau_1,\omega))]dG(\omega)-(1-\alpha)p_{\tau_1}\mathcal{R}_{\tau_1}\end{aligned} \tag{B.41}$$

subject to

$$\mathcal{R}_{\tau_2}-\mathcal{R}_{\tau_1} \geq \mathcal{R}^T(\tau_2,\tau_1;q)+\mathcal{R}^W(\tau_2,\tau_1;q) \geq 0, \tag{B.42}$$

$$\mathcal{R}_{\tau_1} = -\int_{\underline{\omega}}^{\overline{\omega}}[1-G(\omega)]C_\theta(q(\tau_1,\hat{\omega});\Theta(\tau_1,\hat{\omega}))\Theta_\omega(\tau_1,\hat{\omega})d\hat{\omega} \geq 0, \tag{B.43}$$

$$\mathcal{R}_{\tau_2} \geq -\int_{\underline{\omega}}^{\overline{\omega}}[1-G(\omega)]C_\theta(q(\tau_2,\hat{\omega});\Theta(\tau_2,\hat{\omega}))\Theta_\omega(\tau_2,\hat{\omega})d\hat{\omega} \geq 0. \tag{B.44}$$

The last two constraints result from using the results of part (d) of Theorem 3.2 (equation (3.33)) along with the result of Lemma B.1, and setting $\mathcal{R}_{\tau_i,\underline{\omega}} \geq 0$, for $i = 1, 2$. We note that constraint (B.43) is written as an equality (binding) constraint (unlike (B.44)) since for any value of $\mathcal{R}_{\tau_1}$ that does not bind (B.43), we can reduce the value of $\mathcal{R}_{\tau_1}$ by $\epsilon > 0$ without violating any other constraints and improve the objective function (B.41).

In the following, we determine the solution to the above optimization problem for the specific example considered in Section 3.7. We have, $C(q;\theta) = \theta q$, $\Theta(\tau_1, \omega) = (1-\omega)$, $\Theta(\tau_1, \omega) = (1-\omega)^2$, $\mathcal{V}(q) = q - \frac{1}{2}q^2$, and $G(\omega) = \omega$ for $\omega \in [0, 1]$.

From (B.43), we obtain,

$$\mathcal{R}_{\tau_1} = \int_0^1 q(\tau_1, \omega)(1-\omega)d\omega. \tag{B.45}$$

It can be shown that constraints (B.44) is binding at the optimal solution.[2] When (B.44) is binding, we obtain,

$$\mathcal{R}_{\tau_2} = \int_0^1 q(\tau_2, \omega)2(1-\omega)^2 d\omega. \tag{B.46}$$

Moreover, from (B.42), we obtain

$$\begin{aligned}
\mathcal{R}_{\tau_2} - \mathcal{R}_{\tau_1} &\geq \int_0^1 \int_{\omega}^{1-(1-\omega)^2} q(\tau_1, \hat{\omega}) d\hat{\omega} d\omega \\
&= \int_0^1 \int_{1-\sqrt{1-\hat{\omega}}}^{\hat{\omega}} q(\tau_1, \hat{\omega}) d\omega d\hat{\omega} \\
&= \int_0^1 \left[\sqrt{1-\hat{\omega}} - (1-\hat{\omega})\right] q(\tau_1, \hat{\omega}) d\hat{\omega},
\end{aligned} \tag{B.47}$$

where the first equality results from (B.12), and the second equality results from changing the order of integration.

[2]To show this, one can solve the optimization problem (B.41) relaxing the constraint (B.44), and show that the optimal solution violates the constraint (B.44).

Using (B.45) and (B.46), we can rewrite (B.47) as,

$$\int_0^1 \left[q(\tau_2,\omega)2(1-\omega)^2 - q(\tau_1,\omega)\sqrt{1-\omega}\right] d\omega \geq 0. \tag{B.48}$$

Substituting (B.45) and (B.46) in (B.41), we can rewrite the optimization problem as,

$$\begin{aligned}\max_{q(\cdot,\cdot),t(\cdot,\cdot)} p_{\tau_2}\int_0^1 \Bigg[&\mathcal{V}(q(\tau_2,\omega)) - (1-\omega)^2 q(\tau_2,\omega) \\ &- (1-\alpha)2(1-\omega)^2 q(\tau_2,\omega)\Bigg] dG(\omega) \\ + p_{\tau_1}\int_0^1 \Bigg[&\mathcal{V}(q(\tau_1,\omega)) - (1-\omega)q(\tau_1,\omega) \\ &- (1-\alpha)(1-\omega)q(\tau_1,\omega)\Bigg] dG(\omega)\end{aligned} \tag{B.49}$$

subject to

$$\int_0^1 \left[q(\tau_2,\omega)2(1-\omega)^2 - q(\tau_1,\omega)\sqrt{1-\omega}\right] d\omega \geq 0,$$

where we replaced constraint (B.42) by (B.48).

The Lagrangian for the above optimization problem is given by,

$$\begin{aligned}&p_{\tau_2}\int_0^1 \left[\mathcal{V}(q(\tau_2,\omega)) - \left[(3-2\alpha)(1-\omega)^2 - 2\lambda(1-\omega)^2\right] q(\tau_2,\omega)\right] dG(\omega) \\ &+ p_{\tau_1}\int_0^1 \left[\mathcal{V}(q(\tau_1,\omega)) - \left[(2-\alpha)(1-\omega) + \lambda\sqrt{1-\omega}\right] q(\tau_1,\omega)\right] dG(\omega).\end{aligned}$$

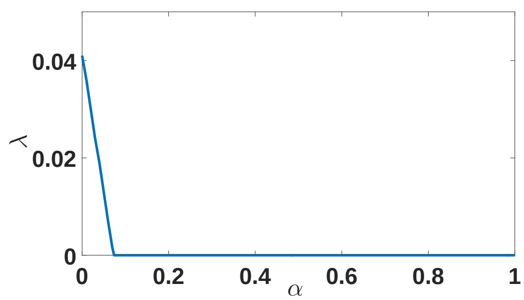
Maximizing the integrands point-wise with respect to $q(\tau_i,\omega)$, $i = 1, 2$, using the first order conditions and setting $\mathcal{V}(q) = q - \frac{1}{2}q^2$, we obtain,

$$q(\tau_2, \omega) = \max\{1-(3-2\alpha-2\lambda)(1-\omega)^2, 0\}, \tag{B.50}$$
$$q(\tau_1, \omega) = \max\{1-(2-\alpha)(1-\omega)-\lambda\sqrt{1-\omega}, 0\}. \tag{B.51}$$

The value of λ must be such that,

$$\lambda \int_0^1 \left[q(\tau_2, \omega)2(1-\omega)^2 - q(\tau_1, \omega)\sqrt{1-\omega}\right] d\omega = 0. \tag{B.52}$$

By numerical evaluation, $\lambda = 0$ for $\alpha \geq 0.07$ and $\lambda > 0$ for $\alpha \leq 0.07$.



Therefore, for $\alpha = 0.5$, we have,

$$q(\tau_2, \omega) = \max\{1-2(1-\omega)^2, 0\}, \tag{B.53}$$
$$q(\tau_1, \omega) = \max\{1-1.5(1-\omega), 0\}. \tag{B.54}$$

Using Lemma B.1, the payment functions $t(\tau_i, \omega)$ for $i = 1, 2$, is given by,

$$t(\tau_2,\omega) = C(q(\tau_2,\omega);\Theta(\tau_2,\omega)) + \int_{\underline{\omega}}^{\omega} C_\theta(q(\tau_2,\hat{\omega});\Theta(\tau_2,\omega))\Theta_\omega(\tau_2,\hat{\omega})d\hat{\omega}, \tag{B.55}$$
$$t(\tau_1,\omega) = C(q(\tau_1,\omega);\Theta(\tau_1,\omega)) + \int_{\underline{\omega}}^{\omega} C_\theta(q(\tau_1,\hat{\omega});\Theta(\tau_1,\omega))\Theta_\omega(\tau_1,\hat{\omega})d\hat{\omega}. \tag{B.56}$$

Appendix C

Proofs of Chapter 4

Proof of Theorem 4.1. Given a public information disclosure mechanisms $(\mathcal{M}, \psi)$, define $\bar{W}(m)$ as the expected social welfare conditioned on the realization of message m. Using (4.12) and (4.13), we have

$$\begin{aligned}\bar{W}(m) &= \left(\frac{1}{2}+\frac{1}{2}\Delta_m\right)\left(a-\frac{1}{2}-\frac{1}{2}\Delta_m\right)+\left(\frac{1}{2}-\frac{1}{2}\Delta_m\right)\left(\mathbb{E}\{\theta|m\}-\frac{1}{2}+\frac{1}{2}\Delta_m\right)\\ &= \left(\frac{a+\mathbb{E}\{\theta|m\}-1}{2}\right).\end{aligned}$$

Therefore,

$$W=\mathbb{E}\{\bar{W}(M)\}=\mathbb{E}\left\{\frac{a+\mathbb{E}\{\theta|M\}-1}{2}\right\}=\left(\frac{a+\mathbb{E}\left\{\mathbb{E}\{\theta|M\}\right\}-1}{2}\right)=\left(\frac{a+\mu-1}{2}\right),$$

where the last equality holds by the smoothing property of conditional expectation.

□

Proof of Lemma 4.1. We have $x^r(\theta) = 1 - x^s(\theta)$. Thus,

$$U^\sigma(s,\theta) = a - x^s(\theta) = a - 1 + x^r(\theta),$$
$$U^\sigma(r,\theta) = \theta - x^r(\theta) = \theta - 1 + x^s(\theta).$$

Obedience condition (4.18) is satisfied if and only if,

$$\sum_{\theta\in\Theta} p_\theta x^s(\theta)(a - x^s(\theta)) \geq \sum_{\theta\in\Theta} p_\theta x^s(\theta)(\theta - x^r(\theta))$$
$$\Leftrightarrow$$
$$\sum_{\theta\in\Theta} p_\theta x^s(\theta)(a - \theta + 1 - 2x^s(\theta)) \geq 0$$
$$\Leftrightarrow$$
$$\sum_{\theta\in\Theta} p_\theta x^s(\theta)(\frac{1}{2} + \frac{\Delta_\theta}{2} - x^s(\theta)) \geq 0.$$

Similarly, obedience condition (4.19) is satisfied if and only if,

$$\sum_{\theta\in\Theta} p_\theta x^r(\theta)(\theta - x^r(\theta)) \geq \sum_{\theta\in\Theta} p_\theta x^r(\theta)(a - x^s(\theta))$$
$$\Leftrightarrow$$
$$\sum_{\theta\in\Theta} p_\theta x^r(\theta)(\theta - a + 1 - 2x^r(\theta)) \geq 0$$
$$\Leftrightarrow$$
$$\sum_{\theta\in\Theta} p_\theta x^r(\theta)(\frac{1}{2} - \frac{\Delta_\theta}{2} - x^r(\theta)) \geq 0.$$

□

Proof of Theorem 4.2. By Lemma 4.1, the efficient routing policy in implementable if and only if

$$\begin{aligned}
&\mathbb{E}\left\{\left(\frac{1}{2}+\frac{\Delta_\theta}{4}\right)\left(\left(\frac{1}{2}+\frac{\Delta_\theta}{2}\right)-\left(\frac{1}{2}+\frac{\Delta_\theta}{4}\right)\right)\right\}\geq 0,\\
&\text{and}\\
&\mathbb{E}\left\{\left(\frac{1}{2}-\frac{\Delta_\theta}{4}\right)\left(\left(\frac{1}{2}-\frac{\Delta_\theta}{2}\right)-\left(\frac{1}{2}-\frac{\Delta_\theta}{4}\right)\right)\right\}\geq 0,\\
&\Leftrightarrow\\
&\mathbb{E}\left\{2\Delta_\theta+\Delta_\theta^2\right\}\geq 0,\\
&\text{and}\\
&\mathbb{E}\left\{-2\Delta_\theta+\Delta_\theta^2\right\}\geq 0,\\
&\Leftrightarrow\\
&2\Delta+\Delta^2+\sigma^2\geq 0,\\
&\text{and}\\
&-2\Delta+\Delta^2+\sigma^2\geq 0.
\end{aligned}$$

By combining the last two inequalities, we establish the result. $\square$

Appendix D

Proofs of Chapter 5

Proof of Lemma 5.1. For any given consistent CIB assessment (σ^*, γ), let g^* denote the behavioral strategy profile constructed according to (5.17). In the following, we construct recursively a belief system μ that is consistent with g^* and satisfies (5.18).

For $t = 1$, we have $P_1^i = Y_1^i$ and $C_1 = Z_1$. Define

$$\mu_t^i(h_1^i)(x_1, p_1^{-i}) := \frac{\mathbb{P}\{y_1, z_1 | x_1\}\eta(x_1)}{\sum_{\hat{x}_1 \in \mathcal{X}_1} \mathbb{P}\{p_1^i, z_1 | \hat{x}_1\}\eta(\hat{x}_1)}. \tag{D.1}$$

For $t > 1$, if $\mathbb{P}^{g^*}_{\mu^i_{t-1}}\{h_t^i | h_{t-1}^i\} > 0$ (*i.e.* no deviation from g^*_{t-1} at $t-1$), define μ_t^i recursively by Bayes' rule,

$$\mu_t^i(h_t^i)(x_t, p_t^{-i}) := \frac{\mathbb{P}^{g^*}_{\mu^i_{t-1}}\{h_t^i, x_t, p_t^{-i} | h_{t-1}^i\}}{\mathbb{P}^{g^*}_{\mu^i_{t-1}}\{h_t^i | h_{t-1}^i\}}. \tag{D.2}$$

For $t > 1$, if $\mathbb{P}^{g^*}_{\mu^i_{t-1}}\{h_t^i | h_{t-1}^i\} = 0$ (*i.e.* there is a deviation from g^*_{t-1} at $t-1$), define μ_t^i as,

$$\mu_t^i(h_t^i)(x_t, p_t^{-i}) := \frac{|\mathcal{P}_t|}{|\mathcal{S}_t|} \frac{\gamma_t(c_t)(x_t, s_t)}{\sum_{\hat{s}_t^{-i} \in \mathcal{S}_t^{-i}} \gamma(c_t)(x_t, \hat{s}_t^{-i}, s_t^i)}, \tag{D.3}$$

where $s_t^j = l_t^j(p_t^j, c_t)$ for all $j \in \mathcal{N}$.

At $t = 1$, (5.18) holds by construction from (D.1).

For $t > 1$,

$$\begin{aligned}
\mathbb{P}^{g^*}\{S_t^{-i}|h_t^i\} &= \mathbb{P}^{g^*}\{S_t^{-i}|p_t^i, c_t\} \\
&= \mathbb{P}^{g^*}\{S_t^{-i}|s_t^i, c_t\} \\
&= \frac{\mathbb{P}^{g^*}\{S_t^{-i}, s_t^i|c_t\}}{\mathbb{P}^{g^*}\{s_t^i|c_t\}} \\
&= \frac{\pi_t(S_t^{-i}, s_t^i)}{\sum_{\hat{s}_t^{-i} \in \mathcal{S}_t^{-i}} \pi_t(\hat{s}_t^{-i}, s_t^i)} = \mathbb{P}\{S_t^{-i}|s_t^i, \pi_t\}
\end{aligned}$$

where the second equality follows from (5.12). Therefore, (5.18) holds for all $t \in \mathcal{T}$. □

Proof of Theorem 5.1. We prove the result by induction.

For $t = 1$ the result holds since the agents have not taken any action yet. Suppose that (5.19) holds for $t - 1$. Then,

$$\begin{aligned}
\mathbb{P}^{g^*,g^i}\{x_t, h_t^{-i}|h_t^i\} &= \sum_{x_{t-1} \in \mathcal{X}_{t-1}} \mathbb{P}^{g^*,g^i}\{x_t, x_{t-1}, h_t^{-i}|h_t^i\} \\
&= \sum_{x_{t-1} \in \mathcal{X}_{t-1}} \mathbb{P}^{g^*,g^i}\{x_t, x_{t-1}, h_{t-1}^{-i}, a_{t-1}^{-i}, y_t^{-i}|h_{t-1}^i, a_{t-1}^i, y_t^i, z_t\} \\
&= \sum_{x_{t-1} \in \mathcal{X}_{t-1}} \mathbb{P}\{y_t^{-i}|x_t, a_{t-1}\}\mathbb{P}^{g^*,g^i}\{x_t, x_{t-1}, h_{t-1}^{-i}, a_{t-1}^{-i}|h_{t-1}^i, a_{t-1}^i, y_t^i, z_t\} \\
&= \sum_{x_{t-1} \in \mathcal{X}_{t-1}} \Big[\mathbb{P}\{y_t^{-i}|x_t, a_{t-1}\}\mathbb{P}\{x_t|x_{t-1}, a_{t-1}\} \\
&\qquad \mathbb{P}^{g^*,g^i}\{x_{t-1}, h_{t-1}^{-i}, a_{t-1}^{-i}|h_{t-1}^i, a_{t-1}^i, y_t^i, z_t\}\Big] \\
&= \sum_{x_{t-1} \in \mathcal{X}_{t-1}} \Big[\mathbb{P}\{y_t^{-i}|x_t, a_{t-1}\}\mathbb{P}\{x_t|x_{t-1}, a_{t-1}\}g_{t-1}^{*-i}(h_{t-1}^{-i})(a_{t-1}^{-i}) \\
&\qquad \mathbb{P}^{g^*,g^i}\{x_{t-1}, h_{t-1}^{-i}|h_{t-1}^i, a_{t-1}^i, y_t^i, z_t\}\Big]
\end{aligned} \tag{D.4}$$

Consider the term $\mathbb{P}^{g^*,g^i}\{x_{t-1}, h_{t-1}^{-i} | h_{t-1}^i, a_{t-1}^i, y_t^i, z_t\}$ in the expression above. We have,

$$\begin{aligned}
\mathbb{P}^{g^*,g^i}\{x_{t-1}, h_{t-1}^{-i} | h_{t-1}^i, a_{t-1}^i, y_t^i, z_t\} &= \frac{\mathbb{P}^{g^*,g^i}\{x_{t-1}, h_{t-1}^{-i}, y_t^i, z_t | h_{t-1}^i, a_{t-1}^i\}}{\mathbb{P}^{g^*,g^i}\{y_t^i, z_t | h_{t-1}^i, a_{t-1}^i\}} \\
&= \sum_{a_{t-1}^{-i}, x_t} \mathbb{P}\{y_t^i, z_t, x_t | a_{t-1}, x_{t-1}\} g_{t-1}^{*-i}(h_{t-1}^{-i})(a_{t-1}^{-i}) \times \\
&\quad \frac{\mathbb{P}^{g^*,g^i}\{x_{t-1}, h_{t-1}^{-i} | h_{t-1}^i, a_{t-1}^i\}}{\mathbb{P}^{g^*,g^i}\{y_t^i, z_t | h_{t-1}^i, a_{t-1}^i\}} \\
&= \sum_{a_{t-1}^{-i}, x_t} \mathbb{P}\{y_t^i, z_t | a_{t-1}, x_t\} \mathbb{P}\{x_t, | x_{t-1}, a_{t-1}\} g_{t-1}^{*-i}(h_{t-1}^{-i})(a_{t-1}^{-i}) \times \\
&\quad \frac{\mathbb{P}^{g^{*-i}}\{x_{t-1}, h_{t-1}^{-i} | h_{t-1}^i\}}{\sum_{\hat{a}_{t-1}^{-i}, \hat{x}_{t-1}} \mathbb{P}^{g^*,g^i}\{y_t^i, z_t, \hat{a}_{t-1}^{-i}, \hat{x}_{t-1} | h_{t-1}^i, a_{t-1}^i\}},
\end{aligned} \tag{D.5}$$

where the last equality follows from the induction hypothesis (5.19) for $t-1$. We can write the term in the denominator of (D.5) as

$$\begin{aligned}
\mathbb{P}^{g^*,g^i}\{y_t^i, z_t, \hat{a}_{t-1}^{-i}, \hat{x}_{t-1} | h_{t-1}^i, a_{t-1}^i\} &= \sum_{\hat{h}_{t-1}^{-i}, \hat{x}_t} \mathbb{P}^{g^*,g^i}\Big\{y_t^i, z_t, \hat{a}_{t-1}^{-i}, \hat{x}_{t-1}, \hat{h}_{t-1}^{-i}, \hat{x}_t | h_{t-1}^i, a_{t-1}^i\Big\} \\
&= \sum_{\hat{h}_{t-1}^{-i}, \hat{x}_t} \Big[\mathbb{P}\{y_t^i, z_t | \hat{x}_t, a_{t-1}^i, \hat{a}_{t-1}^{-i}\} \mathbb{P}\{\hat{x}_t | \hat{x}_{t-1}, a_{t-1}^i, \hat{a}_{t-1}^{-i}\} \\
&\quad g_t^{*-i}(\hat{h}_{t-1}^{-i})(\hat{a}_{t-1}^{-i}) \mathbb{P}^{g^*,g^i}\Big\{\hat{x}_{t-1}, \hat{h}_{t-1}^{-i} | h_{t-1}^i, a_{t-1}^i\Big\}\Big] \\
&= \sum_{\hat{h}_{t-1}^{-i}, \hat{x}_t} \Big[\mathbb{P}\{y_t^i, z_t | \hat{x}_t, a_{t-1}^i, \hat{a}_{t-1}^{-i}\} \mathbb{P}\{\hat{x}_t | \hat{x}_{t-1}, a_{t-1}^i, \hat{a}_{t-1}^{-i}\} \\
&\quad g_t^{*-i}(\hat{h}_{t-1}^{-i})(\hat{a}_{t-1}^{-i}) \mathbb{P}^{g^{*-i}}\Big\{\hat{x}_{t-1}, \hat{h}_{t-1}^{-i} | h_{t-1}^i, a_{t-1}^i\Big\}\Big] \\
&= \mathbb{P}^{g^{*-i}}\Big\{y_t^i, z_t, \hat{a}_{t-1}^{-i}, \hat{x}_{t-1} | h_{t-1}^i, a_{t-1}^i\Big\},
\end{aligned} \tag{D.6}$$

where the last equality follows from the induction hypothesis (5.19) for $t-1$.

Substituting (D.6) in (D.5),

$$\mathbb{P}^{g^*,g^{-i}}\big\{x_{t-1}, h^{-i}_{t-1} | h^i_{t-1}, a^i_{t-1}, y^i_t, z_t\big\} = \mathbb{P}^{g^{*-i}}\big\{x_{t-1}, h^{-i}_{t-1} | h^i_{t-1}, a^i_{t-1}, y^i_t, z_t\big\}\,. \tag{D.7}$$

Combining (D.4) and (D.7), we obtain

$$\mathbb{P}^{g^*,g^i}\left\{x_t, h^{-i}_t \middle| h^i_t\right\} = \mathbb{P}^{g^{*-i}}\left\{x_t, h^{-i}_t \middle| h^i_t\right\},$$

which establishes the induction step for t.

$\square$

To provide the proof for Theorems 5.2 and 5.4, we need the following result.

Lemma D.1. *Given a CIB strategy profile σ^* and update rule ψ consistent with σ^*,*

$$\mathbb{P}^{\sigma^*}_{\psi}\{S_{t+1}, \Pi_{t+1} | p_t, c_t, a_t\} = \mathbb{P}^{\sigma^*}_{\psi}\{S_{t+1}, \Pi_{t+1} | s_t, \pi_t, a_t\}. \tag{D.8}$$

for all s_t, π_t, a_t.

Proof of Lemma D.1. We have,

$$\mathbb{P}^{\sigma^*}_{\psi}\{s_{t+1},\pi_{t+1}|p_t,c_t,a_t\} = \mathbb{P}^{\sigma^*}_{\psi}\{s_{t+1},\pi_{t+1}|p_t,c_t,a_t,\pi_t\}$$

$$\overset{\text{Define } \hat{\mathcal{Z}}:=\{z_{t+1}:\pi_{t+1}=\psi_{t+1}(\pi_t,z_{t+1})\}\text{; see (5.16)}}{=}$$

$$= \sum_{z_{t+1}\in\hat{\mathcal{Z}}_t} \mathbb{P}^{\sigma^*}_{\psi}\{s_{t+1},\psi_{t+1}(\pi_t,z_{t+1})|p_t,c_t,a_t,\pi_t\}$$

$$=$$

$$\sum_{y_{t+1},x_{t+1},x_t,z_{t+1}\in\hat{\mathcal{Z}}} \mathbb{P}^{\sigma^*}_{\psi}\{s_{t+1},\pi_{t+1},y_{t+1},z_{t+1},x_{t+1},x_t|s_t,c_t,a_t,\pi_t\}$$

$$\overset{\text{by system dynamics (5.1) and (5.2)}}{=}$$

$$\sum_{y_{t+1},x_{t+1},x_t,z_{t+1}\in\hat{\mathcal{Z}}} \Big[\mathbb{P}^{\sigma^*}_{\psi}\{s_{t+1},\pi_{t+1}|s_t,c_t,a_t,\pi_t,y_{t+1},z_{t+1},x_{t+1},x_t\}\mathbb{P}\{y_{t+1},z_{t+1}|a_t,x_{t+1}\}$$
$$\mathbb{P}\{x_{t+1}|x_t,a_t\}\mathbb{P}^{\sigma^*}\{x_t|s_t,c_t,a_t\}\Big]$$

$$\overset{\text{Define } \hat{\mathcal{Y}}(z_{t+1}):=\{y_{t+1}:s^j_{t+1}=\phi^j_{t+1}(s^j_t,\{y^j_{t+1},z_{t+1},a^j_t\}),\forall j\}}{=}$$

$$\sum_{x_{t+1},x_t,z_{t+1}\in\hat{\mathcal{Z}},y_{t+1}\in\hat{\mathcal{Y}}(z_{t+1})} \Big[\mathbb{P}^{\sigma^*}_{\psi}\{s_{t+1},\pi_{t+1}|s_t,\pi_t,a_t,y_{t+1},z_{t+1}\}\mathbb{P}\{y_{t+1},z_{t+1}|a_t,x_{t+1}\}$$
$$\mathbb{P}\{x_{t+1}|x_t,a_t\}\mathbb{P}^{\sigma^*}_{\psi}\{x_t|s_t,c_t\}\Big]$$

$$\overset{\text{by Bayes' rule}}{=}$$

$$\sum_{z_{t+1}\in\hat{\mathcal{Z}},,y_{t+1}\in\hat{\mathcal{Y}}(z_{t+1}),x_{t+1},x_t} \Big[\mathbb{P}^{\sigma^*}_{\psi}\{s_{t+1},\pi_{t+1}|s_t,\pi_t,a_t,y_{t+1},z_{t+1}\}\mathbb{P}\{y_{t+1}\in\hat{\mathcal{Y}},z_{t+1}\in\hat{\mathcal{Z}}|a_t,x_{t+1}\}$$
$$\mathbb{P}\{x_{t+1}|x_t,a_t\}\frac{\mathbb{P}^{\sigma^*}_{\psi}\{x_t,s_t|c_t\}}{\mathbb{P}^{\sigma^*}_{\psi}\{s_t|c_t\}}\Big]$$

$$=$$

$$\sum_{z_{t+1}\in\hat{\mathcal{Z}},,y_{t+1}\in\hat{\mathcal{Y}}(z_{t+1}),x_{t+1},x_t} \Big[\mathbb{P}^{\sigma^*}_{\psi}\{s_{t+1},\pi_{t+1}|s_t,\pi_t,a_t,y_{t+1},z_{t+1}\}\mathbb{P}\{y_{t+1},z_{t+1}|a_t,x_{t+1}\}$$
$$\mathbb{P}\{x_{t+1}|x_t,a_t\}\frac{\pi_t(x_t,s_t)}{\sum_{\hat{x}_t}\pi_t(\hat{x}_t,s_t)}\Big]$$

$$=$$

$$\mathbb{P}^{\sigma^*}_{\psi}\{s_{t+1},\pi_{t+1}|s_t,\pi_t,a_t\}.$$

$\square$

Proof of Theorem 5.2. Consider a "super dynamic system" as the collection of the original dynamic system along with agents $-i$ who play according to CIB assessment (σ^*, γ). We establish the claim of Theorem 5.2 in two steps: (i) we show that the super dynamic system is a POMDP, and (ii) we show that $\{\Pi_t, S_t^i\}$ is an information state for agent i when he faces the super dynamic system with the original utility $u_{1:T}^i(\cdot, \cdot)$. Therefore, without loss of optimality, agent i can choose his best response from the class of strategies that depend on the information state $\{\Pi_t, S_t^i\}$, *i.e.* the class of CIB strategies.

To establish step (i), consider $\tilde{X} := \{X_t, \Pi_t, S_t, \Pi_{t-1}, S_{t-1}\}$ as the system state at t for the super dynamic system. Agent i's observation at time t is given by $\tilde{Y}_t^i := \{Y_t^i, Z_t\}$. To show that the super dynamic system is a POMDP, we need to show that it satisfies the following properties:
(a) it has a controlled Markovian dynamics, that is,

$$\mathbb{P}_\psi^{\sigma*}\{\tilde{x}_{t+1}|\tilde{x}_{1:t}, a_{1:t}^i, \tilde{y}_{1:t}^i\} = \mathbb{P}_\psi^{\sigma*}\{\tilde{x}_{t+1}|\tilde{x}_t, a_t^i\}, \quad \forall t \in \mathcal{T}, \tag{D.9}$$

(b) agent i's observation $\tilde{Y}_t^i$ is a function of system state $\tilde{X}_t$ along with the previous action A_{t-1}^i, that is,

$$\mathbb{P}_\psi^{\sigma*}\{\tilde{y}_t^i|\tilde{x}_{1:t},, a_{1:t-1}^i, \tilde{y}_{1:t-1}^i\} = \mathbb{P}_\psi^{\sigma*}\{\tilde{y}_t^i|\tilde{x}_t, a_{t-1}^i\}, \quad \forall t \in \mathcal{T}, \tag{D.10}$$

(c) agent i's instantaneous utility at $t \in \mathcal{T}$ can be written as a function $\tilde{u}_t(\tilde{x}_t, a_t^i)$ of system state $\tilde{X}_t$ along with his action A_t^i, that is,

$$\mathbb{E}_\psi^{\sigma^*}\left\{u_t^i(X_t, A_t)|\tilde{x}_{1:t}, a_{1:t}^i, \tilde{y}_{1:t}^i\right\} = \tilde{u}_t(\tilde{x}_t, a_t^i), \quad \forall t \in \mathcal{T}. \tag{D.11}$$

Condition (a) is true because,

$$\mathbb{P}^{\sigma*}_{\psi}\{\tilde{x}_{t+1}|\tilde{x}_{1:t}, a^i_{1:t}, \tilde{y}^i_{1:t}\}$$

$$=$$

$$\mathbb{P}^{\sigma*}_{\psi}\{x_{t+1}, \pi_{t+1}, s_{t+1}, \pi_t, s_t|x_{1:t}, \pi_{1:t}, s_{1:t}, y^i_{1:t}, z_{1:t}, a^i_{1:t}\}$$

$$=$$

$$\sum_{a_t^{-i}, z_{t+1}, y_{t+1}} \mathbb{P}^{\sigma*}_{\psi}\{x_{t+1}, \pi_{t+1}, s_{t+1}, a_t^{-i}, z_{t+1}, y_{t+1}|x_{1:t}, \pi_{1:t}, s_{1:t}, y^i_{1:t}, z_{1:t}, a^i_{1:t}\}$$

$$\overset{\text{by system dynamics (5.1) and (5.2)}}{=}$$

$$\sum_{a_t^{-i}, z_{t+1}, y_{t+1}} \Big[\mathbb{P}^{\sigma*}_{\psi}\{\pi_{t+1}, s_{t+1}|x_{1:t}, \pi_{1:t}, s_{1:t}, y^i_{1:t}, z_{1:t}, a^i_{1:t}, a_t^{-i}, , x_{t+1}, z_{t+1}, y_{t+1}\}$$
$$\mathbb{P}\{z_{t+1}, y_{t+1}|x_{t+1}, a_t\}\mathbb{P}\{x_{t+1}|x_t, a_t\}\sigma_t^{*-i}(\pi_t, s_t^{-i})(a_t^{-i})\Big]$$

$$\overset{\text{Define } \hat{\mathcal{Z}}:=\{z_{t+1}:\pi_{t+1}=\psi_{t+1}(\pi_t,z_{t+1})\} \text{ and } \hat{\mathcal{Y}}(z_{t+1}):=\{y_{t+1}:s^j_{t+1}=\phi^j_{t+1}(s^j_t,\{y^j_{t+1},z_{t+1},a^j_t\}),\forall j\}}{=}$$

$$\sum_{a_t^{-i}, z_{t+1}\in\hat{\mathcal{Z}}, y_{t+1}\in\hat{\mathcal{Y}}(z_{t+1})} \Big[\mathbb{P}^{\sigma^*}_{\psi}\{z_{t+1}, y_{t+1}|s_t, \pi_t, y_{t+1}, z_{t+1}, a_t\}\mathbb{P}\{y_{t+1}, z_{t+1}|x_{t+1}, a_t\}\mathbb{P}\{x_{t+1}|x_t, a_t\}$$
$$\sigma_t^{*-i}(\pi_t, s_t^{-i})(a_t^{-i})\Big]$$

$$=$$

$$\mathbb{P}^{\sigma*}_{\psi}\{x_{t+1}, \pi_{t+1}, s_{t+1}|x_t, \pi_t, s_t, a^i_t\}$$

$$=$$

$$\mathbb{P}^{\sigma*}_{\psi}\{\tilde{x}_{t+1}|\tilde{x}_t, a^i_t\}.$$

Condition (b) is true because

$$\mathbb{P}_{\psi}^{\sigma*}\{\tilde{y}_t^i|\tilde{x}_{1:t},,a_{1:t-1}^i,\tilde{y}_{1:t-1}^i\}$$
$$=$$
$$\mathbb{P}_{\psi}^{\sigma*}\{y_t^i,z_t|x_{1:t},\pi_{1:t},s_{1:t},y_{1:t-1}^i,z_{1:t-1},a_{1:t-1}^i\}$$
$$=$$
$$\sum_{a_{t-1}^{-i}}\mathbb{P}_{\psi}^{\sigma*}\{y_t^i,z_t,a_{t-1}^{-i}|x_{1:t},\pi_{1:t},s_{1:t},y_{1:t-1}^i,z_{1:t-1},a_{1:t-1}^i\}$$
$$=$$
$$\sum_{a_{t-1}^{-i}}\mathbb{P}_{\psi}^{\sigma*}\{y_t^i,z_t|x_{1:t},\pi_{1:t},s_{1:t},y_{1:t-1}^i,z_{1:t-1},a_{1:t-1}^i,a_{t-1}^{-i}\}\sigma_t^{*-i}(\pi_{t-1},s_{t-1})(a_{t-1}^{-i})$$
$$\overset{\text{by system dynamics (5.2)}}{=}$$
$$\sum_{a_{t-1}^{-i}}\mathbb{P}_{\psi}^{\sigma*}\{y_t^i,z_t|x_t,a_{t-1}^i,a_{t-1}^{-i}\}\sigma_t^{*-i}(\pi_{t-1},s_{t-1})(a_{t-1}^{-i})$$
$$=$$
$$\mathbb{P}_{\psi}^{\sigma*}\{y_t^i,z_t|x_t,\pi_{t-1},s_{t-1},a_{t-1}^i\}$$
$$=$$
$$\mathbb{P}_{\psi}^{\sigma*}\{\tilde{y}_t^i|\tilde{x}_t,a_{t-1}^i\}.$$

Condition (c) is true because

$$\begin{aligned}
\mathbb{E}_{\psi}^{\sigma^*}\big\{u_t^i(X_t,A_t)|\tilde{x}_{1:t},a_{1:t}^i,\tilde{y}_{1:t}^i\big\} &= \mathbb{E}_{\psi}^{\sigma^*}\big\{u_t^i(X_t,A_t)|x_t,\pi_t,s_t,\tilde{x}_{1:t-1},a_{1:t}^i,\tilde{y}_{1:t}^i\big\}\\
&= \mathbb{E}_{\psi}^{\sigma^*}\big\{u_t^i(X_t,(a_t^i,\sigma^{*-i}(\pi_t,s_t^{-i})))|x_t,\pi_t,s_t,\tilde{x}_{1:t-1},a_{1:t}^i,\tilde{y}_{1:t}^i\big\}\\
&= u_t^i(x_t,(a_t^i,\sigma^{*-i}(\pi_t,s_t^{-i})))\\
&:= \tilde{u}(\tilde{x}_t,a_t^i)
\end{aligned}$$

To establish step (ii), that it, to show that $\{\Pi_t,S_t^i\}$ is an information state for agent i, we need to prove that (1) it can be updated recursively at t, *i.e.* it can be determined using $\{\Pi_{t-1},S_{t-1}^i\}$ and $\{\tilde{Y}_t^i,A_t^i\}=\{Y_t^i,Z_t,A_t^i\}$, (2) agent i's belief about

$\{\Pi_{t+1}, S^i_{t+1}\}$ conditioned on $\{\Pi_t, S^i_t, A^i_t\}$ is independent of H^i_t, and (3) it is sufficient to evaluate the agent i's instantaneous utility at t for every action $a^i_t \in \mathcal{T}$, for all $t \in \mathcal{T}$.

Condition (1) is satisfied since $\Pi_t = \psi_t(\Pi_{t-1}, Z_t)$ and $S^i_t = \phi_t(S^i_{t-1}, \{Y^i_t, Z_t, A^i_t\})$ for $t \in \mathcal{T}\backslash\{1\}$; see part (i) of Definition 5.5 and (5.16).

To prove condition (2) let

$$g^{*j}_t(h^j_t) = \sigma^{*j}_t(l(h^j_t), \gamma_t(c_t)), \tag{D.12}$$

for all $j \in \mathcal{N}$ and $t \in \mathcal{T}$. Then condition (2) is satisfied since

$$\mathbb{P}_{\psi}^{\sigma^*}\{s_{t+1}^i, \pi_{t+1} | h_t^i, a_t^i\}^1$$

$$=$$

$$\sum_{h_t^{-i}, a_t^{-i}} \mathbb{P}_{\psi}^{\sigma^*}\{s_{t+1}^i, \pi_{t+1}, h_t^{-i}, a_t^{-i} | h_t^i, a_t^i\}$$

$$\overset{\text{by Theorem 5.1 and (D.12)}}{=}$$

$$\sum_{h_t^{-i}, a_t^{-i}} \mathbb{P}_{\psi}^{\sigma^*}\{s_{t+1}^i, \pi_{t+1} | h_t, a_t\} \mathbb{P}^{g^{*-i}}\{h_t^{-i} | h_t^i\} g_t^{*-i}(h_t^{-i})(a_t^{-i})$$

$$=$$

$$\sum_{h_t^{-i}, a_t^{-i}, s_{t+1}^{-i}} \mathbb{P}_{\psi}^{\sigma^*}\{s_{t+1}^i, \pi_{t+1}, s_{t+1}^{-i} | h_t, a_t\} \mathbb{P}^{g^{*-i}}\{h_t^{-i} | h_t^i\} g_t^{*-i}(h_t^{-i})(a_t^{-i})$$

$$\overset{\text{by Lemma D.1 and } s_t^{-i} = l_t^{-i}(h_t^{-i}) \text{ (see Definition 5.5)}}{=}$$

$$\sum_{h_t^{-i}, a_t^{-i}, s_{t+1}^{-i}} \mathbb{P}_{\psi}^{\sigma^*}\{s_{t+1}, \pi_{t+1} | s_t, \pi_t, a_t\} \mathbb{P}^{g^{*-i}}\{s_t^{-i} | h_t^i\} g_t^{*-i}(h_t^{-i})(a_t^{-i})$$

$$\overset{\text{by part (ii) of Lemma 5.1}}{=}$$

$$\sum_{s_t^{-i}, a_t^{-i}, s_{t+1}^{-i}} \mathbb{P}_{\psi}^{\sigma^*}\{s_{t+1}, \pi_{t+1} | s_t, \pi_t, a_t\} \mathbb{P}\{s_t^{-i} | s_t^i, \pi_t\} g_t^{*-i}(h_t^{-i})(a_t^{-i})$$

$$\overset{\text{by (D.12)}}{=}$$

$$\sum_{h_t^{-i}, a_t^{-i}, s_{t+1}^{-i}} \mathbb{P}_{\psi}^{\sigma^*}\{s_{t+1}^i, \pi_{t+1}, s_{t+1}^{-i} | h_t, a_t\} \mathbb{P}\{s_t^{-i} | s_t^i, \pi_t\} \sigma_t^{*-i}(\pi_t, s_t^{-i})(a_t^{-i})$$

$$=$$

$$\mathbb{P}_{\psi}^{\sigma^*}\{s_{t+1}^i, \pi_{t+1} | s_t^i, \pi_t, a_t^i\}$$

.

[1]Note that $\pi_t(\tilde{x}_t, \tilde{s}_t) = \mathbb{P}^{\sigma^*}\{\tilde{x}_t, \tilde{s}_t | c_t\}$, that is, given σ^*, π_t is a function of c_t.

To prove condition (3), we need to show that for all $a_t^i \in \mathcal{A}_t^i$,

$$\mathbb{E}^{g^{*-i}}\{u_i^t(X_t, A_t^{-i}, a_t^i)|h_t^i\} = \mathbb{E}^{g^{*-i}}\{u_i^t(X_t, A_t^{-i}, a_t^i)|\pi_t, s_t^i\}, \tag{D.13}$$

for all h_t^i, π_t, s_t^i, $t \in \mathcal{T}$.

By Lemma 5.1,

$$\mathbb{P}^{g^{*-i}}\{s_t^{-i}|h_t^i\} = \mathbb{P}\{s_t^{-i}|\pi_t, s_t^i\}. \tag{D.14}$$

Setting $\pi_t = \gamma(c_t)$, and $A_t^{-i} = \sigma^{*-i}(\pi_t, S_t^{-i})$,

$$\begin{aligned}
\mathbb{E}_\psi^{\sigma^*}\{u_t^i(X_t, A_t^{-i}, a_t^i)|h_t^i\} &= \mathbb{E}_\psi^{\sigma^*}\{u_t^i(X_t, \sigma_t^{*-i}(\pi_t, S_t^{-i}), a_t^i)|h_t^i\} \\
&= \mathbb{E}_\psi^{\sigma^*}\left\{\mathbb{E}^{\sigma^*}\{u_t^i(X_t, \sigma_t^{*-i}(\pi_t, S_t^{-i}), a_t^i)|S_t^{-i}, \pi_t, s_i^t, h_t^i\}\middle|h_t^i\right\} \\
&= \mathbb{E}_\psi^{\sigma^*}\left\{\mathbb{E}^{\sigma^*}\{u_t^i(X_t, \sigma_t^{*-i}(\pi_t, S_t^{-i}), a_t^i)|S_t^{-i}, \pi_t, s_t^i, c_t\}\middle|h_t^i\right\} \\
&= \mathbb{E}_\psi^{\sigma^*}\left\{\mathbb{E}^{\sigma^*}\{u_t^i(X_t, \sigma_t^{*-i}(\pi_t, S_t^{-i}), a_t^i)|S_t^{-i}, \pi_t, s_t^i, c_t\}\middle|h_t^i\right\} \\
&= \mathbb{E}_\psi^{\sigma^*}\{u_t^i(X_t, \sigma_t^{*-i}(\pi_t, S_t^{-i}), a_t^i)|\pi_t, s_t^i\}.
\end{aligned} \tag{D.15}$$

The first equality above is by substituting $A_t^{-i} = \sigma_t^{*-i}(\pi_t, S_t^{-i})$. The second equality follows from the smoothing property of conditional expectation. The third equality holds by condition (iii) of Definition 5.5. The fourth equality holds since for every x_t, s_t, π_t, c_t,

$$\mathbb{P}\{x_t|s_t, \pi_t, c_t\} = \frac{\mathbb{P}\{x_t, s_t|\pi_t, c_t\}}{\mathbb{P}\{s_t, \pi_t, c_t\}} = \frac{\pi_t(x_t, s_t)}{\sum_{\hat{x}_t} \pi_t(\hat{x}_t, s_t)} = \mathbb{P}\{x_t|s_t, \pi_T\}.$$

The last equality is true by (D.14). By (D.15) we prove condition (3) for $\{\Pi_t, S_t^i\}$ to be an information state, and thus establish the result of Theorem 5.2.

□

Proof of Theorem 5.3. Let (σ^*, ψ) denote a solution of the dynamic program. We note that the CIB update rule ψ is consistent with σ^* by requirement (5.26). Therefore, we only need to show that the CIB assessment (σ^*, ψ) is sequentially rational. To prove it, we use the *one-shot deviation principle* for dynamic games with asymmetric information [51]. To state the one-shot deviation, we need the following definitions.

Definition D.1 (One-shot deviation). *We say $\tilde{g}^i$ is a one-shot deviation from g^{*i} if there exists a unique $h^i_t \in \mathcal{H}^i$ such that $\tilde{g}^i_t(h^i_t) \neq g^{*i}_t(h^i_t)$, and $\tilde{g}^i_\tau(h^i_\tau) \neq g^{*i}_\tau(h^i_\tau)$ for all $h^i_\tau \neq h^i_t$, $h^i_\tau \in \mathcal{H}^i$.*

Definition D.2 (Profitable one-shot deviation). *Consider an assessment (g^*, μ). We say $\tilde{g}^i$ is a profitable one-shot deviation for agent i if $\tilde{g}^i$ is a one-shot deviation from g^{*i} at h^i_t such that $\tilde{g}^i_t(h^i_t) \neq g^{*i}_t(h^i_t)$, and*

$$\mathbb{E}^{(g^{*-i},\tilde{g}^i)}_\mu \left\{ \sum_{\tau=t}^{T} u^i_\tau(X_\tau, A_\tau) \middle| h^i_t \right\} > \mathbb{E}^{(g^{*-i},g^{*i})}_\mu \left\{ \sum_{\tau=t}^{T} u^i_\tau(X_\tau, A_\tau) \middle| h^i_t \right\}$$

One-shot deviation principle [51]: A consistent assessment (g^*, μ) is a PBE if and only if there exists no agent that has a profitable one-shot deviation.

Below, we show that the consistent CIB assessment (σ^*, ψ) satisfies the sequential rationality condition using the one-shot deviation principle.

Consider an arbitrary agent $i \in \mathcal{N}$, time $t \in \mathcal{T}$, and history realization $h^i_t \in \mathcal{H}^i_t$. Agent i has a profitable one-shot deviation at h^i_t only if

$$\sigma^{*i}_t(\pi_t, s^i_t) \notin \underset{\tilde{g}^i_t(h^i_t) \in \Delta(\mathcal{A}^i_t)}{\arg\max}\ \mathbb{E}^{\sigma^*}_\pi \left\{ \bar{U}^i_t((\sigma^*(\pi_t, S_t)), S_t, \pi_t, V_{t+1}, \psi_{t+1}) \middle| h^i_t \right\}.$$

Given $(\pi_t, V_{t+1}, \psi_{t+1}, \sigma^*_t)$, the expected value of the function $\bar{U}^i_t$ conditioned on h^i_t is only a function of s^i_t, agent i's belief about S^{-i}_t, as well as agent i's strategy $\tilde{g}^i_t(h^i_t)$. Agent i's belief about S^{-i}_t given h^i_t is only a function of s^i_t and π_t (see (5.18)). Therefore, any solution to the maximization problem above can be written as a function of π_t and s^i_t, that is, it is a CIB strategy $\tilde{\sigma}^i_t(\pi_t, s^i_t)$ for agent i. Consequently, agent i has a profitable one-shot deviation only if

$$\sigma_t^{*i}(\pi_t, s_t^i) \notin \underset{\tilde{\sigma}_t^i(\pi_t, s_t^i) \in \Delta(\mathcal{A}_t^i)}{\arg\max} \ \mathbb{E}_\pi^{\sigma^*} \left\{ \bar{U}_t^i((\sigma^*(\pi_t, S_t)), S_t, \pi_t, V_{t+1}, \psi_{t+1}) \middle| \pi_t, s_t^i \right\}.$$

By (5.25), σ_t^* is BNE of the stage game $G_t(\pi_t, V_{t+1}, \psi_{t+1})$ at t (see also (5.21)), *i.e.*

$$\sigma_t^{*i}(\pi_t, s_t^i) \in \underset{\tilde{\sigma}_t^i(\pi_t, s_t^i) \in \Delta(\mathcal{A}_t^i)}{\arg\max} \ \mathbb{E}_\pi^{\sigma^*} \left\{ \bar{U}_t^i((\sigma^*(\pi_t, S_t)), S_t, \pi_t, V_{t+1}, \psi_{t+1}) \middle| \pi_t, s_t^i \right\}.$$

Consequently, there exists no profitable deviation from $\sigma_t^{*i}(\pi_t, s_t^i)$ at h_t^i. Therefore, there exists no agent that has a profitable one-shot deviation. Hence, by one-shot deviation principle, the consistent CIB assessment (σ^*, ψ) is sequentially rational, and thus, it is a CIB-PBE. □

Proof of Theorem 5.4. Consider an arbitrary strategy profile g. We prove the existence of CIB strategy profile that is equivalent to g by construction.

With some abuse of notation, let $\sigma^i(\Pi_t, S_t^i, \omega)$ denote agent i's strategy using the public randomization device ω. We construct a CIB strategy profile σ_t that has the following properties:

(a) the induced distribution on $\{\Pi_{t+1}, S_{t+1}\}$ under σ coincides with one under g, *i.e.*

$$\mathbb{P}^{\sigma_{1:t}} \{\pi_{t+1}, s_{t+1}\} = \mathbb{P}^{g_{1:t}} \{\pi_{t+1}, s_{t+1}\}. \tag{D.16}$$

(b) the continuation payoff for all the agents under σ is the same as that under g, *i.e.*

$$\mathbb{E}^g \left\{ \sum_{\tau=t}^{T} u_\tau^i(X_\tau, g_\tau(H_\tau)) \right\} = \mathbb{E}^\sigma \left\{ \sum_{\tau=t}^{T} u_\tau^i(X_\tau, \sigma_\tau(\Pi_\tau, S_\tau, \omega)) \right\}, \quad \forall i \in \mathcal{N}. \tag{D.17}$$

We prove condition (a) with forward induction and condition (b) by backward induction. We note that condition (a) is satisfied for $t = 1$, since at $t = 1$ no action has been taken. Moreover, condition (b) is satisfied for $t = T + 1$ since there is no future.

Assume that condition (a) is satisfied from 1 to t, $t \in \mathcal{T}$. We construct σ_t below such that condition (a) is satisfied at $t + 1$.

For every $t \in \mathcal{T}$ and $i \in \mathcal{N}$, let K_t^i denote the dimension of $\mathcal{H}_t^i \subseteq \mathbb{R}^{K_t^i}$, *i.e.* $H_t^i = \{H_t^{i,1}, H_t^{i,2}, ..., H_t^{i,K_t^i}\}$. Let $F^g_{H_t^{i,k}|\Pi_t,S_t^i}(\cdot|\cdot,\cdot)$, $1 \le k \le K_t^i$, denote the cumulative distribution function of $H_t^{i,k}$ conditional on Π_t and S_t^i when the agents play according to g. Define $R_t^{i,k} := F^g_{H_t^{i,k}|\Pi_t,S_t^i}(H_t^{i,k}|\Pi_t, S_t^i)$. The random variable $R_t^{i,k}$, $1 \le k \le K_T^i$, is uniformly distributed on $[0, 1]$, and it is independent of Π_t and S_t^i. Let $R_t^i := (R_t^{i,1}, R_t^{i,2}, ..., R_t^{i,K_t^i})$. We note that H_t^i can be written as a function of (Π_t, S_t^i, R_t^i) as

$$\begin{aligned} H_t^i &= F^{-1}_{R_t^i|\Pi_t,S_t^i}(R_t^i, \Pi_t, S_t^i) \\ &:= (F^{-1}_{R_t^{i,1}|\Pi_t,S_t^i}(R_t^{i,1}, \Pi_t, S_t^i), F^{-1}_{R_t^{i,2}|\Pi_t,S_t^i}(R_t^{i,2}, \Pi_t, S_t^i), ..., F^{-1}_{R_t^{i,K_t^i}|\Pi_t,S_t^i}(R_t^{i,K_t^i}, \Pi_t, S_t^i)), \end{aligned}$$

where

$$F^{-1}_{R_t^{i,k}|\Pi_t,S_t^i}(R_t^{i,k}, \Pi_t, S_t^i) := \inf\{(\hat{h}_t^{i,k} \in \mathcal{H}_t^{i,k} : F^g_{H_t^{i,k}|\Pi_t,S_t^i}(\hat{h}_t^{i,k}|\Pi_t, S_t^i) \ge R_t^{i,k}\},$$

for $1 \le k \le K_t^i$.

We show below that R_t^i is independent of S_t.

Lemma D.2. *The random variable R_t^i, $i \in \mathcal{N}$, is independent of Π_t and S_t for all $t \in \mathcal{T}$.*

Proof of Lemma D.2. To prove that R_t^i is independent of Π_t and S_t, we need to show that $\mathbb{P}^g\{R_t^i \in \tilde{\mathcal{R}}_t^i | \Pi_t \in \tilde{\Pi}_t, S_t^{-i} \in \tilde{\mathcal{S}}_t^{-i}, S_t^i \in \tilde{\mathcal{S}}_t^i\} = \mathbb{P}^g\{R_t^i \in \tilde{\mathcal{R}}_t^i\}$ for all $\tilde{\Pi}_t \in \mathcal{B}(\Delta(\mathcal{X}_t \times \mathcal{S}_t))$[2], $\tilde{\mathcal{S}}_t^{-i} \in 2^{\mathcal{S}_t^{-i}}$[3], $\tilde{\mathcal{S}}_t^i \in 2^{\mathcal{S}_t^i}$, and $\tilde{\mathcal{R}}_t^i \in \mathcal{B}([0, 1])$.

[2]For any uncountable set Λ, $\mathcal{B}(\Lambda)$ denotes the Borel σ-field on Λ.
[3]For any finite set Λ, 2^Λ denotes the power set of Λ.

First, we note that

$$\mathbb{P}^g\{s_t^{-i}|h_t\} = \mathbb{P}^g\{s_t^{-i}|s_t, c_t\} = \frac{\mathbb{P}^g\{s_t^{-i}, s_t^i|c_t\}}{\mathbb{P}^g\{s_t^i|c_t\}} = \frac{\pi_t(s_t^{-i}, s_t^i)}{\sum_{\hat{s}_t^{-i}} \pi_t(\hat{s}_t^{-i}, s_t^i)} = \mathbb{P}\{s_t^{-i}|s_t, \pi_t\}. \tag{D.18}$$

Therefore, we can write

$$\begin{aligned}
\mathbb{P}^g\{S_t^{-i} \in \tilde{\mathcal{S}}_t^{-i}|\Pi_t \in \tilde{\Pi}_t, S_t^i \in \tilde{\mathcal{S}}_t^i\} &\overset{\text{by (D.18)}}{=} \mathbb{P}^g\{S_t^{-i} \in \tilde{\mathcal{S}}_t^{-i}|\Pi_t \in \tilde{\Pi}_t, S_t^i \in \tilde{\mathcal{S}}_t^i, R_t^i \in \tilde{\mathcal{R}}_t^i\} \\
&= \frac{\mathbb{P}^g\{S_t^{-i} \in \tilde{\mathcal{S}}_t^{-i}, R_t^i \in \tilde{\mathcal{R}}_t^i|\Pi_t \in \tilde{\Pi}_t, S_t^i \in \tilde{\mathcal{S}}_t^i\}}{\mathbb{P}^g\{R_t^i \in \tilde{\mathcal{R}}_t^i|\Pi_t \in \tilde{\Pi}_t, S_t^i \in \tilde{\mathcal{S}}_t^i\}} \\
&= \frac{\mathbb{P}^g\{S_t^{-i} \in \tilde{\mathcal{S}}_t^{-i}, R_t^i \in \tilde{\mathcal{R}}_t^i|\Pi_t \in \tilde{\Pi}_t, S_t^i \in \tilde{\mathcal{S}}_t^i\}}{\mathbb{P}^g\{R_t^i \in \tilde{\mathcal{R}}_t^i\}} \\
&= \frac{\mathbb{P}^g\{R_t^i \in \tilde{\mathcal{R}}_t^i|\Pi_t \in \tilde{\Pi}_t, S_t^{-i} \in \tilde{\mathcal{S}}_t^{-i}, S_t^i \in \tilde{\mathcal{S}}_t^i\}\mathbb{P}^g\{S_t^{-i} \in \tilde{\mathcal{S}}_t^{-i}|\Pi_t \in \tilde{\Pi}_t, S_t^i \in \tilde{\mathcal{S}}_t^i\}}{\mathbb{P}^g\{R_t^i \in \tilde{\mathcal{R}}_t^i\}} \\
&\iff \\
\mathbb{P}^g\{R_t^i \in \tilde{\mathcal{R}}_t^i\} &= \mathbb{P}^g\{R_t^i \in \tilde{\mathcal{R}}_t^i|\Pi_t \in \tilde{\Pi}_t, S_t^{-i} \in \tilde{\mathcal{S}}_t^{-i}, S_t^i \in \tilde{\mathcal{S}}_t^i\}.
\end{aligned} \tag{D.19}$$

The first equality holds since H_t^i is uniquely determined by (Π_t, S_t^i, R_t^i) and vice versa. The second equality is by Bayes' rule. The third equality is true since R_t^i is independent of Π_t and S_t^i. The fourth equality is true by Bayes' rule. By (D.19), R_t^i is independent of Π_t and S_t for all $i \in \mathcal{N}$.

□

Using the result of Lemma D.2, we construct a CIB strategy profile σ_t equivalent to g_t as follows. Let $\hat{R}_t^{1:N}(\omega)$ denote a random vector the agents construct using the public randomization device ω that has an identical joint cumulative distribution to that of $R_t^{1:N}$. Note that by Lemma D.2, the distribution of $R_t^{1:N}$ is independent of S_t and Π_t.

Define,

$$\sigma_t^i(\Pi_t, S_t^i, \omega) := g_t^i(F^{-1}_{R_t^i|S_t^i,\Pi_t}(\hat{R}_t^i(\omega), \Pi_t, S_t^i)). \tag{D.20}$$

Then,

$$\begin{aligned}\mathbb{P}^{g_{1:t}}\{\pi_{t+1}, s_{t+1}|H_t\} =&\mathbb{P}^{g_{1:t}}\{\pi_{t+1}, s_{t+1}|\Pi_t, S_t, R_t\} \\ \overset{distribution}{=}&\mathbb{P}^{g_{1:t}}\{\pi_{t+1}, s_{t+1}|\Pi_t, S_t, \hat{R}_t\} \\ =&\mathbb{P}^{\sigma_{1:t}}\{\pi_{t+1}, s_{t+1}|\Pi_t, S_t, \hat{R}_t\}.\end{aligned}$$

Taking the expectation of the left and right hand sides with respect to ω and R_t, respectively, and using the fact that $\hat{R}(\omega)$ and R_t are independent of S_t and Π_t (Lemma D.2), we obtain

$$\mathbb{P}^{\sigma_{1:t}}\left\{\pi_{t+1}, s_{t+1}|\Pi_t, S_t\right\} = \mathbb{P}^{g_{1:t}}\left\{\pi_{t+1}, s_{t+1}|\Pi_t, S_t\right\} \quad w.p.1. \tag{D.21}$$

By induction hypothesis, we have $\mathbb{P}^{\sigma_{1:t-1}}\left\{\pi_t, s_t\right\} = \mathbb{P}^{g_{1:t-1}}\left\{\pi_t, s_t\right\}$. Therefore, taking the expectation from both sides of (D.21) with respect to Π_t, S_t, we establish that condition (a) holds for time $t+1$.

Next, assume that condition (b) is satisfied from $t+1$ to T, $t \in \mathcal{T}$. We prove below that condition (b) is satisfied at t.

We have,

$$\begin{aligned}\mathbb{E}^{g}\{u_t^i(X_t, A_t)|H_t\} =&\mathbb{E}^{g}\{u_t^i(X_t, A_t)|\Pi_t, S_t, R_t\} \\ \overset{distribution}{=}&\mathbb{E}^{g}\{u_t^i(X_t, A_t)|\Pi_t, S_t, \hat{R}_t\} \\ =&\mathbb{E}^{\sigma}\{u_t^i(X_t, A_t)|\Pi_t, S_t, \hat{R}_t\}\end{aligned}$$

Therefore, using condition (a) at time t, $i.e$ $\mathbb{P}^{\sigma_{1:t-1}}\{s_t, \pi_t\} = \mathbb{P}^{g_{1:t-1}}\{s_t, \pi_t\}$, the induction hypothesis on condition (b) for $t+1$, and the fact that R_t and $\hat{R}_t$ are identically distributed and independent of Π_t and S_t, we obtain

$$\mathbb{E}^g \left\{ \sum_{\tau=t}^{T} u_\tau^i(X_\tau, g_\tau(H_\tau)) \right\} = \mathbb{E}^\sigma \left\{ \sum_{\tau=t}^{T} u_\tau^i(X_\tau, \sigma_\tau(\Pi_\tau, S_\tau, \omega)) \right\},$$

where the last equality is true by induction hypothesis on (b) at $t+1$. $\square$

Proof of Theorem 5.5. The proof of Theorem 5.5 follows from an argument similar to that of Theorem 5.3 using the one-shot deviation principle.

Let (σ_s^*, ψ_s) denote a solution of the dynamic programming given by (5.34)-(5.36). The CIB assessment (σ_s^*, ψ_s) is consistent by construction. We prove that (σ_s^*, ψ_s) satisfies sequential rationality by using the one-shot deviation principle.

Agent i has a profitable one-shot deviation at h_t^i only if

$$\sigma_s^{*i}(\pi_t, s_t^i) \notin \underset{\tilde{g}_t^i(h_t^i) \in \Delta(\mathcal{A}_t^i)}{\arg\max} \; \mathbb{E}_\pi^{\sigma^*} \left\{ \bar{U}_s^i((\sigma_s^*(\pi_t, S_t)), S_t, \pi_t, V_s, \psi_s) \middle| h_t^i \right\}. \tag{D.22}$$

Given $(\pi_t, V_s, \psi_s, \sigma_s^*)$, the expected value of the function $\bar{U}_s^i$ conditioned on h_t^i is only a function of s_t^i, agent i's belief about S_t^{-i}, as well as agent i's strategy $\tilde{g}_t^i(h_t^i)$. By part (ii) of Lemma 5.1, agent i's belief about S_t^{-i} given h_t^i is only a function of s_t^i and π_t. Therefore, any solution to the maximization problem above can be written as a function of π_t and s_t^i. Consequently, agent i has a profitable one-shot deviation only if

$$\sigma_s^{*i}(\pi_t, s_t^i) \notin \underset{\tilde{\sigma}_s^i(\pi_t, s_t^i) \in \Delta(\mathcal{A}_t^i)}{\arg\max} \; \mathbb{E}_\pi^{\sigma^*} \left\{ \bar{U}_s^i((\sigma^*(\pi_t, S_t)), S_t, \pi_t, V_s, \psi_s) \middle| \pi_t, s_t^i \right\}. \tag{D.23}$$

By (5.34), σ_s^* is BNE of the stage game $G_s(\pi_t, V_s, \psi_s)$ at t, *i.e.*

$$\sigma_s^{*i}(\pi_t, s_t^i) \in \underset{\tilde{\sigma}_s^i(\pi_t, s_t^i) \in \Delta(\mathcal{A}_t^i)}{\arg\max} \; \mathbb{E}_\pi^{\sigma^*} \left\{ \bar{U}_s^i((\sigma^*(\pi_t, S_t)), S_t, \pi_t, V_s, \psi_s) \middle| \pi_t, s_t^i \right\}. \tag{D.24}$$

That is, there exists no profitable deviation from $\sigma_s^{*i}(\pi_t, s_t^i)$ at h_t^i. Therefore, by one-shot deviation principle, the consistent CIB assessment (σ_s^*, ψ_S) is sequentially rational, and thus, is a PBE. $\square$

Proof of Theorem 5.6. The proof of Theorem 5.6 follows from an argument similar to that of Theorem 5.3. □

Proof of Lemma 5.2. We prove below that if $V_{t+1}(\cdot, s_{t+1})$ is continuous in π_{t+1}, then the dynamic program has a solution at stage t, $t \in \mathcal{T}$; that is, there exists at least one σ_t^* such that $\sigma_t^* \in BNE_t(V_{t+1}, \psi_{t+1})$, where ψ_{t+1} is consistent with σ_t^*.

For every π_t, define a perturbation of the stage game $\mathbf{G_t}(\pi_\mathbf{t}, \mathbf{V_{t+1}}, \psi_\mathbf{t+1})$ by restricting the set of strategies of each agent to mixed strategies that assign probability of at least $\epsilon > 0$ to every action $a_t^i \in \mathcal{A}_t^i$ of agent $i \in \mathcal{N}$; for every agent $i \in \mathcal{N}$ we denote this class of ϵ-restricted strategies by $\Sigma_t^{\epsilon,i}$ and $\Sigma_t^\epsilon := \Sigma_t^{\epsilon,1} \times \ldots \times \Sigma_t^{\epsilon,N}$. In the following we prove that, for every $\epsilon > 0$, the corresponding perturbed stage game has an equilibrium $\sigma_t^{*,\epsilon}$ along with a consistent update rule ψ_{t+1}^ϵ.

We note that when the agents' equilibrium strategies are perfectly mixed strategies, then the update rule ψ_{t+1}^ϵ is completely determined via Bayes' rule. Therefore, for every strategy profile $\sigma_t^{*,\epsilon} \in \Sigma_t^\epsilon$ we can write $\psi_{t+1}^\epsilon := \beta_{t+1}(\sigma^{*,\epsilon})$, where $\beta_{t+1}(\sigma^{*,\epsilon})$ is Bayes' rule where $\sigma^{*,\epsilon}$ is utilized (see (5.13))

For every agent $i \in \mathcal{N}$, define a best response correspondence $BR_t^{i,\epsilon} : \Sigma_t^\epsilon \rightrightarrows \Sigma_t^{\epsilon,i}$ as

$$BR_t^{i,\epsilon}(\sigma_t^{*,\epsilon}) := \Big\{ \sigma_t^i \in \Sigma_t^{\epsilon,i} : \sigma_t^i(\pi_t, s_t^i) \in \arg\max_{\hat{\sigma}_t^i \in \Sigma_t^{\epsilon,i}} \mathbb{E}^{\sigma_t^{*,\epsilon}, \sigma_t^i} \{ \bar{U}_t^i(A_t, S_t, \pi_t, V_{t+1}, \beta_{t+1}(\sigma_t^*)) | s_t^i, \pi_t \}, \quad \forall \pi_t, s_t^i \Big\}, \tag{D.25}$$

which determines the set of all agent i's best responses within the class of ϵ-restricted strategies assuming that agents $-i$ are playing $\sigma_t^{*-i,\epsilon}$ and the update rule $\psi_{t+1}^\epsilon = \beta_{t+1}(\sigma_t^{*,\epsilon})$.

For every $i \in \mathcal{N}$ and $\sigma^{*,\epsilon} \in \Sigma_t^\epsilon$, we prove below that $BR_t^{i,\epsilon}(\sigma_t^*)$ is non-empty, convex, closed, and upper hemicontinuous.

We note that

$$\mathbb{E}^{\sigma_t^{*,\epsilon},\sigma_t^i}\{\bar{U}_t^i(A_t,S_t,\pi_t,V_{t+1},\beta_{t+1}(\sigma_t^{*,\epsilon}))|s_t^i,\pi_t\}$$
$$=$$
$$\sum_{a_t^i}\mathbb{E}^{\sigma_t^*,A_t^i=a_t^i}\{\bar{U}_t^i(A_t,S_t,\pi_t,V_{t+1},\beta_{t+1}(\sigma_t^*))|s_t^i,\pi_t\}\sigma_t^i(\pi_t,s_t^i)(a_t^i)$$
$$=$$
$$\sum_{a_t^i}\tilde{U}_{\pi_t,s_t^i}^{\sigma^*}(a_t^i)\sigma_t^i(\pi_t,s_t^i)(a_t^i),$$

where $\tilde{U}_{\pi_t,s_t^i}^{\sigma^{*,\epsilon}}(a_t^i):=\mathbb{E}^{\sigma_t^{*,\epsilon},A_t^i=a_t^i}\{\bar{U}_t^i(A_t,S_t,\pi_t,V_{t+1},\beta_{t+1}(\sigma_t^{*,\epsilon}))|s_t^i,\pi_t\}$.

Therefore, for every π_t, s_t^i, we have $\sigma_t^i(\pi_t,s_t^i)\in\arg\max_{\alpha\in\Delta(\mathcal{A}_t^i):\alpha(a_t^i)\geq\epsilon,\forall a_t^i}\sum_{a_t^i}\tilde{U}_{\pi_t,s_t^i}^{\sigma^{*,\epsilon}}$ $(a_t^i)\alpha(a_t^i)$. We note that $\max_{\alpha\in\Delta(\mathcal{A}_t^i):\alpha(a_t^i)\geq\epsilon,\forall a_t^i}\sum_{a_t^i}\tilde{U}_{\pi_t,s_t^i}^{\sigma^{*,\epsilon}}(a_t^i)\alpha(a_t^i)$ is a linear program, thus, by Theorem 16 of [93], the set of agent i's best responses $BR_t^{i,\epsilon}(\sigma_t^{*,\epsilon})$ is closed and convex. If V_{t+1} is continuous in π_{t+1} then V_{t+1} is continuous in agent i's strategy σ_t^i. Moreover, the instantaneous utility u_t^i is continuous in agent i's strategy σ_t^i. Therefore, $\bar{U}_t^i$, given by (5.20), is continuous in agent i' strategy σ_t^i. Therefore, by the maximum theorem [95] the set of i's best responses in upper hemicontinuous in $\sigma_t^{*,\epsilon}$ and non-empty.

Consequently, we establish that for every $i\in\mathcal{N}$, $BR_t^{i,\epsilon}(\sigma_t^*)$ is closed, convex, upper hemicontinuous, and non-empty for every $\sigma_t^*\in\Sigma_t^\epsilon$. Define $BR_t^\epsilon:=\times_{i\in\mathcal{N}}BR_t^{i,\epsilon}$ where $\times$ denotes the Cartesian product. The correspondence $BR_t^\epsilon(\sigma_t^*)$ is closed, convex, upper hemicontinuous, and non-empty for every $\sigma_t^*\in\Sigma_t^\epsilon$ since $BR_t^{i,\epsilon}(\sigma_t^*)$ is closed, convex, upper hemicontinuous, and non-empty for every $\sigma_t^{*i}\in\Sigma_t^{i,\epsilon}$ for all $i\in\mathcal{N}$. Therefore, by Kakutani's fixed-point theorem [23, Corollary 15.3], the correspondence BR_t^ϵ has a fixed point. Therefore, every perturbed stage game has an equilibrium $\sigma_t^{*,\epsilon}$ along with a consistent update rule $\psi_{t+1}=\beta_{t+1}(\sigma_t^{*,\epsilon})$.

Now consider the sequence of these perturbed games when $\epsilon\to 0$. Since the set of agents' strategies is compact, there exists a subsequence of these perturbed games whose equilibrium strategies converge, say to σ_t^*. Similarly let ψ_{t+1}^* denote the convergence point of $\beta_{t+1}(\sigma_T^{*,\epsilon})$. We note that ψ_{t+1}^* is consistent with σ_t^* since

$\beta_{t+1}(\sigma^*)$ (*i.e.* Bayes' rule) is continuous in σ_t^*. We show below that for every agent $i \in \mathcal{N}$, σ_t^{*i} is a best response for him given V_{t+1}, ψ_{t+1}^* when he chooses his strategy from the unconstrained class of CIB strategies.

As we proved above, the set of agent i's best responses $BR_t^i(\sigma_t^*)$ is upper hemicontinuous and closed given ψ_{t+1}^*. Therefore, $\sigma_t^*(\pi_t, \cdot)$ is also a best response for agent i in the stage game $\mathbf{G_t}(\pi_\mathbf{t}, \mathbf{V_{t+1}}, \psi^*_\mathbf{t+1})$. Consequently, $\sigma_t^* \in BNE_t(V_{t+1}, \psi_{t+1}^*)$ where ψ_{t+1}^* is consistent with σ_t^*. □

Proof of Theorem 5.7. We have a Bayesian zero-sum game with finite state and action spaces. By [34, Theorem 1] the equilibrium payoff is a continuous function of the agents' common prior/belief. Using this result, we prove, by backward induction, that every stage of the dynamic program described by (5.25)-(5.27), has a solution and V_t is continuous in π_t for all t.

For $t = T + 1$ the dynamic program has a solution trivially since the agents have utility for time less than or equal to T. Moreover, $V_{T+1}(.,.) = 0$ is trivially continuous in π_{T+1}.

For $t \leq T$, assume that V_{t+1} is continuous in π_{t+1}. Then, by Lemma 5.2 the dynamic program has a solution at t. We note that the continuation game from t to T is a dynamic zero-sum game with finite state and actions spaces. Therefore, as we argued above, by [34, Theorem 1] the agents' equilibrium payoff at t (*i.e.* V_t) is unique and is continuous in the agents' common prior given by π_t.

Therefore, by induction we establish the assertion of Theorem 5.7. □

Proof of Lemma 5.3. Assume that $\psi_{1:T}$ is independent of σ. Then, the evolution of Π_t is independent of σ^* and known a priori. As a result, we can ignore the consistency condition (5.26) in the dynamic program. Given ψ_{t+1}, the stage game $\mathbf{G_t}(\pi_\mathbf{t}, \mathbf{V_{t+1}}, \psi_\mathbf{t+1})$ is a static game of incomplete information with finite actions (given by $\mathcal{A}_t^{1:N}$) and finite types (given by $\mathcal{S}_t^{1:N}$) for every π_t. Therefore, by the standard existence results for finite games [43, Theorem 1.1], there exists an equilibrium for the stage game $BNE_t(V_{t+1}, \psi_{t+1})$. Consequently, the correspondence $BNE_t(V_{t+1}, \psi_{t+1})$ is non-empty for every $t \in \mathcal{T}$, thus, the dynamic programming given by (5.25-5.27) has a solution. □

Bibliography

[1] D. Acemoglu, A. Makhdoumi, A. Malekian, and A. Ozdaglar. Informational Braess' paradox: The effect of information on traffic congestion. *arXiv preprint arXiv:1601.02039*, 2016.

[2] E. Akyol, C. Langbort, and T. Başar. Information-theoretic approach to strategic communication as a hierarchical game. *Proceedings of the IEEE*, 105(2):205–218, 2017.

[3] B. Allaz and J. Vila. Cournot competition, forward markets and efficiency. *Journal of Economic theory*, 59(1):1–16, 1993.

[4] R. Arnott, A. De Palma, and R. Lindsey. Does providing information to drivers reduce traffic congestion? *Transportation Research Part A: General*, 25(5):309–318, 1991.

[5] R. Aumann, M. Maschler, and R. Stearns. *Repeated games with incomplete information*. MIT press, 1995.

[6] T. Başar. Decentralized multicriteria optimization of linear stochastic systems. *IEEE Transactions on Automatic Control*, 23(2):233–243, 1978.

[7] T. Başar. Two-criteria LQG decision problems with one-step delay observation sharing pattern. *Information and Control*, 38(1):21–50, 1978.

[8] T. Başar and Y. Ho. Informational properties of the nash solutions of two stochastic nonzero-sum games. *Journal of Economic Theory*, 7(4):370–387, 1974.

[9] T. Başar and G. Olsder. *Dynamic Noncooperative Game Theory*, volume 23. Siam, 1999.

[10] R. Balakrishna, M. Ben-Akiva, J. Bottom, and S. Gao. Information impacts on traveler behavior and network performance: State of knowledge and future directions. In *Advances in dynamic network modeling in complex transportation systems*, pages 193–224. Springer, 2013.

[11] X. Ban, Y. Li, and J. Margulici. Optimal use of changeable message signs for displaying travel times. 2009.

[12] M. Battaglini and R. Lamba. Optimal dynamic contracting: the first-order approach and beyond. *working paper, Available at SSRN 2172414*, 2015.

[13] P. Battigalli. Strategic independence and perfect Bayesian equilibria. *Journal of Economic Theory*, 70(1):201–234, 1996.

[14] M. Ben-Akiva, A. de Palma, and I. Kaysi. The impact of predictive information on guidance efficiency: An analytical approach. In *Advanced methods in transportation analysis*, pages 413–432. Springer, 1996.

[15] D. Bergemann and S. Morris. Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics*, 11(2):487–522, 2016.

[16] D. Bergemann and S. Morris. Information design and Bayesian persuasion information design, Bayesian persuasion, and Bayes correlated equilibrium. *The American Economic Review*, 106(5):586–591, 2016.

[17] D. Bergemann and S. Morris. Information design: A unified perspective. 2017.

[18] K. Bimpikis, S. Ehsani, and M. Mostagir. Designing dynamic contests. In *16th ACM conference on Economics and Computation (EC)*, pages 281–282, 2015.

[19] E. Bitar and K. Poolla. Selling wind power in electricity markets: The status today, the opportunities tomorrow. In *American Control Conference (ACC)*. IEEE, 2012.

[20] E. Bitar, R. Rajagopal, P. Khargonekar, K. Poolla, and P. Varaiya. Bringing wind energy to market. *IEEE Transactions on Power Systems*, 2012.

[21] E. Bitar and Y. Xu. On incentive compatibility of deadline differentiated pricing for deferrable demand. In *52nd IEEE Conference on Decision and Control (CDC)*. IEEE, 2013.

[22] P. Bolton and C. Harris. Strategic experimentation. *Econometrica*, 67(2):349–374, 1999.

[23] K. Border. *Fixed point theorems with applications to economics and game theory*. Cambridge university press, 1989.

[24] S. Borenstein. The trouble with electricity markets: understanding California's restructuring disaster. *Journal of Economic Perspectives*, pages 191–211, 2002.

[25] T. Borgers, D. Krahmer, and R. Strausz. *An introduction to the theory of mechanism design*. Oxford University Press, 2015.

[26] K. Brunhuber. Shortcut-finding app Waze creating residential traffic headaches, February 29 2016. [Online; posted 29-Februar-2016].

[27] P. Cardaliaguet, C. Rainer, D.h Rosenberg, and N. Vieille. Markov games with frequent actions and incomplete information-the limit case. *Mathematics of Operations Research*, 2015.

[28] H. Chao and R. Wilson. Multi-dimensional procurement auctions for power reserves: Robust incentive-compatible scoring and settlement rules. *Journal of Regulatory Economics*, 22(2):161–183, 2002.

[29] Ch. Chen, A. Skabardonis, and P. Varaiya. A system for displaying travel times on changeable message signs. *University of California, Berkeley*, 94720:1720, 2003.

[30] J. Cochran, L. Bird, J. Heeter, and D. Arent. Integrating variable renewable energy in electric power markets. best practices from international experience. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2012.

[31] P. Courty and L. Hao. Sequential screening. *The Review of Economic Studies*, 2000.

[32] P. Dasgupta, P. Hammond, and E. Maskin. The implementation of social choice rules: Some general results on incentive compatibility. *The Review of Economic Studies*, pages 185–216, 1979.

[33] A. De Palma, R. Lindsey, and N. Picard. Risk aversion, the value of information, and traffic equilibrium. *Transportation Science*, 46(1):1–26, 2012.

[34] E. Einy, O. Haimanko, and B. Tumendemberel. Continuity of the value and optimal strategies when common priors change. *International Journal of Game Theory*, 41(4), 2012.

[35] J. Ely. Beeps. *The American Economic Review*, 107(1):31–53, 2017.

[36] R. Emmerink, P. Nijkamp, P. Rietveld, and J. Van Ommeren. Variable message signs and radio traffic information: An integrated empirical analysis of drivers' route choice behaviour. *Transportation Research Part A: Policy and Practice*, 30(2):135–153, 1996.

[37] J. Escobar and J. Toikka. Efficiency in games with Markovian private information. *Econometrica*, 81(5):1887–1934, 2013.

[38] P. Eső and B. Szentes. Optimal information disclosure in auctions and the handicap auction. *The Review of Economic Studies*, 74(3):705–731, 2007.

[39] M. Fahrioglu and F. Alvarado. Designing incentive compatible contracts for effective demand management. *IEEE Transactions on Power Systems*, 15(4):1255–1260, 2000.

[40] F. Farhadi, H. Tavafoghi, D. Teneketzis, and J. Golestani. A dynamic incentive mechanism for security in networks of interdependent agents. In *7th EAI International Conference on Game Theory for Networks*, 2017.

[41] J. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer Science & Business Media, 2012.

[42] F. Forges. Repeated games of incomplete information: non-zero-sum. *Handbook of Game Theory*, 1:109–154, 1992.

[43] D. Fudenberg and J. Tirole. *Game theory. 1991.* Cambridge, Massachusetts, 1991.

[44] F. Gensbittel and J. Renault. The value of Markov chain games with incomplete information on both sides. *Mathematics of Operations Research*, 40(4):820–841, 2015.

[45] J. Guirola. Miami police fight back against Waze app, January 29 2015. [Online; posted 29-January-2015].

[46] A. Gupta, A. Nayyar, C. Langbort, and T. Başar. Common information based Markov perfect equilibria for linear-Gaussian games with asymmetric information. *SIAM J. Control Optim.*, 52(5):3228–3260, 2014.

[47] M. Halac, N. Kartik, and Q. Liu. Optimal contracts for experimentation. *The Review of Economic Studies*, 83(3):1040–1091, 2016.

[48] M. Halac, N. Kartik, and Q. Liu. Contests for experimentation. *Journal of Political Economy*, 125(5):000–000, 2017.

[49] T. Haring, J. Mathieu, and G. Andersson. Decentralized contract design for demand response. In *10th International Conference on the European Energy Market (EEM)*, pages 1–8. IEEE, 2013.

[50] J. Harsanyi. Games with incomplete information played by Bayesian players, i–iii. *Management science*, 50(12_supplement):1804–1817, 2004.

[51] E. Hendon, H. Jacobsen, and B. Sloth. The one-shot-deviation principle for sequential rationality. *Games and Economic Behavior*, 12(2):274–282, 1996.

[52] S. Hendrix. Traffic-weary homeowners and Waze are at war, again. guess who's winning?, June 5 2016. [Online; posted 5-June-2016].

[53] Y.i Ho. Team decision theory and information structures. *Proceedings of the IEEE*, 68(6):644–654, 1980.

[54] J. Hörner and A. Skrzypacz. Learning, experimentation and information design. Technical report, Working Paper, 2016.

[55] J. Hörner, T. Sugaya, S. Takahashi, and N. Vieille. Recursive methods in discounted stochastic games: An algorithm for $\delta \to 1$ and a folk theorem. *Econometrica*, 79(4):1277–1318, 2011.

[56] NERC Intermittent and Variable Generation Task Force. Accommodating high levels of variable generation. Technical report, North American Electric Reliability Corp.(NERC), 2009.

[57] T. Jamasb and M. Pollitt. Electricity market reform in the European Union: review of progress toward liberalization & integration. *The Energy Journal*, pages 11–41, 2005.

[58] A. Kagiannas, D. Askounis, and J. Psarras. Power generation planning: a survey from monopoly to competition. *International journal of electrical power & energy systems*, 26(6):413–421, 2004.

[59] E. Kamenica and M. Gentzkow. Bayesian persuasion. *The American Economic Review*, 101(6):2590–2615, 2011.

[60] G. Keller and S. Rady. Strategic experimentation with poisson bandits. *Theoretical Economics*, 5(2):275–311, 2010.

[61] G. Keller, S. Rady, and M. Cripps. Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68, 2005.

[62] C. Klessmann, C. Nabe, and K. Burges. Pros and cons of exposing renewables to electricity market risks a comparison of the market integration approaches in Germany, Spain, and the UK. *Energy Policy*, 2008.

[63] Daniel Krähmer and Roland Strausz. Ex post information rents in sequential screening. *Games and Economic Behavior*, 90:257–273, 2015.

[64] I. Kremer, Y. Mansour, and M. Perry. Implementing the wisdom of the crowd. *Journal of Political Economy*, 122(5):988–1012, 2014.

[65] W. Krichene, J. Reilly, S. Amin, and A. Bayen. Stackelberg routing on parallel networks with horizontal queues. *IEEE Transactions on Automatic Control*, 59(3):714–727, 2014.

[66] V. Krishna. *Auction theory*. Academic press, 2009.

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

[68] J. Laffont and J. Tirole. *A Theory of Incentives in Procurement and Regulation*. MIT press, 1993.

[69] Y. Leng, L. Rudolph, A. Pentland, J. Zhao, and H. Koutsopolous. Managing travel demand: Location recommendation for system efficiency based on mobile phone data. *arXiv preprint arXiv:1610.06825*, 2016.

[70] L. Li, C. Langbort, and J. Shamma. Solving two-player zero-sum repeated Bayesian games. *arXiv preprint arXiv:1703.01957*, 2017.

[71] L. Li and J. Shamma. Efficient strategy computation in zero-sum asymmetric repeated games. *arXiv preprint arXiv:1703.01952*, 2017.

[72] W. Lin and E. Bitar. Forward electricity markets with uncertain supply: Cost sharing and efficiency loss. In *53rd IEEE Conference on Decision and Control (CDC)*, 2014.

[73] R. Lindsey, T. Daniel, E. Gisches, and A. Rapoport. Pre-trip information and route-choice decisions with stochastic travel conditions: Theory. *Transportation Research Part B: Methodological*, 67:187–207, 2014.

[74] J. Liu, S. Amin, and G. Schwartz. Effects of information heterogeneity in Bayesian routing games. *arXiv preprint arXiv:1603.08853*, 2016.

[75] L. Hansen M. Dewatripont and S. Turnovsky. *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, volume 1. Cambridge University Press, 2003.

[76] A. Mahajan and M. Mannan. Decentralized stochastic control. *Annals of Operations Research*, 241(1-2):109–126, 2016.

[77] A. Mahajan and D. Teneketzis. Optimal design of sequential real-time communication systems. *IEEE Transactions on Information Theory*, 55(11):5317–5338, 2009.

[78] H. Mahmassani and R. Jayakrishnan. System performance and user response under real-time information in a congested traffic corridor. *Transportation Research Part A: General*, 25(5):293–307, 1991.

[79] G. Mailath and L. Samuelson. *Repeated Games and Reputations*. Oxford university press Oxford, 2006.

[80] A. Mas-Colell, M.D. Whinston, and J.R. Green. *Microeconomic Theory*, volume 1. Oxford university press New York, 1995.

[81] E.c Maskin and J.n Tirole. Markov perfect equilibrium: I. observable actions. *Journal of Econonomic Theory*, 100(2):191–219, 2001.

[82] J. Mathieu, T. Haring, J. Ledyard, and G. Andersson. Residential demand response program design: Engineering and economic perspectives. In *European Energy Market (EEM), 2013 10th International Conference on the*, pages 1–8. IEEE, 2013.

[83] I. Milchtaich. Congestion games with player-specific payoff functions. *Games and Economic Behavior*, 13(1):111–124, 1996.

[84] P. Milgrom and I. Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, pages 583–601, 2002.

[85] P. Milgrom and R. Weber. Distributional strategies for games with incomplete information. *Mathematics of Operations Research*, 10(4):619–632, 1985.

[86] D. Miller. Robust collusion with private information. *The Review of Economic Studies*, 79(2):778–811, 2012.

[87] R. Myerson. Multistage games with communication. *Econometrica*, 1986.

[88] R. Myerson and P. Reny. Open sequential equilibria of multi-stage games with infinite sets of types and actions. *Notes*, 2015.

[89] A. Nayyar, A. Gupta, C. Langbort, and T. Başar. Common information based Markov perfect equilibria for stochastic games with asymmetric information: Finite games. *IEEE Transactions on Automatic Control*, 59(3):555–570, March 2014.

[90] A. Nayyar, A. Mahajan, and D. Teneketzis. Optimal control strategies in delayed sharing information structures. *IEEE Transactions on Automatic Control*, 56(7):1606–1620, 2011.

[91] A. Nayyar, A. Mahajan, and D. Teneketzis. Decentralized stochastic control with partial history sharing: A common information approach. *IEEE Transactions on Automatic Control*, 58(7):1644–1658, 2013.

[92] A. Nayyar, K. Poolla, and P. Varaiya. A statistically robust payment sharing mechanism for an aggregate of renewable energy producers. In *European Control Conference (ECC)*, 2013.

[93] A. Nedich. Optimization i. *Lecture Notes*, 2008.

[94] L. Nelson. New traffic apps may be pushing cars into residential areas, January 5 2015. [Online; posted 5-January-2015].

[95] E. Ok. *Real Analysis with Economic Applications*, volume 10. Princeton University Press, 2007.

[96] M. J Osborne and A. Rubinstein. *A Course in Game Theory.* MIT press, 1994.

[97] Y. Ouyang, H. Tavafoghi, and D. Teneketzis. Dynamic oligopoly games with private Markovian dynamics. In *54th IEEE Conference on Decision and Control (CDC)*, 2015.

[98] 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*, 2017.

[99] A. Ozdaglar and R. Srikant. Incentives and pricing in communication networks. *Algorithmic Game Theory*, pages 571–591, 2007.

[100] R. Radner. Team decision problems. *The Annals of Mathematical Statistics*, 33(3):857–881, 1962.

[101] J. Renault. The value of Markov chain games with lack of information on one side. *Math. Oper. Res.*, 31(3):490–512, 2006.

[102] J. Renault. The value of repeated games with an informed controller. *Mathematics of Operations Research*, 37(1):154–179, 2012.

[103] G. Richards. Is Googles Waze app making traffic worse?, January 21 2016. [Online; posted 21-January-2016].

[104] R. Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 2(1):65–67, 1973.

[105] R. Rosenthal. Arbitration of two-party disputes under uncertainty. *The Review of Economic Studies*, 45(3):595–604, 1978.

[106] T. Roughgarden. Routing games. *Algorithmic Game Theory*, 18:459–484, 2007.

[107] A. Rubinstein. *Modeling Bounded Rationality.* MIT press, 1998.

[108] A. Sinha and A. Anastasopoulos. Structured perfect Bayesian equilibrium in infinite horizon dynamic games with asymmetric information. *American Control Conference*, 2016.

[109] S. Sorin. *A First Course on Zero-sum Repeated Games*, volume 37. Springer Science & Business Media, 2002.

[110] R. Spiegler. *Bounded Rationality and Industrial Organization*. Oxford University Press, 2011.

[111] T. Sugaya. Efficiency in Markov games with incomplete and private information. *working paper*, 2012.

[112] C. Swamy. The effectiveness of Stackelberg strategies and tolls for network congestion games. *ACM Transactions on Algorithms (TALG)*, 8(4):36, 2012.

[113] C. Tan and P. Varaiya. Interruptible electric power service contracts. *Journal of Economic Dynamics and Control*, 1993.

[114] W. Tang and R. Jain. Stochastic resource auctions for renewable energy integration. In *49th Annual Allerton Conference on Communication, Control, and Computing*, pages 345–352. IEEE, 2011.

[115] W. Tang and R. Jain. Market mechanisms for buying random wind. *To appear in IEEE Transactions on Sustainable Energy*, 2016.

[116] H. Tavafoghi, Y. Ouyang, and D. Teneketzis. On stochastic dynamic games with delayed sharing information structure. In *55th IEEE Conference on Decision and Control (CDC)*, pages 7002–7009. IEEE, 2016.

[117] H. Tavafoghi and D. Teneketzis. Optimal energy procurement from a strategic seller with private renewable and conventional generation. *arXiv:1401.5759*.

[118] H. Tavafoghi and D. Teneketzis. Optimal contract design for energy procurement. In *52nd Annual Allerton Conference on Communication, Control, and Computing*, pages 62–69. IEEE, 2014.

[119] H. Tavafoghi and D. Teneketzis. Sequential contracts for uncertain electricity resources. *10th Workshop on the Economics of Networks, Systems and Computation (NetEcon)*, 2015.

[120] H. Tavafoghi and D. Teneketzis. Dynamic market mechanisms for wind energy. *arXiv:1608.04143*, 2016.

[121] H. Tavafoghi and D. Teneketzis. Informational incetives in congestion games. In *to appear in 55th Annual Allerton Conference on Communication, Control, and Computing*. IEEE, 2017.

[122] H. Tavafoghi and D. Teneketzis. Multi-dimensional forward contracts under uncertainty for electricity markets. *IEEE Transactions on Control of Network Systems*, 2017.

[123] D. Teneketzis. On the structure of optimal real-time encoders and decoders in noisy communication. *IEEE Transactions on Information Theory*, 52(9):4017–4035, 2006.

[124] J. Thai, N. Laurent-Brouty, and A. Bayen. Negative externalities of gps-enabled routing applications: A game theoretical approach. In *19th International Conference on Intelligent Transportation Systems (ITSC)*, pages 595–601. IEEE, 2016.

[125] Report to the US Congress U.S. Department of Energy. 2010 Smart Grid System Report. February 2012.

[126] P. Varaiya, F. Wu, and J. Bialek. Smart operation of smart grid: Risk-limiting dispatch. *Proceedings of the IEEE*, 2011.

[127] D. Vasal and A. Anastasopoulos. Signaling equilibria for dynamic LQG games with asymmetric information. In *55th IEEE Conference on Decision and Control (CDC)*, pages 6901–6908. IEEE, 2016.

[128] W. Vickrey. Congestion theory and transport investment. *The American Economic Review*, 59(2):251–260, 1969.

[129] R. Vohra. *Mechanism Design: a Linear Programming Approach*, volume 47. Cambridge University Press, 2011.

[130] M. Vrakopoulou, J. Mathieu, and G. Andersson. Stochastic optimal power flow with uncertain reserves from demand response. In *47th Hawaii International Conference on System Sciences, 2014*. IEEE, 2014.

[131] J. Walrand and P. Varaiya. Optimal causal coding-decoding problems. *IEEE Transactions on Information Theory*, 29(6):814–820, 1983.

[132] J. Watson. Perfect Bayesian equilibrium: general definitions and illustrations. *working paper*, 2016.

[133] R. Wilson. *Nonlinear Pricing*. Oxford University Press, 1997.

[134] R. Wilson. Architecture of power markets. *Econometrica*, 70(4):1299–1340, 2002.

[135] H. Witsenhausen. Separation of estimation and control for discrete time systems. *Proceedings of the IEEE*, 59(11):1557–1566, 1971.

[136] H. Witsenhausen. On the structure of real-time source coders. *The Bell System Technical Journal*, 58(6):1437–1451, 1979.

[137] S. Yüksel and T. Başar. *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Springer Science & Business Media, 2013.

[138] S. Zamir. Repeated games of incomplete information: Zero-sum. *Handbook of Game Theory*, 1:109–154, 1992.

[139] Y. Zhao, J. Qin, R. Rajagopal, A. Goldsmith, and V. Poor. Wind aggregation via risky power markets. *IEEE Transactions on Power Systems*, 2014.