ESSAYS ON NETWORK ECONOMICS

by

Majid Mahzoon

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

Doctor of Philosophy in Economics

Tepper School of Business
Carnegie Mellon University
Pittsburgh, PA

April 2023

Dissertation Committee:
Ali Shourideh (chair)
Maryam Saeedi (chair)
James A. Best
Luca Rigotti

Contents

Chapter 1

Hierarchical Bayesian Persuasion: Importance of Vice Presidents

1.1 Abstract

In this paper, I study strategic information transmission in a hierarchy. Information is transmitted through a chain of agents to a decision maker whose action affects all agents' payoffs. Each agent in the chain can conceal all or part of the information she receives. I prove that it is possible to focus on simple equilibria, where only the first agent ever conceals information. This allows me to formulate the hierarchical communication as a direct one between the initial sender and the decision maker subject to recursively defined incentive compatibility constraints imposed by the intermediaries. In the binary-action case, regardless of the number of intermediaries, at most four agents determine the amount of information communicated to the decision maker. In this case, the results in this paper underscore the importance of choosing a pivotal vice president for maximizing the payoff of the decision maker, who is better off by appointing a like-minded but stubborn vice president. Moreover, I provide necessary and sufficient conditions for inefficient equilibria.

1.2 Introduction

On January 28, 1986, the space shuttle Challenger exploded shortly after takeoff, killing all seven crew members aboard. In a meeting the night before the launch, NASA engineers had recommended against the launch. Their supervisor, however, challenged this recommendation, overruled their decision, and approved the launch. The engineers' concerns had not been communicated beyond a certain management level, mainly because of the management structure at NASA, which proved to be fatal in this scenario:

> Failures in communication [...] resulted in a decision to launch 51-L based on incomplete and sometimes misleading information, a conflict between engineering data and management judgments, and a NASA management structure that permitted internal flight safety problems to bypass key Shuttle managers. [...] in the launch preparation for 51-L relevant concerns of Level III

> NASA personnel and element contractors were not [...] adequately communicated to the NASA Level I and II management responsible for the launch.[1]

Organizations, such as NASA, rely on hierarchies to transmit the information vital for improved decision making. However, it is very common that decision makers do not know what is happening on the ground. This is because too much information is filtered along the hierarchy, possibly because of the strategic interests of the middle managers.

Communication failures can lead to suboptimal decisions as in the Challenger disaster. How can decision makers (e.g., the CEO of NASA) avoid such failures and facilitate the most effective communication? In this paper, I address this question by studying communication in a hierarchical organization. To that end, it is important to understand how the hierarchy configuration (i.e., the rank and preferences of each middle manager) affects the information communicated to the decision maker.

I find that the decision maker can enhance information communication and avoid inefficiencies by appointing a *like-minded* vice president who is *more stubborn* than the decision maker. More precisely, the decision maker is better off by choosing a vice president who ex ante has similar action preferences but requires more information to change her mind. Such a vice president acts as a gatekeeper who would pass to the decision maker only the messages that are sufficiently informative for decision making purposes. This characteristic of the vice president motivates the middle managers to conceal less information if they are willing to persuade the decision maker toward an action other than the status quo.

Therefore, here I investigate the outcome of intermediated communication from a low-ranked agent to the decision maker (e.g., CEO or president). I model this kind of hierarchical communication in the framework of Bayesian persuasion developed by Kamenica and Gentzkow (2011). There is a chain of agents from an initial sender to the final receiver (i.e., the decision maker). The initial sender, who is the only agent with direct access to the state, intends to persuade the decision maker to take some action by choosing how much of her information to reveal. However, the agent does not have direct access

[1] Report of the Presidential Commission on the Space Shuttle Challenger Accident, Chapter V

to the decision maker and can reveal information only to the next agent in the chain; the next agents in the chain (i.e., the intermediaries) do the same successively. Therefore, each intermediary can conceal all or part of the information she receives from the previous agent but cannot reveal more information than what she receives, because she does not have direct access to the state. The sequential nature of the game and the assumption that each intermediary, when choosing her revelation strategy, observes only the revelation strategies of the preceding agents and not the information revealed by them justifies using the subgame perfect equilibrium as the solution concept.

First, I prove a version of the revelation principle for this setting: It is without loss of generality to focus on a simple class of equilibria, where, on the equilibrium path, all intermediaries pass all the information they receive to the next agent in the hierarchy. This implies that the initial sender chooses how much information will be revealed to the decision maker, as in the single-sender scenario in Kamenica and Gentzkow (2011), subject to that choice being incentive compatible for all the intermediaries. This helps me formulate the problem of finding the equilibrium outcome recursively, which implies that solving a hierarchical persuasion game is equivalent to solving a single-sender persuasion game subject to recursively defined incentive compatibility constraints. Considering the restriction to simple equilibria, one may wonder if the incentive compatibility constraints can be summarized in the condition that no intermediary prefers to conceal more information. I show that although this simple condition is sufficient for incentive compatibility, it is not necessary.

Then, I focus on the binary-action case where the decision maker will choose between the *status quo* action and the *alternative* action. The status quo action is the decision maker's uninformed action, i.e., the action he takes without any information. In the special case where the state space is also binary, I fully characterize the equilibrium outcome. In this case, the preferences of the agents can be summarized along two dimensions: *stance* and *stubbornness*. Stance divides the agents into *conformists* and *contrarians* depending on whether they prefer the same action as the decision maker in each state or not.[2] An agent is more stubborn if more information is required to dissuade that agent from her uninformed action (i.e., the action she prefers without any information). These observations let me categorize all the possible hierarchy configurations into four types:

1. If there exists an agent who would provide no information in a single-sender game with the decision maker or there exist two agents with highly opposed preferences,[3] no information will be communicated to the decision maker.

2. If all the agents are of the same stance as the decision maker, full information will be communicated to the decision maker.

   Otherwise, the rank and preferences of the most stubborn agent among those who have the same stance and uninformed action as the decision maker,[4] called the *pivotal agent*, determine the outcome.

3. If there exists an agent of opposite stance who is higher ranked (i.e., closer to the decision maker) than the pivotal agent, no information will be communicated to the decision maker.

4. If the pivotal agent is higher ranked than all the agents of opposite stance, some information will be communicated to the decision maker. The more stubborn the pivotal player is, the more information will be communicated to the decision maker.

The non-monotonicity of the pivotal agent's preferences over decision maker's information drives this result: If the decision maker is sufficiently informed (i.e., more than a threshold) when taking the alternative action, a more informed decision maker is preferred; otherwise, an uninformed one is preferred.[5] In other words, the pivotal

[2]For simplicity, here I consider agents who always prefer the same action (i.e., extremists) as the contrarians. In Section 1.7, I will look at these extremists more closely.

[3]Two agents are said to have highly opposed preferences if there exists no belief under which they both prefer the alternative action.

[4]The uninformed action of these agents is the status quo.

[5]This is the case for all the agents who have the same stance and uninformed action as the decision maker. Since the pivotal agent is the most stubborn one, her threshold is higher than those of the other agents.

agent would pass information only if given that information, the decision maker would take the action preferred by the pivotal agent. Therefore, in the last type of hierarchy described above, the game is equivalent to one where the pivotal agent is making decisions. More generally, regardless of the hierarchy configuration, the decision maker prefers to *delegate* the decision rights to the pivotal agent.

The more stubborn the pivotal player is, the higher is her informativeness threshold below which she prefers an uninformed decision maker. This characteristic motivates all the lower-ranked agents to conceal less information if they are willing to persuade the decision maker to take the alternative action. However, in the presence of higher-ranked agents of opposite stance, the pivotal agent would preemptively provide no information because those agents would prevent the decision maker from being sufficiently informed. In this case, the equilibrium is inefficient because there are more informative outcomes that are preferred by all the agents.

The above categorization of hierarchy configurations has an implication for the choice of vice president. If the decision maker could add an intermediary of his choice at the end of the hierarchy (i.e., the vice president), the decision maker would optimally choose this intermediary such that in the new longer hierarchy, the vice president is the pivotal agent. The optimal choice of the vice president is (almost) independent of the hierarchy configuration and eliminates any inefficiencies in communication. The intuition behind the important role of a pivotal vice president in enhancing information communication along a hierarchy is as follows:

- The vice president is the closest agent to the decision maker. Because agents move sequentially, every agent is better off by taking into account the preferences of the higher-ranked agents in the hierarchy. Therefore, all the agents are better off taking into account the preferences of the vice president.

- A pivotal vice president has the same stance and uninformed action as the decision maker. Thus, an agent who intends to persuade the decision maker to take the alternative action must first persuade the vice president. Otherwise, the vice president will not pass any information to the decision maker, being sure that the decision maker will take the action favored by the vice president (i.e., the status quo).

- A pivotal vice president is very stubborn. This implies that more information is required to change her mind and persuade her to prefer the alternative action.

Next, I let the state space be the interval $[0, 1]$ keeping the action space binary. Focusing on simple equilibria implies that, without loss of generality, the game can be considered a binary-state, binary-action one from the intermediaries' point of view. By choosing how much information to reveal, the initial sender implicitly pins down the preferences of the intermediaries in this subgame whose outcome was described above. In general, solving for the optimal choice of the initial sender is not straightforward. I fully characterize the equilibrium outcome in the special case where the prior is uniform, the differential payoff from the actions is linear in the state, and one of the actions yields zero payoff to the initial sender. The linearity assumption implies that the agents' preferences depend only on the posterior mean.

The results in this special case are very similar to those in the binary-state case. Defining a binary-state game corresponding to the original game, I show that no information is communicated to the decision maker in the original game if the same occurs in this binary-state game. Therefore, the conditions resulting in the no-information equilibrium outcome in the binary-state case remain valid in this more general case. However, if these conditions do not hold, the relative preferences of two specific agents in the hierarchy determine the efficiency of the equilibrium and whether any information is communicated to the decision maker. Every equilibrium outcome is equivalent to one which simply distinguishes between two intervals $[0, x]$ and $[x, 1]$, where x is determined by the preferences of the initial sender and two other agents called the pivotal agents. These two pivotal agents are the most stubborn agents among those who prefer the same action as the decision maker in the extreme states (i.e., states 0 and 1).

One implication of this result is that if the decision maker could add an intermediary of his choice at the end of the hierarchy (i.e., the vice president), the decision maker would optimally choose this intermediary such that, in the new longer hierarchy, the vice president is one of the pivotal agents. The optimal choice of vice president depends on the hierarchy configuration and eliminates any inefficiency in communication. Moreover, if the decision maker could add two intermediaries of his choice at the end of the hierarchy, the decision maker would optimally choose these intermediaries such that, in the new longer hierarchy, they are the pivotal agents. The order of these two intermediaries does not matter. Moreover, the decision maker's optimal choice of intermediaries is (almost) independent of the hierarchy configuration and eliminates any inefficiencies in communication.

1.2.1 Related Literature

This paper contributes to the literature on information design and Bayesian persuasion; see Aumann et al. (1995), Kamenica and Gentzkow (2011), and Bergemann and Morris (2016a,b) as well as surveys by Bergemann and Morris (2019) and Kamenica (2019). I apply these tools to a setting with multiple senders who move sequentially, to study how information revelation depends on the order and preferences of the senders.

Ambrus et al. (2013) study the same problem under cheap talk communication, whereas I model it in the framework of Bayesian persuasion. In fact, the authors extend the classic model of Crawford and Sobel (1982) to investigate intermediated communication and show that the set of pure-strategy equilibrium outcomes does not depend on the order of intermediaries and intermediation cannot improve information in these equilibria. Studying cheap talk intermediation networks more complex than the hierarchical ones, Laclau et al. (2020) investigate when it is possible to robustly implement all equilibrium outcomes of the direct communication game on a communication network. In my setting, I find that the equilibrium outcome depends on the order of intermediaries and intermediation can improve information in equilibrium.

I also contribute to the literature on information transmission within organizations. Dessein (2002), Rantakari (2008), and Alonso et al. (2008) consider a decision maker who is trying to elicit information from the agents. All these papers use the cheap talk model of communication, whereas I use Bayesian persuasion. In particular, considering the single-sender scenario, Dessein (2002) identifies cases in which the decision maker optimally delegates control to an intermediary rather than keeping authority or delegating to the sender. I show that regardless of the number of intermediaries, when facing a binary decision, the decision maker prefers to delegate the decision rights to the pivotal agent.

Mathevet and Taneva (2020) consider a designer who disseminates information to the agents. Their concept of vertical communication is similar to my hierarchical communication model. They use the cheap talk model of communication along the hierarchy after the designer commits to a signal structure for communicating with the most-informed agent in the hierarchy. Moreover, the authors assume that each agent in the hierarchy takes an action which affects the utilities of all the agents, whereas in my model, only the final receiver (i.e., the decision maker) is taking an action.

Kamenica and Gentzkow (2016) and Gentzkow and Kamenica (2017) study Bayesian persuasion with multiple senders, assuming that every sender has direct access to the state and can choose a signal that is arbitrarily correlated with those of the other senders. The authors focus on pure-strategy simultaneous-move equilibria and show that greater competition (e.g., adding senders) tends to increase the amount of information revealed. Their result does not hold in my setting, where the initial sender is the only one with direct access to the state and senders move sequentially designing their experiments before observing any signal realizations;[6] adding senders may increase or decrease the amount of information revealed.

The most closely related papers to mine are those by Li and Norman (2021) and Wu (2021), who study the sequential version of the Bayesian persuasion game with multiple senders considered by Kamenica and Gentzkow (2016) and Gentzkow and Kamenica (2017). The rich signal space considered by Li and Norman (2021) and the model considered by Wu (2021) imply that every sender has direct access to the state and observes not only the experiments designed by the preceding senders but also the corresponding signal realizations, when designing her own experiment, and the receiver observes all designed experiments and all signal realizations; that is, each sender may decide only

[6] Li and Norman (2018) show that adding senders can result in loss of information if any of the following assumptions is violated: (i) information can be arbitrarily correlated, (ii) senders reveal information simultaneously, or (iii) senders play pure strategies.

whether to provide more information. In contrast, in my setting, the initial sender is the only one with direct access to the state and each sender observes only the experiments designed by the preceding senders; that is, each sender may decide only whether to provide less information.

Li and Norman (2021) show that it is without loss of generality to restrict attention to a finite set of vertex beliefs. The authors introduce the notion of one-step equilibria, which is the same as simple equilibria in my setting, and formulate the incentive compatibility constraints of the intermediaries to characterize those equilibria. Investigating consultation with multiple experts, they show that adding a sender (i.e., expert) who moves first cannot reduce informativeness in equilibrium; although this question is not relevant in my setting, the opposite holds. In fact, in my setting, which models communication in hierarchical organizations, it is more relevant to study the effect of adding a sender who moves last (i.e., the vice president); my results suggest that adding an appropriate such sender can increase informativeness and efficiency in equilibrium.

Wu (2021) uses recursive concavification to characterize the full set of subgame perfect equilibrium paths. The author also introduces the notion of silent equilibria, which is the same as simple equilibria in my setting. Kuang et al. (2019) consider the binary-state, binary-action case of the sequential models considered in Li and Norman (2021) and Wu (2021), where all senders except for the initial one care about their reputations measured by whether the receiver's action is consistent with their recommendation. One of their results implies the importance of the vice president from the initial sender's point of view, whereas in my setting, the vice president is important from the receiver's point of view.

In simultaneous and independent work, Arieli et al. (2022) also study the hierarchical Bayesian persuasion problem. The authors characterize the initial sender's optimal value using constrained concavification, whereas I focus on the receiver's welfare. The notion of constrained concavification is similar to my recursive formulation. However, the concavification method is not easy to work with; my characterization of recursive incentive compatibility constraints transforms the problem into a simpler and more tractable one. Doing so allows me to derive interesting results in the binary-action case and provide insights about the importance of vice presidents in hierarchical organizations.

The rest of the paper is organized as follows. In section 1.3, I illustrate the importance of vice presidents via a simple example. Section 1.4 sets up the model. In section 1.5, I prove that it is without loss of generality to focus on the set of simple equilibria., where the initial sender is the only agent who may conceal information. Section 1.6 includes the recursive formulation of incentive compatibility constraints, restates the hierarchical persuasion problem as a single-sender one, and investigates a simple sufficient condition for the incentive compatibility conditions. In section 1.7, I fully characterize the equilibrium in the binary-state, binary-action case, and in a special case of the general binary-action case, and discuss the efficiency of equilibria and importance of vice presidents. Section 1.8 concludes the paper. Some of the proofs are relegated to the appendix.

1.3 A Simple Example

Consider a hierarchy of three agents, as shown in Figure 1.1, consisting of the initial sender (agent 1), an intermediary (agent 2), and the receiver (R). All the agents share a uniform prior belief about the binary state of the world $\omega \in \{0, 1\}$. First, agent 1 designs an experiment to send a signal $s_1 \in \{0, 1\}$ to agent 2 about the state ω. Then, agent 2, observing the experiment chosen by agent 1 but not the signal realization, designs another experiment to send a signal $s_2 \in \{0, 1\}$ to the receiver about signal s_1. Finally, the receiver observes the experiments designed by the agents and the signal realization s_2, and chooses a binary action $a \in \{0, 1\}$.

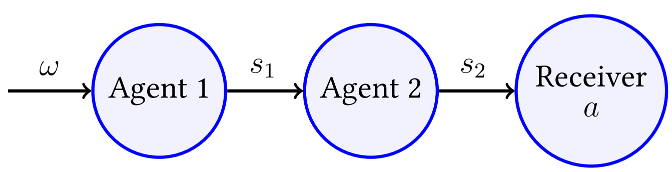


Figure 1.1: A hierarchy consisting of three agents

Suppose that with full information about the state, agent 1 and the receiver would prefer the action that matches the state. However, in the absence of full information,

- agent 1 would prefer action $a = 1$ if and only if she believes that the state is $\omega = 1$ with probability greater than μ_1, where $0 < \mu_1 < 0.5$;

- the receiver would prefer action $a = 1$ if and only if he believes that the state is $\omega = 1$ with probability greater than $\mu_R = 0.3$;

Moreover, suppose agent 2 would prefer action $a = 0$ regardless of the state.

Agent 1 prefers to provide the receiver with as much information as possible (i.e., send signal $s_1 = 0$ in state $\omega = 0$ and signal $s_1 = 1$ in state $\omega = 1$). Agent 2, however, prefers to send signal $s_2 = 0$ as often as possible (i.e., always send signal $s_2 = 0$ when $s_1 = 0$ and sometimes send signal $s_2 = 0$ when $s_1 = 1$). Agent 2 does this as long as, after receiving signal $s_2 = 0$, the receiver believes the state is $\omega = 1$ with probability less than or equal to $\mu_R = 0.3$.

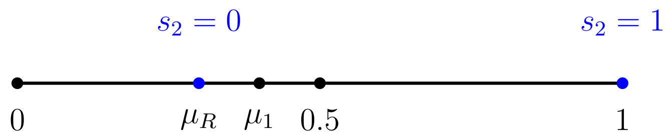
- **Case 1**: Suppose $\mu_1 = 0.4$. In equilibrium, as shown in Figure 1.2a, the receiver would believe that the state is $\omega = 1$ for sure after observing signal $s_2 = 1$ and that the state is $\omega = 1$ with probability $\mu_R = 0.3$ after observing signal $s_2 = 0$.

- **Case 2**: Suppose $\mu_1 = 0.2$. In equilibrium, as shown in Figure 1.2b, agent 1 will provide no information to agent 2. This is because if agent 1 provides any information, Case 1 will follow, and she prefers action $a = 1$ even if she believes that the state is $\omega = 1$ with probability $\mu_R = 0.3$. Therefore, no information will be revealed to the receiver. Note that this equilibrium is inefficient, because both agent 1 and agent 2 prefer the receiver to take action $a = 0$ in some cases.

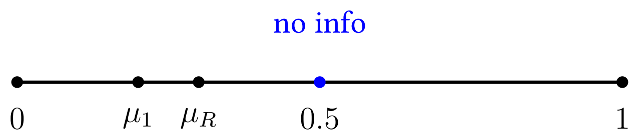
Suppose the receiver adds an agent (agent 3) between himself and agent 2, as shown in Figure 1.3, with the following characteristics:

- with full information about the state, agent 3 would prefer the action that matches the state;

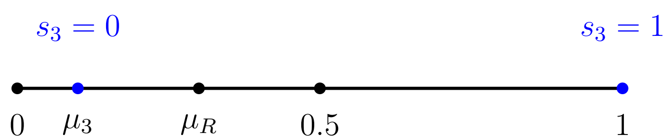
- in the absence of full information, agent 3 would prefer action $a = 1$ if and only if she believes that the state is $\omega = 1$ with probability greater than $\mu_3 = 0.1$.



(a) In the absence of agent 3 (Case 1: $\mu_1 > \mu_R$)



(b) In the presence of agent 3 (Case 2: $\mu_1 < \mu_R$)



(c) In the presence of agent 3

Figure 1.2: The blue dots represent the possible beliefs of the receiver in equilibrium.

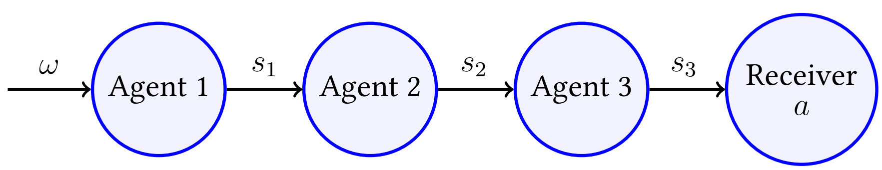


Figure 1.3: A hierarchy consisting of four agents: Compared to Figure 1.1, the receiver has chosen to add agent 3 between himself and agent 2.

In this longer hierarchy, first agent 1 and agent 2 design their experiments. Then agent 3, observing the experiments designed by agent 1 and agent 2 but not the signal realizations, designs an experiment to send a signal $s_3 = \{0, 1\}$ to the receiver about signal s_2. Note that the receiver will now observe signal s_3 instead of signal s_2.

If agent 2 sends signal $s_2 = 0$ as often as she did in the absence of agent 3, agent 3 would prefer action $a = 1$ regardless of signal realization s_2. This is because, whatever signal realization s_2 turns out to be, agent 3 believes that the state is $\omega = 1$ with probability greater than $\mu_3 = 0.1$. Therefore, agent 3 will send no information to the receiver, who will then take action $a = 1$. This is the worst thing that could happen to agent 2, who always prefers action $a = 0$. Therefore, agent 2 has to back down and send signal $s_2 = 0$ less often when $s_1 = 1$. In equilibrium, as shown in Figure 1.2c, agent 3 would believe that the state is $\omega = 1$ for sure after observing signal $s_2 = 1$ and that the state is $\omega = 1$ with probability $\mu_3 = 0.1$ after observing signal $s_2 = 0$. Agent 3 would then reveal all her information to the receiver.

Therefore, in the presence of agent 3, more information is revealed to the receiver in equilibrium. In fact, this would have been the case for any choice of μ_3 such that $0 < \mu_3 < \min(\mu_1, \mu_R)$. Note, however, that if agent 3 were added between agent 1 and agent 2, the equilibrium would have been the same as the one in the absence of agent 3.

Observation: The decision maker can increase the amount of information she receives and eliminate any inefficiency in communication by appointing a vice president who ex ante has similar action preferences but is more stubborn than the decision maker (i.e., $\mu_3 < \mu_R$).

1.4 Model

There are n players (she) interested in the action taken by a receiver R (he). Each player can try to influence the receiver's action by sending a message to the next player in a hierarchical manner. Players' and receiver's payoffs depend on the state of the world $\omega \in \Omega$ and the action taken by the receiver $a \in \mathcal{A}$ where the action space $\mathcal{A}$ is assumed to be finite. In other words, each player's (and receiver's) utility is given by a function $u_i(\omega, a)$, where $i = 1, \ldots, n, R$. All the players are expected utility maximizers.

The players and the receiver are uncertain about the state of the world ω, and their common prior belief is represented by a probability density function f.[7] However, they can obtain information in a hierarchical manner, as shown in Figure 1.4. Each player $i = 1, \ldots, n$, successively sends a signal $s_i \in \mathcal{S}_i$ to player $i + 1$, where $n + 1 = R$; in other words, each player i, successively, designs (commits to) an experiment over $\mathcal{S}_{i-1}$,

[7] If prior belief is a finite-support distribution, then I assume f represents the probability mass function and replace all integrals with sums. The corresponding cumulative distribution function is represented by F.

where $\mathcal{S}_0 = \Omega$, including a signal space $\mathcal{S}_i$ and a signal structure $\pi_i : \mathcal{S}_{i-1} \to \Delta(\mathcal{S}_i)$. In the remainder of the paper, I denote an experiment $(\mathcal{S}, \pi)$ simply by its signal structure π. Let Π and Π_i denote the set of all experiments over Ω and $\mathcal{S}_i$, respectively.

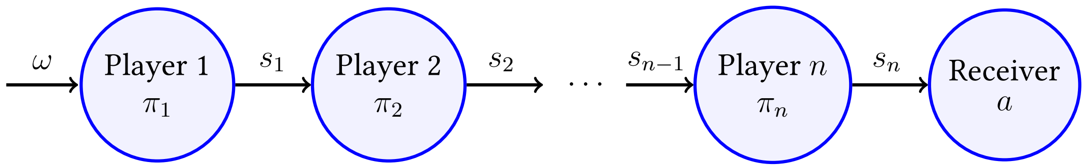


Figure 1.4: A hierarchy consisting of n players and the receiver

When designing her experiment, each player $i = 1, \ldots, n$ observes the experiments designed by the preceding players $\{\pi_j\}_{j=1}^{i-1}$, but not the signal realizations. After each player has designed an experiment, first the state of the world ω, and then, successively, signals $s_1, \ldots, s_n$ are realized according to the designed experiments. The receiver observes the experiments designed by all the players and a signal s_n sent by player n; he updates his belief accordingly and chooses an action $a \in \mathcal{A}$ to maximize his expected utility. Note that player 1 is the only player with direct access to the state, and player n is the only player with direct access to the receiver.

In summary, the game consists of three stages. In the first stage, each player, successively, designs an experiment given the experiments chosen by the preceding players in order to maximize her expected utility. In the second stage, the state of the world is realized and the communication takes place in a hierarchical manner according to the designed experiments. In the last stage, the receiver updates his belief according to the designed experiments and the received signal, and decides what action to take to maximize his expected utility.

For ease of exposition, I assume $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_n$ are all finite. As will be seen in Corollary 2, this assumption is without loss of generality, because $\mathcal{A}$ is finite. Therefore, each π_i for $i = 1, \ldots, n$ is a finite-support conditional distribution, and π_i represents the conditional probability mass function.

1.5 Simple Equilibria

Given the experiments designed by all the players $(\pi_1, \ldots, \pi_n)$ and the signal s_n received from player n, the receiver's expected utility and, thus, optimally chosen action a^* depend only on his posterior belief represented by the probability density function μ.[8]

Assumption 1. *If the receiver is indifferent among a set of actions A_μ given his posterior belief μ, he will take player n's favorite action, i.e., $\mathbb{E}_\mu u_n(\omega, a^*(\mu)) \geq \mathbb{E}_\mu u_n(\omega, a), \ \forall a \in A_\mu$.*[9]

Therefore, I can represent the receiver's optimal strategy by $a^* : \Delta(\Omega) \to \mathcal{A}$. Given the posterior belief of the receiver μ, let $v_i(\mu) = \mathbb{E}_\mu u_i(\omega, a^*(\mu))$ for $i = 1, 2, \ldots, n, R$ represent the expected utility of a player or of the receiver.

The distribution of the signal observed by player i, that is, s_{i-1}, depends on the experiments designed by all the preceding players. The aggregate experiment observed by player i is denoted by $\pi^i : \Omega \to \Delta(\mathcal{S}_{i-1})$ where

$$\pi^i(s_{i-1}|\omega) = \sum_{s'_1} \ldots \sum_{s'_{i-2}} \pi_1(s'_1|\omega)\pi_2(s'_2|s'_1) \ldots \pi_{i-1}(s_{i-1}|s'_{i-2}). \tag{1.1}$$

For ease of exposition, I write $\pi^i \equiv \pi_1 \circ \ldots \circ \pi_{i-1}$. After observing the preceding players' designed experiments, player i's expected utility, and thus her optimally chosen experiment, depend only on π^i, not the individual experiments designed by the preceding players $(\pi_1, \ldots, \pi_{i-1})$. Note that by designing an experiment, each player simply garbles the aggregate experiment designed by the preceding players, that is, $\pi^{i+1} \equiv \pi^i \circ \pi_i$ is less Blackwell-informative than π^i.

Given the receiver's optimal strategy a^*, every profile of experiments designed by the players $(\pi_1, \ldots, \pi_n)$ induces a conditional distribution over the actions taken by the receiver $\pi : \Omega \to \Delta(\mathcal{A})$ called the outcome of the game. Without loss of generality, let us

[8] If the posterior belief is a finite-support distribution, then I assume μ represents the probability mass function.

[9] I use this assumption to ensure that in the next section player n's problem in (1.9) is well-defined.

assume $\mathcal{S}_n = \mathcal{A}$ and the receiver's optimal action is equal to the signal s_n sent by player n[10]: $a^*(\mu_{s_n}) = s_n,\ \forall s_n \in \mathcal{S}_n$. Therefore, the outcome of the game is given by

$$\pi(a|\omega) = \sum_{s_1'} \dots \sum_{s_{n-1}'} \pi_1(s_1'|\omega)\pi_2(s_2'|s_1') \dots \pi_n(a|s_{n-1}'),\ \forall a \in \mathcal{A}, \forall \omega \in \Omega, \tag{1.2}$$

which is the aggregate experiment observed by the receiver. For ease of exposition, I write $\pi \equiv \pi_1 \circ \dots \circ \pi_n$[11]. It is now clear that the receiver's posterior belief is given by the probability density function μ, where

$$\mu(\omega|s_n;\pi_1,\dots,\pi_n) = \frac{f(\omega)\pi(s_n|\omega)}{\int_\Omega f(\omega')\pi(s_n|\omega')d\omega'}. \tag{1.3}$$

Furthermore, every outcome of the game π induces a distribution of posteriors for the receiver τ_π. I sometimes refer to τ_π as the outcome of the game. Since $\mathcal{S}_n = \mathcal{A}$ is finite, τ_π is a finite-support distribution, and I let τ_π represent the probability mass function. Note that the above assumption implies that each signal s_n sent by player n induces a distinct posterior belief for the receiver μ_{s_n}. Therefore,

$$\tau_\pi(\mu_a) = \int_\Omega \sum_{s_1'} \dots \sum_{s_{n-1}'} f(\omega)\pi_1(s_1'|\omega)\pi_2(s_2'|s_1') \dots \pi_n(a|s_{n-1}')d\omega = \int_\Omega f(\omega)\pi(a|\omega)d\omega. \tag{1.4}$$

Definition 1 (Subgame Perfect Equilibria)**.** A subgame perfect equilibrium of the game $\sigma^* = (\pi_1^*, \sigma_2^*, \dots, \sigma_n^*)$ is defined by an experiment $\pi_1^* \in \Pi$ for player 1 and a function (strategy) $\sigma_i^* : \Pi \to \Pi_{i-1}$ for each player $i = 2, \dots, n$ such that

- given $\{\sigma_j^*\}_{j=2}^n$, π_1^* maximizes $\mathbb{E}v_1(\mu)$, and
- for $i = 2, \dots, n$, given $\{\sigma_j^*\}_{j=i+1}^n$, $\sigma_i^*(\pi^i)$ maximizes $\mathbb{E}v_i(\mu)$, $\forall \pi^i \in \Pi$.

[10]This is the case because of Assumption 1. The proof is the same as that of Proposition 1 in Bergemann and Morris (2019).

[11]Note that $\pi = \pi^{n+1}$.

For a subgame perfect equilibrium of the game, let π_i^* denote the experiment chosen by player i on the equilibrium path. The corresponding equilibrium outcome and equilibrium distribution of posteriors for the receiver are denoted by $\pi^* \equiv \pi_1^* \circ \ldots \circ \pi_n^*$ and τ^*, respectively.

Now, I state the first result which is reminiscent of the revelation principle.

Proposition 1 (Simple Equilibria)**.** *Under Assumption 1, for every subgame perfect equilibrium of the game σ^* with equilibrium outcome π^*, there exists an outcome-equivalent subgame perfect equilibrium σ^{**} where player 1's experiment is given by $\pi_1^{**} = \pi^*$ and all other players use the following strategy:*

$$\sigma_i^{**}(\pi^i) = \begin{cases} \mathcal{I} & \textit{if}\, \pi^i = \pi^* \\ \sigma_i^*(\pi^i) & \textit{Otherwise} \end{cases} \tag{1.5}$$

where $\mathcal{I}$ represents the full-revelation experiment, i.e., $\pi \circ \mathcal{I} = \pi, \ \forall \pi \in \Pi$.

This equilibrium is simple in the sense that on the equilibrium path, all the players except for the first one pass all the information they receive to the next player.

Proof. If I prove that the strategy profile σ^{**} is indeed an equilibrium strategy profile, then it is straightforward to see that this is outcome-equivalent to the equilibrium strategy profile σ^*.

Given the experiments chosen by the preceding players and the strategies of the succeeding players $(\pi_1, \ldots, \pi_{i-1}, \sigma_{i+1}, \ldots, \sigma_n)$, by choosing experiment π_i, player i essentially chooses the outcome of the game, where her choice set is a feasible subset of Π. More explicitly, her choice set includes all outcomes of the game π such that $\pi \equiv \pi_1 \circ \ldots \circ \pi_{i-1} \circ \pi_i \circ \sigma_{i+1}(\pi^{i+1}) \circ \ldots \circ \sigma_n(\pi^n)$ for some $\pi_i \in \Pi_{i-1}$.

- Consider player n. The only change in her strategy compared to σ_n^* is when she encounters a profile of experiments chosen by the preceding players such that $\pi^n = \pi^*$. In this case, she can induce any outcome π less Blackwell-informative than $\pi^* = \pi_1^* \circ \ldots \circ \pi_n^*$. In the equilibrium corresponding to σ^*, when she can induce any outcome π less Blackwell-informative than $\pi_1^* \circ \ldots \circ \pi_{n-1}^*$, including π^*, she chooses to induce π^*. Obviously, the choice set is now smaller but includes the optimal choice of a larger choice set;[12] therefore, it is still optimal for player n to induce π^*; that is, $\sigma_n^{**}(\pi^*) = \mathcal{I}$ is optimal.

- Consider player i for $i = 2, \ldots, n-1$. The only change in her strategy compared to σ_i^* is when she encounters a profile of experiments chosen by the preceding players such that $\pi^i = \pi^*$. In this case, she can induce any outcome π such that $\pi = \pi_1^* \circ \ldots \circ \pi_{i-1}^* \circ \pi_i \circ \sigma_{i+1}^{**}(\pi^{i+1}) \circ \ldots \circ \sigma_n^{**}(\pi^n)$ by choosing π_i, where π_i is strictly less Blackwell-informative than $\pi_i^* \circ \ldots \circ \pi_n^*$; she can also induce $\pi = \pi^*$. In the equilibrium corresponding to σ^*, when she can induce any outcome π of the form $\pi_1^* \circ \ldots \circ \pi_{i-1}^* \circ \pi_i \circ \sigma_{i+1}^*(\pi^{i+1}) \circ \ldots \circ \sigma_n^*(\pi^n)$ by choosing π_i, she chooses to induce π^*. Obviously, the choice set is now smaller but includes the optimal choice of a larger choice set; therefore, it is still optimal for player i to induce π^*; that is, $\sigma_i^{**}(\pi^*) = \mathcal{I}$ is optimal.

- Consider player 1. By choosing π_1, she can either choose any outcome π such that $\pi = \pi_1 \circ \sigma_2^{**}(\pi^2) \circ \ldots \circ \sigma_n^{**}(\pi^n) \neq \pi^*$, or choose $\pi = \pi^*$. In the equilibrium corresponding to σ^*, when she can induce any outcome π of the form $\pi_1 \circ \sigma_2^*(\pi^2) \circ \ldots \circ \sigma_n^*(\pi^n)$ by choosing π_1, she chooses to induce π^*. Obviously, the choice set is the same; therefore, it is still optimal for player 1 to induce π^*; that is, $\pi_1^{**} = \pi^*$ is optimal.

■

I can apply the same argument off the equilibrium path as well. Given a subgame perfect equilibrium of the game σ^*, let $\sigma_i^*(h_j, \pi_j, \ldots, \pi_{i-1})$ represent the equilibrium strategy of player $i \geq j$ following a history of the game $h_j = (\pi_1, \ldots, \pi_{j-1})$ (alternatively, in the subgame starting from player j). Similarly, let $\pi_i^*(h_j)$ denote the experiment chosen by player $i \geq j$ on the equilibrium path of the continuation game (alternatively, subgame).

[12] Blackwell information ranking is transitive.

Corollary 1. *Under Assumption 1, for every subgame perfect equilibrium of the game σ^* with equilibrium outcome π^*, there exists an outcome-equivalent subgame perfect equilibrium σ^{**} where following any history of the game $h_j = (\pi_1, \ldots, \pi_{j-1})$ (alternatively, in the subgame starting from player j), player j's strategy is given by $\sigma_j^{**}(h_j) = \pi_j^*(h_j) \circ \pi_{j+1}^*(h_j) \circ \ldots \circ \pi_n^*(h_j)$ and all succeeding players use the following strategy:*

$$\sigma_i^{**}(h_j, \pi_j, \ldots, \pi_{i-1}) = \begin{cases} \mathcal{I} & \text{if } \pi_j \circ \pi_{j+1} \circ \ldots \circ \pi_{i-1} = \\ & \quad \pi_j^*(h_j) \circ \pi_{j+1}^*(h_j) \circ \ldots \circ \pi_n^*(h_j) \\ \pi_i^*(h_i) \circ \pi_{i+1}^*(h_i) \circ \ldots \circ \pi_n^*(h_i) & \text{Otherwise} \end{cases} \tag{1.6}$$

where $h_i = (h_j, \pi_j, \ldots, \pi_{i-1})$.

This equilibrium is simple in the sense that following any history of the game, all the players except for the first one to play pass all the information they receive to the next player.

Being able to focus on simple equilibria implies that on the equilibrium path, the first player recommends to the receiver which action to take, all the other players pass that recommendation to the next player, and the receiver takes the recommended action.[13]

Corollary 2. *Under Assumption 1, it is without loss of generality to assume $\mathcal{S}_1 = \ldots = \mathcal{S}_n = \mathcal{A}$.*

1.6 Incentive Compatibility Constraints and Recursive Formulation

Given Proposition 1, in order to find the equilibrium outcome of the game π^* or, equivalently, the equilibrium distribution of posteriors for the receiver τ^*, I can focus on the "simple" equilibria where the first player chooses the outcome and the others pass the information to the next player. However, the chosen outcome must be Bayes plausible, and

[13] Similarly, following any history of the game, the first player to play recommends to the receiver which action to take, and so on.

it must be incentive compatible for the other players to pass that information to the next player. In this section, I first formally characterize these conditions and then formulate the problem as a set of recursive optimization problems.

Consider a pair of experiments $\pi, \pi' \in \Pi$ and their corresponding distributions of posteriors $\tau, \tau' \in \Delta(\Delta(\Omega))$. Experiment π is more Blackwell-informative than experiment π' if and only if τ' is smaller than τ in the convex order: $\tau' \leq_{cx} \tau$. This means for all convex functions $\phi : \Delta(\Omega) \to \mathbb{R}$[14],

$$\sum_{\mu \in supp(\tau')} \phi(\mu)\tau'(\mu) \leq \sum_{\mu \in supp(\tau)} \phi(\mu)\tau(\mu) \tag{1.7}$$

Now, I formally characterize the necessary conditions for the equilibrium outcome of the game:

- The equilibrium outcome of the game τ^* must be Bayes plausible, i.e., $\sum_\mu \mu\tau^*(\mu) = f$.[15] Let Γ_0 denote the set of Bayes-plausible distributions of posteriors.

- As mentioned before, by designing an experiment, the last player simply garbles the aggregate experiment designed by the preceding players. Therefore, in addition to the above condition, the last player should not prefer any outcome less Blackwell-informative than the equilibrium outcome π^*. Equivalently, she should not prefer any distribution of posteriors τ' that is smaller than the equilibrium distribution of posteriors τ^* in the convex order, i.e.,

$$\tau^* \in \Gamma_0, \quad \sum_\mu \tau'(\mu)v_n(\mu) \leq \sum_\mu \tau^*(\mu)v_n(\mu), \ \forall \tau' \leq_{cx} \tau^*. \tag{1.8}$$

  Let $\Gamma_n \subseteq \Gamma_0$ denote the set of distributions of posteriors τ satisfying (1.8)[16].

  Accordingly, I can define a value function representing the highest attainable expected utility of player n as a function of the distribution of posteriors τ^n corre-

[14]In the remainder of the paper, for ease of exposition, I write $\sum_\mu$ instead of $\sum_{\mu \in supp(\tau)}$.
[15]The proof is the same as that of Proposition 1 in Bergemann and Morris (2019).
[16]Note that $\tau^* \in \Gamma_0$ and $\tau' \leq_{cx} \tau^*$ imply that $\tau' \in \Gamma_0$.

sponding to the aggregate experiment π^n designed by the preceding players:

$$V_n(\tau^n) = \max_{\tau' \leq_{cx} \tau^n} \sum_{\mu} \tau'(\mu) v_n(\mu). \tag{1.9}$$

This implies that $\Gamma_n = \{\tau \in \Gamma_0 | \sum_\mu \tau(\mu) v_n(\mu) = V_n(\tau)\}$.

- Again, by designing an experiment, the second-to-last player simply garbles the aggregate experiment designed by the preceding players. Moreover, the only outcomes the second to last player can induce are those in Γ_n. Therefore, in addition to (1.8), the second to last player should not prefer any "induceable" outcome less Blackwell-informative than the equilibrium outcome π^*. Equivalently, she should not prefer any distribution of posteriors $\tau' \in \Gamma_n$ that is smaller than the equilibrium distribution of posteriors τ^* in the convex order, i.e.,

$$\tau^* \in \Gamma_n, \quad \sum_{\mu} \tau'(\mu) v_{n-1}(\mu) \leq \sum_{\mu} \tau^*(\mu) v_{n-1}(\mu), \ \forall \tau' \in \Gamma_n \text{ s.t. } \tau' \leq_{cx} \tau^*. \tag{1.10}$$

  Let $\Gamma_{n-1} \subseteq \Gamma_n$ denote the set of distributions of posteriors τ satisfying (**??**).

  Accordingly, given V_n, I can define a value function representing the highest attainable expected utility of player $n-1$ as a function of the distribution of posteriors τ^{n-1} corresponding to the aggregate experiment π^{n-1} designed by the preceding players:

$$\begin{gathered} V_{n-1}(\tau^{n-1}) = \max_{\tau' \leq_{cx} \tau^{n-1}} \sum_{\mu} \tau'(\mu) v_{n-1}(\mu) \text{ subject to} \\ \sum_{\mu} \tau'(\mu) v_n(\mu) = V_n(\tau'). \end{gathered} \tag{1.11}$$

  This implies that $\Gamma_{n-1} = \{\tau \in \Gamma_n | \sum_\mu \tau(\mu) v_{n-1}(\mu) = V_{n-1}(\tau)\}$.

- Continuing like this, now consider player i for $i = 2, \ldots, n-2$. Yet again, by designing an experiment, player i simply garbles the aggregate experiment designed by the preceding players. Moreover, the only outcomes player i can induce are those in Γ_{i+1}. Therefore, in addition to $\tau^* \in \Gamma_{i+1}$, player i should not prefer any "induceable" outcome less Blackwell-informative than the equilibrium outcome π^*. Equivalently, she should not prefer any distribution of posteriors $\tau' \in \Gamma_{i+1}$ that is smaller than the equilibrium distribution of posteriors τ^* in the convex order, i.e.,

$$\tau^* \in \Gamma_{i+1}, \quad \sum_{\mu} \tau'(\mu) v_i(\mu) \leq \sum_{\mu} \tau^*(\mu) v_i(\mu), \ \forall \tau' \in \Gamma_{i+1} \text{ s.t. } \tau' \leq_{cx} \tau^*. \tag{1.12}$$

Let $\Gamma_i \subseteq \Gamma_{i+1}$ denote the distributions of posteriors τ satisfying (1.12).

Accordingly, given V_{i+1}, I can define a value function representing the highest attainable expected utility of player i as a function of the distribution of posteriors τ^i corresponding to the aggregate experiment π^i designed by the preceding players:

$$\begin{aligned} V_i(\tau^i) = \max_{\tau' \leq_{cx} \tau^i} & \sum_{\mu} \tau'(\mu) v_i(\mu) \text{ subject to} \\ & \sum_{\mu} \tau'(\mu) v_{i+1}(\mu) = V_{i+1}(\tau'). \end{aligned} \tag{1.13}$$

This implies that $\Gamma_i = \{\tau \in \Gamma_{i+1} | \sum_{\mu} \tau(\mu) v_i(\mu) = V_i(\tau)\}$.

Now, I formally formulate the problem as a set of recursive optimization problems.

Proposition 2 (Equilibrium Distribution of Posteriors)**.** *Under Assumption 1, the equilibrium distribution of posteriors for the receiver is given by*

$$\tau^* \in \operatorname*{argmax}_{\tau \in \Gamma_2} \sum_{\mu} \tau(\mu) v_1(\mu). \tag{1.14}$$

Solving a hierarchical persuasion game is equivalent to solving a single-sender persuasion game subject to recursively-defined incentive compatibility constraints.

Proof. As mentioned before, Proposition 1 allows me to focus on the "simple" equilibria where the first player chooses the outcome τ^* and the others pass the information to the next player, as long as τ^* is Bayes plausible and it is incentive compatible for the other players to pass the information to the next player. Based on the discussion so far, these conditions are equivalent to $\tau^* \in \Gamma_2$. ■

1.6.1 Can Incentive Compatibility Constraints be Simplified?

When formulating the incentive compatibility constraints, I mentioned that the last player should not prefer any outcome less Blackwell-informative than the equilibrium outcome, whereas player $i = 2, \ldots, n-1$ should not prefer any induceable outcome less Blackwell-informative than the equilibrium outcome. One may wonder whether it is possible to simplify the incentive compatibility constraints and say player $i = 2, \ldots, n$ should not prefer any outcome less Blackwell-informative than the equilibrium outcome. To that end, let $\tilde{\Gamma}$ denote the Bayes-plausible distributions of posteriors τ_π such that no player (not considering player 1) prefers any outcome less Blackwell-informative than π:

$$\tilde{\Gamma} = \{\tau \in \Gamma_0 | \sum_{\mu} \tau(\mu) v_i(\mu) \geq \sum_{\mu} \tau'(\mu) v_i(\mu), \forall \tau' \leq_{cx} \tau, \forall i = 2, \ldots, n\}. \tag{1.15}$$

The next proposition illustrates that while this simple condition implies incentive compatibility and is thus a sufficient condition, it is stricter than needed and thus not a necessary condition.

Proposition 3. *Under Assumption 1,* $\tilde{\Gamma} \subseteq \Gamma_2$*, but* $\tilde{\Gamma} \neq \Gamma_2$.

Proof. See Appendix C. ■

1.7 Binary-Action Games

I now consider a common special case where the action space is binary: $\mathcal{A} = \{0, 1\}$.

Based on Corollary 2, it is without loss of generality to assume $\mathcal{S}_1 = \ldots = \mathcal{S}_n = \{0, 1\}$. Thus, player 1's experiment can be written as $\pi_1 : \Omega \to [0, 1]$, where $\pi_1(\cdot)$ represents the conditional probability of sending signal 1. Similarly, player i's experiment can be written as $\pi_i : \{0,1\} \to [0,1]$ for $i = 2, \ldots, n$. Following Proposition 1 and Corollary 2, every subgame perfect equilibrium of the game is outcome-equivalent to the one where the first player recommends to the receiver what action to take and all the other players pass that recommendation to the next player.

Player 1 chooses an experiment π_1, or, equivalently, a Bayes-plausible distribution of posteriors $\tau_1 \in \Delta(\Delta(\Omega))$, where $supp(\tau_1) = \{q_0, q_1\}$.[17] From this point on, the game is a binary-state, binary-action one. In other words, given the experiment chosen by player 1, the game played by players $2, 3, \ldots, n$ is a binary-state one where the state space is given by $\hat{\Omega} = supp(\tau_1)$.

I now consider the binary-state, binary-action game.

1.7.1 Binary-State Binary-Action Games

Let $\Omega = \mathcal{A} = \{0,1\}$. Each player's and receiver's utility is represented by four numbers: $u_i^{00} = u_i(\omega = 0, a = 0)$, $u_i^{10} = u_i(\omega = 1, a = 0)$, $u_i^{01} = u_i(\omega = 0, a = 1)$, and $u_i^{11} = u_i(\omega = 1, a = 1)$ for $i = 1, \ldots, n, R$. Let $p \in (0,1)$ and $\mu \in [0,1]$ represent the common prior belief and receiver's posterior belief that $\omega = 1$, respectively, and let μ_i represent the posterior belief at which player i is indifferent between the two actions:

$$(1-\mu_i)u_i^{00} + \mu_i u_i^{10} = (1-\mu_i)u_i^{01} + \mu_i u_i^{11} \Rightarrow \mu_i = \frac{u_i^{00} - u_i^{01}}{(u_i^{00} - u_i^{01}) + (u_i^{11} - u_i^{10})}. \quad (1.16)$$

All I need to know about player i is her indifference posterior belief μ_i[18] and her preferred action at state $\omega = 1$[19]. As a result, players can be categorized into different types, as follows:

1. 0-Extremists: If $\mu_i > 1$ or $\mu_i < 0$ and player i prefers $a = 0$ at $\omega = 1$, she will always prefer $a = 0$.

[17] If player 1 chooses to provide no information, $q_0 = q_1$ and this leads to the no-information outcome.
[18] For simplicity, I assume $\mu_i \notin \{0, p, 1\}$ and $\mu_i \neq \mu_j$ if $i \neq j$.
[19] The argument behind this claim can be found in Appendix A.

2. 1-Extremists: If $\mu_i > 1$ or $\mu_i < 0$ and player i prefers $a = 1$ at $\omega = 1$, she will always prefer $a = 1$.

3. Conformists: If $0 < \mu_i < 1$ and player i prefers $a = 1$ at $\omega = 1$, she will prefer $a = 0$ if $\mu < \mu_i$ and $a = 1$ if $\mu > \mu_i$, as shown in Figure 1.5a.

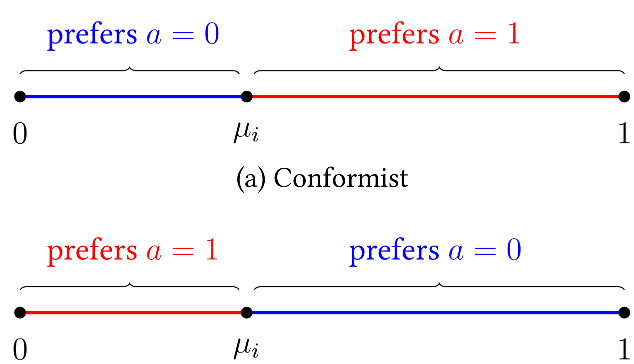
   (a) If $0 < \mu_i < p$, player i is said to be biased[20] toward $a = 1$; the lower μ_i, the higher the bias, as shown in Figure 1.6.

   (b) If $p < \mu_i < 1$, player i is said to be biased toward $a = 0$; the higher μ_i, the higher the bias.

4. Contrarians: If $0 < \mu_i < 1$ and player i prefers $a = 0$ at $\omega = 1$, she will prefer $a = 0$ if $\mu > \mu_i$ and $a = 1$ if $\mu < \mu_i$, as shown in Figure 1.5b.

   (a) If $0 < \mu_i < p$, player i is said to be biased toward $a = 0$; the lower μ_i, the higher the bias, as shown in 1.6.

   (b) If $p < \mu_i < 1$, player i is said to be biased toward $a = 1$; the higher μ_i, the higher the bias.



(a) Conformist

(b) Contrarian

Figure 1.5: Action preferences of conformists and contrarians

[20] A player is biased toward action a if she prefers that action given the common prior belief p.

For example, consider the following utility function

$$u_i(\omega, a) = -(a - \omega)^2 - \alpha_i a, \tag{1.17}$$

where α_i represents player i's willingness (bias) to take action $a = 0$. Then, $\mu_i = \frac{1+\alpha_i}{2}$. In this example, if $\alpha_i > 1$, player i is a 0-extremist, and if $\alpha_i < -1$, she is a 1-extremist; otherwise, she is a conformist, because she would prefer to match the state.

Similarly, consider the following utility function

$$u_i(\omega, a) = (a - \omega)^2 - \alpha_i(1 - a), \tag{1.18}$$

where α_i represents player i's willingness (bias) to take action $a = 1$. Then, $\mu_i = \frac{1+\alpha_i}{2}$. In this example, if $\alpha_i > 1$, player i is a 1-extremist, and if $\alpha_i < -1$, she is a 0-extremist; otherwise, she is a contrarian, because she would prefer not to match the state.

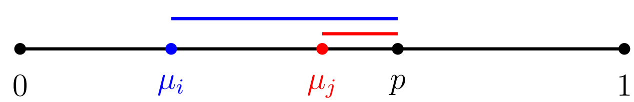


Figure 1.6: The lengths of the blue and red segments represent the biases of players i and j, respectively. Player i is higher biased than player j.

If the receiver is a 0-extremist or a 1-extremist, he will always take action $a = 0$ or $a = 1$, respectively, regardless of the experiments chosen by the players, and thus all the players are indifferent among all experiments.

Suppose the receiver is a conformist who is biased toward $a = 1$, i.e., $0 < \mu_R < p$. Define the following sets of players:

- $A = \{i : i$ is a conformist with $0 < \mu_i < \mu_R\}$
- $B = \{i : i$ is a conformist with $p < \mu_i < 1\}$
- $C = \{i : i$ is a contrarian with $\mu_R < \mu_i < 1\}$
- $D = \{i : i$ is a contrarian with $0 < \mu_i < \mu_R\}$

- $E_0 = \{i : i \text{ is a } 0\text{-extremist}\}$
- $E_1 = \{i : i \text{ is a } 1\text{-extremist}\}$

The next proposition implies that the relative location of two specific players is relevant in determining the equilibrium outcome of the binary-state, binary-action game:

- Let A^* represent the player with the highest bias among those in A: $\mu_{A^*} = \min_{i \in A} \mu_i$.[21] If $A = \emptyset$, let $A^* = R$. Figure 1.7 illustrates an example.

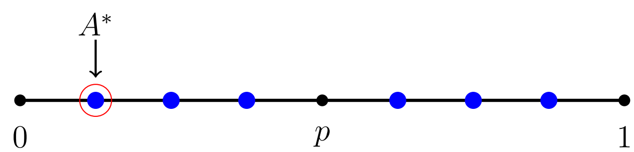


Figure 1.7: The blue dots represent the indifference beliefs of all the conformists in the hierarchy including the receiver. The ones to the right of p correspond to the conformists biased toward $a = 0$ and those to the left of p correspond to the conformists biased toward $a = 1$.

- If $D \cup E_0 \neq \emptyset$, let E^* represent the player closest to the receiver among those in $D \cup E_0$: $E^* = \max_{i \in D \cup E_0} i$. Figure 1.8 illustrates an example.

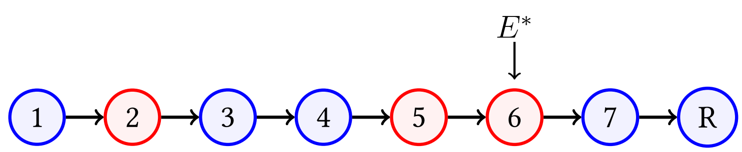


Figure 1.8: The red circles represent players in $D \cup E_0$ of a hierarchy with $n = 7$.

Proposition 4. *Let $\Omega = \mathcal{A} = \{0, 1\}$ and suppose the receiver is a conformist who is biased toward $a = 1$[22] and Assumption 1 holds. If there exists a 1-extremist or a contrarian with $\mu_i > \mu_{A^*}$,[23] the equilibrium is characterized by the no-information outcome.[24] Otherwise,*

[21]Equivalently, A^* is the player with the highest bias toward $a = 1$ among conformists, including the receiver.

[22]$0 < \mu_R < p$.

[23]$C \cup E_1 \neq \emptyset$ or $\exists i \in D$ such that $\mu_i > \mu_{A^*}$.

[24]The no-information outcome is characterized by $supp(\tau^*) \subset (\mu_R, 1]$ if $a^*(\mu_R) = 0$ or $supp(\tau^*) \subset [\mu_R, 1]$ if $a^*(\mu_R) = 1$, which is determined by Assumption 1. However, under the assumption that $\mathcal{S}_n = \mathcal{A}$, the no-information outcome is equivalent to $supp(\tau^*) = p$.

a. *if all the players are conformists,*[25] *the equilibrium outcome* τ^* *is characterized by* $supp(\tau^*) = \{0, 1\}$ *(full-information outcome).*

b. *if there are players who are not conformists and* $A^* < E^*$*, the equilibrium is characterized by the no-information outcome.*

c. *if there are players who are not conformists and* $A^* > E^*$*, the equilibrium outcome* τ^* *is characterized by* $supp(\tau^*) = \{\mu_{A^*}, 1\}$.[26]

In equilibrium, if all the players are conformists, the receiver gets full information. Otherwise,[27] *the location of the pivotal player* A^* *determines whether the receiver gets any information, and if so, her bias* μ_{A^*} *determines the amount of information communicated to the receiver, as shown in Figure 1.9; the higher the bias of* A^**, the more information is communicated.*

Proof. See Appendix A. ■

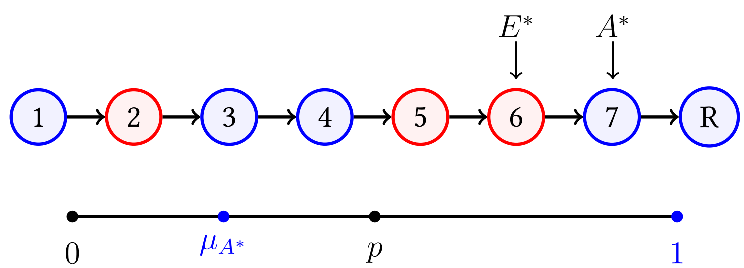


Figure 1.9: A hierarchy with $n = 7$ and its equilibrium distribution of posteriors: The blue dots represent the support of the equilibrium distribution of posteriors.

Another specific player will be of importance in the discussion of the next results: Let D^* represent the player with the lowest bias among those in D with $\mu_i < \mu_{A^*}$: $\mu_{D^*} = \max_{i \in D: \mu_i < \mu_{A^*}} \mu_i$. If $\{i \in D : \mu_i < \mu_{A^*}\} = \emptyset$, let $\mu_{D^*} = 0$.

[25] $D \cup E_0 = \emptyset$.

[26] Note that Assumption 1 is needed here if $A^* = R$; otherwise, there would be no equilibrium.

[27] This is the case except for some rather uninteresting cases resulting in no information for the receiver, as explained in what follows.

The no-information equilibrium outcome could emerge because of three reasons:

1. Existence of players who would provide no information in a single-sender game: This is the case if there exists a 1-extremist or a contrarian with $\mu_i > \mu_R$, i.e., $C \cup E_1 \neq \emptyset$. 1-extremists always prefer action $a = 1$; by providing no information, they make sure that the receiver, who is biased toward action $a = 1$, takes action $a = 1$. If the equilibrium is not the no-information outcome, it means sometimes the receiver takes action $a = 0$. However, the set of beliefs at which a contrarian with $\mu_i > \mu_R$ prefers action $a = 0$ and the set of beliefs at which the receiver takes action $a = 0$ are disjoint, as shown in Figure 1.10. Therefore, no information is provided in equilibrium.

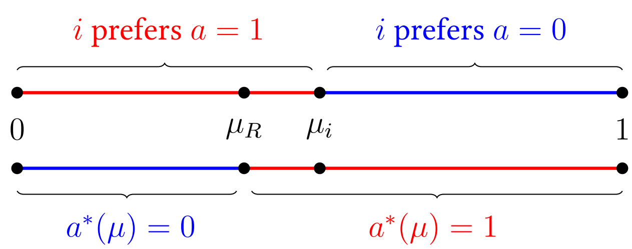


Figure 1.10: The blue segments do not overlap, that is, there does not exist a posterior belief at which both player i and the receiver prefer $a = 0$.

2. Existence of players with highly opposed preferences: Assuming $C \cup E_1 = \emptyset$, this is the case if there exists $i \in D$ such that $\mu_i > \mu_{A^*}$. As mentioned before, if the equilibrium is not the no-information outcome, it means sometimes the receiver takes action $a = 0$. However, the set of beliefs at which a contrarian with $\mu_i > \mu_{A^*}$ prefers action $a = 0$ and the set of beliefs at which A^* prefers action $a = 0$ are disjoint, as shown in Figure 1.11. Therefore, no information is provided in equilibrium.

3. Location of the pivotal player: Assuming $C \cup E_1 = \emptyset$ and that there exists no $i \in D$ such that $\mu_i > \mu_{A^*}$, this is the case if $A^* < E^*$, i.e., E^* is closer to the receiver than is the pivotal player A^*, as shown in Figure 1.12. The reason is that

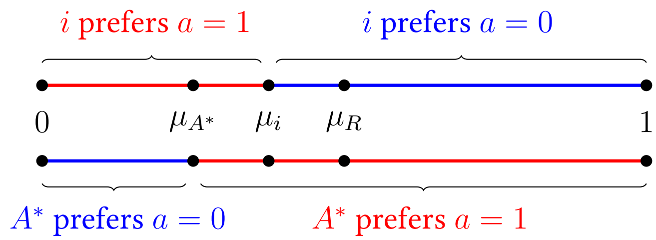


Figure 1.11: The blue segments do not overlap, that is, there does not exist a posterior belief at which both player i and A^* prefer $a = 0$.

when designing her experiment, E^*, whose preferences are somewhat opposed to those of A^*, considers only the incentive compatibility constraints of the succeeding players, not those of the preceding players, such as A^*. Knowing this, A^* preemptively provides no information, because she understands that the presence of E^* closer to the receiver guarantees an outcome in which A^* prefers action $a = 1$. This equilibrium is inefficient, because all the players prefer every outcome τ with $\min\,(supp(\tau)) \in [\mu_{D^*}, \mu_{A^*}]$.

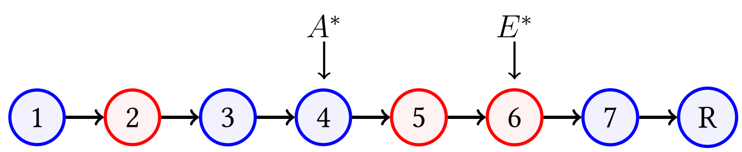


Figure 1.12: The blue circles represent the conformists while the red ones represent players in $\{i \in D : \mu_i < \mu_{A^*}\} \cup E_0$ of a hierarchy with $n = 7$.

The next corollary shows that the configuration 3 above is the only one that causes inefficiency in communication.

Corollary 3. *Let $\Omega = \mathcal{A} = \{0, 1\}$ and suppose the receiver is a conformist who is biased toward $a = 1$ and Assumption 1 holds. The equilibrium is inefficient if and only if the conditions in part (b) Proposition 4 hold.*

One question that may arise is if the receiver, who could be the CEO of a firm or the president of a government, can get more information and thus increase his payoff by assigning a vice president and if so, how. The next corollary shows that the answer is affirmative.

Corollary 4. *Let $\Omega = \mathcal{A} = \{0, 1\}$ and suppose the receiver is a conformist who is biased toward $a = 1$ and Assumption 1 holds. If the receiver could add a player of his choice at the end of the hierarchy, he would choose a conformist biased toward $a = 1$ with higher bias than A^* but lower bias than D^*.*[28] *The receiver's utility would then be increasing in the bias of the added player and the new equilibrium would be efficient. Moreover, the receiver would not benefit from adding more players.*

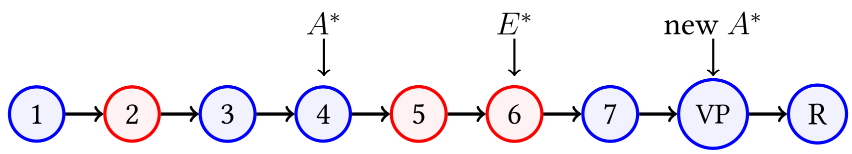


Figure 1.13: Adding a vice president to a hierarchy with $n = 7$

The intuition behind this result is as follows:

- The vice president is the closest player to the receiver. Because players move sequentially, every player is better off by taking into account the preferences of the players closest to the receiver (i.e., those who are higher-ranked). Therefore, all the players are better off taking into account the preferences of the vice president.

- If the receiver follows the prescription in the corollary, the vice president ex ante has similar action preferences to those of the receiver; they are both conformists biased toward $a = 1$. This implies that a player who intends to persuade the receiver to take action $a = 0$ must first persuade the vice president toward action $a = 0$. Otherwise (i.e., if $\min\ supp(\tau) > \mu_{VP}$), the vice president will not pass any information to the receiver, being sure that the receiver will take the action favored by the vice president (i.e., $a = 1$).

[28] $\mu_{D^*} < \mu_i < \mu_{A^*}$.

- By behaving as prescribed in the corollary, the receiver makes the vice president the new pivotal player, as shown in Figure 1.13; the vice president is the conformist with the highest bias toward $a = 1$. This implies that more information is required to persuade the vice president toward action $a = 0$.

- As mentioned before, the higher the bias of the pivotal player, the more information is communicated.

To generalize Proposition 4 to other types of receiver, I call conformist the opposite type of contrarian and vice versa. Similarly, I call bias toward $a = 1$ the opposite of bias toward $a = 0$ and vice versa. Also, when two players are of opposite type and bias, I compare their biases in the same way as if they were of the same type and bias.

Define the following sets of players:

- $A = \{i : i$ is of the same type and bias as the receiver but with higher bias$\}$

- $B = \{i : i$ is of the same type as the receiver but with opposite bias$\}$

- $C = \{i : i$ is of opposite type to the receiver but with the same bias or lower opposite bias$\}$

- $D = \{i : i$ is of opposite type to the receiver but with higher opposite bias$\}$

- $E_0 = \{i : i$ is an extremist of opposite bias to the receiver$\}$

- $E_1 = \{i : i$ is an extremist of the same bias as the receiver$\}$

As in Proposition 4, the relative location of two specific players is relevant in determining the equilibrium outcome of the binary-state, binary-action game:

- Let A^* represent the player with the highest bias among those in A; if $A = \emptyset$, let $A^* = R$.

- If $D \cup E_0 \neq \emptyset$, let E^* represent the player closest to the receiver among those in $D \cup E_0$: $E^* = \max_{i \in D \cup E_0} i$.

Proposition 5. *Let $\Omega = \mathcal{A} = \{0, 1\}$ and suppose the receiver is not an extremist and Assumption 1 holds. If $C \cup E_1 \neq \emptyset$ or if there exists $i \in D$ such that i is less biased than A^*, the equilibrium is characterized by the no-information outcome.*[29] *Otherwise,*

a. if $D \cup E_0 = \emptyset$,[30] *the equilibrium outcome τ^* is characterized by $supp(\tau^*) = \{0, 1\}$ (full-information outcome).*

b. if $D \cup E_0 \neq \emptyset$[31] *and $A^* < E^*$, the equilibrium is characterized by the no-information outcome.*

c. if $D \cup E_0 \neq \emptyset$ and $A^* > E^*$, the equilibrium outcome τ^* is characterized by $supp(\tau^*) = \{\mu_{A^*}, a\}$, where $a = a^*(p)$ if the receiver is a conformist and $a = 1 - a^*(p)$ if he is a contrarian.[32]

In equilibrium, if all the players are of the same type as that of the receiver, he gets full information. Otherwise,[33] *the location of the pivotal player A^* determines whether the receiver gets any information, and if so, her bias μ_{A^*} determines the amount of information communicated to the receiver; the higher the bias of A^*, the more information is communicated.*

Proof. The proof is similar to that of Proposition 4 presented in Appendix A. ■

Another specific player will be of importance in the discussion of the next results: Let D^* represent the player with the lowest bias among those in D with higher bias than A^*. If $\{i \in D : i \text{ has higher bias than } A^*\} = \emptyset$, let $\mu_{D^*} = 1 - a^*(p)$ if the receiver is a conformist and $\mu_{D^*} = a^*(p)$ if he is a contrarian.

[29]The no-information outcome is characterized by (i) $supp(\tau^*) \subset (\mu_R, 1]$ or $supp(\tau^*) \subset [\mu_R, 1]$, or (ii) $supp(\tau^*) \subset [0, \mu_R)$ or $supp(\tau^*) \subset [0, \mu_R]$, whichever includes p; open or closed interval is determined by Assumption 1. However, under the assumption that $\mathcal{S}_n = \mathcal{A}$, the no-information outcome is equivalent to $supp(\tau^*) = p$.

[30]All the players are of the same type as the receiver.

[31]There are players who are not of the same type as the receiver.

[32]Note that Assumption 1 is needed here if $A^* = R$; otherwise, there would be no equilibrium.

[33]This is the case except for some rather uninteresting cases resulting in no information for the receiver, as explained in what follows.

The no-information equilibrium outcome could emerge because of three reasons:

1. Existence of players who would provide no information in a single-sender game: This is the case if there exists an extremist of the same bias as the receiver or a player of the opposite type to that of the receiver but with the same bias or lower opposite bias $\mu_R < \mu_i < 1$, i.e., $C \cup E_1 \neq \emptyset$.

2. Existence of players with highly opposed preferences: Assuming $C \cup E_1 = \emptyset$, this is the case if there exists $i \in D$ such that i is less biased than A^*.

3. Location of the pivotal player: Assuming $C \cup E_1 = \emptyset$ and that there exists no $i \in D$ such that i is more biased than A^*, this is the case if $A^* < E^*$, i.e., E^* is closer to the receiver than is the pivotal player P. This equilibrium is inefficient, because all the players prefer every outcome τ with $\min\left(supp(\tau)\right) \in [\mu_{D^*}, \mu_{A^*}]$ or $\max\left(supp(\tau)\right) \in [\mu_{A^*}, \mu_{D^*}]$, whichever is a well-defined interval. The equilibrium is inefficient if and only if this condition holds.

Corollary 5. *Let $\Omega = \mathcal{A} = \{0, 1\}$ and suppose the receiver is not an extremist and Assumption 1 holds. The equilibrium is inefficient if and only if the conditions in part (b) of Proposition 5 hold.*

Corollary 6. *Let $\Omega = \mathcal{A} = \{0, 1\}$ and suppose the receiver is not an extremist and Assumption 1 holds. If the receiver could add a player of his choice at the end of the hierarchy, he would choose one of the same type and bias as himself but with higher bias than A^* and lower bias than D^*. The receiver's utility would then be increasing in the bias of the added player and the new equilibrium would be efficient. Moreover, the receiver would not benefit from adding more players.*

1.7.2 General Binary-Action Games

Let $\Omega = [0, 1]$. Given the experiment π_1 chosen by player 1, Proposition 5 characterizes the equilibrium outcome of the binary-state, binary-action subgame starting from player

2. By choosing π_1, player 1 not only chooses the state space of the following subgame $supp(\tau_1) = \{q_0, q_1\}$ but also, implicitly, the four numbers representing the utilities of the subsequent players,[34] and thus their types.

In the general binary-action game, for each player, there is a utility function corresponding to each action: $u_i(\omega, 0)$ and $u_i(\omega, 1)$. Let $\Delta u_i(\omega) = u_i(\omega, 1) - u_i(\omega, 0)$ and suppose $\Delta u_i(\omega) = \alpha_i \omega + \beta_i$ for some $\alpha_i, \beta_i \in \mathbb{R}$. For example, it could be the case that $u_i(\omega, 0) = 0$ and $u_i(\omega, 1) = \alpha_i \omega + \beta_i$. Let $\omega_i = -\frac{\beta_i}{\alpha_i}$ represent the state, or, equivalently, the posterior mean, at which i is indifferent between the two actions.

Generally, as mentioned before, player 1 chooses a Bayes-plausible distribution of posteriors $\tau_1 \in \Delta(\Delta(\Omega))$ with $supp(\tau_1) = \{q_0, q_1\}$. From this point on, the game is a binary-state, binary-action one. In the special case where $\Delta u_i(\omega) = \alpha_i \omega + \beta_i$, because action preferences depend only on the posterior mean, I can assume, with some abuse of notation, that player 1 chooses a distribution of posterior means $\tau_1 \in \Delta(\Omega)$,[35] where $supp(\tau_1) = \{m_0, m_1\}$[36], and τ_1 is a mean preserving contraction of F (implying $\mathbb{E}_{\tau_1}[\omega] = \mathbb{E}_F[\omega] = m$).[37] For example, if F is the uniform distribution over $[0, 1]$, a distribution of posterior means τ with $supp(\tau) = \{m_0, m_1\}$ gives a mean preserving contraction of F (with the Bayes-plausible probabilities, i.e., $\tau(m_0)m_0 + (1 - \tau(m_0))m_1 = m$) if and only if $m_1 - m_0 \leq 0.5$.[38]

Generally, the shape of the utility functions u_i and choice of the first player τ_1 determine the type of each subsequent player (conformist, contrarian, or extremist) in the following binary-state, binary-action subgame. In the special case where $\Delta u_i(\omega) = \alpha_i \omega + \beta_i$, all I need to know about player $i = 2, \ldots, n$ is her indifference posterior mean

[34] $u_i^{lk} = \mathbb{E}_{q_l}[u_i(\omega, k)],\ \forall l, k \in \{0, 1\}$

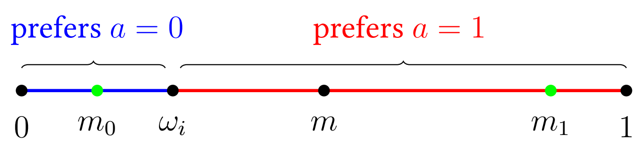
[35] Because τ_1 is a finite-support distribution, τ_1 represents the probability mass function.

[36] Recall that in the binary-state, binary-action games, each player was represented by the posterior belief μ_i (or equivalently, posterior mean) at which she was indifferent between the two actions. In this special case, in the binary-state, binary-action game starting from player 2, $\mu_i = \frac{\omega_i - m_0}{m_1 - m_0}$. Clearly, $\mu_i > \mu_j$ if and only if $\omega_i > \omega_j$.

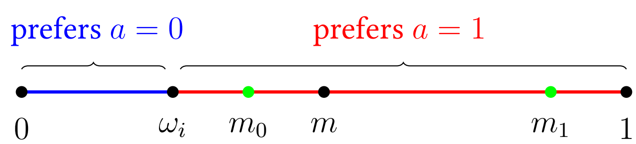
[37] Clearly, $0 \leq m_0 \leq m \leq m_1 \leq 1$. In the special case where $\Delta u_i(\omega) = \alpha_i \omega + \beta_i$, with some abuse of notation, the outcome τ represents a distribution of posterior means.

[38] Note that, generally, if the pair (m_0, m_1) gives a mean preserving contraction of F, any pair (m_0', m_1') also gives a mean preserving contraction of F if and only if $m_0 \leq m_0' \leq m \leq m_1' \leq m_1$.

ω_i[39] and the sign of α_i (or equivalently, her preferred action at $\omega = 1$).[40] If $\omega_i < 0$ or $\omega_i > 1$, player i is called an absolute extremist; that is, regardless of τ_1 chosen by player 1, player i will be an extremist in the subgame starting from player 2. Otherwise, $\alpha_i > 0$ (< 0) implies that player i is a conformist (contrarian). However, given τ_1, and thus, m_0 and m_1, each conformist or contrarian may turn into an extremist in the subgame starting from player 2, as shown in Figure 1.14: this happens if $\omega_i < m_0$ or $\omega_i > m_1$; also, whenever τ_1 turns an a-biased[41] conformist or contrarian into an extremist, she will become an a-extremist. The new extremists are related as follows: if player i with $\omega_i < m$ ($\omega_i > m$) turns into an extremist, all players j with $\omega_j < \omega_i$ ($\omega_j > \omega_i$) turn into extremists as well.



(a) Player i remains a conformist with $\mu_i = \frac{\omega_i - m_0}{m_1 - m_0}$.



(b) Player i turns into a 1-extremist.

Figure 1.14: The distribution of posterior means τ_1 chosen by player 1 with $supp(\tau_1) = \{m_0, m_1\}$ may turn a conformist biased toward $a = 1$ into a 1-extremist in the subgame starting from player 2.

Consider the binary-state, binary-action game starting from player 2, where $\Omega = \{0, 1\}$ and $\mu_i = \omega_i$, for all $i = 2, \ldots, n, R$.[42] I call this game the *reduced* binary-state game corresponding to the general binary-action game.

[39] For simplicity, I assume $\omega_i \notin \{0, m, 1\}$ where $m = \mathbb{E}_F[\omega]$ and $\omega_i \neq \omega_j$ if $i \neq j$.

[40] This is similar to the claim I made in the binary-state, binary-action case, because the subgame starting from player 2 is a binary-state, binary action one.

[41] Bias is defined similar to the binary-state, binary-action case: A player is biased toward action a if she prefers that action given the common prior mean m.

[42] Note that the only extremists of this game are the absolute extremists of the original game.

Lemma 1. *Let $\Delta u_i(\omega) = \alpha_i \omega + \beta_i$ for $i = 2, \ldots, n, R$. If the equilibrium outcome of the reduced binary-state game corresponding to the binary-action game is the no-information outcome, the same holds for the original binary-action game.*[43]

Proof. See Appendix B. ■

In general, the optimal choice of player 1 is not clear based only on $\Delta u_1(\omega)$. To make analysis tractable, let us assume $u_1(\omega, 0) = 0$.[44] Consider the binary-state, binary-action game where $\Omega = \{0, 1\}$ and $\mu_i = \omega_i$, for all $i = 1, \ldots, n, R$. I call this game the binary-state game corresponding to the general binary-action game.

Lemma 2. *Let $\Delta u_i(\omega) = \alpha_i \omega + \beta_i$ for $i = 1, \ldots, n, R$ and $u_1(\omega, 0) = 0$. If the equilibrium outcome of the binary-state game corresponding to the binary-action game is the no-information outcome, the same holds for the original binary-action game*[45].

Proof. See Appendix B. ■

Again, if the receiver is an absolute extremist, he will always take action $a = 0$ or $a = 1$ regardless of the experiments chosen by the players, and thus all the players are indifferent among all experiments.[46]

Suppose receiver is a conformist who is biased toward $a = 1$, i.e., $\alpha_R > 0$ and $0 < \omega_R < m$. Define the sets of players (not including player 1) A, B, C, D, E_0, E_1 as in Proposition 4 by replacing μ_i and μ_R with ω_i and ω_R, respectively. Similarly, define A^* and E^*.

[43]Lemma 1 can be written more generally as follows: If the equilibrium outcome of the reduced binary-state game starting from player $i \geq 2$ is the no-information outcome, the same holds for the binary-action game starting from player $j < i$.

[44]This implies that $u_1(\omega, 1) = \alpha_1 \omega + \beta_1$.

[45]Lemma 2 can be written more generally as follows: If the equilibrium outcome of the binary-state game starting from player i with $u_i(\omega, 0) = 0$ is the no-information outcome, the same holds for the binary-action game starting from player $j \leq i$.

[46]This is clearly true for any binary-action game regardless of the assumptions made in this section.

The next proposition implies that the relative location and relative bias of three specific players in addition to player 1 and E^* are relevant in determining the equilibrium outcome of the binary-action game in the special case where $u_1(\omega, 0) = 0$ and the prior belief F is uniform:

- Let P represent the conformist with the highest bias toward $a = 1$:[47] $\omega_P = \min_{i:\alpha_i>0, 0<\omega_i<1} \omega_i$.
- Let B^* represent the player with the highest bias among those in B: $\omega_{B^*} = \max_{i \in B} \omega_i$. If $B = \emptyset$, let $\omega_{B^*} = m$.
- Let B^*_P represent the player with the highest bias among those in B and closer to the receiver than is P: $\omega_{B^*_P} = \max_{i \in B: i > P} \omega_i$. If $\{i \in B : i > P\} = \emptyset$, let $\omega_{B^*_P} = m$.

Figure 1.15 illustrates an example.

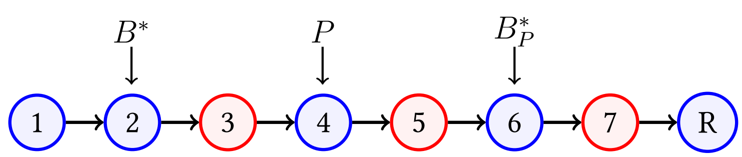


Figure 1.15: The blue circles represent the conformists while the red ones represent extremists and contrarians of a hierarchy with $n = 7$.

Proposition 6. *Let $\Omega = [0, 1]$, $\mathcal{A} = \{0, 1\}$, $\Delta u_i(\omega) = \alpha_i \omega + \beta_i$ for $i = 1, \ldots, n, R$, $u_1(\omega, 0) = 0$ and the prior belief F be the uniform distribution over $[0, 1]$. Suppose the receiver is a conformist who is biased toward $a = 1$,[48] Assumption 1 holds, and the no-information outcome is not the equilibrium outcome of the corresponding binary-state game.*

[47] This implies $P = 1$ or $P = A^*$.
[48] $\alpha_R > 0$ and $0 < \omega_R < m$.

The equilibrium is characterized by the no-information outcome[49] *unless* $\omega_{B_P^*} - \omega_P \leq 0.5$;[50] *in this case,*

a. *if there are players who are not conformists,*[51] $supp(\tau^*) = \{\omega_P, \omega_P + 0.5\}$*;*

b. *if all the players are conformists and* $\omega_{B^*} - \omega_P > 0.5$, $supp(\tau^*) = \{\omega_P, \omega_P + 0.5\}$*;*

c. *if all the players are conformists and* $\omega_{B^*} - \omega_P \leq 0.5$,[52] $supp(\tau^*) = \{m_0, m_0 + 0.5\}$, *where* $m_0 = \min\{\omega_P, \max(\omega_{B^*} - 0.5, 0.5\omega_1)\}$.[53]

Assuming the equilibrium outcome of the corresponding binary-state game is not the no-information outcome, the receiver gets some information if and only if $\omega_{B_P^*} - \omega_P \leq 0.5$, *i.e., there is not a conformist whose preferences are somewhat opposed to those of the pivotal player* P *and is closer to the receiver than is the pivotal player* P*. In this case, the biases of player 1 and of the pivotal players* P *and* B^* *determine the information communicated to the receiver, as shown in Figure 1.16. Every equilibrium is outcome-equivalent to one which simply distinguishes between the two intervals* $[0, x]$ *and* $[x, 1]$ *for some* $x \in [0, \omega_P]$.

Proof. I first state two lemmas which are proved in Appendix B:

Lemma 3. *Let* $\Omega = [0,1]$, $\mathcal{A} = \{0,1\}$ *and* $\Delta u_i(\omega) = \alpha_i \omega + \beta_i$ *for* $i = 2, \ldots, n, R$*. Suppose the receiver is a conformist who is biased toward* $a = 1$*, Assumptions 1 holds, and the no-information outcome is not the equilibrium outcome of the reduced binary-state game.*[54] *Given player 1's choice of distribution of posterior means* τ_1 *with* $supp(\tau_1) = \{m_0, m_1\}$,

[49] The no information outcome is characterized by $supp(\tau^*) \subset (\omega_R, 1]$ if $a^*(\omega_R) = 0$ or $supp(\tau^*) \subset [\omega_R, 1]$ if $a^*(\omega_R) = 1$, which is determined by Assumption 1. However, under the assumption that $\mathcal{S}_n = \mathcal{A}$, the no-information outcome is equivalent to $supp(\tau^*) = m$.

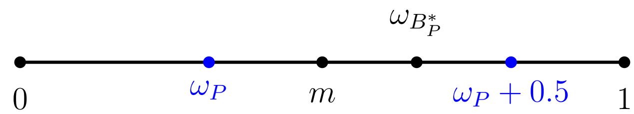
[50] That is, there would be no new 0-extremists closer to the receiver than is P if information is provided: $\omega_i \leq \omega_P + 0.5$ for all players $i > P$.

[51] That is, there are players who are absolute 0-extremists or contrarians with $0 < \omega_i < \omega_P$.

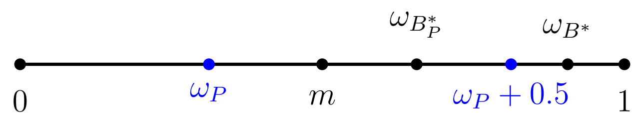
[52] That is, there would be no new 0-extremists if information is provided: $\omega_i \leq \omega_P + 0.5$ for all the players.

[53] Equivalently, $m_0 = \max\{\omega_{B^*} - 0.5, \min(\omega_P, 0.5\omega_1)\}$.

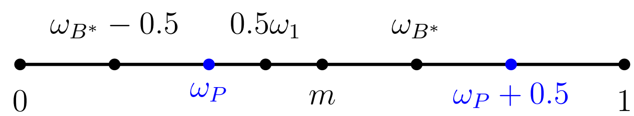
[54] This implies that (not including player 1) there exists no absolute 1-extremist or contrarian with $\omega_{A^*} < \omega_i < 1$, and A^* is closer to the receiver than are all absolute 0-extremists and contrarians with $0 < \omega_i < \omega_{A^*}$.



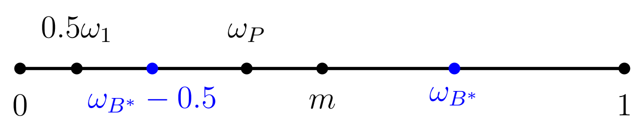
(a) There are players who are not conformists.



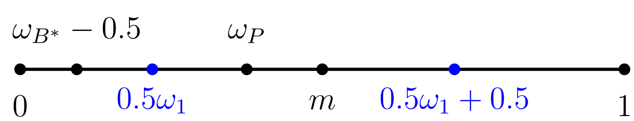
(b) All players are conformists and $\omega_{B^*} - \omega_P > 0.5$.



(c) All players are conformists, $\omega_{B^*} - \omega_P \leq 0.5$, and $0.5\omega_1 > \omega_P$.



(d) All players are conformists, $\omega_{B^*} - \omega_P \leq 0.5$, and $0.5\omega_1 < \omega_{B^*} - 0.5$.



(e) All players are conformists, $\omega_{B^*} - \omega_P \leq 0.5$, and $\omega_{B^*} - 0.5 < 0.5\omega_1 < \omega_P$.

Figure 1.16: The equilibrium outcome if $\omega_{B_P^*} - \omega_P \leq 0.5$: The blue dots represent the support of the equilibrium distribution of posteriors.

the equilibrium of the following subgame starting from player 2 is characterized by the no-information outcome unless:

a. *(i) all the remaining players are conformists,*[55] *(ii)* $m_0 \leq \omega_{A^*}$ *and* $m_1 \geq \omega_{B^*}$*, in which case, the equilibrium outcome* τ^* *of the following subgame starting from player 2 is*

[55] $C \cup D \cup E_0 \cup E_1 = \emptyset$.

characterized by $supp(\tau^*) = \{m_0, m_1\}$ (full-information outcome);[56]

b. *(i) all the remaining players are conformists but $m_1 < \omega_{B^*}$, or there are players (not including player 1) who are not conformists, (ii) $m_0 \leq \omega_{A^*}$, (iii) A^* is closer to the receiver than are all the new 0-extremists, i.e., conformists with $\omega_i > m_1$, in which case, the equilibrium outcome τ^* of the following subgame starting from player 2 is characterized by $supp(\tau^*) = \{\mu_{A^*}m_1 + (1 - \mu_{A^*})m_0 = \omega_{A^*}, m_1\}$.*[57]

Lemma 4. *Let $\Omega = [0, 1]$, $\mathcal{A} = \{0, 1\}$, $\Delta u_i(\omega) = \alpha_i\omega + \beta_i$ for $i = 1, \ldots, n, R$ and $u_1(\omega, 0) = 0$. Suppose the receiver is a conformist who is biased toward $a = 1$. The preferences of different types of player 1 over outcomes τ with $supp(\tau) = \{m_0, m_1\}$ are as follows:*

- *If player 1 is an absolute 1-extremist or a contrarian ($\alpha_1 < 0$) with $\omega_R < \omega_1 < 1$, she would prefer the no-information outcome to all other outcomes.*[58]
- *If player 1 is an absolute 0-extremist, she would prefer outcomes with higher m_1 and higher m_0 as long as the outcome is not the no-information outcome.*
- *If player 1 is a contrarian ($\alpha_1 < 0$) with $0 < \omega_1 < \omega_R$, she would prefer outcomes with higher m_0 and m_1, as long as $m_0 \geq \omega_1$ and the outcome is not the no-information outcome.*[59] *She would also prefer the no-information outcome to all outcomes with $m_0 < \omega_1$.*

[56] Note that if $A^* \neq R$ and $m_0 = \omega_{A^*}$, another possible equilibrium is the no-information outcome. Similarly, if $m_1 = \omega_{B^*}$, in other possible equilibria, player 1's choice of distribution of posterior means (m_0, m_1) would reach the receiver as (m', ω_{B^*}), where $m_0 \leq m' \leq \omega_{A^*}$. However, such outcomes will be ruled out as equilibrium in the original game starting from player 1.

[57] Note that Assumption 1 is needed here if $A^* = R$; otherwise, there would be no equilibrium.

[58] Thus, the equilibrium outcome τ^* is characterized by $supp(\tau^*) \subset (\mu_R, 1]$ if $a^*(\mu_R) = 0$ or $supp(\tau^*) \subset [\mu_R, 1]$ if $a^*(\mu_R) = 1$, which is determined by Assumption 1 (*no-information outcome*). However, under the assumption that $\mathcal{S}_n = \mathcal{A}$, the no-information outcome is equivalent to $supp(\tau^*) = m$. Note that $u_1(\omega, 0) = 0$ is not necessary for this part.

[59] In this case, if $m_0 = \omega_1$, she would be indifferent about m_1.

- *If player 1 is a conformist with* $0 < \omega_1 < \omega_R$*, she would prefer outcomes with lower* m_0 *and higher* m_1 *as long as* $m_0 \leq \omega_1$*.*[60] *She would also prefer the no-information outcome only to outcomes with* $m_0 > \omega_1$*.*

- *If player 1 is a conformist (*$\alpha_1 > 0$*) with* $\omega_R < \omega_1 < m$*, she would prefer lower* m_0 *and higher* m_1 *as long as the outcome is not the no-information outcome.*

- *If player 1 is a conformist (*$\alpha_1 > 0$*) with* $m < \omega_1 < 1$*, she would prefer higher* m_1 *and (i) lower* m_0 *if* $m_1 \geq \omega_1$*, (ii) higher* m_0 *if* $m_1 < \omega_1$*, as long as the outcome is not the no-information outcome.*[61]

Now, I can prove Proposition 6. Note that there are no absolute 1-extremists or contrarians with $\omega_P < \omega_i < 1$ and $P > E^*$. As mentioned before, if F is the uniform distribution over $[0, 1]$, a distribution of posterior means τ with $supp(\tau) = \{m_0, m_1\}$ gives a mean preserving contraction of F (with the Bayes-plausible probabilities, i.e., $\tau(m_0)m_0 + (1 - \tau(m_0))m_1 = m$) if and only if $m_1 - m_0 \leq 0.5$.

Suppose $\omega_{B_P^*} - \omega_P > 0.5$.

- If $P = A^* = R$, every possible choice of m_0 and m_1 by player 1 either is the no-information outcome or turns B_P^* into a 0-extremist, which leads to no-information outcome, by Lemma 3.

- If $P = 1$, every possible choice of m_0 and m_1 by player 1 with $m_0 \leq \omega_1$ turns B_P^* into a 0-extremist, which leads to the no-information outcome, by Lemma 3. Moreover, by Lemma 4, player 1 would prefer the no-information outcome to every possible choice of m_0 and m_1 with $m_0 > \omega_1$, which is not the no-information outcome. Therefore, the only possible equilibrium outcome is the no-information outcome.

[60] In this case, if $m_0 = \omega_1$, she would be indifferent about m_1.

[61] In this case, if $m_1 = \omega_1$, she would be indifferent among all m_0 as long as the outcome is not the no-information outcome.

- If $P = A^* \neq 1$, every possible choice of m_0 and m_1 by player 1 either turns P into a 1-extremist or turns B_P^* into a 0-extremist, both of which lead to the no-information outcome, by Lemma 3.

Suppose $\omega_{B_P^*} - \omega_P \leq 0.5$. Before proceeding, note that given any outcome (other than the no-information outcome) with $m_1 = m_0 + 0.5$, the expected utility of player 1 is given by

$$\begin{aligned}\mathbb{E}_\tau[u_1(\omega, a^*(\omega)] &= \frac{m - m_0}{m_1 - m_0}(\alpha_1 m_1 + \beta_1) \\ &= 2(0.5 - m_0)(\alpha_1 m_0 + 0.5\alpha_1 + \beta_1)\end{aligned}$$

where I used the fact that $\mathbb{E}_F[\omega] = m$. Taking the derivative

$$\frac{\partial \mathbb{E}_\tau[u_1(\omega, a^*(\omega)]}{\partial m_0} = -2\alpha_1 m_0 - \beta_1$$

shows that it is equal to zero at $m_0 = \frac{-\beta_1}{2\alpha_1} = 0.5\omega_1$, where the expected utility of player 1 is maximized if she is a conformist ($\alpha_1 > 0$) and minimized if she is a contrarian ($\alpha_1 < 0$). Moreover, the expected utility of player 1 is increasing in m_0 if she is a 0-extremist.

- Suppose there are players who are not conformists, that is, there are players who are absolute 0-extremists or contrarians with $0 < \omega_i < \omega_P$. Note that $P \neq 1$ because otherwise, $P < E^*$ and the no-information outcome is the equilibrium outcome of the corresponding binary-state game.

  Suppose the only non-conformist is player 1. By Lemma 3, every possible choice of m_0 and m_1 by player 1 with $m_0 \leq \omega_P$ and $m_1 \geq \omega_{B^*}$ leads to the outcome τ with $supp(\tau) = \{m_0, m_1\}$; since by Lemma 4, player 1 would prefer outcomes with higher m_0, she chooses $m_0 = \omega_P$. Similarly, every possible choice of m_0 and m_1 by player 1 with $m_0 \leq \omega_P$ and $\omega_{B_P^*} \leq m_1 < \omega_{B^*}$ leads to the outcome τ with $supp(\tau) = \{\omega_P, m_1\}$; all other possible choices of m_0 and m_1 by player 1 lead to the no-information outcome. Since by Lemma 4, player 1 would prefer outcomes with higher m_1, she chooses $m_1 = m_0 + 0.5$. Therefore, the equilibrium outcome would be $m_0 = \omega_P$ and $m_1 = \omega_P + 0.5$ since $\omega_P + 0.5 \geq \omega_{B_P^*}$.

Suppose there are non-conformists among the players other than player 1. By Lemma 3, every possible choice of m_0 and m_1 by player 1 with $m_0 \leq \omega_P$ and $m_1 \geq \omega_{B_P^*}$ leads to the outcome τ with $supp(\tau) = \{\omega_P, m_1\}$; all other choices of m_0 and m_1 by player 1 lead to the no-information outcome. Since by Lemma 4, player 1 would prefer outcomes with higher m_1, she chooses $m_1 = m_0 + 0.5$. Therefore, the equilibrium outcome would be $m_0 = \omega_P$ and $m_1 = \omega_P + 0.5$ since $\omega_P + 0.5 \geq \omega_{B_P^*}$.

- Suppose all the players are conformists and $\omega_{B^*} - \omega_P > 0.5$. Every possible choice of m_0 and m_1 by player 1 with $m_0 \leq \omega_P$ turns B^* into a 0-extremist, and by Lemma 3, leads to the outcome τ with $supp(\tau) = \{\omega_P, m_1\}$ as long as $m_1 \geq \omega_{B_P^*}$; all other possible choices of m_0 and m_1 by player 1 lead to the no-information outcome. Since by Lemma 4, player 1 would prefer outcomes with higher m_1, she chooses $m_1 = m_0 + 0.5$. Therefore, the equilibrium outcome would be $m_0 = \omega_P$ and $m_1 = \omega_P + 0.5$, since $\omega_P + 0.5 \geq \omega_{B_P^*}$.

- Suppose all the players are conformists and $\omega_{B^*} - \omega_P \leq 0.5$. By Lemma 3, every possible choice of m_0 and m_1 by player 1 with $m_0 \leq \omega_P$ and $m_1 \geq \omega_{B^*}$ leads to the outcome τ with $supp(\tau) = \{m_0, m_1\}$. Similarly, every possible choice of m_0 and m_1 by player 1 with $m_0 \leq \omega_P$ and $\omega_{B_P^*} \leq m_1 < \omega_{B^*}$ leads to the outcome τ with $supp(\tau) = \{\omega_P, m_1\}$; all other possible choices of m_0 and m_1 by player 1 lead to the no-information outcome. Since by Lemma 4, player 1 would prefer outcomes with higher m_1, she chooses $m_1 = m_0 + 0.5$. Her expected utility is maximized at $m_0 = 0.5\omega_1$ if $0.5\omega_1 \leq \omega_P$ and $0.5\omega_1 + 0.5 \geq \omega_{B^*}$; if the first one is violated, she chooses $m_0 = \omega_P$, and if the second one is violated, she chooses $m_0 = \omega_{B^*} - 0.5$ since her expected utility is decreasing in m_0 in the interval $[0.5\omega_1, \omega_P]$. Therefore, the equilibrium outcome would be $m_0 = \min\{\omega_P, \max(\omega_{B^*} - 0.5, 0.5\omega_1)\}$ and $m_1 = m_0 + 0.5$.

■

Another specific player will be of importance in the next results: Let D^{**} represent the player with the lowest bias among those in D with $\omega_i < \omega_P$: $\omega_{D^{**}} = \max_{i \in D: \omega_i < \omega_P} \omega_i$. If $\{i \in D : \omega_i < \omega_P\} = \emptyset$, let $\omega_{D^{**}} = 0$. If player 1 is a contrarian with $0 < \omega_1 < \omega_P$, let $\omega_{D^*} = \max(\omega_{D^{**}}, \omega_1)$; otherwise, let $\omega_{D^*} = \omega_{D^{**}}$.

The no-information equilibrium outcome could emerge because of two main reasons:

1. *Condition 1*: The no-information outcome is the equilibrium outcome of the corresponding binary-state game.

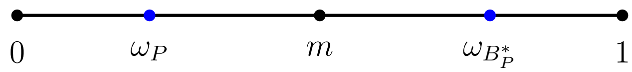
   (a) Existence of players who would provide no information in a single-sender game: This is the case if there exists an absolute 1-extremist or a contrarian ($\alpha_1 < 0$) with $\omega_R < \omega_i < 1$.

   (b) Existence of players with highly opposed preferences: If condition (a) does not hold, this is the case if there exists a contrarian ($\alpha_1 < 0$) with $\omega_P < \omega_i < \omega_R$.

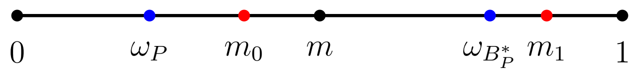
   (c) Location of the pivotal player P: If conditions (a) and (b) do not hold, this is the case if $P < E^*$, i.e., E^* is closer to the receiver than is the pivotal player P. This equilibrium is inefficient since all the players prefer every outcome τ with $\min(supp(\tau)) \in [\omega_{D^*}, \omega_P]$.

2. *Condition 2*: Location and bias of the pivotal player P: If condition (1) does not hold, this is the case if $\omega_{B_P^*} - \omega_P > 0.5$, i.e., there exists a conformist whose preferences are somewhat opposed to those of the pivotal player P and is closer to the receiver as shown in Figure 1.17. This equilibrium is inefficient since all players prefer every outcome τ with $\min(supp(\tau)) \in [\omega_{D^*}, \omega_P]$.

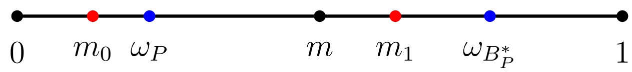
Condition 1(c) and Condition 2 are similar: There exists a player with somewhat opposed preferences to those of the pivotal player P who is closer to the receiver than is the pivotal player P. The next corollary shows that this configuration is the only one which causes inefficiency in communication.



(a) If $\omega_{B_P^*} - \omega_P > 0.5$, either $m_0 > \omega_P$ or $m_1 < \omega_{B_P^*}$ for every choice of m_0 and m_1 such that $m_1 - m_0 \leq 0.5$.



(b) If $m_0 > \omega_P$, P turns into a 1-extremist.



(c) If $m_1 < \omega_{B_P^*}$, B_P^* turns into a 0-extremist closer to the receiver than P.

Figure 1.17: If $\omega_{B_P^*} - \omega_P > 0.5$, no-information equilibrium outcome emerges.

Corollary 7. *Let $\Omega = [0,1]$, $\mathcal{A} = \{0,1\}$, $\Delta u_i(\omega) = \alpha_i \omega + \beta_i$ for $i = 1, \ldots, n, R$, $u_1(\omega, 0) = 0$ and the prior belief F be the uniform distribution over $[0,1]$. Suppose the receiver is a conformist who is biased toward $a = 1$ and Assumption 1 holds. The equilibrium is inefficient if and only if Condition 1(c) or Condition 2 holds.*

Similar to the binary-state case, as the next corollary shows, the receiver can get more information and thus increase her payoff by assigning a suitable vice-president.

Corollary 8. *Let $\Omega = [0,1]$, $\mathcal{A} = \{0,1\}$, $\Delta u_i(\omega) = \alpha_i \omega + \beta_i$ for $i = 1, \ldots, n, R$, $u_1(\omega, 0) = 0$ and the prior belief F be the uniform distribution over $[0,1]$. Suppose the receiver is a conformist who is biased toward $a = 1$ and Assumption 1 holds. If the receiver could add a player of his choice at the end of the hierarchy, depending on the hierarchy configuration, he would choose either a conformist biased toward $a = 1$ with $\omega_{n+1} = \min\{\omega_P, \max(0.5\omega_R, \omega_{D^*})\}$,[62] or a conformist biased toward $a = 0$ with $\omega_{n+1} = \min(0.5\omega_R, \omega_P) + 0.5$. The new equilibrium would be efficient.*[63]

Notice that from the receiver's point of view, $supp(\tau) = \{0.5\omega_R, 0.5\omega_R + 0.5\}$ is

[62] Equivalently, $\omega_{n+1} = \max\{\omega_{D^*}, \min(0.5\omega_R, \omega_P)\}$.

[63] To determine the new equilibrium outcome using Proposition 6, if there are two players with indifference belief ω_P, I consider the closest one to the receiver as the pivotal player P.

equivalent to the full-information outcome. In the following discussion, I assume $\omega_{D^*} = 0$, for simplicity.

If the players have sufficiently opposed preferences,[64], based on Proposition 6, the receiver can try to achieve full information by choosing a vice president with $\omega_{n+1} = 0.5\omega_R$ who is the pivotal player P in the new longer hierarchy. However, this is possible only if $\omega_P \geq 0.5\omega_R$; otherwise, the receiver's choice does not change the pivotal player P in the new longer hierarchy. If $\omega_P < 0.5\omega_R$, the receiver can still avoid possible inefficiencies by choosing a vice president with $\omega_{n+1} = \omega_P$.

However, if all the players have somewhat similar preferences,[65] based on Proposition 6, the receiver can try to achieve full information in two different ways:

- The receiver can choose a vice president with $\omega_{n+1} = 0.5\omega_R$ who is the pivotal player P in the new longer hierarchy. However, this is effective only if $\min\left(supp(\tau^*)\right) \geq 0.5\omega_R$. Otherwise, either $\omega_P < 0.5\omega_R$ and the receiver's choice does not change the pivotal player P in the new longer hierarchy, or $\max\left(\omega_{B^*} - 0.5, 0.5\omega_1\right) < 0.5\omega_R \leq \omega_P$ and the receiver's choice does not change the outcome.

- The receiver can choose a vice president with $\omega_{n+1} = \min\left(0.5\omega_R, \omega_P\right) + 0.5$ who is the pivotal player B^* in the new longer hierarchy. However, this is effective only if $\min\left(supp(\tau^*)\right) = \max\left(\omega_{B^*} - 0.5, 0.5\omega_1\right) < 0.5\omega_R$. Otherwise, either $\omega_{B^*} - 0.5 > 0.5\omega_R$ and the receiver's choice does not change the pivotal player B^* in the new longer hierarchy, or $\omega_{B^*} - 0.5 \leq 0.5\omega_R \leq 0.5\omega_1$ and the receiver's choice does not change the outcome.

Corollary 9. *Let $\Omega = [0,1]$, $\mathcal{A} = \{0,1\}$, $\Delta u_i(\omega) = \alpha_i\omega + \beta_i$ for $i = 1,\ldots,n,R$, $u_1(\omega, 0) = 0$ and the prior belief F be the uniform distribution over $[0,1]$. Suppose the receiver is a conformist who is biased toward $a = 1$ and Assumption 1 holds. If the receiver could add two players of his choice at the end of the hierarchy, he would choose a conformist biased toward $a = 1$ with $\omega_{n+1} = \min\left\{\omega_P, \max\left(0.5\omega_R, \omega_{D^*}\right)\right\}$ and a conformist biased*

[64]That is, either all are not conformists or if they are conformists, $\omega_{B^*} - \omega_P > 0.5$.

[65]That is, all are conformists and $\omega_{B^*} - \omega_P \leq 0.5$.

toward $a = 0$ *with* $\omega_{n+1} = \min(0.5\omega_R, \omega_P) + 0.5$. *The order does not matter and the new equilibrium would be efficient. Moreover, the receiver would not benefit from adding more players.*

Similar to the binary-state case, Proposition 6, as well as Corollary 8 and Corollary 9, can be generalized to other types of the receiver.

1.8 Conclusion

In this paper, I have investigated the outcome of intermediated communication in hierarchical organizations through the lens of Bayesian persuasion. I have shown that a version of the revelation principle holds, that is, it is without loss of generality to assume the only sender who may conceal information along the hierarchy is the initial one. I then used this simplification to show that hierarchical Bayesian persuasion is equivalent to single-sender Bayesian persuasion between the initial sender and the final receiver subject to recursively-defined incentive compatibility constraints dictated by the preferences of the intermediaries.

Applying these results to the case that the decision maker faces a binary decision underscores the importance of vice presidents in hierarchical organizations. I have shown that, in such cases, regardless of the number of intermediaries, a few of them are relevant in determining the outcome.

Given the above result, an interesting question that could be addressed in future research is whether it is possible to pinpoint a few relevant intermediaries if the decision is not binary. Another question to be addressed is the effect of private information on the outcome.

Moreover, this paper can be considered as a first step toward analyzing Bayesian persuasion in networks such as social media. The main difficulty of such an extension lies in analyzing the optimal experiment of a sender who is not aware of the prior of the receiving end.

Bibliography

ALONSO, R., W. DESSEIN, AND N. MATOUSCHEK (2008): "When does coordination require centralization?" *American Economic Review*, 98, 145–179.

AMBRUS, A., E. M. AZEVEDO, AND Y. KAMADA (2013): "Hierarchical cheap talk," *Theoretical Economics*, 8, 233–261.

ARIELI, I., Y. BABICHENKO, AND F. SANDOMIRSKIY (2022): "Bayesian persuasion with mediators," *arXiv preprint arXiv:2203.04285*.

AUMANN, R. J., M. MASCHLER, AND R. E. STEARNS (1995): *Repeated games with incomplete information*, MIT Press.

BERGEMANN, D. AND S. MORRIS (2016a): "Bayes correlated equilibrium and the comparison of information structures in games," *Theoretical Economic*, 11, 487–522.

——— (2016b): "Information design, Bayesian persuasion, and Bayes correlated equilibrium," *American Economic Review*, 106, 586–591.

——— (2019): "Information design: A unified perspective," *Journal of Economic Literature*, 57, 44–95.

CRAWFORD, V. P. AND J. SOBEL (1982): "Strategic information transmission," *Econometrica: Journal of the Econometric Society*, 1431–1451.

DESSEIN, W. (2002): "Authority and communication in organizations," *The Review of Economic Studies*, 69, 811–838.

Gentzkow, M. and E. Kamenica (2017): "Bayesian persuasion with multiple senders and rich signal spaces," *Games and Economic Behavior*, 104, 411–429.

Kamenica, E. (2019): "Bayesian persuasion and information design," *Annual Review of Economics*, 11, 249–272.

Kamenica, E. and M. Gentzkow (2011): "Bayesian persuasion," *American Economics Review*, 101, 2590–2615.

——— (2016): "Competition in persuasion," *The Review of Economic Studies*, 84, 300–322.

Kuang, Z., J. W. Lien, and J. Zheng (2019): "Hierarchical Bayesian persuasion," *Technical report.*

Laclau, M., L. Renou, and X. Venel (2020): "Robust communication on networks," *arXiv preprint arXiv:2007.00457.*

Li, F. and P. Norman (2018): "On Bayesian persuasion with multiple senders," *Economics Letters*, 170, 66–70.

——— (2021): "Sequential persuasion," *Theoretical Economics*, 16, 639–675.

Mathevet, L. and I. Taneva (2020): "Organized information transmission," *Working paper.*

Rantakari, H. (2008): "Governing adaptation," *The Review of Economic Studies*, 75, 1257–1285.

Wu, W. (2021): "Sequential Bayesian persuasion," *Working paper.*

1.9 Appendix A: Binary-State, Binary-Action Games

Bayes-plausibility implies that for any outcome of the game, or equivalently, any distribution of posteriors for the receiver τ with $supp(\tau) = \{q_0, q_1\}$, $\mathbb{E}_\tau[q] = p$. Without loss of generality, let $q_0 \leq q_1$ which implies $0 \leq q_0 \leq p \leq q_1 \leq 1$, $\tau(q_1) = \frac{p-q_0}{q_1-q_0}$, and $\tau(q_0) = \frac{q_1-p}{q_1-q_0}$. Note that in the binary-state, binary-action case, being smaller in the convex order boils down to being a mean preserving contraction. The set of binary-support distributions of posteriors which are mean preserving contractions of τ includes all τ' with $supp(\tau') = \{q_0', q_1'\}$ such that $q_0 \leq q_0' \leq p \leq q_1' \leq q_1$ and $\mathbb{E}_{\tau'}[q'] = p$. In other words, τ is more Blackwell-informative, the lower q_0 is and the higher q_1 is, as shown in Figure 1.18.

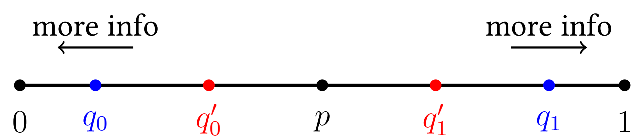


Figure 1.18: Comparison of outcomes

Suppose the receiver is a conformist who is biased toward $a = 1$. Also, for ease of exposition, suppose receiver takes $a = 1$ at $\mu = \mu_R$. Given an outcome τ with $q_0 < \mu_R$, the expected utility of a player is given by

$$\mathbb{E}_\tau[v(q)] = \frac{p-q_0}{q_1-q_0}[q_1 u^{11} + (1-q_1)u^{01}] + \frac{q_1-p}{q_1-q_0}[q_0 u^{10} + (1-q_0)u^{00}]. \tag{1.19}$$

Moreover, given any outcome τ with $q_0 \geq \mu_R$, the expected utility of a player is simply given by

$$\begin{aligned}\mathbb{E}_\tau[v(q)] &= \frac{p-q_0}{q_1-q_0}[q_1 u^{11} + (1-q_1)u^{01}] + \frac{q_1-p}{q_1-q_0}[q_0 u^{11} + (1-q_0)u^{01}] \\ &= \left(\frac{p-q_0}{q_1-q_0}q_1 + \frac{q_1-p}{q_1-q_0}q_0\right)u^{11} + \left(\frac{p-q_0}{q_1-q_0}(1-q_1) + \frac{q_1-p}{q_1-q_0}(1-q_0)\right)u^{01} \\ &= pu^{11} + (1-p)u^{01},\end{aligned} \tag{1.20}$$

which is the same as the expected utility of a player given the outcome τ with $supp(\tau) = \{p\}$, where no information is communicated to the receiver. This makes sense since the receiver would always take action $a = 1$ if $q_0 \geq \mu_R$, and thus I call all such outcomes the no-information outcome.

1. Consider a 1-extremist. She would prefer to persuade the receiver to take action $a = 1$ as often as possible. Therefore she would prefer any outcome τ with $q_0 \geq \mu_R$, i.e., the no-information outcome, to all other outcomes[66].

2. Consider a 0-extremist. She would prefer to persuade the receiver to take action $a = 0$ as often as possible. The receiver would only take action $a = 0$ if his posterior belief is q_0 and $q_0 < \mu_R$. Since $\frac{\partial \tau(q_0)}{\partial q_0} > 0$ and $\frac{\partial \tau(q_0)}{\partial q_1} > 0$, she would prefer outcomes with higher q_1 and higher q_0 as long as $q_0 < \mu_R$. Clearly, her least preferred outcomes are those with $q_0 \geq \mu_R$, i.e., the no-information outcome.

3. Consider a non-extremist. She would prefer any outcome τ with $q_0 \geq \mu_R$ to another outcome with $q_0 < \mu_R$ if and only if she prefers action $a = 1$ at q_0, as shown below:

$$\begin{aligned} pu^{11} + (1-p)u^{01} &> \frac{p - q_0}{q_1 - q_0}[q_1 u^{11} + (1-q_1)u^{01}] + \frac{q_1 - p}{q_1 - q_0}[q_0 u^{10} + (1-q_0)u^{00}] \\ q_0(q_1 - p)(u^{11} - u^{10}) &> (1-q_0)(q_1 - p)(u^{00} - u^{01}) \\ (1-q_0)u^{01} + q_0 u^{11} &> (1-q_0)u^{00} + q_0 u^{10}. \end{aligned} \tag{1.21}$$

[66] Her preferences over the outcomes τ with $q_0 < \mu_R$ do not matter since she prefers the no-information outcome to all such outcomes.

To see her preferences among outcomes τ with $q_0 < \mu_R$, note that

$$\begin{aligned}\frac{\partial \mathbb{E}_\tau[v(q)]}{\partial q_0} &= -\frac{q_1 - p}{(q_1 - q_0)^2}[q_1 u^{11} + (1-q_1)u^{01} - q_0 u^{10} - (1-q_0)u^{00}] \\ &\quad + \frac{q_1 - p}{q_1 - q_0}(u^{10} - u^{00}) \\ &= -\frac{q_1 - p}{q_1 - q_0}(u^{00} - u^{01}) + \frac{q_1 - p}{(q_1 - q_0)^2}[q_1(u^{11} - u^{10}) \\ &\quad - (1-q_0)(u^{00} - u^{01})] \\ &= -\frac{q_1 - p}{(q_1 - q_0)^2}[(1-q_1)u^{01} + q_1 u^{11} - (1-q_1)u^{00} - q_1 u^{10}],\end{aligned} \tag{1.22}$$

which is negative if and only if she prefers action $a = 1$ at q_1. Similarly, note that

$$\begin{aligned}\frac{\partial \mathbb{E}_\tau[v(q)]}{\partial q_1} &= -\frac{p - q_0}{(q_1 - q_0)^2}[q_1 u^{11} + (1-q_1)u^{01} - q_0 u^{10} - (1-q_0)u^{00}] \\ &\quad + \frac{p - q_0}{q_1 - q_0}(u^{10} - u^{00}) \\ &= \frac{p - q_0}{q_1 - q_0}(u^{11} - u^{10}) - \frac{p - q_0}{(q_1 - q_0)^2}[q_1(u^{11} - u^{10}) - (1-q_0)(u^{00} - u^{01})] \\ &= \frac{p - q_0}{(q_1 - q_0)^2}[(1-q_0)u^{00} + q_0 u^{10} - (1-q_0)u^{01} - q_0 u^{11}],\end{aligned} \tag{1.23}$$

which is positive if and only if she prefers action $a = 0$ at q_0. The above observations imply the following:

(a) If player i is a conformist biased toward $a = 1$, she would prefer any outcome τ with $q_0 \geq \mu_R$ only to those with $\mu_i < q_0 < \mu_R$ since she prefers action $a = 1$ at q_0 only if $q_0 > \mu_i$[67]. Among outcomes τ with $q_0 \leq \mu_i$, she prefers ones with lower q_0 and higher q_1 since she prefers action $a = 1$ at q_1 and action $a = 0$ at q_0 if $q_0 < \mu_i$.

(b) If player i is a conformist biased toward $a = 0$, she would prefer any outcome

[67]Note that if $\mu_i > \mu_R$, no outcome with $\mu_i < q_0 < \mu_R$ exists.

τ with $q_0 < \mu_R$ to all those with $q_0 \geq \mu_R$ since she prefers action $a = 0$ at q_0. Among outcomes τ with $q_0 < \mu_R$, she prefers ones with higher q_1 since she prefers action $a = 0$ at q_0; also, she prefers ones with lower q_0 if $q_1 \geq \mu_i$ and higher q_0 if $q_1 < \mu_i$.

(c) If player i is a contrarian biased toward $a = 1$, she would prefer any outcome τ with $q_0 \geq \mu_R$ to all outcomes with $q_0 < \mu_R$ since she prefers action $a = 1$ at q_0[68].

(d) If player i is a contrarian biased toward $a = 0$, she would prefer any outcome τ with $q_0 \geq \mu_R$ only to those with $q_0 < \min(\mu_i, \mu_R)$ since she prefers action $a = 1$ at q_0 only if $q_0 < \mu_i$. Among outcomes τ with $\mu_i \leq q_0 < \mu_R$, she prefers ones with higher q_0 and higher q_1 since she prefers action $a = 0$ at q_1 and action $a = 0$ at q_0 if $q_0 > \mu_i$[69].

The above arguments imply that all I need to know about a player is her type, i.e., her indifference posterior belief μ_i and her preferred action at state $\omega = 1$.

Proof of Proposition 4: Following 1, 3(c), and 3(d), if there exists a 1-extremist or a contrarian with $\mu_i > \mu_R$, the equilibrium outcome would be the no-information outcome since such a player would prefer the no-information outcome to all other outcomes.

Following 3(a) and 3(d), if $A \neq \emptyset$ and there exists a contrarian with $\mu_{A^*} < \mu_i < \mu_R$, the equilibrium outcome would be the no-information outcome since A^* would only prefer the outcomes with $q_0 \leq \mu_{A^*}$ to the no-information outcome while player i would only prefer the outcomes with $\mu_i \leq q_0 < \mu_R$ to the no-information outcome. Note that these two sets of outcomes are disjoint.

Following 3(a) and 3(b), if all players are conformists, the equilibrium outcome would be the full-information outcome where $q_0 = 0$ and $q_1 = 1$ since this is the most-preferred outcome for all the players.

[68] Her preferences over the outcomes τ with $q_0 < \mu_R$ do not matter since she prefers the no-information outcome to all such outcomes.

[69] Note that if $\mu_i > \mu_R$, no outcome with $\mu_i < q_0 < \mu_R$ exists.

Let D^* represent the player with the lowest bias among those in D with $\mu_i < \mu_{A^*}$: $\mu_{D^*} = \max_{i \in D: \mu_i < \mu_{A^*}} \mu_i$. If $\{i \in D : \mu_i < \mu_{A^*}\} = \emptyset$, let $\mu_{D^*} = 0$.

Let A^*_E represent the player with the highest bias among those in A and closer to the receiver than E^*: $\mu_{A^*_E} = \min_{i \in A: i > E^*} \mu_i$. If $\{i \in A : i > E^*\} = \emptyset$, let $A^*_E = R$.

Following 2, 3(a), 3(b), and 3(d), if there exists no 1-extremists or contrarians with $\mu_i > \mu_{A^*}$, but there are players who are not conformists, that is, 0-extremists or contrarians with $\mu_i < \mu_{A^*}$, the equilibrium outcome would either be the no-information outcome or one with $\mu_{D^*} \leq q_0 \leq \mu_{A^*}$ and $q_1 = 1$.

- If $A^* < E^*$, following 2 and 3(d), the best response of E^* to any outcome induced by the preceding players with $\mu_{D^*} \leq q_0 \leq \mu_{A^*}$ and $q_1 = 1$ is to induce the outcome with $q_0 = \mu_{A^*_E}$ and $q_1 = 1$. However, following 3(a), A^* would prefer the no-information outcome to such an outcome since by definition $\mu_{A^*} < \mu_{A^*_E}$. Thus, the only possible equilibrium outcome would be the no-information outcome.

- If $E^* < A^*$, following 2 and 3(d), the best response of any player in $D \cup E_0$ to any outcome induced by the preceding players with $\mu_{D^*} \leq q_0 \leq \mu_{A^*}$ and $q_1 = 1$ is to induce the outcome with $q_0 = \mu_{A^*}$ and $q_1 = 1$ since $A^* = A^*_E$. The best response of all conformists to any outcome induced by the preceding players with $\mu_{D^*} \leq q_0 \leq \mu_{A^*}$ and $q_1 = 1$ is to pass on that information. Moreover, all the players prefer any outcome with $\mu_{D^*} \leq q_0 \leq \mu_{A^*}$ to the no-information outcome. Thus, the equilibrium outcome would be the one with $q_0 = \mu_{A^*}$ and $q_1 = 1$.

■

1.10 Appendix B: General Binary-Action Games

In what follows, the sets A, B, C, D, E_0, and E_1 as well as players A^* and E^* are the same as those in section 1.7.2; that is, the definitions do not include player 1.

Proof of Lemma 1: I prove the result for the case that receiver is a conformist biased toward $a = 1$, i.e., $\alpha_R > 0$ and $0 < \omega_R < m$. The other cases can be proved similarly.

As mentioned before, no-information outcome is the equilibrium outcome of the reduced binary-state game corresponding to the original binary-action game only if one of the following conditions holds:

1. *There exists a 1-extremist or a contrarian with* $\omega_R < \omega_i < 1$ *in the reduced game.*

   Note that the only 1-extremists in the reduced game are the absolute 1-extremists in the original game. In the original game, any choice of m_0 and m_1 by player 1 does not change the type of absolute 1-extremists; it can only turn conformists or contrarians biased toward $a = 1$ into 1-extremists. This implies that contrarians with $m < \omega_i < 1$ either remain unchanged or are turned into 1-extremists. As a result, if there exists a 1-extremist or a contrarian with $m < \omega_i < 1$ in the reduced game, there would exist one in the subgame starting from player 2 in the original game as well, regardless of the choice of player 1, which leads to no-information outcome by Proposition 4.

   Any choice of m_0 and m_1 by player 1 that turns a contrarian with $\omega_R < \omega_i < m$ into an extremist satisfies $m_0 > \omega_R$; the assumption that $\mathcal{S}_n = \mathcal{A}$ implies $m_0 = m_1 = m$, which implies the no-information outcome. As a result, if there exists a contrarian with $\omega_R < \omega_i < m$ in the reduced game, either player 1 chooses the no-information outcome, or there would exist one in the subgame starting from player 2 in the original game as well, which leads to no-information outcome by Proposition 4.

2. $A \neq \emptyset$ *and there exists a contrarian with* $\omega_{A^*} < \omega_i < \omega_R$ *in the reduced game.*

   Consider any choice of m_0 and m_1 by player 1 in the original game. If $m_0 \leq \omega_{A^*}$, $A \neq \emptyset$ and contrarians with $\omega_{A^*} < \omega_i < \omega_R$ remain unchanged with $\omega_{A^*} < \omega_i$. If $m_0 > \omega_{A^*}$, all conformists in A turn into 1-extremists; therefore, A^* turns into a 1-extremist. As a result, if $A \neq \emptyset$ and there exists a contrarian with $\omega_{A^*} < \omega_i < m$ in the reduced game, either (i) $A \neq \emptyset$ and there would exist such a contrarian, or (ii) there would exist a 1-extremist in the subgame starting from player 2 in the original game, both of which lead to no-information outcome by Proposition 4.

3. *$D \cup E_0 \neq \emptyset$ and $A^* < E^*$ in the reduced game.*

   Note that the only 0-extremists in the reduced game are the absolute 0-extremists in the original game. In the original game, any choice of m_0 and m_1 by player 1 does not change the type of absolute 0-extremists; it can only turn conformists or contrarians biased toward $a = 0$ into 0-extremists. This implies that contrarians with $\omega_i < \omega_{A^*}$ either remain unchanged or are turned into 0-extremists. Therefore, for any choice of m_0 and m_1 by player 1, E^* in the subgame starting from player 2 in the original game is the same as E^* in the reduced game.

   If $m_0 > \omega_{A^*}$, A^* turns into a 1-extremist. If $m_0 < \omega_{A^*}$, A^* in the subgame starting from player 2 in the original game is the same as A^* in the reduced game, and thus $A^* < E^*$.

   As a result, if $A^* < E^*$ in the reduced game, either there would exist a 1-extremist, or $A^* < E^*$ in the subgame starting from player 2 in the original game, both of which lead to no-information outcome by Proposition 4. ■

Bayes-plausibility implies that for any outcome of the game, or equivalently, any distribution of posterior means for the receiver τ with $supp(\tau) = \{m_0, m_1\}$, $\mathbb{E}_\tau[\omega] = m$. Without loss of generality, let $m_0 \leq m_1$ which implies $0 \leq m_0 \leq m \leq m_1 \leq 1$, $\tau(m_1) = \frac{m-m_0}{m_1-m_0}$, and $\tau(m_0) = \frac{m_1-m}{m_1-m_0}$. Note that in the binary-action case, being smaller in the convex order boils down to being a mean preserving contraction. The set of binary-support distributions of posterior means which are mean preserving contractions of τ includes all τ' with $supp(\tau') = \{m_0', m_1'\}$ such that $m_0 \leq m_0' \leq m \leq m_1' \leq m_1$ and $\mathbb{E}_{\tau'}[\omega] = m$. In other words, τ is more Blackwell-informative, the lower m_0 is and the higher m_1 is.

Proof of Lemma 4: For ease of exposition, suppose receiver takes $a = 1$ at $\omega = \omega_R$[70]. Given an outcome τ with $m_0 < \omega_R$, the expected utility of player 1 with $u_1(\omega, 0) = 0$

[70]This removes the necessity of Assumption 1.

and $u_1(\omega, 1) = \alpha_1\omega + \beta_1$ is given by

$$\mathbb{E}_\tau[u_1(\omega, a^*(\omega)] = \frac{m - m_0}{m_1 - m_0}(\alpha_1 m_1 + \beta_1). \tag{1.24}$$

Moreover, given any outcome τ with $m_0 \geq \omega_R$, the expected utility of player 1 with $u_1(\omega, 0) = 0$ and $u_1(\omega, 1) = \alpha_1\omega + \beta_1$ is simply given by

$$\begin{aligned}\mathbb{E}_\tau[u_1(\omega, a^*(\omega)] &= \frac{m - m_0}{m_1 - m_0}[\alpha_1 m_1 + \beta_1] + \frac{m_1 - m}{m_1 - m_0}[\alpha_1 m_0 + \beta_1] \\ &= \alpha_1\left(\frac{m - m_0}{m_1 - m_0}m_1 + \frac{m_1 - m}{m_1 - m_0}m_0\right) + \beta_1 \\ &= \alpha_1 m + \beta_1,\end{aligned} \tag{1.25}$$

which is the same as her expected utility given the outcome τ with $supp(\tau) = \{m\}$, where no information is communicated to the receiver. This makes sense since the receiver would always take action $a = 1$ if $m_0 \geq \omega_R$, and thus I call all such outcomes the no-information outcome.

1. Suppose player 1 is an absolute 1-extremist. She would prefer to persuade the receiver to take action $a = 1$ as often as possible. Therefore she would prefer any outcome τ with $m_0 \geq \omega_R$, i.e., the no-information outcome, to all other outcomes[71].

2. Suppose player 1 is an absolute 0-extremist. She would prefer to persuade the receiver to take action $a = 0$ as often as possible. The receiver would only take action $a = 0$ if his posterior mean is m_0 and $m_0 < \omega_R$. Since $\frac{\partial\tau(m_0)}{\partial m_0} > 0$ and $\frac{\partial\tau(m_0)}{\partial m_1} > 0$, she would prefer outcomes with higher m_1 and higher m_0 as long as $m_0 < \omega_R$. Clearly, her least preferred outcomes are those with $m_0 \geq \omega_R$, i.e., the no-information outcome.

3. Suppose player 1 is a non-extremist. She would prefer any outcome τ with $m_0 \geq \omega_R$ to another outcome with $m_0 < \omega_R$ if and only if she prefers action $a = 1$ at m_0,

[71] Her preferences over the outcomes τ with $m_0 < \omega_R$ do not matter since she prefers the no-information outcome to all such outcomes.

as shown below:

$$
\begin{aligned}
\alpha_1 m + \beta_1 &> \frac{m - m_0}{m_1 - m_0}(\alpha_1 m_1 + \beta_1) \\
\beta_1 m_1 - \alpha_1 m_0 m &> \beta_1 m - \alpha_1 m_1 m_0 \\
(m_1 - m)(\alpha_1 m_0 + \beta_1) &> 0.
\end{aligned} \tag{1.26}
$$

To see her preferences among outcomes τ with $m_0 < \omega_R$, note that

$$
\frac{\partial \mathbb{E}_\tau[u_1(\omega, a^*(\omega)]}{\partial m_0} = -\frac{m_1 - m}{(m_1 - m_0)^2}(\alpha_1 m_1 + \beta_1), \tag{1.27}
$$

which is negative if and only if she prefers action $a = 1$ at m_1. Similarly, note that

$$
\begin{aligned}
\frac{\partial \mathbb{E}_\tau[u_1(\omega, a^*(\omega)]}{\partial m_1} &= -\frac{m - m_0}{(m_1 - m_0)^2}(\alpha_1 m_1 + \beta_1) + \frac{m - m_0}{m_1 - m_0}\alpha_1 \\
&= -\frac{m - m_0}{(m_1 - m_0)^2}(\alpha_1 m_0 + \beta_1),
\end{aligned} \tag{1.28}
$$

which is positive if and only if she prefers action $a = 0$ at m_0. The above observations imply the following:

(a) If player 1 is a conformist biased toward $a = 1$, she would prefer any outcome τ with $m_0 \geq \omega_R$ only to those with $\omega_1 < m_0 < \omega_R$ since she prefers action $a = 1$ at m_0 only if $m_0 > \omega_1$[72]. Among outcomes τ with $m_0 \leq \omega_1$, she prefers ones with lower m_0 and higher m_1 since she prefers action $a = 1$ at m_1 and action $a = 0$ at m_0 if $m_0 < \omega_1$.

(b) If player 1 is a conformist biased toward $a = 0$, she would prefer any outcome τ with $m_0 < \omega_R$ to all those with $m_0 \geq \omega_R$ since she prefers action $a = 0$ at m_0. Among outcomes τ with $m_0 < \omega_R$, she prefers ones with higher m_1 since she prefers action $a = 0$ at m_0; also, she prefers ones with lower m_0 if $m_1 \geq \omega_1$ and higher m_0 if $m_1 < \omega_1$.

[72] Note that if $\omega_1 > \omega_R$, no outcome with $\omega_1 < m_0 < \omega_R$ exists.

(c) If player 1 is a contrarian biased toward $a = 1$, she would prefer any outcome τ with $m_0 \geq \omega_R$ to all outcomes with $m_0 < \omega_R$ since she prefers action $a = 1$ at m_0[73].

(d) If player 1 is a contrarian biased toward $a = 0$, she would prefer any outcome τ with $m_0 \geq \omega_R$ only to those with $m_0 < \min(\omega_1, \omega_R)$ since she prefers action $a = 1$ at m_0 only if $m_0 < \omega_1$. Among outcomes τ with $\omega_1 \leq m_0 < \omega_R$, she prefers ones with higher m_0 and higher m_1 since she prefers action $a = 0$ at m_1 and action $a = 0$ at m_0 if $m_0 > \omega_1$[74].

■

Proof of Lemma 2: I prove the result for the case that receiver is a conformist biased toward $a = 1$, i.e., $\alpha_R > 0$ and $0 < \omega_R < m$. The other cases can be proved similarly.

No-information outcome is the equilibrium outcome of the binary-state game corresponding to the original binary-action game only if one of the following conditions holds:

1. The equilibrium outcome of the reduced binary-state game corresponding to the original binary-action game is the no-information outcome. In this case, by Lemma 1, the equilibrium outcome of the original game is the no-information outcome as well.

2. Player 1 is an absolute 1-extremist or a contrarian with $\omega_R < \omega_1 < 1$ in the corresponding binary-state game. Following 1, 3(c), and 3(d) in the proof of Lemma 4, the equilibrium outcome of the original game would be the no-information outcome since player 1 would prefer the no-information outcome to all other outcomes.

3. $A \neq \emptyset$ and player 1 is a contrarian with $\omega_{A^*} < \omega_1 < \omega_R$ in the corresponding binary-state game. In the original game, following 3(d) in the proof of Lemma 4,

[73]Her preferences over the outcomes τ with $m_0 < \omega_R$ do not matter since she prefers the no-information outcome to all such outcomes.

[74]Note that if $\omega_1 > \omega_R$, no outcome with $\omega_1 < m_0 < \omega_R$ exists.

player 1 would only prefer the outcomes with $\omega_1 \leq m_0 < \omega_R$ to the no-information outcome. However, any choice of m_0 and m_1 by player 1 with $m_0 \geq \omega_1$, turns A^* into a 1-extremist in the subgame starting from player 2 in the original game, which leads to no-information outcome by Proposition 4. Therefore, the only possible equilibrium outcome in the original game would be the no-information outcome.

4. $D \cup E_0 \neq \emptyset$, $E^* < A^*$[75], and $\omega_1 < \omega_{A^*}$ in the corresponding binary-state game. By Proposition 4, for any choice of m_0 and m_1 by player 1, the equilibrium outcome of the subgame starting from player 2 would be either the no-information outcome or τ with $supp(\tau) = \{\mu_{A^*}m_1 + (1-\mu_{A^*})m_0, m_1\} = \{\omega_{A^*}, m_1\}$. However, following 3(a) in the proof of Lemma 4, player 1 would prefer the no-information outcome since $\omega_1 < \omega_{A^*}$. Thus, the only possible equilibrium outcome in the original game would be the no-information outcome.

■

Proof of Lemma 3: Since the no-information outcome is not the equilibrium outcome of the reduced binary-state game, not considering player 1, there exists no absolute 1-extremists or contrarians with $\omega_{A^*} < \omega_i$, and A^* is closer to the receiver than all absolute 0-extremists and contrarians with $\omega_i < \omega_{A^*}$.

As mentioned before, given player 1's choice of distribution of posterior means τ_1 with $supp(\tau_1) = \{m_0, m_1\}$, Proposition 4 characterizes the equilibrium outcome of the binary-state, binary-action subgame starting from player 2. Based on Proposition 4, the equilibrium of the subgame starting from player 2 is characterized by no-information outcome unless:

1. all players are conformists: In this case, the equilibrium outcome τ of the binary-state, binary-action game starting from player 2 with $\widehat{\Omega} = \{m_0, m_1\}$ is characterized by $supp(\tau) = \{0, 1\}$ (full-information outcome).

[75] If $A^* < E^*$, the equilibrium outcome of the reduced binary-state game corresponding to the original binary-action game is the no-information outcome, and thus condition 1 holds.

All players in the binary-state, binary-action game starting from player 2 are conformists if and only if in the original game (i) all of them are conformists, and (ii) none of them turn into extremists given the choice of player 1, i.e., $m_0 \leq \omega_{A^*}$ and $m_1 \geq \omega_{B^*}$. If this is the case, τ implies that the the equilibrium outcome τ^* of the subgame starting from player 2 is characterized by $supp(\tau^*) = \{0 \cdot m_1 + 1 \cdot m_0, 1 \cdot m_1 + 0 \cdot m_0\} = \{m_0, m_1\}$.

2. all players are not conformists but the non-conformist players are either 0-extremists or contrarians with $\omega_i < \omega_{A^*}$: In this case, the equilibrium outcome τ of the binary-state, binary-action game starting from player 2 with $\widehat{\Omega} = \{m_0, m_1\}$ is characterized by $supp(\tau) = \{\mu_{A^*}, 1\}$[76].

   All players in the binary-state, binary-action game starting from player 2 satisfy the above condition if and only if in the original game (i) all of them are conformists but at least one conformist biased toward $a = 0$ turns into a 0-extremist given the choice of player 1, i.e., $m_1 < \omega_{B^*}$, or there are players who are not conformists, (ii) none of the conformists biased toward $a = 1$ turn into 1-extremists given the choice of player 1, i.e., $m_0 \leq \omega_{A^*}$, and (iii) A^* is closer to the receiver than all the new 0-extremists, i.e., conformists biased toward $a = 0$ with $\omega_i > m_1$. If this is the case, τ implies that the the equilibrium outcome τ^* of the subgame starting from player 2 is characterized by $supp(\tau^*) = \{\mu_{A^*} \cdot m_1 + (1 - \mu_{A^*}) \cdot m_0, 1 \cdot m_1 + 0 \cdot m_0\} = \{\mu_{A^*}, m_1\}$.

■

1.11 Appendix C: Proof of Proposition 3

The first direction is straightforward. If $\tau \in \tilde{\Gamma}$, it implies that $\tau \in \Gamma_0$ and for every $i = 2, \ldots, n$,

$$\sum_{\mu} \tau(\mu) v_i(\mu) \geq \sum_{\mu} \tau'(\mu) v_i(\mu), \forall \tau' \leq_{cx} \tau, \tag{1.29}$$

[76] As mentioned before, $\mu_{A^*} = \frac{m - m_0}{m_1 - m_0}$.

which, in turn, implies that for every $i = 2, \dots, n$,

$$\sum_{\mu} \tau(\mu) v_i(\mu) \geq \sum_{\mu} \tau'(\mu) v_i(\mu), \ \forall \tau' \in \Gamma_{i+1} \text{ s.t. } \tau' \leq_{cx} \tau^*, \tag{1.30}$$

where $\Gamma_{n+1} = \Gamma_0$. This can be used to show recursively that $\tau \in \Gamma_n, \tau \in \Gamma_{n-1}, \dots, \tau \in \Gamma_2$. Therefore, $\tilde{\Gamma} \subseteq \Gamma_2$.

The following simple counterexample proves the other direction. However, it is based on the notation, concepts, and results introduced in section 1.7.1 and Appendix A. Consider a binary-state, binary-action game and let $n = 3$. Suppose player 2 is a 0-extremist while receiver and player 3 are conformists who are biased toward $a = 1$ with player 3 being higher biased, i.e, $0 < \mu_3 < \mu_R < p$. Note that the receiver takes action $a = 1$ at $\mu = \mu_R$ according to Assumption 1.

First of all, the set of Bayes plausible distributions of posteriors is given by

$$\Gamma_0 = \{\tau | supp(\tau) = \{q_0, q_1\}, 0 \leq q_0 \leq p \leq q_1 \leq 1, \mathbb{E}_\tau[q] = p\}. \tag{1.31}$$

As mentioned in Appendix B, in the binary-state binary-action case, being smaller in the convex order boils down to being a mean preserving contraction. The set of binary-support distributions of posteriors which are mean preserving contractions of τ includes all τ' with $supp(\tau') = \{q'_0, q'_1\}$ such that $q_0 \leq q'_0 \leq p \leq q'_1 \leq q_1$ and $\mathbb{E}_{\tau'}[q'] = p$. In other words, τ is more Blackwell-informative, the lower q_0 is and the higher q_1 is.

Starting with player 3, and considering (1.8),

$$\Gamma_3 = \{\tau \in \Gamma_0 | 0 \leq q_0 \leq \mu_3 \text{ or } \mu_R \leq q_0 \leq p\}. \tag{1.32}$$

Now, continuing to player 2, and considering (**??**),

$$\Gamma_2 = \{\tau \in \Gamma_0 | q_0 = \mu_3 \text{ or } \mu_R \leq q_0 \leq p\}. \tag{1.33}$$

However,

$$\tilde{\Gamma} = \{\tau \in \Gamma_0 | \mu_R \leq q_0 \leq p\}. \tag{1.34}$$

This is due to the fact that while player 2 prefers experiments with $q_0 = \mu_3$ to all less Blackwell-informative ones in Γ_3, i.e., those with $\mu_R \leq q_0 \leq p$, there exist less-Blackwell informative experiments in Γ_0, namely those with $\mu_3 < q_0 < \mu_R$, which she prefers even more. ■

Chapter 2

Indicator Choice in Pay-for-Performance

with Ali Shourideh and Ariel Zetlin-Jones

2.1 Abstract

We study the classic principal-agent model when the signal observed by the principal is chosen by the agent. We fully characterize the optimal information structure from an agent's perspective in a general moral hazard setting with limited liability. Due to endogeneity of the contract chosen by the principal, the agent's choice of information is non-trivial. We show that the agent's problem can be mapped into a geometrical game between the principal and the agent in the space of likelihood ratios. We use this representation result to show that coarse contracts are sufficient: The agent can achieve her best with binary signals. Additionally, we can characterize conditions under which the agent is able to extract the entire surplus and implement the first-best efficient allocation. Finally, we show that when effort and performance are one-dimensional, under a general class of models, threshold signals are optimal. Our theory can thus provide a rationale for coarseness of contracts based on the bargaining power of the agent in negotiations.

2.2 Introduction

The use of pay-for-performance contracting is a cornerstone of modern employment contracts. Executive compensation is often indexed in part to company performance metrics including growth in the company's stock price. Employment contracts for top athletes frequently involve performance bonuses for specific outcomes such as goals scored for soccer players or the number of touch-downs for players in the NFL. When employment contracts feature such performance pay incentives, employers and employees must agree to a set of performance indicators during contract negotiations.[1] Given this observation, what are the incentives of employees and employers to negotiate on performance indicators? If the employees have a role in choosing the performance metrics, what metrics

[1] Bebchuk and Fried (2004) discuss the various issues with the negotiations process between the CEOs and boards and possible issues arising from choosing particular performance indicators, i.e., vesting stocks. In soccer, news outlets often describe the process in which players and soccer clubs agree on what performance measure to use. See for example this article in Daily Mail which describes several soccer players in the Premier League negotiating over the relevant performance indicator as a base for pay.

would they choose? Finally, how does the choice of indicators interact with the ultimate productive efficiency of the firm?

In this paper, we answer these questions by considering the problem of indicator design in the textbook moral hazard problem with limited liability. More specifically, we consider the standard principal-agent problem of Holmström (1979) in which an agent has quasi-linear preferences and must be paid a non-negative wage. The *performance technology* is one that maps costly effort, e, by the agent into a distribution of some performance measure x. Before the principal offers a compensation contract, the agent chooses an *indicator*, a possibly random signal s of x where the principal can only offer a contract that is contingent on the indicator, s. Once the indicator is chosen, the principal and the agent play the textbook moral hazard game.[2]

In this environment, one might conjecture that more information would lead to more efficient outcomes. While this is true under certain circumstances – see the informativeness principle of Holmström (1979), and its extension by Chaigneau et al. (2019) – information can often be detrimental to the agent. To see the intuition for this observation, suppose that the performance technology is degenerate at $x = e$ and thus the agent can use her effort as the indicator. Then, revealing all information by choosing $s = x = e$ would give the principal the ability to fully capture all the surplus generated by the effort. On the other hand, choosing a fully uninformative indicator leads to no surplus for the agent as the principal is unable to incentivize the agent to contribute effort when signals are fully uninformative. This points to a trade-off in the problem of indicator design by the agent.

In this model, we show three main results: first, we provide a geometric interpretation of the indicator design game in the space of likelihoods; the probability of a particular performance/signal realization for an arbitrary effort relative to the effort that the agent would like to implement. Under this interpretation, the agent's indicator design problem is equivalent to a geometric game where the agent chooses a convex set (of likelihood

[2]Throughout the paper, we assume that the agent commits to the indicator s while she cannot commit to the effort level e. A justification for this assumption is that employment contracts are often enforced by courts and thus hard to break.

ratios) and the principal chooses a point within that set. Our geometric interpretation provides a tractable formulation to understand how the agent's choice of indicator influences the resulting compensation scheme offered by the principal. Second, we provide conditions under which the agent is able to choose the indicator in such a way to implement the first-best efficient effort and capture all the surplus created by her effort. Finally, we consider a specific case where performance measure x has a continuous distribution in real numbers and show that under certain conditions optimal indicator structure takes the form of monotone or hump-shaped thresholds signals.

The reason that the agent is sometimes able to capture all of the surplus and implement the efficient outcome can be easily understood when performance technology is fully informative, i.e., $x = e$. In this case, suppose the agent chooses an indicator with two realizations: high and low. By choosing the probability of high and low indicators appropriately for every $x = e$, the agent is able to force the principal's hand. Note that if the principal wants to implement a given effort level $\hat{e}$, he must compensate the agent following a realization of the high signal. The amount of compensation the principal must deliver is increasing as the cost of the desired effort level $\hat{e}$ rises relative to the cost of any other effort and is increasing in the likelihood of observing a high signal from some other effort relative to that of observing a high signal from effort $\hat{e}$. By choosing the signal probabilities appropriately, the agent is then able to make the off-path likelihood of the high signal sufficiently large so that the principal must give up all the surplus in order to implement the desired effort level $\hat{e}$.

In general, our geometric interpretation allows us to show that the agent can implement any desired effort level with a "coarse" information structure that has at most two signal realizations. Additionally, this interpretation allows us to provide conditions under which even when the performance technology is stochastic, it is possible for the agent to capture all the surplus and implement the first-best outcome. Our sufficient condition amounts to checking whether a particular point in the likelihood space belongs to convex hull of the likelihood functions implied by the performance technology. This type of sufficient condition is easy to evaluate using existing convex hull algorithms.

Beyond our technical contributions, our results shed light on the debate on the efficiency of incentive contracts in executive compensation. Bebchuk and Fried (2004) have argued that the standard agency model of shareholder maximization is at odds with the data since negotiations often happen between the CEO and the board whose incentives are not necessarily aligned with that of the shareholders – in fact they claim that CEOs and compensation committees often trade favors at a cost to shareholders. In our model, and consistent with Bebchuk and Fried (2004)'s interpretation, we endow the agent with full bargaining power over the choice of performance pay indicators. While the agent captures all of the gains from trade under this assumption, her choices also maximize total surplus. In contrast, standard models of moral hazard with exogenous performance technology often feature inefficiencies in the sense that providing incentives to the agent often entails reductions in total surplus. In this sense, we find that optimal design of performance pay indicators may help to reduce inefficiencies within the firm. Further investigation of the data on negotiations and choice of indicators would be a good test of our theory.

2.2.1 Related Literature

Our paper is related to several strands of the literature on contracting and information design.

With respect to the moral hazard literature, our key innovation is to consider the problem of choosing indicators whereupon contracts are based on showing that this can remove all inefficiencies. In the classical model, Innes (1990) and Poblete and Spulber (2012) consider moral hazard with limited liability and a risk-neutral agent, and show that simple debt contracts are optimal. Carroll (2015) as well as Walton and Carroll (2022) consider the moral hazard problem with limited liability, where the principal has non-Bayesian uncertainty about the production technology and wishes to maximize her worst-case payoff. [3]

[3]It is needless to say that a rather large body of work has considered models with risk-averse agents and risk-neutral principals.

While often information design incentives are ignored in moral hazard, Holmström (1979) and later Chaigneau et al. (2019) are exceptions. They investigate the comparative statics of changing the performance technology on the payoffs; by the informativeness principle, the more informative the output is about the effort, the lower wage the principal needs to pay.

Perhaps, the closest paper to ours is that of Garrett et al. (2020). They consider a model in which the agent can design the performance technology and cost function – they refer to this as technology design – and show that the agent-optimal design involves only binary distributions. In contrast, in our model, the performance technology is fixed and the agent chooses an indicator of this performance, i.e., an information structure to garble the principal's observations. Moreover, our paper has a technical contribution by reformulating the problem in terms of likelihood ratios.

Georgiadis and Szentes (2020) study moral hazard with limited liability and a risk-averse agent where the principal continuously observes signals about the agent's effort at a constant marginal cost. They show that the principal-optimal information-acquisition strategy is a two-threshold policy. Barron et al. (2020) consider an agent who can costlessly add mean-preserving noise to the output. This is also the case in our model; however, in our setting, the noisy output is just used for contracting purposes and does not change the principal's payoff compared to the original output. Moreover, in our model, first the agent chooses the information structure, then the principal offers the contract, and finally the agent chooses her effort, whereas in Barron et al. (2020), first the principal offers a contract and then the agent chooses the effort and the information structure.

Another related strand of the moral hazard literature focuses on career concerns introduced by Holmström (1979) and changes in information structure observed by both the principal and the agent. In this context Holmström (1999) shows that noisy performance signals are beneficial for incentives. Similarly Dewatripont et al. (1999) show that more informative signals in the sense of Blackwell do not necessarily increase incentives. In contrast in our model, indicator design or information design can be used as a tool to reshuffle surplus between the agent and the principal while achieving efficiency.

In the context of team production and moral hazard, Halac et al. (2021) show that the principal can benefit from private contract offers by leveraging rank uncertainty: Each agent is informed only of her own bonus and a ranking distribution; each agent's bonus makes work dominant if higher-rank agents work. Interestingly, in our setting, it is the agent that can use uncertainty about performance for the principal to improve efficiency.

We also contribute to the growing literature on incentives in Bayesian Persuasion. Several papers, including Boleslavsky and Kim (2018), Rosar (2017), Perez-Richet and Skreta (2022), Ball (2019), Saeedi and Shourideh (2020), and Zapechelnyuk (2020) have considered the effect of incentives in the Bayesian persuasion problem where a third party designs an information structure, and a "sender" determines the distribution of the underlying state by exerting a costly effort. From a technical perspective, our problem is different from this class of models. This is partly due to the fact that for the ex-post incentive to exert effort by the agent (after the choice of indicator and contract), the distribution of the signals – or alternatively in the language of Kamenica and Gentzkow (2011), the distribution of posteriors – off the equilibrium path is also relevant. By casting the problem in the space of likelihood functions – as opposed to beliefs as in Kamenica and Gentzkow (2011) – we can characterize its solution using the geometric game. This technique can be used in other information design problems in which on- and off-path beliefs are involved.

The rest of the paper is organized as follows: Section 2.3 describes the key insight about full surplus extraction and first-best implementation in a simple example. Section 2.4 describes the basic model. Section 2.5 describes the geometric interpretation of the indicator choice game between the principal and the agent. Section 2.6 provides a sufficient condition for efficient surplus extraction by the agent. Section 2.7 describes optimality of threshold signals. Section 2.8 concludes.

2.3 A Simple Example

In this section, we use a basic environment to illustrate the main mechanisms at work in our model. Consider the basic textbook model of moral hazard. A principal (he) is hiring an agent (she) to perform a task whose output $x \in \{0, 1\}$ is collected by the principal. The agent chooses how much effort to put in to perform the task. She can either choose the low effort e_L or a costly high effort e_H whose cost is given by $c > 0$.

Choosing the high effort leads to output $x = 1$ with certainty while choosing the low effort leads to output $x = 1$ with probability $p < 1$ and $x = 0$ with probability $1 - p$. We assume that the total surplus under high effort $1 - c$ is higher than that under low effort p and thus it is efficient to implement the high effort.

In this standard principal-agent model with moral hazard, principal observes the output but not the effort of the agent. He can compensate the agent for each output realization but cannot make these payments negative, i.e., he is subject to limited liability. If the principal sets wage w when output is high, then the agent's incentive compatibility constraint is

$$w - c \geq p \cdot w.$$

Hence, as long as $w \geq \frac{c}{1-p}$, the principal can implement the high effort. If the principal is to choose $w = \frac{c}{1-p}$, then his payoff is $1 - \frac{c}{1-p} > 0$ while the agent's payoff is $\frac{c}{1-p} - c = \frac{p}{1-p}c > 0$. Moreover, for the principal to prefer implementing the high effort to low, we must have that $1 - \frac{c}{1-p} \geq p$ or $c \leq (1-p)^2$.

Now suppose that the agent can control principal's information about the output. Specifically, suppose that before the contracting stage, the agent can design a device that can potentially hide the output of the project. More specifically, suppose that the agent can choose an information structure or a Blackwell experiment that probabilistically maps output $x \in \{0, 1\}$ to a signal $S = \{L, H\}$ which is observed by the principal. The principal observes the signal $s = H$ with probability π_H when $x = 1$ is realized and observes $s = H$ with probability π_L if $x = 0$ is realized, where $\pi_L < \pi_H$.

Since the principal can only observe the signal designed by the agent, he will compensate her only when $s = H$. If this compensation is w, then the agent's incentive compatibility constraint is

$$\pi_H \cdot w - c \geq (\pi_H p + (1-p)\,\pi_L) \cdot w.$$

Hence, the principal is able to implement high effort when

$$w \geq \frac{c}{(\pi_H - \pi_L)\,(1-p)}.$$

When minimizing the wage, the expected cost of compensating the agent for the principal is $\pi_H w = \frac{c}{\left(1-\frac{\pi_L}{\pi_H}\right)(1-p)}$. Therefore, as long as $p \leq 1 - \frac{c}{\left(1-\frac{\pi_L}{\pi_H}\right)(1-p)}$, the principal finds it profitable to implement the high effort. This inequality can be rewritten as

$$\frac{c}{(1-p)^2} \leq 1 - \frac{\pi_L}{\pi_H}.$$

This, in turn, implies that when $\frac{c}{(1-p)^2} \leq 1$, the agent can find a signal structure (π_L, π_H) such that the above holds with equality. Under such an information structure, the payoff of the principal is p, what he can achieve without any costly effort, and the agent captures the rest of the surplus, $1 - p - c$. In other words, giving the agent the ability to choose an information structure enables her to guarantee the highest value of the surplus under an efficient level of effort. Intuitively, the change of the information structure allows the agent to induce an arbitrary high value of the wage by increasing the likelihood π_L/π_H, and capture the entire surplus.

The above example illustrates that agent's freedom to choose the information structure, based on which she will be paid, can be extremely powerful. A few natural questions arise: When can the agent capture the efficient level of surplus? What information structure should be used by the agent to achieve her desired outcome? In what follows, we provide a characterization of the optimal signal structure for the agent as well as her ability to extract surplus from the principal.

2.4 Model

Our general model builds upon the textbook moral hazard problem. Consider a principal employing an agent to perform a task whose output is represented by $x \in X$, where X is finite. The agent chooses effort $e \in E = \{e_1, \cdots, e_m\}$ to perform the task, where E is finite. The agent's effort choice induces a probability distribution $f(x|e)$ over the outcome space X, where $\sum_x f(x|e) = 1, \forall e \in E$. We refer to $f(\cdot|\cdot)$ as the performance technology. Effort is costly to the agent; the cost of exerting effort e is given by $c(e)$ for some real-valued function $c : E \to \mathbb{R}_+$.

Throughout the analysis, we assume that $e_1 \in E$ represents the effort with the lowest cost; for simplicity, let $c(e_1) = 0$. The principal's payoff from realization of output x is given by $g(x)$ for some real-valued function $g : X \to \mathbb{R}$. The principal cannot observe the agent's effort and thus cannot offer a contract contingent on the agent's effort; he can only offer contracts contingent on observable outcomes.

The point of departure from the textbook moral hazard model is that the agent may influence the principal's information about the output by choosing an information structure (S, π). Here, S is a signal space and $\pi(\cdot|x) : X \to \Delta(S)$ is a stochastic mapping from the output space X to the signal space, where $\sum_s \pi(s|x) = 1,\ \forall x \in X$. The principal only observes the signal $s \in S$ generated from this information structure and can thus only offer a contract contingent on this signal realization. Therefore, the principal's choice of contract can be represented by a real-valued function $w : S \to \mathbb{R}_+$ where $w(s)$ is the wage paid to the agent when signal s is realized. One interpretation of this contractual restriction is that the task output is not observable to the principal but the agent may verifiably disclose information about the output. An alternative interpretation is that the output is observable but the agent has the option to choose the performance measure based on which she will be paid.

Note that as in the simple example, we have assumed that agent enjoys limited liability, i.e., the contract offered to her by the principal guarantees a non-negative wage regardless of the effort she puts in. The principal's payoff u_P is equal to the payoff from output less the wages paid to the agent. The agent's payoff u_A is equal to the wage she receives from the principal minus the cost of her effort. Both the principal and the agent are assumed to maximize expected utility. Notice that principal can always implement the zero-cost effort, i.e. e_1, by offering $w(s) = 0,\ \forall s \in S$. Therefore, his outside option is to implement e_1 and obtain $\underline{u}_P = \sum_x g(x)f(x|e_1)$.

The timing of the game is as follows:

- Agent chooses an information structure (S, π). To ease the exposition, we assume the signal space S is finite and simply represent the information structure by π. [4]
- Observing the information structure (S, π) chosen by the agent, the principal offers the agent a contract $w : S \to \mathbb{R}_+$ contingent on the realized signal.
- Observing the contract w offered by the principal (and the information structure (S, π) she has chosen), the agent chooses how much effort e to exert.
- Given the effort e chosen by the agent, output x is realized according to $f(x|e)$ and then signal $s \in S$ is realized according to $\pi(s|x)$.
- Payoffs are realized where agent's payoff is $u_A = w(s) - c(e)$, and the principal's payoff is $u_P = g(x) - w(s)$.

To summarize, the game is played sequentially in three stages. In the first stage, the agent chooses the information structure π to (partially) inform the principal about the realized output. In the second stage, the principal offers the agent a contract w contingent on the signal realization. In the last stage, the agent chooses her effort e.

The equilibrium concept is the standard subgame perfect equilibrium (SPE). The agent's strategy $(\pi, \sigma_e(\pi, w))$ consists of a choice of information structure π and a choice of effort $\sigma_e(\pi, w)$ as a function of π and the contract w offered by the principal. The principal's strategy $\sigma_w(\pi)$ is the contract he offers to the agent as a function of the information structure π chosen by the agent.

Definition 1. An SPE of the game $\sigma^* = (\pi^*, \sigma_e^*(\pi, w), \sigma_w^*(\pi))$ consists of a strategy for the agent $(\pi^*, \sigma_e^*(\pi, w))$ and a strategy for the principal $\sigma_w^*(\pi)$ such that:

[4] Later, we show this restriction to finite signal spaces is without loss of generality.

- $\sigma_e^*(\pi, w)$ maximizes the agent's expected utility for every information structure π and every principal's choice of contract w.

- $\sigma_w^*(\pi)$ maximizes the principal's expected utility for every agent's choice of information structure π given agent's equilibrium effort strategy $\sigma_e^*(\pi, w)$.

- π^* maximizes the agent's expected utility given principal's equilibrium strategy $\sigma_w^*(\pi)$ and her own equilibrium effort strategy $\sigma_e^*(\pi, w)$.

We let $p(\cdot|e) : E \to \Delta(S)$ represent the resulting stochastic mapping from the effort space E to the signal space S induced by a given information structure (S, π) and the underlying probability distribution of outcomes given effort. Note, for any level of effort e and any signal realization s, $p(s|e) = \sum_x f(x|e)\pi(s|x)$ and $\sum_s p(s|e) = 1,\ \forall e \in E$. Conditional on a given information structure (S, π), both the principal and the agent use the stochastic mapping $p\,(\cdot|\cdot)$ to evaluate their expected payoffs.

We now formulate the problems the principal and the agent solve beginning with the agent's choice of effort. The agent's expected utility is $U_A = \sum_s p(s|e)w(s) - c(e)$. Therefore, the agent's problem in the last stage of the game (where π and w have previously been chosen in the preceding stages) is

$$\max_e \sum_s p(s|e)w(s) - c(e). \tag{IC-A}$$

The principal's expected utility is $U_P = \sum_x g(x)f(x|e) - \sum_s p(s|e)w(s)$ if he chooses to implement effort e. It is convenient to define $\mathbb{E}[g(x)|e] = \sum_x g(x)f(x|e)$. For a given desired effort level e, the principal chooses a wage schedule that solves

$$\begin{aligned} \min_w \ & \sum_s p(s|e)w(s) \text{ s.t.} \\ & \sum_s p(s|e)w(s) - c(e) \geq \sum_s p(s|\hat{e})w(s) - c(\hat{e}),\ \forall \hat{e} \in E, \\ & w(s) \geq 0,\ \forall s \in S. \end{aligned} \tag{2.1}$$

The constraints represent the agent's incentive compatibility and a set of limited liability constraints. Notice that we have not imposed a participation constraint for the agent. This is because the assumption $c(e_1) = 0$ together with limited liability implies that setting $e = e_1$ guarantees a non-negative payoff for the agent. Let $W(e, \pi)$ represent the optimal value in (**??**). This is the minimum expected wage the principal must pay the agent to implement effort e given the information structure π chosen by the agent.

Given $W(e, \pi)$, the problem of the principal is to choose an effort level to maximize her expected utility, or,

$$\max_{e} \mathbb{E}\left[g\left(x\right)|e\right] - W(e, \pi). \tag{2.2}$$

It is possible to express the above as an incentive compatibility constraint and thus write the agent's problem in the first stage of the game as

$$\max_{e,\pi} W(e, \pi) - c(e)$$

subject to the principal's incentive compatibility constraint

$$\mathbb{E}\left[g\left(x\right)|e\right] - W(e, \pi) \geq \mathbb{E}\left[g\left(x\right)|\hat{e}\right] - W(\hat{e}, \pi), \ \forall \hat{e} \in E. \tag{IC-P}$$

2.4.1 Remarks on the Environment

It is useful to discuss various interpretations of the model as well as our key assumptions.

Performance Measure as Information Structure

In the textbook model of moral hazard, the principal cannot observe the agent's effort. He therefore uses some imperfect signal of effort to incentivize the agent. If he can observe the output, he offers an output-contingent contract; this makes the output the performance measure for the agent. In our model, the principal does not observe the output but does observe a signal, which may be correlated with effort, and he offers a signal-contingent contract. As a result, the signal is the relevant performance measure for the agent. By choosing the information structure, the agent influences the set of observables that will ultimately dictate her compensation, which we interpret as the agent choosing her performance measure.

We make the assumption that the principal cannot offer output-contingent contracts. As we have described, while a conventional interpretation of this assumption is that the principal cannot observe the output directly, an alternative interpretation is that this restriction arises during the negotiations between the agent and the principal in choosing contractual performance measures.

Commitment

We assume that the agent commits to an information structure in the first stage of the game. Our interpretation of the information structure as a contractual performance measure provides a natural justification of the commitment assumption. When signing a contract, all parties are aware of and agree on the probabilistic nature of the chosen performance measure as a function of the output. The contractible nature of the performance measure makes the commitment assumption necessary.

Comparison to the Literature on Bayesian Persuasion

As in the Bayesian persuasion literature, we can write the problem in terms of the distribution of posteriors induced by the information structure. In the Bayesian persuasion, every choice of information structure induces a distribution of posteriors, where the whole distribution matters: not only the support, but also the probabilities. In our setting, every choice of information structure induces a set of distributions of posteriors, one for each choice of effort. These distributions are related through their supports: given the support of one distribution, the supports of the others are pinned down. As we will discuss in section 2.5, the key sufficient statistic about the choice of information structure is the distribution of the likelihood ratios and these are determined by the supports of the above distributions. Therefore, in our model, only the support of the distribution of posteriors matters.

2.5 A Geometric Analysis of the Game

We now characterize the equilibrium outcomes of the game. To do so, we first describe the set of effort levels that are implementable by some information structure. We then show a "coarse"-ness result. That is, we show that it is without loss of generality to restrict the agent's choice of information structures to binary structures, where the set of signals has only two discrete points. Using such information structures, in the next section, we derive sufficient conditions such that the first-best level of effort is implementable. When these sufficient conditions are satisfied, we show that the agent chooses an information structure that implements the first-best effort level and extracts the entire surplus.

While it is possible to work with zero probability events and define likelihood ratios – by describing how division by 0 is defined – in order to avoid complications, we make the following assumption:

Assumption 1. *The performance technology is full support, i.e.,* $\forall x \in X, \forall e \in E, f(x|e) > 0$.

This assumption implies that all the likelihood ratios below are well-defined.

Implementable Effort. To characterize the set of implementable effort levels, we use backward induction and first re-cast the problem of the principal geometrically. Specifically, we show that the likelihood ratios for each signal realization s for any effort level e relative to the desired implementable level e^* are sufficient statistics to solve the principal's problem. In other words, we argue that any desired effort level to implement e^* and any information structure (S, π) give rise to a (geometric) space of possible likelihood ratios and that the principal's optimal choice of compensation schemes may be reduced to choosing a point in this space of likelihood ratios.

More formally, we define an implementable effort level e^* as follows:

Definition 2. An effort level e^* is *implementable* if there exists an information structure (S, π) such that e^* is a solution to the principal's problem (2.2).

Given this definition, an implementable effort level e^* must satisfy the two incentive compatibility constraints in (IC-P) and (IC-A) for the principal and the agent, where the agent's incentive compatibility constraint must hold for all possible histories including those following a deviation by the principal that involves recommending an alternative level of effort.

Let e^* represent some effort level the agent would like to implement (in the first stage of the game). To motivate the relevance of likelihood ratios, consider the agent's interim incentive compatibility constraint when the principal only pays the agent following a signal realization s:

$$p\left(s|e^*\right) w\left(s\right) - c\left(e^*\right) \geq p\left(s|\hat{e}\right) w\left(s\right) - c\left(\hat{e}\right), \forall \hat{e} \in E.$$

These constraints imply that for any effort level $\hat{e}$ where signal s is less likely than under e^*, i.e., $p\left(s|\hat{e}\right) < p\left(s|e^*\right)$, the wage must satisfy

$$w\left(s\right) \geq \frac{c\left(e^*\right) - c\left(\hat{e}\right)}{p\left(s|e^*\right) - p\left(s|\hat{e}\right)}, \tag{2.3}$$

and if s is more likely under $\hat{e}$ than under e^* then

$$\frac{c\left(\hat{e}\right) - c\left(e^*\right)}{p\left(s|\hat{e}\right) - p\left(s|e^*\right)} \geq w\left(s\right). \tag{2.4}$$

The first set of constraints (2.3) imply that the expected wage must satisfy

$$p\left(s|e^*\right) w\left(s\right) \geq \max_{\hat{e}:p(s|e^*)>p(s|\hat{e})} \frac{c\left(e^*\right) - c\left(\hat{e}\right)}{1 - \frac{p(s|\hat{e})}{p(s|e^*)}}.$$

In other words, the expected cost of implementing e^* for the principal (and its implicit benefit for the agent) is determined by the likelihood of signal s. The second set of constraints (2.4) place an upper bound on the wages the principal may deliver while respecting incentives of the agent. As we show below, this restriction can also be formulated in terms of likelihood ratios.

The above illustration only holds under the assumption that the principal only compensates the agent following a single signal s. We now show a more general version of this analysis for arbitrary compensation schemes. To this end, consider an arbitrary wage schedule $w(s)$ chosen by the principal – for any effort $e_i \in E$ chosen by the principal on- or off- equilibrium path . We may write the expected wage paid to the worker as

$$\begin{aligned}\sum_s w(s)\,p(s|e_i) &= \sum_s w(s)\,p(s|e^*)\,\frac{p(s|e_i)}{p(s|e^*)}\\ &= \sum_s w(s)\,p(s|e^*)\cdot\sum_s \frac{w(s)\,p(s|e^*)}{\sum_s w(s)\,p(s|e^*)}\frac{p(s|e_i)}{p(s|e^*)}\\ &= \sum_s w(s)\,p(s|e^*)\cdot\sum_s \alpha_s\frac{p(s|e_i)}{p(s|e^*)},\end{aligned}$$

where $\sum_s \alpha_s = 1$. Since the weights α_s do not depend on e_i, we may write the agent's interim incentive compatibility constraint as

$$\sum_s w(s)\,p(s|e^*)\cdot\sum_s \alpha_s\frac{p(s|e_i)}{p(s|e^*)} - c(e_i) \geq \sum_s w(s)\,p(s|e^*)\cdot\sum_s \alpha_s\frac{p(s|e_j)}{p(s|e^*)} - c(e_j)\,.$$

Therefore, if we define $\ell_i = 1 - \sum_s \alpha_s p(s|e_i)/p(s|e^*)$, this incentive constraint may be written as

$$\sum_s w(s)\,p(s|e^*)\cdot[\ell_j - \ell_i] \geq c(e_i) - c(e_j)\,, \forall j = 1,\cdots,|E|\,. \tag{2.5}$$

Writing the incentive constraint in this manner reveals that the choice of likelihood ratios, ℓ_j's, is sufficient to characterize payoffs of the principal and the agent and that the choice of contract $w(s)$ by the principal may be decomposed into a choice of $\sum_s w(s)\,p(s|e^*)$ as well as the choice of $\{\alpha_s\}$, or alternatively, ℓ_j's. In other words, if we represent the likelihood ratio profile with the following vector

$$\mathbf{l} = \left(\ell_1,\cdots,\ell_{|E|}\right),$$

then $\mathbf{l} \in$ convex hull $\left(\left\{\left(1-\frac{p(s|e_1)}{p(s|e^*)},\cdots,1-\frac{p\left(s|e_{|E|}\right)}{p(s|e^*)}\right)\right\}_{s\in S}\right) = \mathrm{co}\,(\mathbf{p})$. Thus the choice of the principal can be summarized by the overall level of the compensation $\sum_s p\left(s|e^*\right)w\left(s\right) = \overline{w}$ together with an element of the set co $(\mathbf{p})$. We thus have the following lemma:

Lemma 1. *Consider an implementable effort e^* and its associated information structure $\mathbf{p} = \{p\left(\cdot|\cdot\right)\}$. Then the cost to the principal from implementing any effort e_i is given by*

$$W(e_i, \mathbf{p}) = \min_{\mathbf{l}\in co(\mathbf{p})\cap\Omega_i}(1-\ell_i)\cdot\max_{j:\ell_j\geq\ell_i}\frac{c(e_i)-c(e_j)}{\ell_j-\ell_i}, \tag{2.6}$$

where

$$\Omega_i = \left\{\mathbf{l}\in\mathbb{R}^{|E|} : \max_{j:\ell_j\geq\ell_i}\frac{c\left(e_i\right)-c\left(e_j\right)}{\ell_j-\ell_i} \leq \min_{j:\ell_j<\ell_i}\frac{c\left(e_i\right)-c\left(e_j\right)}{\ell_j-\ell_i}\right\}. \tag{2.7}$$

The above lemma states that the problem of the principal can be reduced to choosing an effort level as well as a point in the convex hull of likelihood ratios – its intersection with the convex cone in (2.7). Note that not all likelihood ratio profiles $\mathbf{l}$ are feasible in the sense that there may be no compensation scheme that satisfies the agent's interim incentive compatibility constraint. As in our motivating example above, the likelihood ratios must permit a wage $w(s)$ that lies between the lower and upper bounds in (2.7). The convex cone Ω_i defines the set of likelihood ratio profiles that admit incentive compatible compensation schemes.

These results reveal that by choosing an information structure, the agent effectively determines the convex hull co $(\mathbf{p})$ and then the principal chooses a point that is in the intersection of co $(\mathbf{p})$ and the convex cone Ω_i. This result by itself does not make the analysis more tractable as it does not immediately describe the set of feasible convex hulls co$(\mathbf{p})$. The following proposition provides such a characterization. Note that, similar to co $(\mathbf{p})$, the convex hull co $(\mathbf{f})$ is the convex hull created by the points

$$\left\{\left(1-\frac{f\left(x|e_1\right)}{f\left(x|e^*\right)},\cdots,1-\frac{f\left(x|e_{|E|}\right)}{f\left(x|e^*\right)}\right)\right\}_{x\in X}.$$

Proposition 1. *For any information structure (S,π) with $|S|<\infty$, its associated co $(\mathbf{p})$ is a subset of co $(\mathbf{f})$ that contains the origin $\mathbf{0} = (0,\cdots,0)$. Additionally, for any convex* *subset C of* co $(\mathbf{f})$ *that contains the origin and has a finite set of extreme points, there exists an information structure* (S,π) *such that* co $(\mathbf{p}) = C$.

Proof. Let (S,π) be an information structure. Then,

$$\begin{aligned}\frac{p(s|e_i)}{p(s|e^*)} &= \frac{\sum_x f(x|e_i)\pi(s|x)}{\sum_x f(x|e^*)\pi(s|x)} = \sum_x \frac{f(x|e^*)\pi(s|x)}{\sum_x f(x|e^*)\pi(s|x)}\frac{f(x|e_i)}{f(x|e^*)} \\ &= \sum_x \beta_s(x)\frac{f(x|e_i)}{f(x|e^*)}, \forall i = 1,\cdots,|E|,\end{aligned}$$

where $\sum_x \beta_s(x) = 1$; $\beta_s(x)$ is the principal's posterior probability of x after observing s. The above implies that

$$\left(1-\frac{p(s|e_1)}{p(s|e^*)},\cdots,1-\frac{p(s|e_{|E|})}{p(s|e^*)}\right) = \sum_x \beta_s(x)\left(1-\frac{f(x|e_1)}{f(x|e^*)},\cdots,1-\frac{f(x|e_{|E|})}{f(x|e^*)}\right).$$

Hence, the left hand side of the above is a member of co $(\mathbf{f})$ for all $s \in S$. As a result co $(\mathbf{p}) \subset$ co $(\mathbf{f})$. Moreover, we have

$$\sum_s p(s|e^*)\left(1-\frac{p(s|e_i)}{p(s|e^*)}\right) = 1-\sum_s p(s|e_i) = 0, \forall i = 1,\cdots,|E|.$$

This implies that the convex set co$(\mathbf{p})$ includes the origin.

Now consider an arbitrary convex set $C \subset$ co $(\mathbf{f})$ that contains the origin with each of its members being of the form $\mathbf{z} = \left(z_1,\cdots,z_{|E|}\right)$. Let S be the set of extreme points of C. Then, since $\mathbf{0} \in C$, by Caratheodory theorem – see Rockafellar (1970) – there must exist $\{\tau_{\mathbf{z}}\}_{\mathbf{z}\in S}$ such that $\sum_{\mathbf{z}\in S}\tau_{\mathbf{z}} = 1$ and

$$0 = \sum_{\mathbf{z}\in S}\tau_{\mathbf{z}} z_i, \forall i = 1,\cdots,|E|. \tag{2.8}$$

By definition of co $(\mathbf{f})$, we must have $z_i \leq 1$ for all $\mathbf{z} \in C$. Moreover, since $\mathbf{z} \in$ co $(\mathbf{f})$, there must exist a subset $Y \subset X$ whose members are linearly independent together with

$\beta_{\mathbf{z}}(x)$ such that

$$\sum_{x\in Y}\beta_{\mathbf{z}}(x)=1,\beta_{\mathbf{z}}(x)\geq 0, z_i=\sum_{x\in Y}\beta_{\mathbf{z}}(x)\left[1-\frac{f(x|e_i)}{f(x|e^*)}\right],\forall i=1,\cdots,|E|.$$

Replacing the above in (2.8) leads to

$$0=\sum_{x\in Y}\sum_{\mathbf{z}\in S}\tau_{\mathbf{z}}\beta_{\mathbf{z}}(x)\left[1-\frac{f(x|e_i)}{f(x|e^*)}\right],\forall i=1,\cdots,|E|.$$

Since the points in Y are linearly independent and we also know that

$$0=\sum_{x\in Y}f(x|e^*)\left[1-\frac{f(x|e_i)}{f(x|e^*)}\right],\forall i=1,\cdots,|E|,$$

we must have that

$$f(x|e^*)=\sum_{\mathbf{z}}\tau_{\mathbf{z}}\beta_{\mathbf{z}}(x).$$

Let us define

$$\pi(\mathbf{z}|x)=\frac{\beta_{\mathbf{z}}(x)\tau_{\mathbf{z}}}{\sum_{\mathbf{z}\in S}\beta_{\mathbf{z}}(x)\tau_{\mathbf{z}}}=\frac{\beta_{\mathbf{z}}(x)\tau_{\mathbf{z}}}{f(x|e^*)}.$$

Then under $\pi(\mathbf{z}|x)$, we have

$$\begin{aligned}1-\frac{p(\mathbf{z}|e_i)}{p(\mathbf{z}|e^*)}&=1-\frac{\sum_x\frac{\beta_{\mathbf{z}}(x)\tau_{\mathbf{z}}}{f(x|e^*)}f(x|e_i)}{\sum_x\frac{\beta_{\mathbf{z}}(x)\tau_{\mathbf{z}}}{f(x|e^*)}f(x|e^*)}\\&=1-\frac{\sum_x\beta_{\mathbf{z}}(x)\tau_{\mathbf{z}}\frac{f(x|e_i)}{f(x|e^*)}}{\tau_{\mathbf{z}}\sum_x\beta_{\mathbf{z}}(x)}\\&=1-\sum_x\beta_{\mathbf{z}}(x)\frac{f(x|e_i)}{f(x|e^*)}=z_i,\forall i=1,\cdots,|E|,\end{aligned}$$

which concludes the proof. ■

Proposition 1 implies that any convex subset of $\operatorname{co}(\mathbf{f})$ that contains the origin can be chosen by the agent as $\operatorname{co}(\mathbf{p})$. This implies that instead of a choice of $\mathbf{p}$, we can focus on a choice of a convex subset of $\text{co}\left(\mathbf{f}\right)$ that contains the origin. Thus the agent-optimal information structure that implements e^* can be thought of as the equilibrium of the following game:

- **Stage 1.** The agent chooses a finite set of points L inside the convex set $\text{co}\left(\mathbf{f}\right)$ such that the convex hull of these points $\text{conv}\left(L\right)$ includes the origin.

- **Stage 2.** The principal chooses an effort level $e_i \in E$ and a point $\ell \in \text{conv}\left(L\right) \cap \Omega_i$ to maximize $\mathbb{E}\left[g\left(x\right)|e_i\right] - \left(1-\ell_i\right) \cdot \max_{j:\ell_j > \ell_i} \frac{c(e_i)-c(e_j)}{\ell_j - \ell_i}$.

- **Stage 3.** The choice of principal in stage 2 coincides with e^*.

The following example illustrates the construction of $\text{co}\left(\mathbf{f}\right)$ and the equilibrium response of the principal.

Example 1. Suppose that $X = \{x_1 = 0, x_2 = 1, x_3 = 2\}$ and $E = \{e_1, e_2, e_3\}$, where $c\left(e_1\right) = 0$, $c\left(e_2\right) = 0.1$, and $c\left(e_3\right) = 0.3$. The performance technology is given by

$$f = \begin{bmatrix} 0.35 & 0.50 & 0.15 \\ 0.05 & 0.50 & 0.45 \\ 0.10 & 0.15 & 0.75 \end{bmatrix},$$

where $f_{ij} = f\left(x_j|e_i\right)$. Moreover, the principal's payoff from realization of output x is equal to x, that is, $g(x) = x$. This yields

$$\mathbb{E}\left[g(x)|e_1\right] = 0.8,\ \mathbb{E}\left[g(x)|e_2\right] = 1.4,\ \mathbb{E}\left[g(x)|e_3\right] = 1.65.$$

Therefore, the first-best is to implement effort e_3. If we set $e^* = e_3$, then since $1 - f\left(x|e_3\right)/f\left(x|e^*\right) = 0$, we can embed the set $\text{co}\left(\mathbf{f}\right)$ in $\mathbb{R}^2$. The gray area in Figure 2.1 represents this set with a_i being the point associated with $x_i \in X$.

To understand the geometry of the game between the principal and the agent, consider a signal structure with 4 realizations given by $S = \{s_1, s_2, s_3, s_4\}$ and

$$\pi = \begin{bmatrix} 0.5 & 0.2 & 0.1 & 0.2 \\ 0.2 & 0.3 & 0.3 & 0.2 \\ 0.05 & 0.05 & 0.1 & 0.8 \end{bmatrix},$$

where $\pi_{ij} = \pi(s_j|x_i)$. We can use $p(s|e) = \sum_{x \in X} \pi(s|x) f(x|e)$ to construct the likelihood ratios. The green area in Figure 2.1 illustrates the convex set co($\mathbf{p}$).

To better illustrate the incentives in the geometric game between the principal and the agent, suppose the agent reveals x to the principal. As we have illustrated above, the choice of contract by the principal is equivalent to the choice of a point in the convex set co $(\mathbf{f})$.

To think about the incentives of the principal, if the principal is to implement e_3, he chooses $\mathbf{l} \in$ co $(\mathbf{f})$ to minimize its cost given by

$$\max\left\{\frac{c(e_3) - c(e_2)}{\ell_2}, \frac{c(e_3) - c(e_1)}{\ell_1}\right\} = \max\left\{\frac{0.2}{\ell_2}, \frac{0.3}{\ell_1}\right\}$$

The lower contour sets associated with the above cost function are convex cones in the shape of positive orthants – the shaded blue area highlighted in the right panel of Figure 2.1. For this example, the above cost function is minimized at $\mathbf{l}^*$. On the other hand, if the principal is to implement e_2, he chooses $\mathbf{l} \in$ co $(\mathbf{f})$ to minimize the cost given by $(1 - \ell_2)\frac{0.1}{\ell_1 - \ell_2}$ when $\ell_1 \geq \ell_2$. This cost is minimized at a_3. For e_3 to be implementable, the principal's payoff associated with e_2 – implied by the principal choosing a_3 – should be lower than his payoff associated with e_3– implied by the principal choosing$\mathbf{l}^*$. The set of likelihood ratios that satisfy this condition are highlighted by the blue shaded area in the right panel of Figure 2.1. Note that $\mathbf{l}^*$ is not contained in this area. As a result, by choosing to reveal x to the principal, the agent is unable to implement e_3. One can show that e_2 is implementable under full revelation of x.

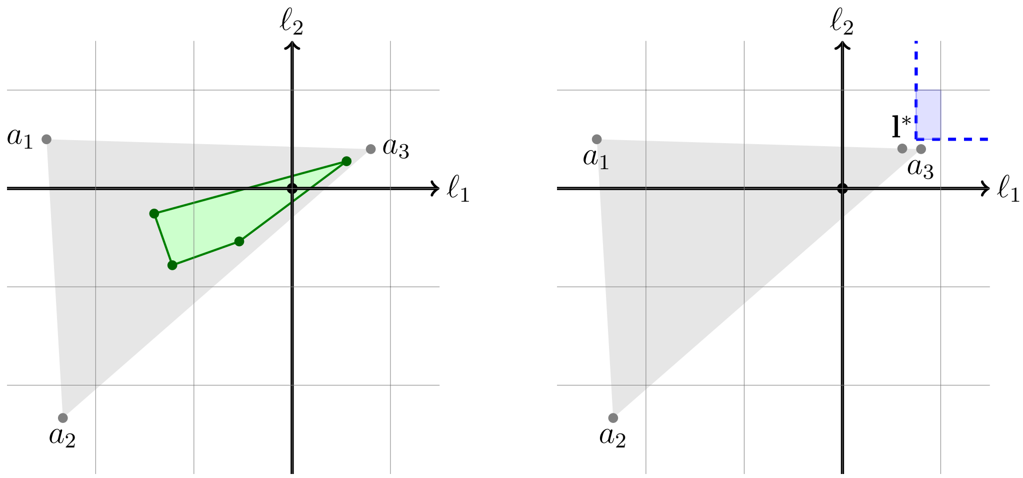


Figure 2.1: The geometric representation of information structures in the space of likelihood ratios. The left panel shows an example with 4 signals; the right panel shows the incentives of the principal when the agent reveals x.

2.5.1 Binary Information Structures

We now exploit our geometric interpretation to show that we may restrict the agent to choose information structures with at most two signals without loss of generality.

Proposition 2. *If e^* is implementable by some information structure (S, π) and delivers expected wage $W(e^*, \pi)$ to the agent, then e^* is also implementable by a binary information structure $\left(\hat{S}, \hat{\pi}\right)$ with $|\hat{S}| = 2$ and $W(e^*, \hat{\pi}) = W(e^*, \pi)$.*

Proof. Suppose the effort level e^* is implementable by (S, π) and consider the optimization problem in (2.6). Let $\mathbf{l}^* \in \mathrm{co}(\mathbf{p})$ be the optimal choice for the principal when choosing e^*. Given the definition of $\mathrm{co}(\mathbf{p})$ and the geometric game described above, there must exist $\{\alpha_s\}_{s\in S}$ such that $\alpha_s \geq 0$, $\sum_{s\in S} \alpha_s = 1$ and

$$\ell_i^* = 1 - \sum_{s\in S} \alpha_s \frac{p(s|e_i)}{p(s|e^*)}, \forall i = 1, \cdots, |E|.$$

Let the information structure $\left(\hat{S}, \hat{\pi}\right)$ be defined as $\hat{S} = \{L, H\}$ and

$$\hat{\pi}\left(H|x\right) = \sum_{s} \beta_s \pi\left(s|x\right), \hat{\pi}\left(L|x\right) = 1 - \hat{\pi}\left(H|x\right),$$

where

$$\beta_s = \frac{\alpha_s / p\left(s|e^*\right)}{\sum_s \alpha_s / p\left(s|e^*\right)}.$$

The probability function given this information structure satisfies

$$\begin{aligned}
1 - \frac{\hat{p}\left(H|e_i\right)}{\hat{p}\left(H|e^*\right)} &= 1 - \frac{\sum_x \hat{\pi}\left(H|x\right) f\left(x|e_i\right)}{\sum_x \hat{\pi}\left(H|x\right) f\left(x|e^*\right)}, \\
&= 1 - \frac{\sum_{s\in S} \frac{\alpha_s}{p(s|e^*)} \sum_x \pi\left(s|x\right) f\left(x|e_i\right)}{\sum_{s\in S} \frac{\alpha_s}{p(s|e^*)} \sum_x \pi\left(s|x\right) f\left(x|e^*\right)} \\
&= 1 - \frac{\sum_{s\in S} \frac{\alpha_s}{p(s|e^*)} p\left(s|e_i\right)}{\sum_{s\in S} \frac{\alpha_s}{p(s|e^*)} p\left(s|e^*\right)} \\
&= 1 - \sum_{s\in S} \frac{\alpha_s}{p\left(s|e^*\right)} p\left(s|e_i\right) = \ell_i^*, \forall i = 1, \cdots, |E|.
\end{aligned}$$

The above implies that if the principal is to choose e^*, $\mathbf{l}^*$ is feasible under the new information structure $\left(\hat{S}, \hat{\pi}\right)$, i.e., $\mathbf{l}^* \in \mathrm{co}\left(\hat{\mathbf{p}}\right)$. Hence, in order to establish our claim, it is sufficient to show that for any alternative $e_i \neq e^*$, $W\left(e_i, \hat{\pi}\right) \geq W\left(e_i, \pi\right)$. To show this, it is sufficient to show that

$$\left(1 - \frac{\hat{p}\left(L|e_1\right)}{\hat{p}\left(L|e^*\right)}, \cdots, 1 - \frac{\hat{p}\left(L|e_{|E|}\right)}{\hat{p}\left(L|e^*\right)}\right) \in \mathrm{co}\left(\mathbf{p}\right)$$

This would imply that $\mathrm{co}\left(\hat{\mathbf{p}}\right)$ is a subset of $\mathrm{co}\left(\mathbf{p}\right)$ since $\mathrm{co}\left(\hat{\mathbf{p}}\right)$ is the line that connects the above point to $\mathbf{l}^*$.

We have

$$\begin{aligned}
1-\frac{\hat{p}\left(L|e_i\right)}{\hat{p}\left(L|e^*\right)} &= 1-\frac{\sum_x\left(1-\hat{\pi}\left(H|x\right)\right)f\left(x|e_i\right)}{\sum_x\left(1-\hat{\pi}\left(H|x\right)\right)f\left(x|e^*\right)} \\
&= 1-\frac{\sum_x\sum_s\left(1-\beta_s\right)\pi\left(s|x\right)f\left(x|e_i\right)}{\sum_x\sum_s\left(1-\beta_s\right)\pi\left(s|x\right)f\left(x|e^*\right)} \\
&= 1-\frac{\sum_s\left(1-\beta_s\right)p\left(s|e_i\right)}{\sum_s\left(1-\beta_s\right)p\left(s|e^*\right)} \\
&= 1-\sum_s\frac{\left(1-\beta_s\right)p\left(s|e^*\right)}{\sum_s\left(1-\beta_s\right)p\left(s|e^*\right)}\frac{p\left(s|e_i\right)}{p\left(s|e^*\right)},
\end{aligned}$$

which establishes the claim. This concludes the proof. ■

We can describe the intuition behind the above proof graphically. Consider an implementable effort e^* and suppose that its associated information structure is as depicted in Figure 2.2; the green area represents the convex hull corresponding to this information structure, $\text{co}\left(\mathbf{p}\right)$. Suppose that point b is the point of optimality for the principal in $\text{co}\left(\mathbf{p}\right)$ if he were to implement e^*. The red line represents the new information structure $\hat{\pi}$ – since there are only two signal realizations, this has to be a line. If the principal is to implement e^*, since $b \in \text{co}\left(\hat{\mathbf{p}}\right)$ and b is chosen under π, b remains optimal. Moreover, since $\text{co}\left(\hat{\mathbf{p}}\right) \subset \text{co}\left(\mathbf{p}\right)$, for any other effort $e_i \neq e^*$, minimized cost under $\hat{\pi}$ must be at least as high as that under π. This, in turn, implies that e^* is implementable under $\hat{\pi}$ and implements the same outcome as under π.

The above observations imply that if the principal prefers to implement effort e^* under the information structure π, he would prefer the same under the binary information structure $\hat{\pi}$ and the expected wage he has to pay to achieve this is the same under both information structures. In what follows, we will use this result to characterize the optimal information structures.

Example 1 (Continued). Recall Example 1 in which we argued that the principal – under full revelation – chooses to implement e_2 and thus his choice is inefficient (since total surplus under e_2 is lower than that under e_3.)

It still remains to be seen whether e_3 is implementable and how well the agent can

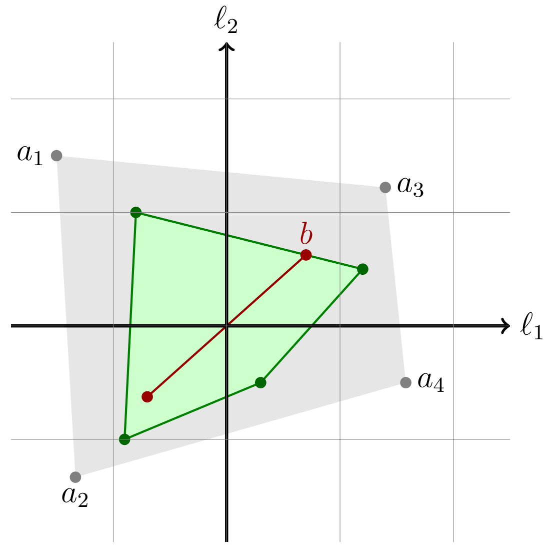


Figure 2.2: The intuition for the construction of a two-point signal

do by choosing an information structure. By Proposition 2, we can focus on information structures that only have two support points. Note that since the expected output under e_1 is $\mathbb{E}[g(x)|e_1] = 0.8$ and its cost is 0, the principal can always guarantee $\mathbb{E}[g(x)|e_1] = 0.8$.

Now consider the point $\hat{\mathbf{l}}$ satisfying

$$\frac{0.2}{\hat{\ell}_2} = \frac{0.3}{\hat{\ell}_1} = 0.85$$

At this point, the payoff to the principal of implementing each effort level is given by

$$\begin{aligned}
e_3 &: 1.65 - \max\left\{\frac{0.2}{\hat{\ell}_2}, \frac{0.3}{\hat{\ell}_1}\right\} = 0.8 \\
e_2 &: 1.40 - \left(1 - \hat{\ell}_2\right) \cdot \frac{0.1}{\hat{\ell}_1 - \hat{\ell}_2} = 1.40 - 0.65 = 0.75 \\
e_1 &: 0.8 - 0 = 0.8
\end{aligned}$$

This implies that at $\hat{\mathbf{l}}$, e_3 maximizes the payoff of the principal. Thus if $\hat{\mathbf{l}} \in \mathrm{co}\,(\mathbf{f})$, we

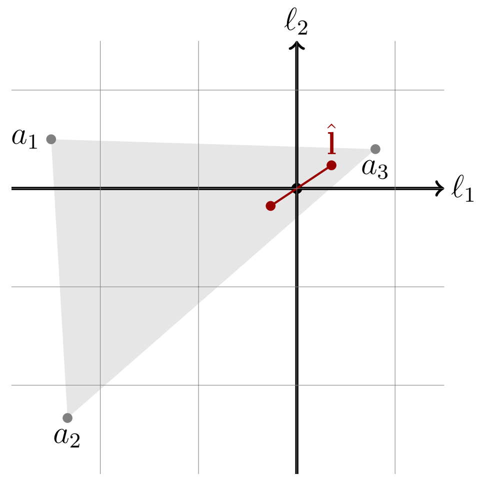


Figure 2.3: Agent-optimal information structure in Example 1

can choose an information structure that implements e_3. In this example, this is indeed the case. Figure 2.3 depicts the point $\hat{\mathbf{l}}$ as well as a two-point information structure that implements it; the red line going through $\hat{\mathbf{l}}$. If we set $S = \{L, H\}$ and

$$\pi\left(H|x\right) = \begin{cases} \frac{27}{83}, & \text{if } x = x_1 \\ \frac{15}{83}, & \text{if } x = x_2 \\ \frac{41}{83}, & \text{if } x = x_3 \end{cases}, \pi\left(L|x\right) = 1 - \pi\left(H|x\right),$$

it can readily be checked that (S, π) is associated with the one depicted in Figure 2.3.

The above example illustrates the agent's power the choice of the information structure provides. By choosing the aforementioned information structure, not only the agent is able to implement e_3, i.e., the efficient level of effort, but also she is able to capture the entire surplus. In what follows, we show that under some conditions on the performance technology, this is always possible.

2.6 Efficient Surplus Extraction

In this section, we provide sufficient conditions on the performance technology $f(\cdot|\cdot)$ and cost function $c(\cdot)$ so that the agent is able to implement the first-best effort and extract the entire surplus.

Let e^* be the first-best level of effort that satisfies

$$e^* \in \arg\max_{e\in E} \quad \mathbb{E}\left[g(x)|e\right] - c(e).$$

Suppose that the agent wishes to implement e^* and capture the entire surplus given by $\mathbb{E}\left[g(x)|e^*\right] - c(e^*) - \mathbb{E}\left[g(x)|e_1\right]$. For this to occur, we need to choose $\mathrm{co}(\mathbf{p})$ and $\mathbf{l}^* \in \mathrm{co}(\mathbf{p})$ that satisfy

$$\begin{aligned}\mathbb{E}\left[g(x)|e_1\right] &= \mathbb{E}\left[g(x)|e^*\right] - \max_{i:\ell_i^*\geq 0}\frac{c(e^*)-c(e_i)}{\ell_i^*}\\ &\geq \max_{e_j\in E,\mathbf{l}\in\mathrm{co}(\mathbf{p})\cap\Omega_j}\mathbb{E}\left[g(x)|e_j\right] - (1-\ell_j)\max_{i:\ell_i>\ell_j}\frac{c(e_j)-c(e_i)}{\ell_i-\ell_j}\end{aligned}$$

Consider the likelihood vector $\mathbf{l}^*$ that satisfies the following property:

$$\mathbb{E}\left[g(x)|e^*\right] - \mathbb{E}\left[g(x)|e_1\right] = \frac{c(e^*)-c(e_i)}{\ell_i^*}, \forall i = 1,\cdots,|E|.$$

In words, this is a point in which the agent is indifferent among all the efforts – see the reformulation of (IC-A) in (2.5). Moreover, the principal's payoff is his outside option and thus if the principal chooses to implement e^* and chooses $\mathbf{l}^* \in \mathrm{co}(\mathbf{p})$, the agent captures the entire surplus. In the following proposition, we show that if $\mathbf{l}^* \in \mathrm{co}(\mathbf{f})$, we can construct an information structure – and its associated $\mathrm{co}(\mathbf{p})$ – in which $\mathbf{l}^*$ is the best choice for the principal and e^* is the best level of effort.

Theorem 1. *Suppose that $\mathbf{l}^* \in co(\mathbf{f})$. Then e^* is implementable and there exists an information structure (S,π) for which the agent's payoff is $U_A = \mathbb{E}\left[g(x)|e^*\right] - \mathbb{E}\left[g(x)|e_1\right] - c(e^*)$ and the principal's payoff is $U_P = \mathbb{E}\left[g(x)|e_1\right]$, i.e., the agent can capture the entire* *surplus.*

Proof. First, we note that by Lemma 2 in Appendix A, $\mathbf{0} \in \text{co}\left(\mathbf{f}\right)$ is an interior point of this convex set, where interiority is defined in an appropriatley defined subspace of $\mathbb{R}^m$. We define the following convex hull (and by Proposition 1, its associated information structure):

$$\text{co}\left(\mathbf{p}\right) = \left\{\lambda \mathbf{l}^* + \left(1-\lambda\right)\left(-\alpha \mathbf{l}^*\right) : \forall \lambda \in \left[0,1\right]\right\},$$

where $\alpha > 0$ is such that $-\alpha \mathbf{l}^* \in \text{co}\left(\mathbf{f}\right)$. Such an α always exists due to $\mathbf{0}$ being an interior point. Since $\text{co}\left(\mathbf{p}\right)$ is a line through $\mathbf{l}^*$ and the origin, all of its members must satisfy

$$\forall i,j, \quad \frac{\ell_i}{\ell_j} = \frac{c\left(e^*\right) - c\left(e_i\right)}{c\left(e^*\right) - c\left(e_j\right)} = \frac{\ell_i^*}{\ell_j^*} \Rightarrow \frac{\ell_i}{\ell_1} = \frac{c\left(e^*\right) - c\left(e_i\right)}{c\left(e^*\right)}, \forall i. \tag{2.9}$$

This implies that the cost of choosing any point $\mathbf{l} \in \text{co}\left(\mathbf{p}\right)$ to implement any effort $e_j \in E$ is given by

$$\begin{aligned}
\left(1-\ell_j\right) \max_{i:\ell_i \geq \ell_j} \frac{c\left(e_j\right) - c\left(e_i\right)}{\ell_i - \ell_j} &= \left(1 - \frac{c\left(e^*\right) - c\left(e_j\right)}{c\left(e^*\right)} \ell_1\right) \frac{c\left(e^*\right)}{\ell_1} \\
&= \frac{c\left(e^*\right)}{\ell_1} + c\left(e_j\right) - c\left(e^*\right).
\end{aligned}$$

Since choice of $\mathbf{l}$ under e_j must be a member of the cone Ω_j, we must have $\ell_j \leq \ell_1$ and this combined with (2.9) implies that $\ell_1 \geq 0$. Thus the above expression is minimized at $\mathbf{l}^*$. Hence, the highest payoff of the principal from choosing e_j is given by

$$\begin{aligned}
\mathbb{E}\left[g\left(x\right)|e_j\right] - c\left(e_j\right) - \frac{c\left(e^*\right)}{\ell_1^*} + c\left(e^*\right) = \\
\mathbb{E}\left[g\left(x\right)|e_j\right] - c\left(e_j\right) + c\left(e^*\right) - \mathbb{E}\left[g\left(x\right)|e^*\right] + \mathbb{E}\left[g(x)|e_1\right]
\end{aligned}$$

The above is maximized at $e_j = e^*$, which implies that e^* can be implemented with the chosen information structure. Moreover, the payoff of the principal is given by $\mathbb{E}\left[g\left(x\right)|e_1\right]$. This concludes the proof. ∎

The proof of the above theorem emphasizes the power of the agent when $\mathbf{l}^* \in \text{co}\left(\mathbf{f}\right)$.

By being able to control the information structure, the agent can control the wage that is needed for the principal to implement his desired effort. By choosing $\mathbf{l}^*$, the agent is forcing the principal to fully compensate the agent for her cost of effort. This implies that the payoff of the principal becomes total surplus shifted by a constant. Hence, it is optimal to choose the first best level of effort.

While Theorem 1 provides sufficient conditions for implementability of the first-best effort and full surplus extraction, it is not immediately evident what it imposes on the structure of the model. In what follows, we try to shed light on this.

Almost Perfect Performance Technology. Suppose that the performance technology satisfies the following property

$$\begin{aligned} X = E \subset \mathbb{R}_+, e_1 = 0, \\ f(e_j|e_j) = 1 - (m-1)\varepsilon, \forall j = 1, \cdots, m, \\ f(e_i|e_j) = \varepsilon, \forall i \neq j, \end{aligned}$$

where $1/(m-1) > \varepsilon > 0$. In words, the above performance technology puts probability ε on $e_i \neq e_j$ if e_j is chosen. As ε converges to 0, it converges to a setting where observing $x \in X$ fully reveals e, i.e., *a perfect performance technology*. Note that for ε small enough, the first-best level of effort is given by

$$e_j \in \arg\max_{e \in E} e - c(e)$$

Suppose that in the above $e_j > e_1 = 0$. Moreover, as ε converges to 0, the point $\mathbf{l}^*$ converges to

$$\ell_i^* = \frac{c(e_j) - c(e_i)}{e_j}$$

Finally, note that for any ε, the set $\text{co}(\mathbf{f})$ is the convex hull of the following likelihood ratios:

$$1-\frac{f(e_k|e_i)}{f(e_k|e_j)}=\begin{cases}1-\frac{\varepsilon}{1-(m-1)\varepsilon}, & \text{if } k=j, k\neq i\\ 0, & \text{if } k=i=j\\ 0, & \text{if } k\neq j, k\neq i\\ m-\frac{1}{\varepsilon}, & \text{if } k\neq j, k=i\end{cases}$$

As ε converges to 0, the above set gets larger and the point associated with $k=j$ converges to

$$\left(\underbrace{1,\cdots,1}_{j-1\text{ times}},0,1,\cdots,1\right),$$

while the points associated with $k\neq j$ converge to

$$\left(\underbrace{0,\cdots,0}_{k-1\text{ times}},-\infty,0,\cdots,0\right).$$

This implies that as ε converges to 0, $\text{co}(\mathbf{f})\to(-\infty,1]^{j-1}\times\{0\}\times(-\infty,1]^{m-j}$. This is depicted in Figure 2.4. Since $\ell_i^*\leq\frac{c(e_j)}{e_j}<1$, for ε small enough, we must have that $\mathbf{l}^*\in\text{co}(\mathbf{f})$. We thus have the following corollary to Theorem 1:

Corollary 1. *Let $f(\cdot|\cdot)$ be an almost perfect performance technology satisfying*

$$\begin{aligned}X=E\subset\mathbb{R}_+,e_1=0,\\ f(e_j|e_j)\leq1-(m-1)\varepsilon,\forall j=1,\cdots,m,\\ f(e_i|e_j)\geq\varepsilon,\forall i\neq j.\end{aligned}$$

Then there exists $\overline{\varepsilon}$ such that for all $\varepsilon\leq\overline{\varepsilon}$, the agent can implement the first-best level of effort and capture the entire surplus.

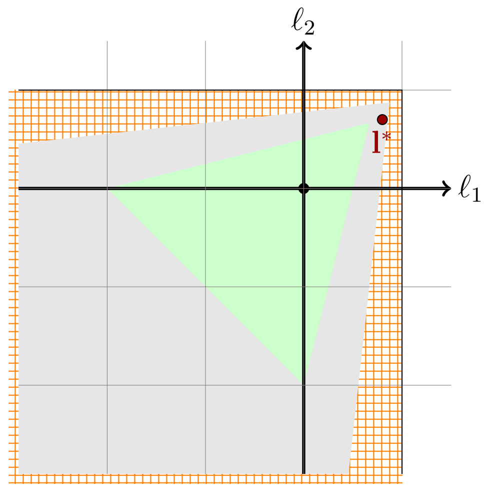


Figure 2.4: Almost Perfect Performance Technology for $m = 3$. The grey and green shaded areas represent $\mathrm{co}\,(\mathbf{f})$ as ε becomes smaller. As $\varepsilon \to 0$, $\mathrm{co}\,(\mathbf{f})$ converges to the shaded quarter-space south-west of $(1, 1)$.

2.7 Continuous Effort and Output

In this section, we consider a version of the model from Section 2.4 where the effort space and the output space are continuous. Applying a first-order approach, we derive sufficient conditions such that the optimal indicator structure takes the form of monotone or hump-shaped threshold signals. In other words, we derive conditions such that the optimal information structure is characterized by an indicator with at most two thresholds.

Consider a principal employing an agent to perform a task whose output is represented by a real number $x \in X = [0, 1]$. The agent chooses effort $e \in E = [0, 1]$ to perform the task. The agent's effort choice induces a probability distribution $f(x|e)$ over the output space X, whose density function $f(x|e)$ is assumed to be twice differentiable with respect to e for every $x \in X$. Effort is costly to the agent; the cost of effort e is given by $c(e)$ for some real-valued function $c : E \to \mathbb{R}_+$ where $c'(e) \geq 0,\ c''(e) > 0,\ \forall e \in E$. For simplicity, let $c(0) = 0$. The timing and payoff functions of the game are otherwise the same as in the model in Section 2.4. To proceed, we assume that the first-order approach is valid for all the optimization problems faced by the agent and the principal.

We characterize this continuous version of the game in the same manner as the discrete case. Given the result in Proposition 2, one can use an approximation argument to show that optimal signal is binary, i.e., $S = \{L, H\}$. As a result, let $p(e) : E \to [0, 1]$ represent the probability of $s = H$ induced by a given information structure (S, π) and the underlying probability distribution of outcomes given effort $f(x|e)$. Note, for any level of effort e, $p(e) = \int_0^1 f(x|e)\,\pi(H|x)\,dx$. Differentiating the stochastic mapping p with respect to e yields

$$p'(e) = \int_0^1 f_e(x|e)\,\pi(H|x)\,dx, \ \forall e \in E.$$

The agent's problem in the last stage of the game, given an information structure and a compensation scheme, is

$$\max_{e \in E} \ p(e)\,w(H) + [1 - p(e)]\,w(L) - c(e)$$

The first-order condition characterizing the agent's optimal effort is

$$p'(e)\,[w(H) - w(L)] = c'(e).$$

For any desired level of effort, the principal chooses a compensation scheme that solves

$$\begin{aligned} &\min_{w} p(e)\,w(H) + (1 - p(e))\,w(L) \\ &\text{s.t } p'(e)\,[w(H) - w(L)] = c'(e)\,, \\ &\quad w(H)\,, w(L) \geq 0. \end{aligned} \tag{2.10}$$

If $\lambda(e, \pi)$ denotes the Lagrange multiplier on the agent's incentive compatibility constraint in (2.10), then

$$\lambda(e,\pi)=\min\left\{\frac{p\left(e\right)}{p'\left(e\right)},-\frac{1-p\left(e\right)}{p'\left(e\right)}\right\}.$$

Exactly as in the discrete case, the above reveals that the agent's choice of information structure, which induces a set of likelihood ratios dependent on the stochastic mapping p, determines the principal's shadow cost of incentivizing a given effort level.

If we define the expected payoff to the principal from a given effort level e as $\mathbb{E}\left[g(x)|e\right]= g(x)f(x|e)dx$, then the principal's optimal choice of effort solves

$$\max_{e}\mathbb{E}\left[g\left(x\right)|e\right]-\lambda\left(e,\pi\right)c'\left(e\right)$$

with associated optimality condition

$$\frac{\partial\mathbb{E}\left[g(x)|e\right]}{\partial e}-\frac{\partial\lambda(e,\pi)}{\partial e}c'(e)-\lambda(e,\pi)c''(e)=0.$$

Using the above sequence of optimality conditions, we obtain the following optimization problem that describes the agent's choice of effort and information structure in the first stage of the game:

$$\begin{aligned}
&\max_{e,\pi}\lambda(e,\pi)c'(e)-c(e)\text{ s.t.}\\
&\quad\frac{\partial\mathbb{E}[g(x)|e]}{\partial e}-\frac{\partial\lambda(e,\pi)}{\partial e}c'(e)-\lambda(e,\pi)c''(e)=0,\\
&\quad\frac{1}{\lambda(e,\pi)}=\max\left\{\frac{\int_0^1 f(x|e)\pi(x)dx}{\int_0^1 f_e(x|e)\pi(x)dx},-\frac{1-\int_0^1 f(x|e)\pi(x)dx}{\int_0^1 f_e(x|e)\pi(x)dx}\right\}.
\end{aligned}\tag{2.11}$$

The following assumption allows us to provide a sharp characterization of the optimal information structure:

Assumption 2. *Given any effort $e\in E$, the likelihood $\frac{f_e(x|e)}{f(x|e)}$ is strictly monotone in output x and its derivative $\frac{\partial}{\partial e}\frac{f_e(x|e)}{f(x|e)}$ is a convex function of the likelihood $\frac{f_e(x|e)}{f(x|e)}$.*

Several distribution functions satisfy Assumption 2. Examples include power distributions: $f(x|e) = ex^{e-1}$, and truncated exponential distributions: $f(x|e) = \frac{d}{dx}\frac{e^x-1}{e-1}$.

Our main characterization of the optimal information structure is as follows:

Proposition 3. *Suppose that Assumption 2 holds. Then, the equilibrium information structure is characterized by at most two thresholds in the output space. If the equilibrium information structure has a single threshold, say x^*, then $\pi(H|x) = 1$ if and only if $x \geq x^*$. If the equilibrium information structure has two thresholds, say (x_1^*, x_2^*) then $\pi(H|x) = 1$ if and only if $x \in [x_1^*, x_2^*]$.*

Proof. See Appendix B. ■

Given an effort choice, the agent's optimal choice of information structure requires the agent to choose the probability of being paid at any output such that there is no net marginal benefit from a marginal change in this probability. Generally, it is not possible to satisfy this condition for all the output levels. The agent ends up choosing the extreme probabilities for every output, depending on whether her net marginal benefit is increasing or decreasing in the probability of being paid at that output level. Assumption 2 ensure that the output intervals at which the agent uses either of the extreme probabilities are structured such that the optimal information structure takes the form of monotone or hump-shaped thresholds signals.

Example 2. Let $f(x|e) = 3ex^{3e-1}$. As mentioned earlier, this distribution satisfies Assumption 2. If the effort cost is $c(e) = \frac{e^2}{2}$, the equilibrium information structure is characterized by $x^* = 0.45$ where $\pi(H|x) = 1$ if and only if $x \geq x^*$. The principal pays the agent $w = 0.1048$ if he receives the high signal; otherwise, he pays nothing. This choice implements effort $e^* = 0.2725$ yielding the following payoffs: $U_P = 0.3450$, $U_A = 0.0677$. For comparison, the first-best effort is $e = 0.4708$. In this case, and opposite to that of section 2.6, the agent is unable to implement the efficient outcome.

2.8 Conclusion

In this paper, we have developed the theoretical tool in the design of contracts in principal-agent settings. Part of the contract design is what indicators should be used as contingencies for payments. Despite the importance of this question and its relevance for analysis of contracting decisions, this part of the contracting procedure is less explored.

Our paper has two broad implications. First, our methodology of thinking about likelihood ratios can be applied to other settings in which considerations of communication off the equilibrium path are important. In our setup, unlike other models of communication and information design, the design of information structure for the equilibrium level of effort affects off-path communication and the ability of the principal and the agent to capture surplus. This can arise in other settings with strategic information transmission and our method can be useful for that.

Second, our paper ties the choice of indicators in contracting to the bargaining power of the parties. In the textbook moral hazard problem, principal makes a take-it-or-leave it offer. As a result, a version of the informativeness principle often holds; the principal wishes to use all the information available, if possible, as contingency for payments. In contrast, in our model, the agent has different incentives for choice of indicators. There are some casual observations that are in line with this explanation. For example, contracts in NFL are often extremely detailed and payments to football players are highly contingent on various measures of individual and team outcomes. In contrast, contracts found in the English Premier League are not as detailed. They are often contingent on very coarse personal outcomes such as the number of goals scored reaching a particular threshold. In light of our theory, the level of competition in English/European soccer (in the form of increased player bargaining power) compared to a lack thereof in NFL could be behind this observation. Future work can hopefully shed light on the importance of this channel in the data.

Bibliography

Ball, I. (2019): "Scoring strategic agents," *arXiv preprint arXiv:1909.01888.*

Barron, D., G. Georgiadis, and J. Swinkels (2020): "Optimal contracts with a risk-taking agent," *Theoretical Economics*, 15, 715–761.

Bebchuk, L. A. and J. M. Fried (2004): *Pay without performance: The unfulfilled promise of executive compensation*, Harvard University Press.

Boleslavsky, R. and K. Kim (2018): "Bayesian persuasion and moral hazard," *Available at SSRN 2913669.*

Carroll, G. (2015): "Robustness and linear contracts," *American Economic Review*, 105, 536–563.

Chaigneau, P., A. Edmans, , and D. Gottlieb (2019): "The informativeness principle without the first-order approach," *Games and Economic Behavior*, 113, 743–755.

Dewatripont, M., I. Jewitt, and J. Tirole (1999): "The economics of career concerns, part I: Comparing information structures," *The Review of Economic Studies*, 66, 183–198.

Garrett, D., G. Georgiadis, A. Smolin, and B. Szentes (2020): "Optimal Technology Design," *Available at SSRN 3720594.*

Georgiadis, G. and B. Szentes (2020): "Optimal monitoring design," *Econometrica*, 88, 2075–2107.

Halac, M., E. Lipnowski, and D. Rappoport (2021): "Rank uncertainty in organizations," *American Economic Review*, 111, 757–86.

Holmström, B. (1979): "Moral hazard and observability," *The Bell Journal of Economics*, 74–91.

——— (1999): "Managerial incentive problems: A dynamic perspective," *The Review of Economic Studies*, 66, 169–182.

Innes, R. D. (1990): "Limited liability and incentive contracting with ex-ante action choices," *Journal of economic theory*, 52, 45–67.

Kamenica, E. and M. Gentzkow (2011): "Bayesian persuasion," *American Economics Review*, 101, 2590–2615.

Perez-Richet, E. and V. Skreta (2022): "Test design under falsification," *Econometrica*, 90, 1109–1142.

Poblete, J. and D. Spulber (2012): "The form of incentive contracts: agency with moral hazard, risk neutrality, and limited liability," *The RAND Journal of Economics*, 40, 215–234.

Rockafellar, R. T. (1970): *Convex analysis*, vol. 18, Princeton University Press.

Rosar, F. (2017): "Test design under voluntary participation," *Games and Economic Behavior*, 104, 632–655.

Saeedi, M. and A. Shourideh (2020): "Optimal rating design," *arXiv preprint arXiv:2008.09529*.

Walton, D. and G. Carroll (2022): "A General Framework for Robust Contracting Models," *Econometrica*, 90, 2129–2159.

Zapechelnyuk, A. (2020): "Optimal quality certification," *American Economic Review: Insights*, 2, 161–176.

2.9 Appendix A: Interiority of the Origin

Lemma 2. *Let T represent the lowest-dimensional linear subspace in $\mathbb{R}^{|E|}$ that contains $co\,(\mathbf{f})$. If $f(\cdot|e^*)$ is full-support, then the origin is an interior point of $co\,(\mathbf{f})$ with respect to T.*

Proof. Notice that the origin can be written as a convex combination of the points defining co $(\mathbf{f})$ with weights $f(x|e^*)$:

$$\sum_x f(x|e^*)\left(1 - \frac{f(x|e_i)}{f(x|e^*)}\right) = \sum_x f(x|e^*) - \sum_x f(x|e_i) = 1 - 1 = 0,\ \forall i,$$

where $f(x|e^*) > 0,\ \forall x \in X$ by the full-support assumption. Therefore, the origin is always included in the convex set co $(\mathbf{f})$. Suppose by contradiction that the origin is not an interior point of co $(\mathbf{f})$ with respect to T. By the supporting hyperplane theorem, there exists a hyperplane in the linear subspace T that contains the origin and co $(\mathbf{f})$ is entirely contained in one of the two closed half-spaces bounded by the hyperplane. However, since $f(x|e^*) > 0,\ \forall x \in X$, this is possible only if co $(\mathbf{f})$ is entirely contained in this hyperplane. This is a contradiction because T is the lowest-dimensional linear subspace in $\mathbb{R}^{|E|}$ that contains $co\,(\mathbf{f})$. ■

2.10 Appendix B: Proof of Proposition 3

Proof. Let $\pi(x) = \pi(H|x),\ \forall x \in X$. Without loss of generality, for any desired implementable effort level e, we impose $\int_X f_e(x|e)\pi(x)dx \geq 0$. Consequently, we have

$$\lambda(e, \pi) = \frac{\int_X f(x|e)\pi(x)dx}{\int_X f_e(x|e)\pi(x)dx} \geq 0.$$

We may then write the agent's problem in (2.11) as

$$
\begin{aligned}
\max_{e,\pi}\ & \lambda(e,\pi)c'(e) - c(e) \text{ s.t.} \\
& \frac{\partial E[g(x)|e]}{\partial e} - \frac{\partial \lambda(e,\pi)}{\partial e}c'(e) - \lambda(e,\pi)c''(e) = 0, \\
& \lambda(e,\pi) = \frac{\int_X f(x|e)\pi(x)dx}{\int_X f_e(x|e)\pi(x)dx}, \\
& \lambda(e,\pi) \geq 0,\ 0 \leq \pi(x) \leq 1,\ \forall x \in X.
\end{aligned} \tag{2.12}
$$

We write the Lagrangian corresponding to (2.12), momentarily ignoring the inequality constraints:

$$
\mathcal{L}(e,\pi,\eta) = \lambda(e,\pi)c'(e) - c(e) + \eta\left[\frac{\partial \mathbb{E}[g(x)|e]}{\partial e} - \frac{\partial \lambda(e,\pi)}{\partial e}c'(e) - \lambda(e,\pi)c''(e)\right].
$$

For every output $\tilde{x} \in X$, the agent's optimal information structure in (2.12) must satisfy

$$
\begin{aligned}
& \frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(\tilde{x})} = 0, \text{ if } 0 < \pi(\tilde{x}) < 1, \\
& \frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(\tilde{x})} \geq 0, \text{ if } \pi(\tilde{x}) = 1, \\
& \frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(\tilde{x})} \leq 0, \text{ if } \pi(\tilde{x}) = 0.
\end{aligned} \tag{2.13}
$$

With some algebra, we get

$$
\frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(\tilde{x})} = f(\tilde{x}|e)\left[A_1(e,\pi,\eta) - A_2(e,\pi,\eta)\frac{f_e(\tilde{x}|e)}{f(\tilde{x}|e)} + A_3(e,\pi,\eta)\frac{f_{ee}(\tilde{x}|e)}{f(\tilde{x}|e)}\right] \tag{2.14}
$$

where A_1, A_2, A_3 are some functions independent of $\tilde{x}$. Note that since

$$
\frac{f_{ee}(x|e)}{f(x|e)} = \frac{\partial}{\partial e}\frac{f_e(x|e)}{f(x|e)} + \left(\frac{f_e(x|e)}{f(x|e)}\right)^2,
$$

we can write (2.14) as

$$\frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(\tilde{x})} = f(\tilde{x}|e)\Bigg[A_1(e,\pi,\eta) - A_2(e,\pi,\eta)\frac{f_e(\tilde{x}|e)}{f(\tilde{x}|e)} + \\ A_3(e,\pi,\eta)\left(\frac{f_e(\tilde{x}|e)}{f(\tilde{x}|e)}\right)^2 + A_3(e,\pi,\eta)\frac{\partial}{\partial e}\frac{f_e(\tilde{x}|e)}{f(\tilde{x}|e)}\Bigg]. \tag{2.15}$$

Since $\frac{f_e(x|e)}{f(x|e)}$ is monotone, the function in (2.15) inherits the curvature of $\frac{\partial}{\partial e}\frac{f_e(x|e)}{f(x|e)}$. That is, $\frac{1}{f(x|e)}\frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(x)}$ is a convex or concave function of $\frac{f_e(x|e)}{f(x|e)}$ depending on the sign of $A_3(e,\pi,\eta)$. As a result, the sign of $\frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(x)}$ changes at most twice over the interval X.

Note that if $\frac{\partial \mathcal{L}(e,\pi,\eta)}{\partial \pi(x)}$ is always positive or always negative (for a given effort level e), then the information structure is fully uninformative. Such information structures cannot be incentive compatible as they provide no incentives for the agent to conduct any effort level $e > e_1$. Consequently, the equilibrium information structure has either one or two thresholds. In either case, it follows immediately from (2.13) that if the equilibrium information structure has a single threshold, say x^*, then $\pi(x) = 1$ if and only if $x \geq x^*$. If the equilibrium information structure has two thresholds, say (x_1^*, x_2^*), then $\pi(x) = 1$ if and only if $x \in [x_1^*, x_2^*]$. ■

Chapter 3

Pricing and Mergers in Complex Networks: The Case of Natural Gas Pipelines

with Maryam Saeedi and Ali Shourideh

3.1 Abstract

Natural gas is a large and growing share of U.S. energy consumption; based on data from U.S. Energy Information Administration (EIA), in 2021, it accounted for 32% of the primary energy consumption, 38% of the electricity generation, and about half of home heating in the U.S. The transportation of natural gas from producer to consumer, however, is relatively unexamined by economists. Using an extensive panel of daily data on natural gas flows through interstate pipelines from 2005 to 2016, and applying machine learning methods, we are able to obtain an overall view of the natural gas flows in the network. Specifically, we can identify the flow between every pair of states, which can help us in analyzing this networked market in more depth in future work. In this paper, we take the first steps of analyzing this networked market by studying two-part price schedules used by local distributing companies (LDCs) and performing demand estimation to derive price-elasticity of demand in different sectors: residential, commercial, industrial, and electric utility. Our results suggest that while demand in all these sectors is relatively inelastic with respect to the average price, electric utility is the most elastic sector and industrial sector is the most inelastic one. We then investigate one of the largest mergers among natural gas interstate pipelines. Our results suggest that this merger had a significant effect on natural gas transportation prices even though the two pipelines were not in the same physical market. Furthermore, using an extensive panel of daily data on temperature, natural gas prices, and inventories at various locations across the U.S., from 2009 to 2015, we are able to investigate the role of storage in the natural gas market, quantify its effect on natural gas prices at different geographical locations across the U.S., and evaluate the network effect in this market. We test the standard rational expectations competitive storage model, which is widely used in the literature, to estimate natural gas demand as well as storage costs. We then quantify the effect of temperature, storage, and pipeline congestion on natural gas prices. Our results suggest that storage and temperature are the main factors explaining the price shocks in the natural gas market where abundant storage dampens the effect of cold winter weather on natural gas prices. Finally, we investigate the network effect in the natural gas market, evaluating the effect of

changes in the temperature of one geographical region on all other regions. Our results suggest that a change or shock in one geographical region can have a significant effect on prices even in the farthest regions. This observation can be rationalized by the change of natural gas route in the pipeline network from low-demand regions to high-demand regions.

3.2 Introduction

There is a lack of research on regulation and antitrust policy in complex networked markets where upstream and downstream markets are not easily distinguishable. An example of such a market is the U.S. natural gas market comprising of gas producers, interstate and intrastate pipeline operators, and local distributing companies (LDCs). How does a merger among interstate pipelines affect the natural gas transportation price? The answer to this question is critical to antitrust decision-making as policymakers are concerned about the welfare implications if most of the changes in the price paid by shippers are passed onto final end use consumers.

The natural gas transportation industry has seen considerable consolidation over the last 25 years. However, traditional methods of merger analysis cannot be used off the shelf. In such markets, it is not straightforward to define the relevant market or to identify which market participant (producers, pipelines, storage operators, traders, or shippers) enjoys market power. Consequently, it is an intricate task to determine which pipeline or storage mergers should be allowed, or which pipeline or storage operators should be allowed to offer market-based rates.

Part of the complexity stems from the fact that a particular portion can be used in multiple routes. A pipeline that goes from Louisiana to New York can be used to ship gas along a longer route, such as from Texas to Boston. Merging pipelines would increase concentration along all possible route-markets in which that pipeline could participate. Mergers can look vertical, if a Louisiana-NYC pipeline merges with NYC-Boston, but they can also look horizontal if two separate Louisiana-NYC pipelines merge. Furthermore, a merger anywhere in the network of pipelines can influence prices even in distant parts of the network due to shifting routes as traders minimize cost of transport.

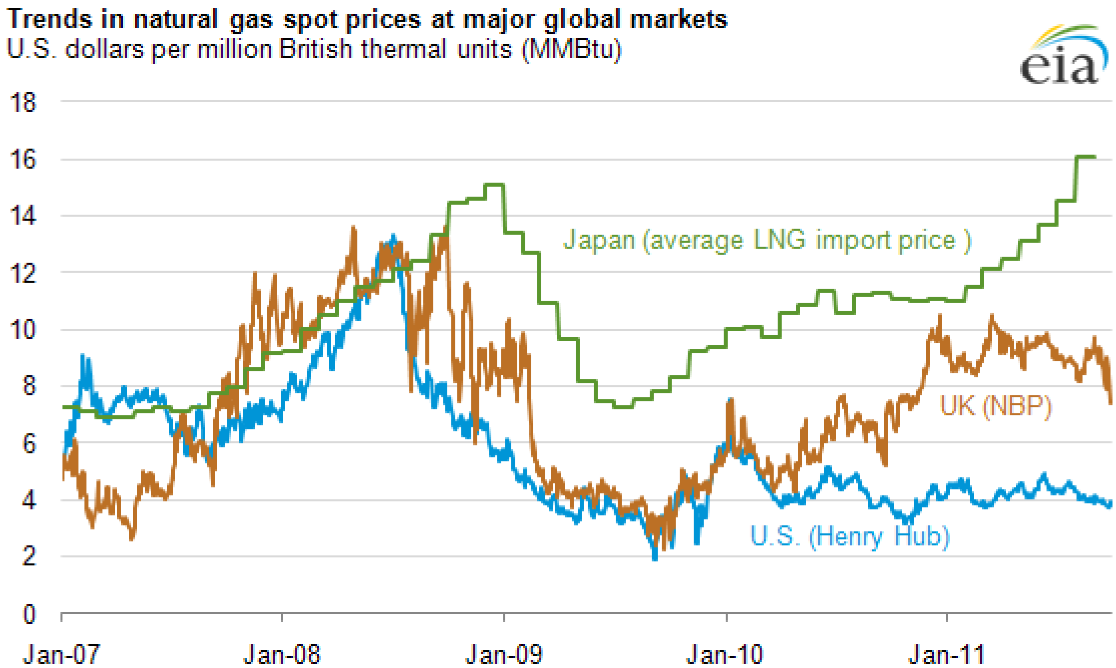
The price effects of a merger are the result of changes in market power and efficiency. On the one hand, the decrease in competition and increase in market power through a merger lead to an increase in price; on the other hand, any possible efficiency gains lead to a decrease in price. A retrospective study of a consummated merger can provide empirical evidence on the price effects of a merger, showing the net effect of these two forces and offering guidance for future merger analysis. As the natural gas market is relatively unexamined by economists, in this paper, we take the first steps of analyzing this complex networked market.

3.2.1 An Overview of Natural Gas Market

Natural gas is one of the principal sources of energy for many of our day-to-day needs and activities and plays a central role in determining our overall economic well-being, from heating and cooling our homes and businesses to determining the cost and composition of goods and services produced in the economy. In the U.S., natural gas is a large and growing share of energy consumption; based on data from EIA, in 2021, it accounted for 32% of the primary energy consumption, 38% of the electricity generation, and about half of home heating in the U.S. Due to its environmental cleanliness compared to coal, along with advances in unconventional sources of supply due to fracking, natural gas is projected to continue to grow as a share of U.S. energy consumption.

Natural gas is mostly transported using pipelines. As a result, its price can be different at different parts of the world or even within a country as a function of distance to the wellhead. In other words, natural gas has a localized market. Natural gas hubs, which tend to be at the heart of gas infrastructure network, are used as central pricing points for the network's natural gas. In some cases, financial derivative contracts are priced off gas delivered at these points as well. Natural gas prices are a function of market supply and demand. Figure 3.1 illustrates the trends in natural gas spot prices at three major markets around the world.

Figure 3.1: Trends in natural gas spot prices at major global markets (Source: EIA website)



The natural gas market, specifically the transportation of natural gas from ultimate producer to consumer is relatively unexamined by economists. However, this market has undergone a process of market deregulation over the last decades. For instance, price controls on wellhead prices were a major feature of the U.S. natural gas market for much of the post-war period but were terminated in 1989.

The most constructive effort led by market participants and the Federal Energy Regulatory Commission (FERC) to further deregulate and revitalize the natural gas industry emerged in the format of the FERC Open Access Rule Orders 436 and 500. The concept of Orders 436 and 500 was to allow pipelines to continue buying gas from producers and selling to end users in the traditional manner, but to also allow producers and end users to enter into contracts with the pipelines to obtain capacity on these transmission systems for their own use. Furthermore, orders 436 and 500 led to the development of various transportation tariff structures by the pipelines, which identify different levels of transportation service. As a result, shippers, which are title holders which transport natural gas on pipelines, could choose between more or less reliable transportation service, as well as transportation service for varying lengths of time, depending on how much they were willing to pay for that service. For example, transportation service for one month which could be interrupted at the pipeline's discretion ("interruptible service") would be less expensive than transportation service for one year which could not be interrupted by the pipeline ("firm service").

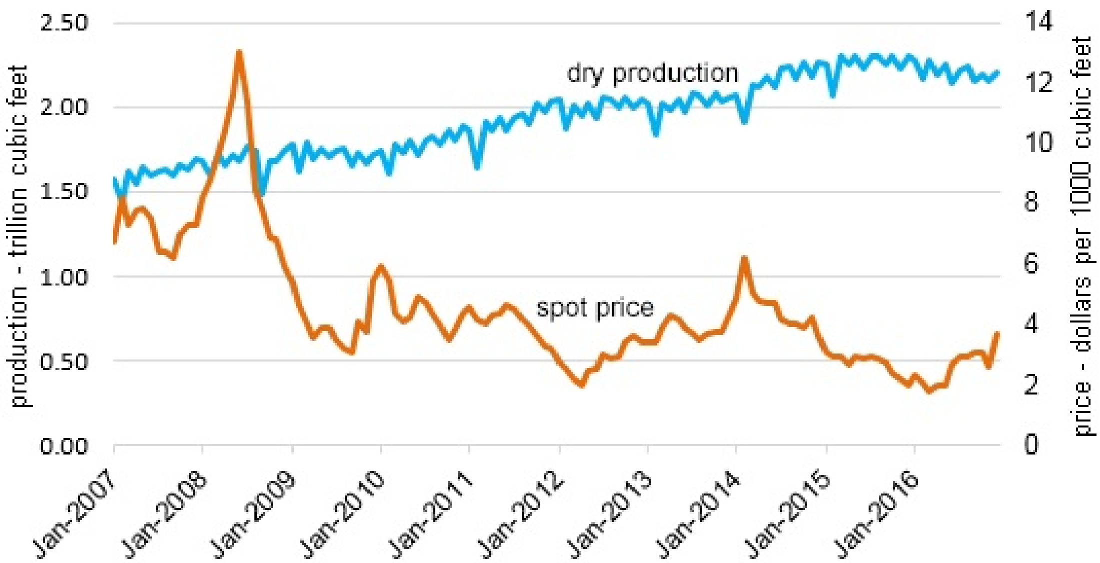
The next step to achieving a more open market environment came on July 31, 1991 when the FERC issued a Notice of Proposed Rule-making (NOPR), and after a lengthy comment and review process, the Mega-NOPR became FERC Final Order 636. The Final Order 636 brought together the open-access concepts in Orders 436 and 500 and, in general, outlined the "unbundling" of services provided by the interstate pipelines. Unbundling is the process of identifying and separating the various services provided by a pipeline and allowing these services to be contracted for, independent of one another. Bundling, on the other hand, is the term used for "rolling-in" all costs incurred by a pipeline for all available services, and charging all shippers equally, regardless of the services actually used by the shippers. Unbundled service calls for the segregation of each service provided by a pipeline, allowing for independent selection and purchase of those services.

Moreover, in states with active "retail choice programs" such as Georgia, Maryland, New York, Ohio, Pennsylvania, and Virginia, customers have a choice between buying natural gas from their LDC and buying natural gas from independent natural gas marketers. The LDC provides and is reimbursed for transportation services but marketers perform the financial transactions, procuring natural gas in the wholesale market and then selling it to final customers.

Due to changing market dynamics and requests from shippers, interstate pipelines were permitted to allow for capacity release among their firm transportation shippers in April of 1992. Capacity release refers only to firm transportation capacity and is the option whereby a firm shipper can assign its firm transportation capacity on a pipeline to a third party. Although deals can be negotiated exclusively between shippers, the preferred method of releasing capacity is through a closed bidding process on the pipeline's EBB (electronic bulletin board) wherein the buyer submits a percentage of maximum firm transportation rate bid to the seller of the firm capacity. The buyer with the highest percentage bid will be assigned that firm capacity.

Most of the natural gas consumed in the U.S. comes from domestic production. As Figure 3.2 illustrates, U.S. dry natural gas production increased from 2006 to 2015, and U.S. natural gas spot prices and consumer prices generally decreased during the same period. However, in 2016, production declined for the first time since 2005, and prices increased toward the end of the year.

Figure 3.2: Monthly U.S. dry natural gas production and monthly average spot price at the Henry Hub (Source: EIA website)



Because of limited alternatives for natural gas consumption or production in the near term, even small changes in supply or demand over a short period can result in large price movements that bring supply and demand back into balance. Three major supply-side factors affect prices: amount of natural gas production, level of natural gas in storage, volumes of natural gas imports and exports. Three major demand-side factors affect prices: variations in winter and summer weather, level of economic growth, availability and prices of competing fuels.

The strength of the economy influences natural gas markets. During periods of economic growth, increases in demand for goods and services from the commercial and industrial sectors may increase natural gas consumption. Economic-related increases in consumption can be particularly strong in the industrial sector, which uses natural gas as a fuel and a feedstock for making many products.

During cold months, natural gas demand for heating by residential and commercial consumers generally increases overall natural gas demand and can put upward pressure on prices. If unexpected cold or severe weather occurs, the effect on prices intensifies because supply is often unable to react quickly to short-term increases in demand. The effect of weather on natural gas prices may be greater if the natural gas transportation (pipeline) system is already operating at full capacity. Natural gas supplies in storage may help to cushion the impact of high demand during cold weather.

To balance seasonal swings in use, natural gas is placed in storage. The level of natural gas in underground storage fields has a large influence on overall supply. Storage helps to meet seasonal and sudden increases in demand, which domestic production and imports might not otherwise meet. When demand is lower, storage absorbs excess domestic production and, sometimes, imports. Natural gas storage levels tend to be highest sometime between the end of October and mid-November and lowest at the end of winter. EIA's Weekly Natural Gas Storage Report shows that most injections occur between April and October and most withdrawals occur between November and March.

High summer temperatures can have direct and indirect effects on natural gas prices. Warm temperatures increase the demand for air conditioning, which generally increases the electric utility sector's demand for natural gas. During high demand periods, natural gas prices on the spot market may increase sharply if natural gas supply sources are relatively low or constrained. Increases in natural gas consumption by the electric utility sector during the summer may lead to smaller-than-normal injections of natural gas into storage and to lower available storage volumes in the winter, which could have an effect on prices.

Some large-volume fuel consumers such as electric utilities and industrial consumers (e.g., iron, steel, and paper mills) can switch between natural gas, coal, and petroleum, depending on the cost of each fuel. When the cost of the other fuels fall, demand for natural gas may decrease, which may reduce natural gas prices. When the cost of competing fuels rise relative to the cost of natural gas, switching from those fuels to natural gas may increase natural gas demand and prices. In 2016, more electricity was generated from natural gas than coal for the first time on record, and natural gas was the largest source of overall electricity generation.

3.2.2 Our Contributions

After obtaining an overall view of the natural gas flows in the network by identifying the flow between every pair of states, we take the first steps of analyzing this complex networked market.

First, we focus on natural gas distribution in the U.S. This is a clear natural monopoly with high fixed costs and low marginal costs. We perform demand estimation to derive price-elasticity of demand in different sectors: residential, commercial, industrial, and electric utility. Our results suggest that while demand in all these sectors is relatively inelastic with respect to the average price, electric utility is the most elastic sector and industrial sector is the most inelastic one.

Second, we investigate the merger of Kinder Morgan Inc. and El Paso Corp., one of the largest mergers among interstate pipelines. Our results suggest that this merger had a significant effect on natural gas transportation prices even though the two pipelines were not in the same physical market.

Third, we estimate the natural gas inverse demand function as well as storage costs using the standard rational expectations competitive storage model, which is widely used in the literature.

Fourth, we investigate and quantify the effect of temperature, storage, and pipeline congestion on natural gas prices at different geographical locations across the U.S. by running some linear regressions. Our results suggest that storage and temperature are the main factors explaining the price shocks in the natural gas market where abundant storage dampens the effect of cold winter weather on natural gas prices.

Finally, we investigate the network effect in the natural gas market, evaluating the effect of changes or shocks in one geographical region on all other regions by modifying the above-mentioned linear regressions. Our results suggest that a change or shock in one geographical region can have a significant effect on prices even in the farthest regions. This observation can be rationalized by the change of natural gas route in the pipeline network from low-demand regions to high-demand regions.

The rest of the paper is organized as follows. Section 2 describes the data sets used in this paper. Section 3 briefly describes the efforts made to build the layout of the interstate pipeline network. Section 4 studies the two-part price schedule used by LDCs, while Section 5 derives the price-elasticity of demand for natural gas in different sectors. Section 6 investigates the effects of a large merger between two interstate pipelines on natural gas transportation prices. Section 7 briefly describes the rational expectations competitive storage model and summarizes the demand estimation results. Section 8 lays out a few linear regressions quantifying the effect of temperature, storage, and pipeline congestion on natural gas prices, while Section 9 derives the network effects using similar regressions. Section 10 concludes the paper and discusses the future work.

3.3 Data

The main data set, which essentially motivated this paper, includes a panel of daily data on more than 11,000 points in the network of interstate pipelines; these are points where a pipeline is connected to a facility such as a gathering system, gas processing plant, compressor, throughput meter, storage, power plant, industrial end user, LDC, or another interstate or intrastate pipeline. The information for each point includes pipeline name, location latitude and longitude, facility type, capacity, scheduled flow, and available capacity. Our sample includes data from from January 2006 through December 2016. This is the data set used in section III for constructing the pipeline network, and will be used further in future work.

The data sets we use in section IV for studying LDCs' two-part price schedule include the number of natural gas consumers for residential, commercial, and industrial sectors by state at the annual level, the natural gas volumes delivered to end use consumers by sector and state at the monthly level, and average price of natural gas delivered to consumers by sector and state at the monthly level. All these data come from the U.S. Department of Energy, Energy Information Administration (EIA) website. For residential and commercial sectors, we use data from January 1990 through December 2016. For industrial sector, we use data from January 2001 to December 2016 as data for industrial consumers are only available from 2001.

The data sets we use in section V for demand analysis include the data used in section IV as well as receipts and average cost of fossil-fuels including coal and natural gas delivered to electric utility sector for electricity generations by state at the monthly level. For electric utility sector, we use data from March 2003 to December 2016. These data also come from the U.S. Department of Energy, Energy Information Administration (EIA) website. Prices, which are available by state, month, and customer class, include all charges paid by end users including transportation costs as well as all federal, state, and local taxes. All dollar values in this section have been normalized to reflect July 2018 prices.

The data sets we use in section VI for merger analysis include data on the ownership history of interstate pipelines and the daily natural gas spot prices at different hubs in the U.S. from 1999 through 2015. The data on spot prices come from the Intercontinental Exchange (ICE).

The data sets we use in subsequent sections include daily natural gas spot prices at 83 hubs across the U.S. and Canada from 1999 through 2015, daily natural gas inventories at 69 storage locations across the U.S. from 2009 through 2015, average daily temperature by state, monthly natural gas production by state from 1997 through 2018, annual state in-flow capacity from other states from 1994 through 2018, and monthly futures prices for delivery at the Henry Hub in Louisiana up to four months from the trade date from 1994 through 2018. The data on spot prices come from the Intercontinental Exchange (ICE), the data on temperature come from National Oceanic and Atmospheric Administration (NOAA), U.S. Department of Commerce, and the rest of the data sets come from the U.S. Energy Information Administration (EIA) website. Therefore, our regressions are run over the range of data from 2009 to 2015. All dollar values in this section have been normalized to reflect July 2018 prices.

Using daily average temperature, we calculate daily heating and cooling degree days: Each day's heating (cooling) degree days is the number of degrees that day's average temperature is below (above) 65 degrees Fahrenheit, which is the temperature below (above) which buildings need to be heated (cooled). Heating (Cooling) degree days is set euql to zero if the average temperature is above (below) 65 degrees Fahrenheit. This is mainly important for the residential and commercial customers that use natural gas primarily for space heating.

In the next sections, we perform our estimations sometimes based on state-wise data, and sometimes based on regional data. When performing state-wise analysis, we will derive each state's inventory by aggregating all inventories in that state; also, we will derive each state's price data by taking the average of all hub prices which are affected by natural gas prices in that state.

To perform regional analysis, we first have to define some geographical regions; this can be done in several different ways. For example, EIA used to divide the U.S. into three storage regions: Producing, Consuming East, and Consuming West. It later modified this division into 5 storage regions: East, Midwest, South Central, Mountain, and Pacific. On the other hand, EIA divides the U.S. into six pipeline regions: Northeast, Midwest, Central, Southeast, Southwest, and West. We will use this last six-region division in the regional analysis of the subsequent sections. Table 3.1 represents the states included in each region.

When performing regional analysis, we will derive each region's inventory and production by aggregating all inventories and productions of the corresponding states; also, we will derive each region's price data by taking the average of all hub prices which are

Table 3.1: U.S. pipeline regions

Region	States
Northeast	Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont, Virginia, West Virginia
Midwest	Illinois, Indiana, Michigan, Minnesota, Ohio, Wisconsin
Central	Colorado, Iowa, Kansas, Missouri, Montana, Nebraska, North Dakota, South Dakota, Utah, Wyoming
Southeast	Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee
Southwest	Arkansas, Louisiana, New Mexico, Oklahoma, Texas
West	Arizona, California, Idaho, Nevada, Oregon, Washington

used in deriving the prices of the corresponding states. However, each region's in-flow capacity is derived by aggregating the in-flow capacity of the corresponding states from all states not in that region.

3.4 Data Cleaning and Network Construction

In this section, our goal is to obtain an overall view of the natural gas flows in the network, and more specifically, identify the natural gas flow between every pair of states. To this end, we clean our main data set including a panel of daily information about more than 11,000 points in the network of interstate pipelines.

We first retrieved the state and county where each point in the data set is located, and then using this information and applying machine learning methods, we could trace the flow of natural gas along each pipeline. As some of the largest pipelines have a very complex layout and it is possible that a pipeline travels in and out of a state or county several times over its route, characterizing the gas flow direction is not trivial from the information of a few nodes on that pipeline. In other words, the order in which gas traverses these points on a pipeline cannot be easily retrieved. We managed to characterize this order using machine learning techniques.

Considering each state (or county if needed) as a single node in the network of interstate pipelines, we calculated the daily flow of natural gas between each pair of nodes (states) using the above-mentioned processed data. This information can help in analyzing the market from a network point of view. We will discuss the lines of research that can be followed using the resulting data set and network layout in Section VII.

3.5 LDCs' Two-part Price Schedule

There are two main consumers of natural gas: (i) LDCs which mostly serve residential, commercial, and industrial customers, and (ii) electric utilities. LDCs are regulated by state utility commissions which set price schedules for each customer class using traditional rate-of-return techniques. Regulators create price schedules to equate total revenues from all customer classes with total operating costs plus an allowed rate of return on the firm's capital expenditures.

In natural gas distribution, the dominant price schedule is a two-part tariff. First, LDCs charge customers a fixed fee, typically levied monthly, that does not depend on the level of consumption. This fee varies by customer class; typically, industrial and commercial customers pay a higher monthly fee than residential customers. In some cases, the fixed fee also varies within customer class based on the level of consumption. Second, customers pay a per-unit charge for each unit of natural gas that is consumed. This price includes a commodity charge for natural gas purchased on their behalf by the LDC. Commodity costs change throughout the year with the LDCs' procurement costs. In addition to commodity costs, most companies also charge customers a per unit "transportation charge" per unit of natural gas consumed. This fee also varies by customer class; typically, residential customers pay a higher transportation charge than commercial and industrial customers.

EIA provides the number of natural gas consumers for residential, commercial, and industrial sectors by state at the annual level, the natural gas volumes delivered to end use consumers by sector and state at the monthly level, and average price of natural gas delivered to consumers by sector and state at the monthly level. For residential and commercial sectors, we use data from January 1990 through December 2016. For industrial sector, we use data from January 2001 to December 2016 as data for industrial consumers are only available from 2001. Computing the monthly revenue of LDCs by sector and state (Revenue = Average Price * Consumption) and following the regression equation:

$$R_{state,t} = \alpha * N_{state,t} + \beta * C_{state,t} + \epsilon_{state,t},$$

we regress the monthly revenue from each sector's sales in each state, R, on the corresponding number of consumers, N, and monthly gas consumption, C, forcing the intercept of the regression equation to zero. The estimated coefficients, α and β, can be interpreted respectively as the fixed and per-unit charges in the two-part price schedule in each state and each sector. In other words, α is the average amount paid in fixed monthly fees and β is the average per-unit price. Tables 3.2 and 3.3 present the results of estimating the two-part price schedule used by LDCs in every state for residential and commercial sectors, respectively. Taking the average over all states, for residential consumers, the average monthly fixed charge is \$11.11, and the average per-unit charge is \$8.05. Similarly, for commercial consumers, the average monthly fixed charge is \$22.39, and the average per-unit charge is \$7.63. As was expected, commercial consumers which have a higher level of consumption face higher fixed charges and lower per-unit charges.

The results of estimating the two-part price schedule used by LDCs for industrial sectors are less accurate including more negative fixed charges which is due to the fact that both fixed and per-unit charges also vary within customer class in every state depending on the level of consumption. This can also be the case for residential and commercial consumers based on the state regulations. However, for the residential sector, the fixed charge is negative only for the state of Maine. The results of estimating the two-part price schedule used by LDCs in every state for industrial sector can be found in Appendix.

The above analysis assumes that, for all sectors and in all states, both fixed and per-unit charges have not changed over the period under study. However, typically changes in commodity costs are passed on relatively quickly to final customers; in fact, every few years, LDCs file a rate case with FERC to increase their rates. Running the previous regression separately for every year or every three years (assuming LDCs file a rate case every three years), lets us observe the changes in fixed and per-unit charges by LDCs

Table 3.2: Two-part tariff estimates ($/Mcf) for the residential sector by state

State	Fixed Charge	Per-unit Charge	R^2	State	Fixed Charge	Per-unit Charge	R^2
AL	13.38	8.70	0.68	NE	11.28	5.90	0.74
AZ	12.34	8.96	0.75	NV	10.36	8.13	0.81
AR	11.97	7.07	0.71	NH	7.40	11.25	0.77
CA	5.76	7.56	0.65	NJ	6.61	8.96	0.78
CO	10.52	5.88	0.76	NM	13.32	5.69	0.66
CT	11.79	11.16	0.87	NY	12.32	9.86	0.81
DE	13.91	9.66	0.77	NC	11.62	9.43	0.84
FL	12.38	10.08	0.59	ND	12.52	5.68	0.74
GA	19.85	8.71	0.64	OH	15.91	7.01	0.69
ID	3.10	8.10	0.86	OK	15.55	5.97	0.70
IL	16.40	6.13	0.71	OR	8.13	9.35	0.80
IN	12.41	6.67	0.76	PA	15.02	8.60	0.77
IA	13.94	6.23	0.73	RI	12.97	10.51	0.78
KS	12.80	6.59	0.71	SC	10.01	9.37	0.84
KY	12.97	6.65	0.71	SD	9.51	6.56	0.77
LA	9.94	6.72	0.68	TN	8.25	7.96	0.80
ME	-2.05	14.17	0.88	TX	12.24	6.23	0.73
MD	12.65	9.12	0.79	UT	5.45	7.04	0.82
MA	2.28	11.98	0.82	VT	21.09	9.76	0.71
MI	14.84	6.14	0.68	VA	13.03	9.54	0.83
MN	7.59	7.04	0.77	WA	10.33	8.66	0.75
MS	7.06	7.23	0.72	WV	9.47	7.93	0.77
MO	17.53	6.65	0.70	WI	5.86	7.89	0.80
MT	9.03	6.23	0.71	WY	12.84	5.77	0.68

over the period under study. Tables 3.4 and 3.5 present respectively the results of estimating two-part tariffs used by LDCs in Pennsylvania for the residential sector from 1990 through 2016 assuming LDCs file a rate case every year and every three years. Figures 3.3 and 3.4, corresponding to Tables 3.4 and 3.5, indicate how fixed and per-unit charges in the residential sector have changed in Pennsylvania over the period under study. The results of estimating two-part tariffs used by LDCs in Pennsylvania for the residential sector from 1990 through 2016, assuming LDCs file a rate case every two years, can be found in Appendix.

As can be observed from these tables and figures, the per-unit charge, which includes the commodity cost for natural gas, was increasing from 1990 to 2008, but started decreasing after that. This trend can be explained by the increasing Shale gas production

Table 3.3: Two-part tariff estimates ($/Mcf) for the commercial sector by state

State	Fixed Charge	Per-unit Charge	R^2	State	Fixed Charge	Per-unit Charge	R^2
AL	51.69	8.45	0.56	NE	-13.77	6.32	0.69
AZ	34.26	7.72	0.44	NV	68.97	6.95	0.57
AR	9.53	7.27	0.70	NH	25.70	10.95	0.79
CA	36.36	6.87	0.33	NJ	-69.01	9.53	0.69
CO	42.04	5.21	0.64	NM	47.25	5.27	0.55
CT	34.63	8.40	0.76	NY	-35.42	8.83	0.67
DE	37.18	10.48	0.78	NC	20.12	8.49	0.70
FL	-198.20	12.47	0.46	ND	39.42	5.64	0.72
GA	25.83	8.05	0.63	OH	29.60	7.05	0.70
ID	14.79	7.10	0.77	OK	68.93	5.96	0.60
IL	93.06	6.06	0.64	OR	26.66	7.67	0.66
IN	42.18	6.39	0.72	PA	39.20	8.55	0.75
IA	29.93	6.09	0.70	RI	39.76	9.88	0.77
KS	78.43	4.99	0.49	SC	-8.49	9.45	0.64
KY	43.17	6.84	0.66	SD	30.62	5.80	0.71
LA	-3.97	8.01	0.55	TN	36.65	7.39	0.68
ME	-125.45	14.47	0.91	TX	49.73	5.52	0.40
MD	-37.68	9.73	0.78	UT	4.84	6.45	0.81
MA	33.94	9.65	0.74	VT	63.32	7.18	0.73
MI	76.98	5.72	0.68	VA	16.53	8.29	0.71
MN	9.75	6.70	0.72	WA	62.99	7.13	0.57
MS	8.05	7.28	0.56	WV	84.12	7.25	0.57
MO	49.16	6.86	0.68	WI	-1.74	7.32	0.76
MT	29.00	6.88	0.78	WY	34.02	5.72	0.69

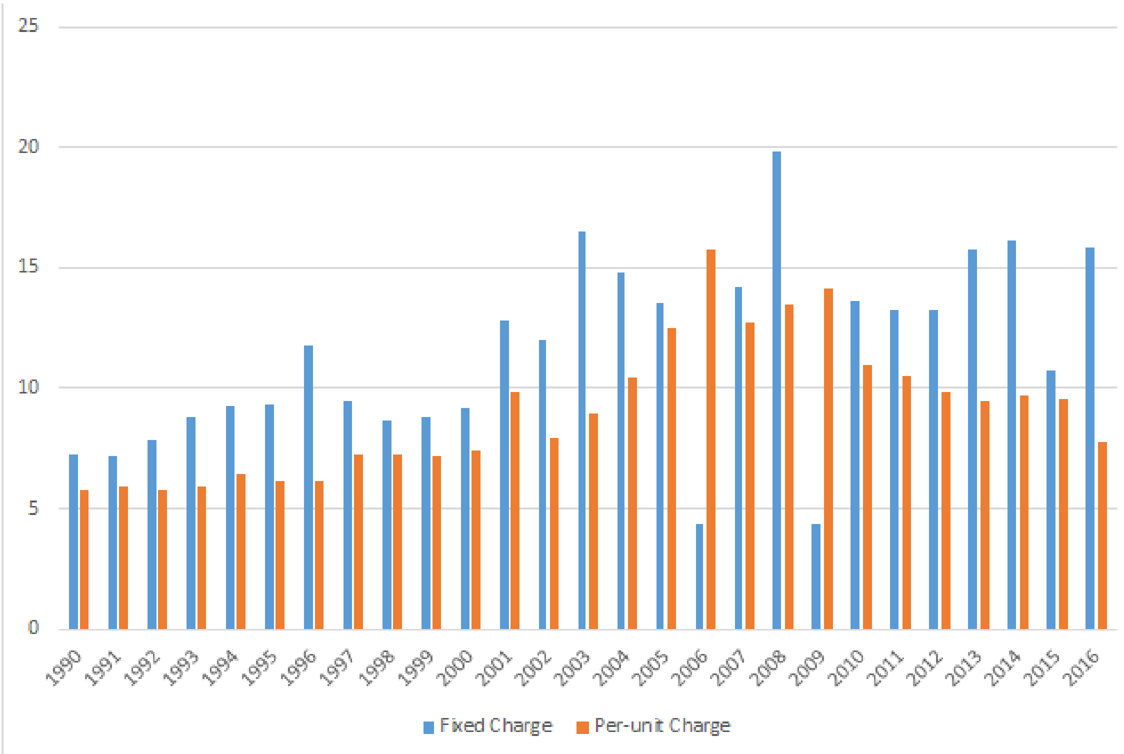
in the U.S. which resulted in lower prices.

Peoples natural gas company, which provides natural gas service to retail gas customers in Pennsylvania, currently charges a fixed charge of $13.95 per residential meter per month (which is close to our estimates) and a transportation charge of $3.133 per Mcf (1000 cubic feet) which is close to our estimates of per-unit charge in Pennsylvania once we add up the natural gas commodity cost (around $4) and taxes as well. Note that this transportation charge causes departure from marginal cost pricing by LDCs; Davis and Muehlegger (2010) estimate that these distortions impose hundreds of millions of dollars of annual welfare loss.

Table 3.4: Two-part tariff estimates ($/Mcf) for the residential sector in Pennsylvania assuming LDCs file a rate case every year

Year	Fixed Charge	Per-unit Charge
1990	7.24	5.78
1991	7.20	5.93
1992	7.89	5.76
1993	8.84	5.90
1994	9.28	6.45
1995	9.32	6.13
1996	11.77	6.14
1997	9.46	7.26
1998	8.67	7.26
1999	8.81	7.22
2000	9.19	7.45
2001	12.83	9.87
2002	12.01	7.91
2003	16.52	8.94
2004	14.84	10.41
2005	13.54	12.49
2006	4.39	15.78
2007	14.21	12.73
2008	19.80	13.49
2009	4.38	14.13
2010	13.60	10.97
2011	13.24	10.54
2012	13.21	9.85
2013	15.74	9.46
2014	16.17	9.71
2015	10.72	9.58
2016	15.85	7.78
R^2	0.96	

Figure 3.3: Trend in two-part tariff charges ($/Mcf) in Pennsylvania assuming LDCs file a rate case every year



3.6 Demand Analysis

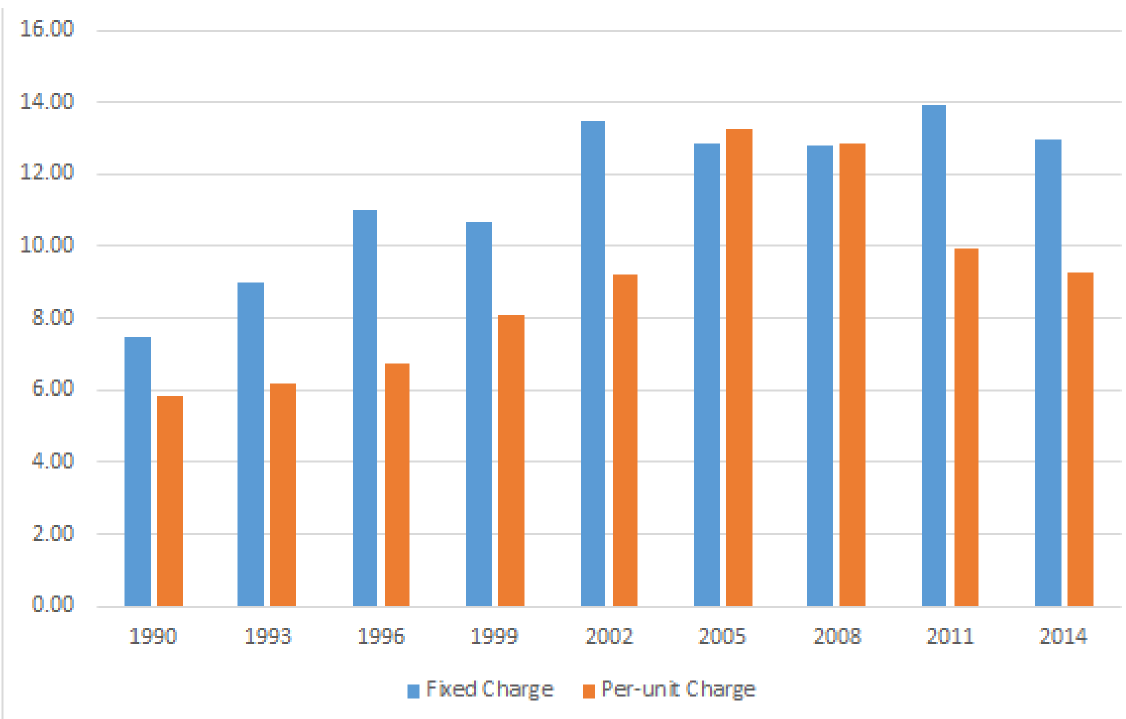
Use of natural gas has two seasonal peaks with consumption patterns predominantly driven by weather. The largest peak occurs during the winter, when cold weather increases the demand for natural gas space heating in the residential and commercial sectors. The smaller peak occurs in the summer, when air conditioning use increases demand for electric power, an increasing portion of which is provided by natural gas-fired generators.

The electric power sector, which surpassed the industrial sector in 2008 and the combined residential and commercial sectors in 2015, has become the largest consumer of natural gas in the U.S. Consumption of natural gas in the power sector peaks in the summer when demand for electricity is highest. A smaller peak occurs during the winter,

Table 3.5: Two-part tariff estimates ($/Mcf) for the residential sector in Pennsylvania assuming LDCs file a rate case every three years

Year	Fixed Charge	Per-unit Charge
1990-1992	7.45	5.82
1993-1995	8.97	6.18
1996-1998	11.01	6.73
1999-2001	10.69	8.11
2002-2004	13.49	9.22
2005-2007	12.86	13.27
2008-2010	12.81	12.84
2011-2013	13.93	9.96
2014-2016	12.95	9.25
R^2	0.97	

Figure 3.4: Trend in two-part tariff charges ($/Mcf) in Pennsylvania assuming LDCs file a rate case every three years



while the spring and fall seasons have the lowest consumption of natural gas for electric power.

Increased deliveries of natural gas to the electric power sector have accounted for much of the growth in total natural gas deliveries. As the share of natural gas used for power generation has increased, so has the interdependence between natural gas supply and infrastructure and electric power generator operations. Residential and commercial natural gas use peaks during the winter as these consumers use natural gas primarily for space heating. Industrial users of natural gas exhibit the least seasonality.

In this paper, we omit electric utility consumers from LDCs' customers since these facilities consume sufficiently large amounts of natural gas such that it is often profitable to build a dedicated line directly to an interstate pipeline, contract with suppliers directly, and bypass the LDC. Firm contracts and interruptible contracts are two broad types of contracts for purchasing natural gas. Transactions between power plants and natural gas suppliers generally include a supply component, which involves an agreement with a fuel producer or marketer to supply the commodity, and a delivery component, which involves an agreement with a pipeline operator to transport the fuel from the producer to the generator. For any transaction, one or both of these components can be firm or interruptible.

Firm contracts provide power plant operators with an agreed-upon capacity for the producer or pipeline to supply natural gas, establishing a high priority for fuel requested by the power plant. The supply or delivery of natural gas cannot be curtailed under a firm contract except under unforeseeable circumstances. Firm contracts are most prevalent in the West and South regions of the United States. In contrast, interruptible contracts are lower-priority fuel supply and transportation arrangements. Under these contracts, the flow of natural gas to a power plant may be stopped or curtailed if firm contract holders use the available capacity or if other interruptible customers outbid the power plant. These contracts are generally set up for short periods, often for next-day delivery. Interruptible contracts are less expensive than firm contracts, reflecting the higher risk of disrupted fuel receipts. Interruptible contracts are most common in the Northeast; in 2016, more natural gas in this region was purchased using interruptible contracts than firm contracts.

Most natural gas-fired power plants in the continental U.S. fulfill all of their fuel requirements through firm contracts that obligate the natural gas producer and the pipeline operator to send the natural gas to the power plant when requested. These power plants reported receiving 71% of the natural gas purchased by power plants in 2016. About 16% of natural gas used for power generation in 2016 was purchased by plants that reported using only interruptible contracts, in which the natural gas supplier or pipeline operator has the option of interrupting the fuel supply for contractually stipulated reasons. The remaining 13% of natural gas was purchased by plants through some mix of firm and interruptible contracts throughout the year.

Now, we perform demand estimation for natural gas by classifying consumption into residential, commercial, industrial, and electric utility sectors, and estimating the price-elasticity of demand for each sector.

In addition to the data for residential, commercial, and industrial sectors explained in the previous sector, EIA also provides receipts and average cost of fossil-fuels including coal and natural gas delivered to electric utility sector for electricity generations by state at the monthly level. For residential and commercial sectors, we use data from January 1990; for industrial sector, we use data from January 2001 to December 2016; and for electric utility sector, we use data from March 2003 to December 2016. Prices, which are available by state, month, and customer class, include all charges paid by end users including transportation costs as well as all federal, state, and local taxes. All dollar values in the article have been normalized to reflect July 2018 prices.

The analysis which follows is performed at the monthly level. Although there is daily variation in prices, short-run variation in natural gas prices is mitigated by the ability of natural gas suppliers to store natural gas. In the U.S., total natural gas storage capacity exceeds eight trillion cubic feet, enough to meet total consumption for several months. With access to storage, natural gas suppliers are able to arbitrage within-month price differences.

For each customer class, we regress the log of monthly consumption on the log of average natural gas prices, state*month-of-year fixed effects and state*year fixed effects.

The state*month-of-year fixed effects allow for unique state-specific seasonal patterns in natural gas consumption and state*year fixed effects allow demand in each state to change flexibly with long-run trends in income or housing growth. With state*month-of-year and state*year fixed effects, we estimate our coefficients off of deviations from mean seasonal patterns in each state.

In addition, we include demand shifters appropriate for each customer class' use of natural gas. For all sectors, we include same-month demeaned heating and cooling degree days. Each month's heating (cooling) degree days is the sum of the number of degrees that each day's average temperature is below (above) 65 degrees Fahrenheit, which is the temperature below (above) which buildings need to be heated (cooled). This is mainly important for the residential and commercial customers that use natural gas primarily for space heating; we could choose to omit heating and cooling degree days for industrial customers who use natural gas for production. Furthermore, we include the spot price of West Texas Intermediate (WTI) crude oil in the demand equation as some industrial customers have the ability to switch between fuel oil and natural gas. Finally, for electric utility sector, we include the spot price of WTI crude oil as well as average cost of coal delivered to electric utility sector in each state in the demand equation as some power plants, based on their type, can use either petroleum, coal, or natural gas as fuel.

We present our demand elasticity estimates in Table 3.6. We estimate that demand for natural gas in all sectors is relatively inelastic. The elasticity point estimates for residential, commercial, industrial, and electric utility users are -0.5132, -0.2407, -0.0996, and -0.6424, respectively. In addition, we find that the cross-price elasticity of industrial demand with respect to the crude oil price is 0.0661.

Our results confirm that the demand curve is indeed sloping downward and suggest that while demand in all these sectors is relatively inelastic with respect to the average price, electric utility is the most elastic sector and industrial sector is the most inelastic one.

To check the robustness of our estimation results, we run the same regression several times with a few changes made every time. First, for residential and commercial demand,

Table 3.6: Price-elasticity of demand by sector

	Residential	Commercial	Industrial	Electric Utility
Log(Price)	-0.5132^{***}	-0.2407^{***}	-0.0996^{***}	-0.6424^{***}
	(0.0104)	(0.0101)	(0.0099)	(0.0468)
Log(WTI Crude Oil Price)			0.0661^{***}	-0.0621
			(0.0091)	(0.0657)
Log(Coal Price)				-0.0650
				(0.1561)
Heating Degree Days	0.0007^{***}	0.0006^{***}	0.0002^{***}	0.0007^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0001)
Cooling Degree Days	-0.0002^{***}	-0.0001	0.0000	0.0029^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0003)
State*Year FE	Yes	Yes	Yes	Yes
State*Month-of-year FE	Yes	Yes	Yes	Yes
R^2	0.9925	0.9891	0.9910	0.9231

we use the WTI crude oil spot price as an instrument for natural gas prices. We present these demand elasticity estimates in Table 3.7.

Table 3.7: Price-elasticity of demand by sector: WTI as instrument

	Residential	Commercial	Industrial	Electric Utility
Log(Price)	-0.3198^{***}	-0.2171^{***}	-0.0996^{***}	-0.6424^{***}
	(0.0329)	(0.0335)	(0.0099)	(0.0468)
Log(WTI Crude Oil Price)			0.0661^{***}	-0.0621
			(0.0091)	(0.0657)
Log(Coal Price)				-0.0650
				(0.1561)
Heating Degree Days	0.0007^{***}	0.0006^{***}	0.0002^{***}	0.0007^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0001)
Cooling Degree Days	-0.0003^{***}	-0.0001	0.0000	0.0029^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0003)
State*Year FE	Yes	Yes	Yes	Yes
State*Month-of-year FE	Yes	Yes	Yes	Yes
R^2	0.9913	0.9887	0.9910	0.9231

Furthermore, we can interact demeaned heating and cooling degree days with prices to allow the elasticity to vary with temperature. We present these demand elasticity estimates in Table 3.8. We estimate that the elasticity of customers is negatively correlated with heating degree days, i.e., consumers respond less to exogenous shifts in price during cold months.

Table 3.8: Price-elasticity of demand by sector: elasticity changing with temperature

	Residential	Commercial	Industrial	Electric Utility
Log(Price)	-0.5133^{***}	-0.2428^{***}	-0.1020^{***}	-0.5870^{***}
	(0.0103)	(0.0101)	(0.0099)	(0.0468)
Log(WTI Crude Oil Price)			0.0693^{***}	-0.1252^{*}
			(0.0092)	(0.0655)
Log(Coal Price)				-0.1065
				(0.1550)
Heating Degree Days	-0.0023^{***}	-0.0002	-0.0003^{**}	0.0005^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.0002)
Cooling Degree Days	-0.0022^{***}	0.0010^{***}	-0.0002	0.0003
	(0.0004)	(0.0004)	(0.0003)	(0.0004)
Log(Price)*Heating Degree Days	0.0003^{***}	0.0001^{***}	0.0001^{***}	0.0002^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0001)
Log(Price)*Cooling Degree Days	0.0002^{***}	-0.0001^{***}	0.0000	0.0015^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0002)
State*Year FE	Yes	Yes	Yes	Yes
State*Month-of-year FE	Yes	Yes	Yes	Yes
R^2	0.9928	0.9891	0.9911	0.9246

For residential, commercial, and industrial sectors, we can also instrument for natural gas prices using heating and cooling degree days in all other states. These demand elasticity estimates can be found in Appendix. We have also done the same experiments separately for each state, but the results are omitted for brevity.

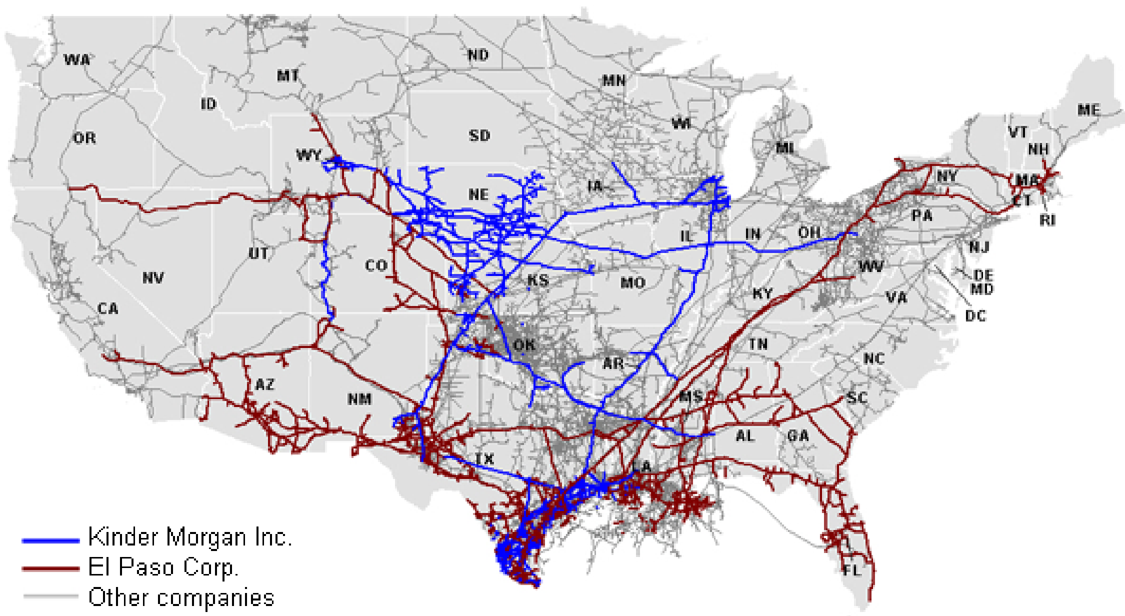
As can be seen in Tables 3.6, 3.7, and 3.8, the qualitative results do not change; i.e., while demand in all these sectors is relatively inelastic with respect to the average price, electric utility is the most elastic sector and industrial sector is the most inelastic one.

3.7 Merger Analysis

The merger of Kinder Morgan Inc. (KMI) and El Paso Corp. (El Paso) which was announced on October 16, 2011 and completed on May 24, 2012, created the U.S. largest natural gas pipeline company. The combined company was operating about 67,000 miles of natural gas pipelines (see the blue and red lines in Figure 3.5), or about 22% the U.S. natural gas pipeline network at the time of merger completion.

El Paso's natural gas pipeline network complements Kinder Morgan's natural gas

Figure 3.5: U.S. natural gas pipeline network, November 2011 (Source: EIA website)



system. By adding El Paso's network to its own for 21.1 billion dollars, KMI increased its access to natural gas markets in the Southwest, Southeast, Northwest, and Northeast.

In this section, we lay out an econometric strategy to show that this merger had a significant effect on natural gas transportation prices, and measure the effect of this merger on transportation prices. The natural gas price at Henry Hub in Louisiana is the benchmark price for natural gas traded in the U.S. and is considered reflective of the commodity value without transportation costs. Therefore, we use the difference in hub prices as a measure of the cost of natural gas transportation along pipelines between the hubs. On the other hand, by acquiring El Paso, KMI increased its market power in the region around the Chicago Citygate Hub. Thus, we focus on the difference in hub prices at the Henry Hub and the Chicago Citygate Hub.

To measure the effect of this merger on transportation prices, we run a linear regression of differential hub prices (Henry-Chicago) on a set of explanatory variables that are relevant for understanding the price impact: lag of differential hub prices, a dummy variable indicating the merger completion date, volume of transactions at the Chicago Citygate Hub (to capture demand at this hub), Shale's share of U.S. oil production (to capture the local supply), and WTI crude oil spot price. The Intercontinental Exchange (ICE) provides data on spot prices at different hubs in the U.S. from 1999 through 2015. In this section, we use daily spot prices at Henry Hub and Chicago Citygate Hub from October 16, 2010 through October 15, 2011 (the day before the merger announcement date) and from May 24, 2012 (merger completion date) to May 23, 2013. The results of this regression are presented in the second column of Table 3.9. The price effects of a merger are the result of changes in market power and efficiency. The regression analysis shows that, for the hub pairs which are located on the route of the merged pipelines, the effect of merger on the cost of transportation is statistically significant and the net effect of the two forces mentioned above, causes prices to increase. To check the robustness of our results, we performed the same regression with different pairs of hubs located on the route of KMI. The qualitative results did not change.

In order to realize whether the effect of merger on transportation prices realized after merger announcement or merger completion, we run the previous regression by adding a dummy variable indicating the merger announcement date and use data from October 16, 2010 to May 23, 2013. The results of this regression are presented in the third column of Table 3.9. As can be observed from Table 3.9, the increase in transportation prices occurred after the merger completion.

The main implication of this exercise is that in complex networked market, such as the natural gas market, mergers can have a significant effect on natural gas transportation prices even though the two merging companies are not in the same physical market.

3.8 Demand and Storage Cost Estimation

In this section, we estimate the natural gas inverse demand function as well as storage costs using the standard rational expectations competitive storage model. We first briefly describe the rational expectations competitive storage model and how it can be used for

Table 3.9: Effects of merger on natural gas transportation prices

	Merger Completion	Merger Announcement and Completion
Merger Announcement Indicator		0.0107
		(0.0103)
Merger Completion Indicator	0.0397^{**}	0.0232^{**}
	(0.0163)	(0.0096)
Hub Differential Lag	0.6837^{***}	0.7049^{***}
	(0.0331)	(0.0283)
Chicago Citygate Hub Volume	0.0004^{***}	0.0003^{***}
	(0.0001)	(0.0001)
Shale Ratio	-0.3548^{**}	-0.3131^{**}
	(0.1617)	(0.1553)
WTI Crude Oil Price	0.0003	0.0001
	(0.0005)	(0.0004)
Year FE	Yes	Yes
R^2	0.72	0.72

estimation, and then summarize the demand estimation results.

3.8.1 Rational Expectations Competitive Storage Model

Following Deaton and Laroque (1992), consider the following simplified model of natural gas. In each period of time, say a month, there is an exogenous production z following a stochastic process characterized by a transition function $\Lambda(z, Z)$, which gives the probability that the production in year $t+1$ is less than or equal to Z when the production in period t is z. The production shocks come from outside the model, are determined by the conditions of production, and are what ultimately drive the behaviour of prices. We assume z follows a first-order linear auto-regressive process.

There are two kinds of actors in the market: final consumers and risk-neutral speculators. The consumers have inverse demand functions that we take to be linear, so that if there was no storage and the production z was consumed in each period, price would be given by

$$P(z) = a + bz, \tag{3.1}$$

where a and $b < 0$ are parameters. Given equation (1), and if consumers were the only demanders in the market, we could directly infer the behaviour of prices from the behaviour of the production. In particular, prices will follow an $AR(1)$ if the production follows an $AR(1)$.

Inventory holders demand the commodity in order to transfer it into the next period, and will do so whenever they expect to make a profit over the storage and interest costs. The technology of storage takes the simple form of proportional deterioration during storage, so that if I units are stored, $(1-\delta)I$ units are available at the beginning of the next period; speculators also have to pay a real interest rate r per period on the value of commodities stored. Define x, the availability or amount on hand as the sum of the production and inherited inventories, so that

$$x_t = (1-\delta)I_{t-1} + z_t. \tag{3.2}$$

The crucial non-linearity in this specification is introduced by the requirement that inventories be non-negative. Given this, together with the assumptions that speculators are risk neutral and have rational expectations, current and expected future prices must satisfy

$$p_t = \max\left[P(x), \frac{1-\delta}{1+r}\mathbb{E}_t p_{t+1}\right]. \tag{3.3}$$

The second term in brackets is the expected value of one unit stored, after allowing for storage costs, conditioned on the speculator's current information, assumed to be the current production and the current availability. The first term is the price that would hold if current availability were sold to consumers, and this will be the actual price if it is greater than the net expected future price. This is the situation in which speculators would hold negative inventories were it possible to do so. If the net expected price is greater than the price when no inventories are held, speculators exploit the arbitrage opportunities and drive up the current price until current price and their expectation of the net future price are equal.

Provided certain conditions are met, most notably that $r + \delta > 0$, there exists a solution to equation (3) in terms of a price function where $f(x, z)$ is the unique monotone decreasing in its first argument solution to the functional equation

$$f(x,z) = \max\left(\frac{1-\delta}{1+r}\int f(z' + (1-\delta)(x - P^{-1}(f(x,z))), z')\Lambda(z, dz'), P(x)\right). \quad (3.4)$$

Except in very special cases, this equation does not permit an analytical solution, and must be solved numerically.

When the production process follows a linear auto-regression

$$z_{t+1} = \rho z_t + \epsilon_{t+1}, \quad (3.5)$$

for $-1 < \rho < 1$, and the ϵ's are i.i.d., specializing from equations (4) to (5) gives

$$f(x,z) = \max\Big[G(x,z), P(x)\Big], \quad (3.6)$$

where $G(x,z)$ is the expected price net of storage costs,

$$G(x,z) = \frac{1-\delta}{1+r}\int f(\epsilon + \rho z + (1-\delta)(x - P^{-1}(f(x,z))), \epsilon + \rho z)d\Phi(\epsilon), \quad (3.7)$$

and Φ is the uni-variate distribution of ϵ.

Once the functions have been computed, they can be used to characterize the behavior of the stochastic processes. In particular, note that, by equation (2), availability evolves according to

$$x_{t+1} = (1-\delta)(x_t - P^{-1}(f(x_t, z_t))) + z_{t+1}, \quad (3.8)$$

and the production z, evolves according to the transition function $\Lambda(z, Z)$, which in our case means that it follows the auto-regression (5). We have a bi-variate stochastic process in the vector (x_t, z_t), both elements of which are assumed unobserved, and with which there is the associated observable price given by $f(x,z)$.

In the formulation described above, apart from the distribution of the production innovations, there are five free parameters: the depreciation rate δ, the real interest rate r, the slope and intercept of the inverse demand function a and b, and the auto-regressive coefficient ρ.

It turns out that, when productions and stocks are assumed unobserved and when the demand function is linear, it is impossible to recover separately the demand function and the mean and standard errors of the disturbances. Thus, we normalize the innovation of the production process to be of mean zero and of variance one; the standard deviation of the production z is thus $1/\sqrt{1-\rho^2}$.

For simplicity, we have taken the distribution to be normal and also fixed the real interest rate at the annual level of 2%.

Our general approach is to use the theory to construct one-period-ahead conditional expectations and variances of prices, and to match these theoretical conditional moments to the actual data. There are two elements in this process: the construction of the moments and the matching to the data. The latter is done using pseudo maximum likelihood estimation (PMLE) (see Gourieroux et al., 1984). Write the one-period-ahead conditional mean and variance of price as $m(p_t)$ and $s(p_t)$, so that

$$\begin{aligned} m(p_t) &= \mathbb{E}(p_{t+1}|p_t) \\ s(p_t) &= \mathbb{V}(p_{t+1}|p_t) \end{aligned} \tag{3.9}$$

and define (twice the log of) the pseudo-likelihood function by

$$2\ln L = \sum_{t=1}^{T-1} \ln l_t = -(T-1)\ln(2\pi) - \sum_{t=1}^{T-1} \ln s(p_t) - \sum_{t=1}^{T-1} \frac{(p_{t+1} - m(p_t))^2}{s(p_t)} \tag{3.10}$$

This function is maximized with respect to the parameters of the model. Equation (10) would yield the exact likelihood function if prices were normally and heteroscedastically distributed conditional on lagged price. When there is storage, we would not expect prices to be normally distributed, even if the productions were normal, but the parameter estimates will nevertheless be consistent, in spite of the fact that both the conditional expectation and conditional variance are non-differentiable in p_t so that the likelihood may be non-differentiable in the structural parameters (see Laroque and Salanie, 1994).

In order to calculate the conditional means and variances (9), we follow a strategy as follows. For each price in the data p_t, we calculate a set of values of (x_t, z_t) that are consistent with that price. For each element of the set, we use the transition function to calculate the transition probabilities (densities) to each possible (x_{t+1}, z_{t+1}) in period $t+1$, and thus to the corresponding price $p_{t+1} = f(x_{t+1}, z_{t+1})$. In order to put relative weights on these various p_{t+1}, we assume that the data in period t are drawn from the unique invariant limit distribution of the bi-variate process on productions and availability. We compute this distribution and derive from it the conditional probabilities associated with each possible (x_t, z_t) consistent with p_t. Weighting the associated values of p_{t+1} yields the conditional expectation and variance of p_{t+1}, for insertion in the PML criterion.

3.8.2 Estimation Results

Note that while we assume productions and stocks are unobserved in the PMLE, but since we actually observe this data, we first estimate the auto-regressive coefficient ρ using monthly production data from January 2009 through December 2015. This yields $\rho = 0.87$.

Then we use the PMLE to estimate the rest of the parameters (a, b, δ) using monthly price data. The results are shown in Table 3.10. Note that for this estimation, we convert our daily prices to average monthly prices.

Table 3.10: Demand and storage cost estimates

Region	a	b	δ
United States	4.1382	-0.9557	0.5610
Northeast	5.3163	-1.5740	0.4405
Midwest	4.4570	-1.0632	0.5317
Central	4.2379	-0.9757	0.4360
Southeast	3.7478	-0.8102	0.3158
Southwest	4.0165	-0.8343	0.4214
West	3.9307	-0.7747	0.3724

One downside of the results in Table 3.10 is that the estimated storage costs are much higher than expected. Because of these unreasonable storage cost estimates, we also considered a specification of the storage model with positive constant marginal cost k as well as proportional deterioration following Cafiero et al. (2011). Note that considering the storage cost as a constant proportional deterioration or shrinkage of the stock implies that the marginal cost of storage is high when stocks are low since the price is decreasing in stocks. This would change equation (3) to

$$p_t = \max \left[P(x), \frac{1-\delta}{1+r} \mathbb{E}_t p_{t+1} - k \right], \tag{3.11}$$

but our estimation results did not change qualitatively and the estimates of the constant marginal cost k turned out to be negligibly small compared to the estimated δ.

As mentioned above, Deaton and Laroque (1992) did not use data on production or stocks in their estimation approach; we also just used production data to estimate the auto-regressive coefficient. However, we can use monthly production and storage data as well as monthly data on price of futures contracts for delivery in the following month to estimate the inverse demand function and storage costs using the standard rational expectations competitive storage model. Considering equation (3), we can use futures contract price for the next month delivery F_t^{t+1} as the expectation of price of the next period, that is, $F_t^{t+1} = \mathbb{E}_t p_{t+1}$. Thus, minimizing the mean-squared error

$$MSE = \sum_{t=1}^{T-1} \left(p_t - \max \left[a + bx, \frac{1-\delta}{1+r} F_t^{t+1} \right] \right)^2, \tag{3.12}$$

we can estimate the parameters of the model. However, it turns out the model does not identify any stock-out and thus parameters a and b cannot be estimated. The storage cost is estimated to be $\delta = 0.0135$ which is much more reasonable compared to what we got from the other approach.

3.9 Natural Gas Price as a Function of Temperature, Storage, and Pipeline Congestion

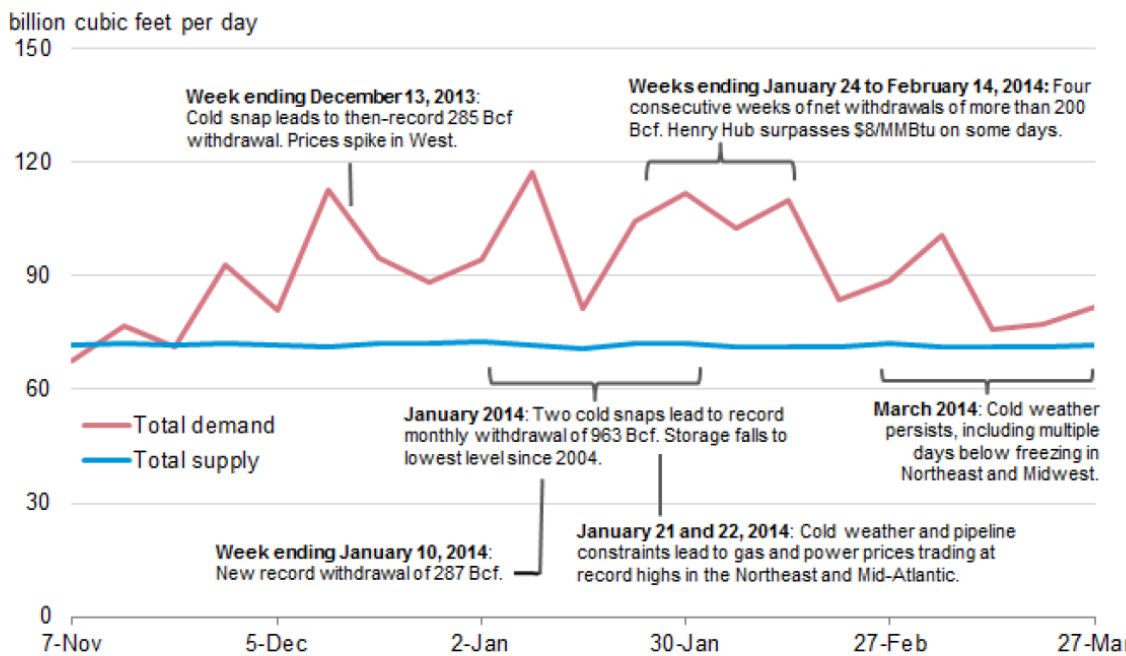
In this section, we investigate and quantify the effect of temperature, storage, and pipeline congestion on natural gas prices at different geographical locations across the U.S. by running some linear regressions.

As mentioned before, during cold months, natural gas demand for heating by residential and commercial consumers generally increases overall natural gas demand and can put upward pressure on prices. If unexpected cold or severe weather occurs, the effect on prices intensifies because supply is often unable to react quickly to short-term increases in demand. The effect of weather on natural gas prices may be greater if the natural gas transportation (pipeline) system is already operating at full capacity. Natural gas supplies in storage may help to cushion the impact of high demand during cold weather.

Also, high summer temperatures can have direct and indirect effects on natural gas prices. Warm temperatures increase the demand for air conditioning, which generally increases the electric utility sector's demand for natural gas. During high demand periods, natural gas prices on the spot market may increase sharply if natural gas supply sources are relatively low or constrained. Increases in natural gas consumption by the electric utility sector during the summer may lead to smaller-than-normal injections of natural gas into storage and to lower available storage volumes in the winter, which could have an effect on prices.

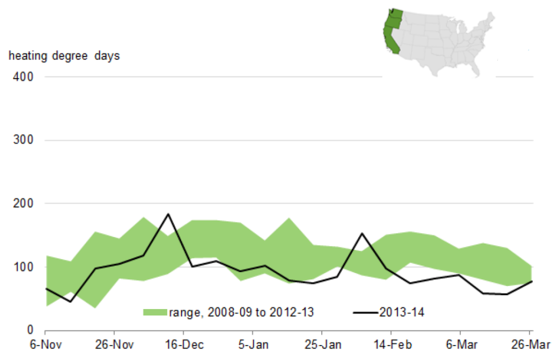
One extreme example that clearly illustrates the effect of temperature and storage on natural gas prices is the 2013-14 winter defined by record high natural gas consumption and storage withdrawals, as waves of bitterly cold weather repeatedly swept across the United States. Natural gas markets in the U.S. began the winter season (November 2013 to March 2014) with a robust gas inventory. As Figure 3.6 illustrates, cold snaps starting in mid-November 2013 drove up demand for natural gas, particularly from residential and commercial consumers. Three consecutive months of 10-year record storage withdrawals (November 2013 through January 2014) pushed inventories to their lowest January levels since 2004. Even after storage fell to a 10-year low, the cold snaps continued, pushing the Henry Hub spot price to $8.00 per million British thermal units (MMBtu) on some days in February. Sustained colder-than-normal temperatures through March pushed storage to its lowest level since 2003 and kept prices volatile, despite higher production.

Figure 3.6: Natural gas average weekly supply/demand balance (Source: EIA website)

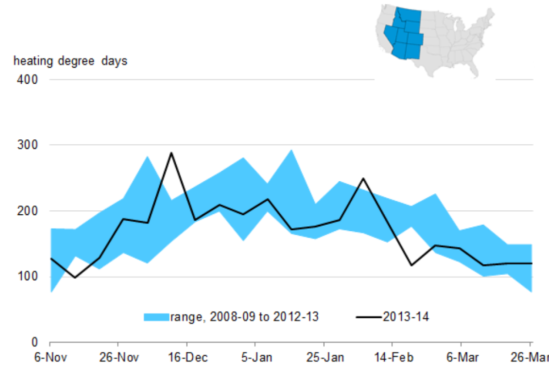


Temperatures east of the Rockies were persistently cold, record-setting as much for their sustained nature as for any single event. This cold weather was particularly persistent in the Northeast, the Midwest, and the Mid-continent producing region. Bitterly cold waves swept through these three regions, as well as regions that traditionally have milder winter temperatures, like Texas, the Rockies, and the Pacific Northwest. This cold weather not only caused spikes in natural gas demand, but it also lessened the performance of the energy supply chain. As Figure 3.7 illustrates, every major region of the contiguous United States experienced waves of significantly colder-than-normal temperatures.

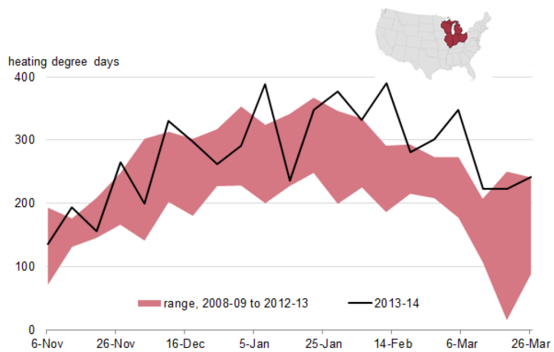
The cold weather had a dramatic effect on U.S. natural gas consumption, which rose



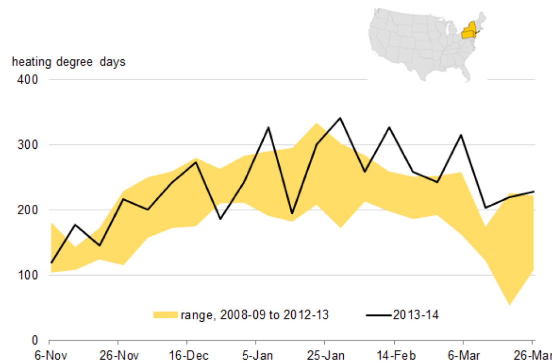
(a) Pacific



(b) Mountain



(c) East North Central



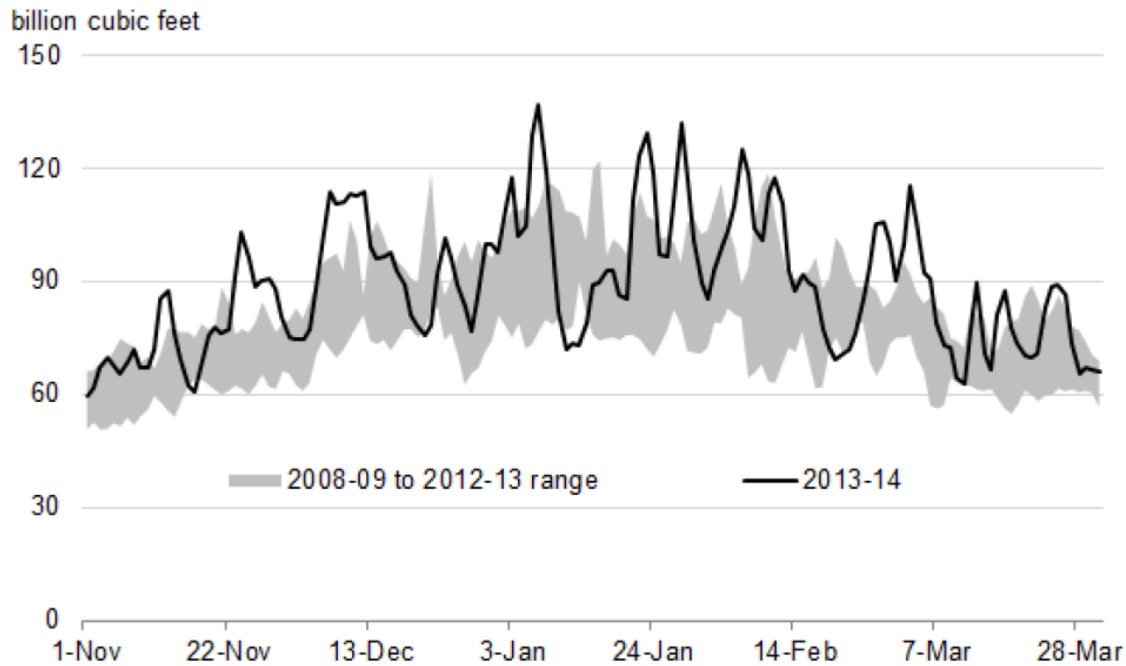
(d) Mid-Atlantic

Figure 3.7: Weekly heating degree days, (Source: EIA website)

to a record per-day average during the 2013-14 winter. As Figure 3.8 illustrates, consumption spiked during the cold weather in mid-December, early and late January, the beginning of February, and the beginning of March. U.S. natural gas consumption reached a same-month record every month from November 2013 to March 2014, according to EIA data going back to 1995. Cold weather drove consumption up by 9% over 2012-13 winter levels, with higher consumption occurring across all major sectors, despite higher prices.

As Figure 3.9 illustrates, natural gas markets in the U.S. began this winter season (November 2013 to March 2014) with relatively high levels of gas inventories. Starting in mid-November 2013, however, U.S. natural gas consumption increased at a greater rate than supply, pushing inventories to their lowest levels since 2004, even as high prices

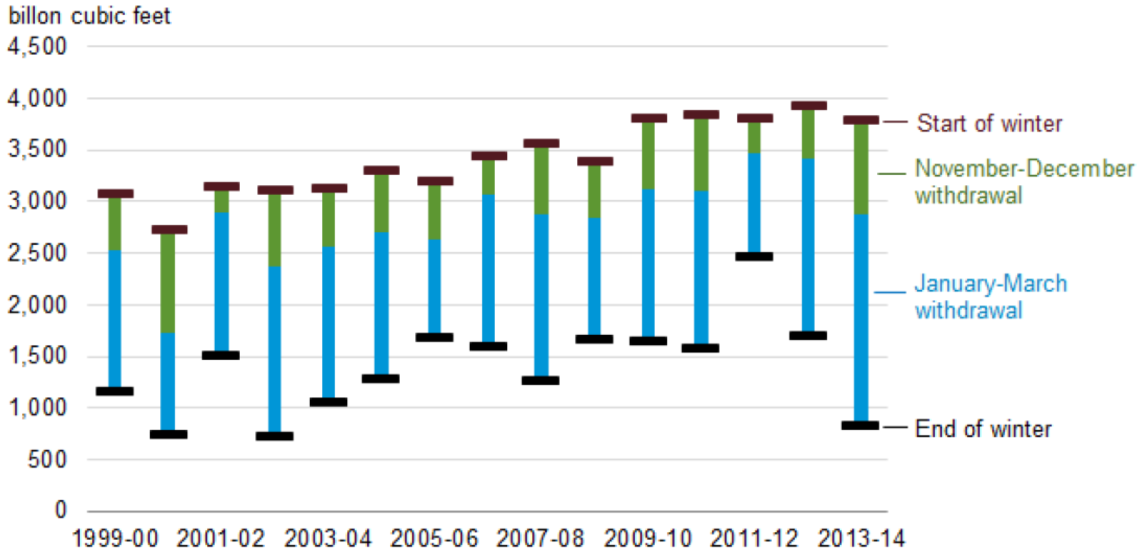
Figure 3.8: Total daily U.S. natural gas consumption (Source: EIA website)



curtailed electric sector demand in some parts of the country. December saw record storage withdrawals in the West region, and a record withdrawal in the Lower 48 states for the week ending on December 13, 2013. Lower 48 gas inventories ended January at less than 2,000 Bcf for the first time since 2005. Heavy withdrawals continued through the first two weeks of February. Inventories, which had already reached their 10-year low by the end of January, dropped below 1,000 Bcf by the end of March for the first time since 2003. November-to-March withdrawals set records in all three storage regions, which depleted storage heading into April.

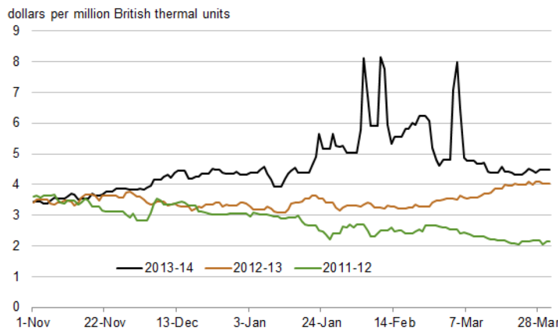
U.S. natural gas markets tightened as consumption outstripped supply and led to record-high storage withdrawals. The national benchmark spot price for natural gas traded at Henry Hub in Erath, Louisiana, averaged $4.63 per million British thermal units (MMBtu) during the 2013-14 winter, 33% higher than the winter of 2012-13, and the highest average winter spot price in four years, according to data from SNL Energy.

Figure 3.9: U.S. natural gas winter gas inventories (Source: EIA website)

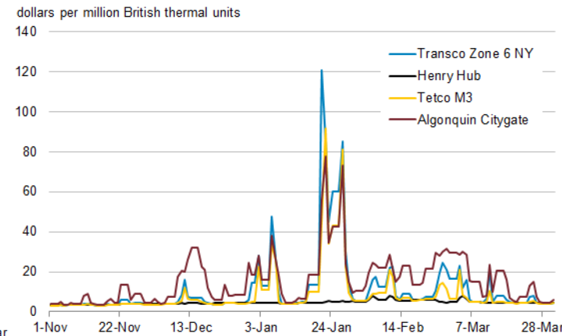


As Figure 3.10a illustrates, the Henry Hub spot price began the 2013-14 winter at less than \$3.50/MMBtu, and increased to more than \$5.00/MMBtu by the end of January, as inventories fell. As withdrawals continued to rise through the beginning of February, so did the Henry Hub spot price, peaking at \$8.15/MMBtu on February 10, according to SNL data. As Figure 3.10b illustrates, prices responded to sharp daily demand increases and continued draw-downs from storage by increasing significantly in the northeastern United States, as occurred in the winter of 2012-13. However, these price spikes were not confined to the Northeast; they also occurred at trading hubs serving consumers in the central and western United States, as Figures 3.10c and 3.10d illustrates. These observations emphasize the effect of temperature and storage on natural gas prices.

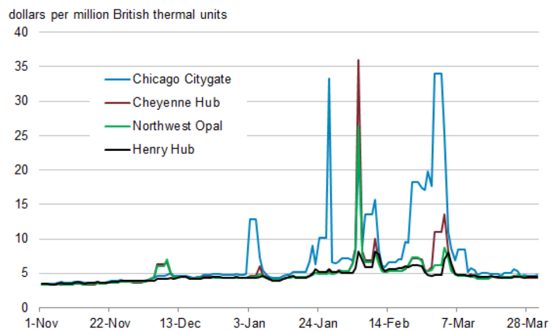
However, in the winter of 2014-15, despite similar cold weather and high consumption, the price increases were not as severe. More specifically, despite somewhat colder weather and higher natural gas consumption in the Northeast in the 2014-15 winter, natural gas wholesale (spot) prices at Algonquin Citygate, with service to Boston, and Transcontinental (Transco) Zone 6 NY, with service to New York City, were less volatile compared to the 2013-14 winter and remained below \$30 and \$40 per million British ther-



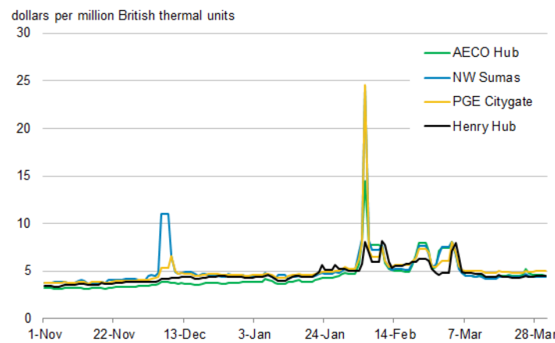
(a) Henry Hub



(b) Eastern US



(c) Central US



(d) Western US

Figure 3.10: Natural gas spot prices (Source: EIA website)

mal units (MMBtu), respectively. New England and New York are pipeline-constrained markets in the Northeast, where prices during peak-demand days during winter often surge to \$20 to \$30/MMBtu and reached record levels in January 2014 of \$78/MMBtu and \$121/MMBtu, respectively.

Based on EIA's Natural Gas Weekly Update, in the winter of 2014-15, there was an increase of natural gas supplied to the New England and New York areas. The increased natural gas supplies came from three different sources, which include domestic pipeline, imported liquefied natural gas (LNG) that is regasified and then sent out from the importing terminal, and pipeline imports from Canada. From the end of 2013-14 winter to the beginning of 2014-15 winter, two pipeline expansions came online and added about 1 billion cubic feet per day (Bcf/d) of supply to the Northeast. This observation emphasizes the effect of pipeline congestion on natural gas prices.

Given the above observations, we now try to quantify the effect of temperature, storage, and pipeline congestion on natural gas prices. First, to investigate the effect of temperature and storage on natural gas prices, we regress daily prices p_t on storage level $Storage_t$, the interaction of storage level and winter indicator, heating degree days HDD_t, interaction of heating degree days and winter indicator, cooling degree days CDD_t, interaction of cooling degree days and winter indicator, interaction of storage level and heating degree days and winter indicator, state(region)*month-of-year fixed effects and state(region)*year fixed effects. The state(region)*month-of-year fixed effects allow for unique state(region)-specific seasonal patterns in natural gas consumption and state(region)*year fixed effects allow demand in each state(region) to change flexibly with long-run trends in income or housing growth. The first two columns of Table 3.11 present the results of estimating the effect of temperature and storage on natural gas prices.

The estimation results in the first two columns of Table 3.11 suggest that all else the same, higher storage levels have a negative effect on price and this effect is even higher in winter. On the other hand, all else equal, higher heating or cooling degree days have a positive effect on price and this effect is even higher in winter; however, the higher the storage level, the lower is the positive effect of heating degree days on price in winter.

We can also investigate the effect of pipeline congestion on natural gas prices by adding the production level $Production_t$ and in-flow capacity $InFlow_t$ to the regressors. Note that it is reasonable to add both production and in-flow capacity (as opposed to just in-flow capacity) since a region or state may not have a large in-flow capacity which leads us to expect high prices for that region or state; however, it is possible that we are facing a producing region or state and thus lower prices than expected. The last two columns of Table 3.11 present the results of estimating the effect of temperature, storage, and pipeline congestion on natural gas prices.

The estimation results in the last two columns of Table 3.11 suggest that all else the same, the in-flow pipeline capacity decreases price; this can be interpreted as higher pipeline utilization (higher pipeline congestion) increases price. Last but not least, the

Table 3.11: Effect of temperature, storage, and pipeline congestion on natural gas prices

	State-wise	Regional	State-wise	Regional
$Production_t$	–	–	-0.2693^{***}	-0.2749^{***}
	–	–	(0.0361)	(0.0324)
$InFlow_t$	–	–	-0.0521^{***}	-0.0692^{***}
	–	–	(0.0119)	(0.0068)
$Storage_t$	-0.0063^{***}	-0.0042^{***}	-0.0057^{***}	-0.0041^{***}
	(0.0005)	(0.0004)	(0.0005)	(0.0004)
$Storage_t * Winter$	-0.0069^{***}	-0.0043^{***}	-0.0072^{***}	-0.0041^{***}
	(0.0006)	(0.0006)	(0.0005)	(0.0006)
HDD_t	0.0175^{***}	0.0431^{***}	0.0171^{***}	0.0425^{***}
	(0.0022)	(0.0054)	(0.0021)	(0.0054)
$HDD_t * Winter$	0.0357^{***}	0.0589^{***}	0.0318^{***}	0.0605^{***}
	(0.0027)	(0.0077)	(0.0026)	(0.0077)
CDD_t	0.0285^{***}	0.0615^{***}	0.0280^{***}	0.0616^{***}
	(0.0028)	(0.0070)	(0.0027)	(0.0069)
$CDD_t * Winter$	0.0066	0.2156^{***}	0.0089	0.2210^{***}
	(0.0201)	(0.0525)	(0.0192)	(0.0523)
$Storage_t * HDD_t * Winter$	-0.0001^{***}	-0.0000^{***}	-0.0001^{***}	-0.0001^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
State*Year FE	Yes	–	Yes	–
Region*Year FE	–	Yes	–	Yes
State*Month-of-year FE	Yes	–	Yes	–
Region*Month-of-year FE	–	Yes	–	Yes
R^2	0.4053	0.5038	0.4293	0.5075

higher the production in a state or region, the lower the price.

3.10 Network Effects

In this section, we investigate the network effect in the natural gas market, evaluating the effect of changes or shocks in one geographical region on all other regions by modifying the linear regressions in the previous section.

In the U.S., natural gas is delivered to different parts of the country using pipelines. All different geographical markets are connected to each other via the large network of interstate and intrastate natural gas pipelines. Therefore, a change or shock anywhere in the network of pipelines can influence prices even in distant parts of the network due to the shift in supply, demand and/or pipeline congestion.

To derive network effects, we only work with regional data and make minor modifications to the regressions in the previous section. In the regressions of the last section, we concatenated the data of all regions, but now, we run the previous regressions six times, each time leaving out the data of one region out and instead include the heating degree days of the excluded region as an extra regressor. Since the heating degree days across regions are highly correlated, we use instead the deviations of heating degree days from the seven-year trend (2009–2015) which are much less correlated across states. Tables 3.12 and 3.13 present the results of estimating the network effects following the second and fourth columns of Table 3.11, respectively.

Table 3.12: Network effects in natural gas price dynamics

Excluded Region	Region 1	Region 2	Region 3	Region 4	Region 5	Region 6
$Storage_t$	-0.0071^{***}	-0.0038^{***}	-0.0043^{***}	-0.0040^{***}	-0.0038^{***}	-0.0036^{***}
	(0.0005)	(0.0005)	(0.0004)	(0.0005)	(0.0005)	(0.0005)
$Storage_t * Winter$	-0.0013^{**}	-0.0042^{***}	-0.0045^{***}	-0.0034^{***}	-0.0059^{***}	-0.0046^{***}
	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0007)	(0.0006)
HDD_t	0.0428^{***}	0.0316^{***}	0.0253^{***}	0.0409^{***}	0.0361^{***}	0.0309^{***}
	(0.0047)	(0.0068)	(0.0062)	(0.0061)	(0.0060)	(0.0062)
$HDD_t * Winter$	0.0695^{***}	0.0501^{***}	0.0586^{***}	0.0687^{***}	0.0389^{***}	0.0537^{***}
	(0.0066)	(0.0091)	(0.0087)	(0.00089)	(0.0091)	(0.0086)
CDD_t	0.0553^{***}	0.0629^{***}	0.0561^{***}	0.0624^{***}	0.0689^{***}	0.0572^{***}
	(0.0057)	(0.0080)	(0.0075)	(0.0083)	(0.0084)	(0.0075)
$CDD_t * Winter$	0.0579	0.2370^{***}	0.2015^{***}	0.2004^{***}	0.2279^{***}	0.1757^{***}
	(0.0400)	(0.0551)	(0.0540)	(0.0690)	(0.0826)	(0.0560)
$Storage_t * HDD_t * Winter$	-0.0002^{***}	-0.0000	-0.0000^{***}	-0.0001^{***}	0.0000	-0.0000^{**}
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$Excluded\ HDD_{dev,t}$	0.0023	0.0265^{***}	0.0339^{***}	0.0055	0.0316^{***}	0.0427^{***}
	(0.0024)	(0.0033)	(0.0039)	(0.0041)	(0.0051)	(0.0043)
Region*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Region*Month-of-year FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.5614	0.4999	0.5186	0.4917	0.4933	0.5059

The estimation results in Tables 3.12 and 3.13 suggest that for four out of the six regions, the deviations of the heating degree days from the seven-year trend have a significant effect on the price of natural gas in the other regions. This effect can be interpreted as network effect which makes the analysis of the natural gas market in the U.S. much more difficult. Lack of attention to this kind of network effects is one of the main pieces that is missing in the regulation and antitrust policy prevailing in the natural gas market.

Table 3.13: Network effects in natural gas price dynamics

Excluded Region	Region 1	Region 2	Region 3	Region 4	Region 5	Region 6
$Production_t$	-0.3395^{***}	-0.2266^{***}	-0.1695^{***}	-0.2737^{***}	-0.6911^{***}	-0.2625^{***}
	(0.0258)	(0.0337)	(0.0341)	(0.0344)	(0.0692)	(0.0339)
$InFlow_t$	-0.0471^{***}	-0.0692^{***}	-0.0684^{***}	-0.0237^{***}	-0.0144^{*}	-0.1986^{***}
	(0.0087)	(0.0086)	(0.0070)	(0.0079)	(0.0087)	(0.0260)
$Storage_t$	-0.0075^{***}	-0.0037^{***}	-0.0042^{***}	-0.0038^{***}	-0.0026^{***}	-0.0034^{***}
	(0.0005)	(0.0005)	(0.0004)	(0.0005)	(0.0005)	(0.0005)
$Storage_t * Winter$	-0.0006	-0.0040^{***}	-0.0042^{***}	-0.0032^{***}	-0.0065^{***}	-0.0044^{***}
	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0007)	(0.0006)
HDD_t	0.0422^{***}	0.0311^{***}	0.0255^{***}	0.0404^{***}	0.0357^{***}	0.0305^{***}
	(0.0046)	(0.0068)	(0.0062)	(0.0061)	(0.0060)	(0.0062)
$HDD_t * Winter$	0.0725^{***}	0.0519^{***}	0.0606^{***}	0.0699^{***}	0.0390^{***}	0.0552^{***}
	(0.0065)	(0.0091)	(0.0087)	(0.00088)	(0.0090)	(0.0086)
CDD_t	0.0559^{***}	0.0632^{***}	0.0561^{***}	0.0624^{***}	0.0687^{***}	0.0574^{***}
	(0.0057)	(0.0080)	(0.0075)	(0.0082)	(0.0084)	(0.0075)
$CDD_t * Winter$	0.0540	0.2411^{***}	0.2066^{***}	0.2084^{***}	0.2422^{***}	0.1823^{***}
	(0.0395)	(0.0549)	(0.0539)	(0.0688)	(0.0821)	(0.0558)
$Storage_t * HDD_t * Winter$	-0.0002^{***}	-0.0000^{*}	-0.0001^{***}	-0.0001^{***}	-0.0000	-0.0000^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$Excluded\ HDD_{dev,t}$	0.0017	0.0265^{***}	0.0326^{***}	0.0052	0.0277^{***}	0.0414^{***}
	(0.0024)	(0.0033)	(0.0039)	(0.0041)	(0.0051)	(0.0043)
Region*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Region*Month-of-year FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.5708	0.5027	0.5201	0.4957	0.4996	0.5096

3.11 Conclusion and Future Work

In this paper, we took the first steps of analyzing complex networked markets by focusing on the U.S. natural gas market. We studied two-part price schedules used by LDCs and performed demand estimation to derive price-elasticity of demand in different sectors: residential, commercial, industrial, and electric utility. Our results suggest that while demand in all these sectors is relatively inelastic with respect to the average price, electric utility is the most elastic sector and industrial sector is the most inelastic one. Having derived the demand curve, in future work, we will obtain the supply curve and characterize the market equilibrium.

In addition, we investigated one of the largest mergers among natural gas interstate pipelines. Our results suggest that this merger had a significant effect on natural gas transportation prices even though the two pipelines were not in the same physical market.

Furthermore, we obtained an overall view of the natural gas flows in the network of interstate pipelines. Specifically, we identified the flow between every pair of states, which can help us in analyzing the natural gas market in more depth in future work from a network point of view.

We also utilized the standard rational expectations competitive storage model to estimate natural gas inverse demand function as well as storage costs. We then quantified the effect of temperature, storage, and pipeline congestion on natural gas price. Our results suggest that these are the main factors explaining the price shocks in the natural gas market. Higher storage levels have a negative effect on price while higher heating or cooling degree days have a positive effect on price, and these effects are even higher in winter; however, the higher the storage level, the lower is the positive effect of heating degree days on price in winter. Moreover, higher pipeline congestion increases price while higher production decreases the price.

Finally, we investigated the network effect in the natural gas market, evaluating the effect of changes or shocks in one geographical region on all other regions. Our results confirm the existence of network effects suggesting that deviations of the heating degree days from the seven-year trend in one region have a significant effect on the price of natural gas in the other regions, even the farthest ones.

Now that we have analyzed the natural gas supply, demand, distribution, storage, and pricing in the U.S., we have to put all these pieces together and try to investigate the regulation and antitrust policy of the natural gas networked market. Which market participants enjoy market power, if any? Which mergers should be allowed and based on what? Which pipeline or storage operators should be allowed to offer market-based rates? These are the questions we will try to address as we continue this line of research.

In future work, in order to characterize which hubs are affected by a merger, we will use the concept of centrality. There are different notions of centrality. Degree centrality measures how connected a node is; closeness centrality measures how easily a node can reach other nodes; betweenness centrality measures how important a node is in terms of connecting other nodes; and finally there is a set of centrality measures (such as Katz prestige, eigenvector centrality, and Bonacich centrality) that measure neighbors' characteristics: how important, central, or influential a node's neighbors are in terms of connecting other nodes. We will most probably use one of the measures in the last set.

As another research direction, one can write down a model for the pipeline transportation price as the outcome of a bargaining process between the pipelines and the shippers. This allows us to observe how mergers change the bargaining power of the different agents in the market. Alternatively, to properly analyze mergers, one can formulate a network game where pipelines non-cooperatively set tariffs and all spot prices, consumption, and trade flows are solved simultaneously. This allows us to measure the impact of a merger in any part of the system on any other part.

Yet another line of research is to follow the literature on price integration in competitive networked markets to estimate the surcharges created by pipeline congestion. In the presence of market power, the estimated congestion surcharges may capture not only the shadow price of capacity but also rents associated with transient market power. However, the surcharge estimates can still highlight the potential value of new transportation infrastructure or inventory reserves. Moreover, while this literature only uses price data, since we have access to flow data, we may be able to disentangle strategic pricing from real congestion surcharges using this additional data.

Bibliography

Birge, John R., Timothy C.Y. Chan, J. Michael Pavlin, AND Ian Yihang Zhu (2022): “Spatial price integration in commodity markets with capacitated transportation networks," *Operations Research*, 70, 1739-1761.

Cafiero, Carlo AND Brian D. Wright (2011): “The empirical relevance of the competitive storage model," *Journal of Econometrics*, 162, 44–54.

Choi, Syngjoo, Andrea Galeotti, AND Sanjeev Goyal (2017): “Trading in networks: Theory and experiments," *Journal of the European Economic Association*, 15, 784–817.

Cremer, Helmuth AND Jean-Jacques Laffont (2002): “Competition in gas markets," *European Economic Review*, 46, 928–935.

Cremer, Helmuth, Farid Gasmi, AND Jean-Jacques Laffont (2003): “Access to pipelines in competitive gas markets," *Journal of Regulatory Economics*, 24.

Davis, Lucas W. AND Erich Muehlegger (2010): “Do Americans consume too little natural gas? An empirical test of marginal cost pricing," *The RAND Journal of Economics*, 41, 791–810.

Deaton, Angus AND Guy Laroque (1992): “On the behaviour of commodity prices," *The Review of Economic Studies*, 59, 1–23.

——— (1995): “Estimating a nonlinear rational expectations commodity price model with unobservable state variables," *Journal of Applied Econometrics*, 10, S9–S40.

Deneckere, Raymond AND Carl Davidson (1985): "Incentives to form coalitions with Bertrand competition," *The RAND Journal of Economics*, 16, 473–486.

Gourieroux, Christian, Alain Monfort, AND Alain Trognon (1984): "Pseudo maximum likelihood methods: Theory," *Econometrica*, 52, 681–700.

Gowrisankaran, Gautam (1999): "A dynamic model of endogenous horizontal mergers," *The RAND Journal of Economics*, 30, 56–83.

Gowrisankaran, Gautam, Aviv Nevo, AND Robert Town (2015): "Mergers when prices are negotiated: Evidence from the hospital industry," *American Economic Review*, 105, 172–203.

Horn, Henrick AND Asher Wolinsky (1988): "Bilateral monopolies and incentives for merger," *The RAND Journal of Economics*, 19, 408–419.

Laroque, Guy, AND Bernard Salanie (1994): "Estimating the canonical disequilibrium model: Asymptotic theory and finite sample properties," *Journal of Econometrics*, 62, 165–210.

Melo, Emerson (2017): "Imperfect price competition in networks," *Job Market Paper*.

Oliver, Matthew E., Charles F. Mason, AND David Finnoff (2014): "Pipeline congestion and basis differentials," *Journal of Regulatory Economics*, 46, 261–291.

Perry, Martin K. AND Robert H. Porter (1985): "Oligopoly and the incentive for horizontal merger," *The American Economic Review*, 75, 219–227.

Salant, Stephen W., Sheldon Switzer, AND Robert J. Reynolds (1983): "Losses from horizontal merger: The effects of an exogenous change in industry structure on Cournot-Nash equilibrium," *The Quarterly Journal of Economics*, 98, 185–199.

Sturm, Fletcher J. (1997): "Trading natural gas: cash futures options and swaps," PennWell Books.

3.12 Appendix

Table 3.14: Two-part tariff estimates ($/Mcf) for the industrial sector by state

State	Fixed Charge	Per-unit Charge	R^2	State	Fixed Charge	Per-unit Charge	R^2
AL	15243.10	2.95	0.06	NE	1089.83	4.19	0.38
AZ	3107.43	6.68	0.34	NV	32393.28	3.56	0.50
AR	25399.51	3.84	0.26	NH	425.68	9.83	0.62
CA	13708.39	-0.85	0.06	NJ	1705.06	6.07	0.27
CO	-2154.82	8.63	0.57	NM	-1590.64	6.64	0.21
CT	-416.22	9.01	0.38	NY	825.57	7.71	0.34
DE	-17321.25	11.64	0.80	NC	10115.80	4.18	0.07
FL	80138.12	1.92	0.25	ND	14166.35	3.01	0.32
GA	21075.03	2.19	0.12	OH	7317.91	5.96	0.16
ID	55521.15	2.22	0.07	OK	-12139.72	12.04	0.35
IL	1984.62	5.13	0.18	OR	681.18	7.10	0.30
IN	4413.24	6.19	0.22	PA	-6202.84	11.10	0.42
IA	10429.70	5.09	0.36	RI	3845.33	8.85	0.55
KS	-330.44	6.41	0.17	SC	49283.89	-4.48	0.13
LA	174101.92	2.54	0.04	TN	14033.02	2.55	0.04
ME	-16421.48	11.45	0.75	TX	39844.32	2.19	0.09
MD	2063.94	8.70	0.47	UT	-866.29	5.97	0.49
MA	604.62	9.01	0.61	VT	-2711.85	7.15	0.51
MI	3215.20	5.26	0.43	VA	14205.53	4.06	0.08
MN	-8673.62	7.16	0.55	WA	-5528.49	11.71	0.40
MS	37774.07	1.93	0.06	WV	29472.60	4.81	0.22
MO	2025.22	7.57	0.35	WI	-2411.45	8.98	0.51
MT	12691.10	4.86	0.40	WY	-855.83	5.87	0.19

??

Table 3.15: Two-part tariff estimates ($/Mcf) for the residential sector in Pennsylvania assuming LDCs file a rate case every two years

Period	Fixed Charge	Per-unit Charge
1990-1991	7.59	5.84
1992-1993	8.81	6.20
1994-1995	10.49	6.14
1996-1997	8.96	7.27
1998-1999	8.88	7.35
2000-2001	11.45	9.01
2002-2003	16.16	9.60
2004-2005	11.96	13.58
2006-2007	17.21	13.08
2008-2009	9.23	12.53
2010-2011	12.72	10.29
2012-2013	15.78	9.61
2014-2015	12.32	8.86
R^2	0.94	

??

Figure 3.11: Trend in two-part tariff charges ($/Mcf) in Pennsylvania assuming LDCs file a rate case every two years

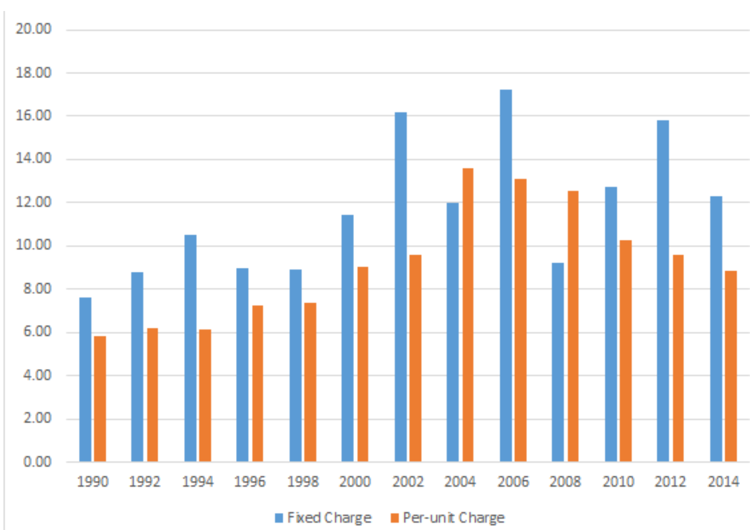


Table 3.16: Price-elasticity of demand by sector: WTI, heating and cooling degree days in all other states as instruments

	Residential	Commercial	Industrial
Log(Price)	-0.2286^{***}	-0.1179^{***}	-0.0314^{*}
	(0.0172)	(0.0181)	(0.0187)
Log(WTI Crude Oil Price)			0.0519^{***}
			(0.0091)
Heating Degree Days	0.0007^{***}	0.0006^{***}	0.0002^{***}
	(0.0000)	(0.0000)	(0.0000)
Cooling Degree Days	-0.0003^{***}	-0.0001^{***}	0.0000
	(0.0000)	(0.0000)	(0.0000)
state*year FE	Yes	Yes	Yes
state*month-of-year FE	Yes	Yes	Yes
R^2	0.9913	0.9887	0.9909

??

Table 3.17: Price-elasticity of demand by sector: WTI, heating and cooling degree days in all other states as instrument; elasticity changing with temperature

	Residential	Commercial	Industrial
Log(Price)	-0.2065^{***}	-0.1097^{***}	-0.0337^{*}
	(0.0182)	(0.0181)	(0.0189)
Log(WTI Crude Oil Price)			0.0508^{***}
			(0.0091)
Heating Degree Days	0.0006^{***}	0.0005	0.0001^{***}
	(0.0000)	(0.0000)	(0.0000)
Cooling Degree Days	-0.0004^{***}	-0.0001^{***}	0.0000
	(0.0000)	(0.0000)	(0.0000)
Log(Price)*Heating Degree Days	0.0000^{***}	0.0000^{***}	0.0000^{***}
	(0.0000)	(0.0000)	(0.0000)
Log(Price)*Cooling Degree Days	0.0000^{***}	0.0000^{***}	0.0000
	(0.0000)	(0.0000)	(0.0000)
state*year FE	Yes	Yes	Yes
state*month-of-year FE	Yes	Yes	Yes
R^2	0.9914	0.9887	0.9909

??