Three Essays in Applied Microeconomic Theory

by

Jooyong Jun

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2009

Doctoral Committee:

Professor Stephen W. Salant, Co-Chair
Associate Professor Uday Rajan, Co-Chair
Assistant Professor Daniel Susman Silverman
Assistant Professor Amitabh Sinha

To Eunjung

Acknowledgements

I extend my sincere thanks to the members of my dissertation committee. I am particularly grateful to my co-chairs, Uday Rajan and Stephen Salant. They have been great advisors in every stage of this dissertation. Their critical comments were greatly helpful for me to sharpen rudimentary ideas into implementable research topics and their insightful advice improves this dissertation significantly. In addition to superb academic advising, they have provided endless encouragement and support over all my years in Michigan. Without their advice and support, even this small achievement would be impossible for me. I also want to give my sincere gratitude to Daniel Silverman and Amitabh Sinha for fruitful feedback and suggestions. The support from the Department of Economics and the University of Michigan is appreciated as well.

The second and third essays of the dissertation are joint work with Kyoung-soo Yoon. His profound knowledge of microeconomic theory was vital in those chapters.

My parents and sister have been ardent supporters at every stage and aspect of my life. Without their support, I would not be able to achieve anything at all. I cannot thank them enough.

Finally, I dedicate this dissertation to my wife, Eunjung Yeo. She guided me to the joy, as well as the pain, of economics research, and has been my best colleague, friend, critic and supporter during my study in Michigan.

Table of Contents

List of Figures

Figure

Chapter 1

Introduction

This dissertation consists of three essays studying various issues in applied microeconomic theory.

The second chapter proposes a model of two-party elections in which voters' preferences over candidates is affected by their perception of candidates' sincerity. Unlike valence, voters' assessments of a candidate's sincerity depend on his platform choice for the current election because of his political baggage. In equilibrium, the following properties are observed. First, when candidates seek to maximize their winning-probability, higher baggage leads to more extreme platform choice from the median, but valence shows *no* influence. Second, the equilibrium outcome is more sensitive to the change in voters' policy preference when candidates seek to maximize their chance of winning than their expected share of votes. Finally, in a tight electoral race between share-maximizing candidates, the equilibrium outcomes that obtain under simultaneous and sequential platform choice coincide and are completely insensitive to a small change in the voter distribution. Our results provide an alternative explanation for the cause and extent of policy extremism and rigidity in political competitions.

The third chapter considers optimal contracts when a principal obtains advice from experts. When advisors have conditionally uncorrelated signals about an unknown true state of nature, aggregating multiple signals can lead to a better outcome than just choosing the single best signal. In this paper, we examine a principal's design of an optimal compensation scheme to screen the type of each expert, or the precision of his signal, for proper aggregation of multiple signals. Under a Gaussian specification of information, we show that there exists a compensation scheme which achieves the first best outcome: each expert is induced to honestly report his posterior on the true state, to truthfully reveal his type, and is paid only his reservation utility. Further, the optimal contract is a linear function with respect to a convex function of the mean squared error. The optimal contract satisfies a single crossing property with respect to the fixed wage component of the compensation and the incentive component which depends on the prediction error. This result comes from the cheap-talk feature of the professional advising, which implies that an agent's payoff solely depends on the transfer from a principal.

The fourth chapter considers the optimal employment problem faced by a monopolistic employer when potential employees are heterogeneous in ability and the marginal contribution of an expert depends not only on his own ability, but also on the abilities of the remaining employees. If an increasing submodular production function of indivisible employees shows increasing difference in marginal production, or decreasing concavity, and the reservation wage schedule is a fraction of the production by single employee, the optimal employment portfolio can be described by a cutoff element: all employees with greater ability than the cutoff level are hired and the rest are not. Moreover, the efficiency in employment is achieved through myopic decisions.

Chapter 2

Political Baggage and Electoral Competition

2.1 Introduction

Voting decisions are based in part on perceptions of candidates' sincerity. When a candidate announces his platform, voters often question whether the candidate is really committed to that platform, and whether he has the expertise or ability to implement it. The candidate's political history influences voters' perception about his sincerity on the announced platform.[1]

For example, in the 2004 U.S. presidential election, a key difference between George W. Bush and John Kerry was their stance on the Iraq war. During his campaign, Kerry suffered from the accusation of flip-flopping over his position on the war. Whether this accusation was correct or not, it influenced voters' evaluation of his commitment to his promised policies, even among voters who were not completely happy with the policies of Bush at that time.

In this paper, the canonical model of two-party electoral competition (e.g. Downs

[1] For convenience, "he" is used as the pronoun for a candidate, and "she" for a voter.

(1957)) is extended to allow voters' preferences to be affected by their assessment of the candidates' sincerity as well as their platform choice and valence. Every candidate has a political history which includes, for example, his historical policy platforms, experiences, name recognition, party affiliation, and successful delivery of past promises. These comprise his political "baggage." A candidate's political baggage can be considered a political liability for him. Because of such baggage, voters may discount a candidate's sincerity if the candidate changes his platform from his historical one. Their belief in a candidate's sincerity will be weaker if the candidate has a larger baggage or changes his platform by a lot. On the other hand, political baggage does not affect the evaluation of a candidate's valence. This simple but non-trivial addition of realism to the canonical two-party electoral competition model leads to dramatic changes in various equilibrium properties, many of which are unique to this model or more pronounced than in the previous literature.

We show the existence of a pure-strategy Nash equilibrium in which candidates' platform choices do not converge. In this equilibrium, the following properties are observed. First, when candidates seek to maximize their probability of winning, the degree of policy extremism is *not* affected by the candidates' valence but by their baggage.[2] Second, the equilibrium outcomes are affected by candidates' electoral objectives: maximizing their probability of winning or expected share of votes. The equilibrium outcome is sided to a favored candidate's historical platform and

[2]We use *policy extremism* to refer to the distance from the median voters' ideal to a candidate's platform and use *policy rigidity* to refer to the inertia in platform choice with the change of the voters' policy preferences distribution.

fluctuates more with respect to a change in the voter distribution when candidates seek to maximize their probability of winning. Finally, when candidates maximize their expected share of votes in a tight electoral competition, the equilibrium outcomes of simultaneous and sequential platform choice coincide and become *completely insensitive* to a small change of the median voter platform location. While equilibrium policy choices fluctuate more with the change of voters' policy preference under winning-probability maximization, policy rigidity is more salient under expected share maximization.

The equilibrium platform choices reflect the trade-off between preserving a candidate's perceived sincerity and aligning his platform with voters. Due to his political baggage, a candidate's perceived sincerity is greater when he stays around his historical platform than when he is close to the median voter's ideal. Suppose that the candidate's current platform is the median voter platform. At this platform, the first derivative of the median voter's utility with respect to platform change is zero. Thus, a marginal departure toward his historical platform has no first-order effect on the median voters' utility from his platform but it strictly increases the evaluation of his sincerity. Thus, the candidate has good reason to move away from the median voter platform, toward his historical platform, if he wants to maximize the median voter's utility.

Note that only the marginal change of perceived sincerity, which is affected by political baggage, affects the equilibrium platform choices. In other words, policy extremism is affected only by the size of political baggage; valence has no influence.[3] In equilibrium, for each candidate, the marginal increase of the median voters' utility from policy preference is equal to the marginal decrease of their utility from his perceived sincerity. The candidate with greater baggage will choose a platform farther from these voters' ideal than his opponent.

Different electoral objectives affect the degree of policy extremism and rigidity in policy choice, topics which are little investigated in the previous literature. Policy extremism becomes more of an issue when a favored candidate seeks to maximize his probability of winning. For a favored candidate, one more likely to earn the majority of votes, the benefit of reaching more voters is lower under winning-probability maximization. He wants to secure his base by offering voters higher perceived sincerity rather than reach more voters. His platform choice is, therefore, closer to his historical platform than under expected share maximization.[4] On the other hand, an unfavored candidate's equilibrium choice shows the opposite pattern. Thus, the equilibrium platform choices when candidates care about winning-probability are closer to the favored candidate's historical policy platform than when they care about the expected share.

Policy rigidity becomes a more salient issue when candidates try to maximize their expected share of votes in a tight electoral race. In this case, the equilibrium platform choices become completely insensitive to a small change in the median voter

[3]Note that the size of baggage by itself does not predict which candidate chooses the more extreme platform.

[4]We refer to a candidate's platform choice as more *stubborn* if it is closer to his historical platform.

platform. Further, the equilibrium outcomes of simultaneous and sequential platform choice coincide. The optimal platform for the Stackelberg leader is actually located at a kink in this case, which happens to be also the Nash equilibrium platform choice. The Nash equilibrium platforms in this case are where each candidate's best response functions turn from strategic substitutes into strategic complements. Note that this kind of insensitivity does not appear when candidates try to maximize winning-probability where attracting the median voters is the dominant strategy.

This result also provides alternative explanations to the cause and extent of policy rigidity as a result of political competition, without assuming influence from special interest groups or continuation of a specific policy. While previous papers such as Coate and Morris (1999) and Grossman and Helpman (1994) provide excellent explanations of policy persistency, defined as continuation of a specific unpopular policy, none of this previous work predicts the emergence of the complete insensitivity of equilibrium platform choices with respect to voters' policy preferences change.

The effect of a candidate's valence on spatial electoral competitions has been widely studied. Ansolabehere and Snyder Jr. (2000), Aragones and Palfrey (2002) and Groseclose (2001) show that the difference in candidates' valence leads to policy platform divergence in equilibrium. In the setting of Kartik and McAfee (2007) and Callander and Wilkie (2007), voters have uncertainty about a candidate's character, equivalent to valence, and interpret the candidate's platform as a signal about his character as well. Unique perfect Bayesian equilibrium appears when a candidate type is interested only in securing office strategically mixes or obscures his platform choices. Overall, as Bernhardt et al. (2008) note, the higher valence candidate chooses a more moderate policy to increase his electability if each candidate's valence is publicly known to voters, as in the first three papers. Conversely, the higher valence candidate chooses a more extreme policy to signal his valence if each candidate's valence is unknown to voters, as in the second two papers. Empirical evidence, however, shows no definite support for either of these predictions.[5]

Our result predicts that candidates' valence does not affect the degree of policy extremism.[6] It is a significant departure from the predictions of the previous literature and possibly explains the non-consensus of the previous empirical studies.

When the evaluation of a candidate's non-policy attributes, like sincerity in this paper, is affected by his platform choice, an equilibrium outcome reflects a compromise between the needs for preserving the level of his non-policy attribute and aligning his policy platform with voters' preferences.[7] For example, Berger et al. (2000), Bernhardt and Ingberman (1985), and Wittman (2007) consider reputational cost proportional to the distance between a reference platform and the one for the current election, with a sequential move game structure and under the objective of

[5]For example, Fiorina (1973) shows evidence supporting the first three papers while Griffin (2006) provides counter evidence and support the second two. See Groseclose (2001) and Bernhardt et al. (2008) for further information.

[6]Note, however, that the level of a candidate's perceived sincerity affects the chance of winning and the margin of victory.

[7]The cost of platform change is not a simple issue because the outcomes often critically depend on it. There are other papers with different assumptions on the cost of platform change such as commitment with no cost (Aragones and Palfrey, 2002), fixed cost (Callander and Wilkie, 2007; Kartik and McAfee, 2007), or cheap-talk (Ottaviani and Sørensen, 2006b). A recent empirical study by DeBacker (2008) provides evidence supporting the cost proportional to the degree of platform change.

expected share maximization. Eyster and Kittsteiner (2007) propose a parliamentary election model where political parties choose their platforms first, and then candidates choose their own platforms. Distance from a candidate's platform to her/his party's incurs cost to the candidate. Our model incorporates these observations and provides additional findings about the equilibrium platform choices of winning probability maximizing candidates, in a sequential game with a different order of platform choices, and in a simultaneous move game, questions not covered in their work.

In our model, candidates' electoral objectives, either maximizing probability of winning or expected share of votes, affect the equilibrium outcome. This result provides a contrast with previous work such as Duggan (2000), Banks and Duggan (2005), and Patty (2002), which takes a probabilistic voting approach and predicts the equivalence of equilibrium platform choices under both objectives. Our model violates one key assumption in these three papers: the independence of utility shocks across voters. Our result is more in line with comparative institutions models such as Persson and Tabellini (1999), Lizzeri and Persico (2001) and Persson et al. (2000), which investigate the effect of political institutions difference on fiscal and redistribution policies.

The rest of this paper is organized as follows. Section 2 outlines the model. Section 3 characterizes equilibrium platform choice under both winning probability maximization and expected share maximization. Section 4 compares equilibrium outcomes, policy rigidity and extremism observed under different electoral objectives.

Section 5 contains a discussion of the results and their applications.

2.2 Model

Consider an election that proceeds in two stages. First, each candidate chooses a platform. At the second stage, voters observe policy platforms. In addition, there may be external news, positive or negative, about each candidate. Each voter then chooses the candidate giving her the higher utility. The candidate with the most votes either wins the election or holds political power proportional to the share of votes.

2.2.1 Players

There are two candidates. An incumbent (i) faces an election against a new entrant (e). Candidate j ($j \in \{i, e\}$) chooses his platform w_j along from a convex policy set $\Theta \subseteq \Re$. A candidate's political history is abstracted by his *historical platform*, $h_j \in \Theta$. For example, the historical platform may represent the candidate's expertise or his platform choices in the past. Each candidate has political baggage, which is a liability if he wants to alter his platform away from his historical platform. The incumbent has larger political baggage compared with an entrant. That is, it is more costly by the incumbent to move away from his historical platform.

Voters are distributed according to a symmetric, atomless distribution $F(\theta)$, with a median m over Θ. A voter has a single-peaked policy preference. Her ideal platform $\theta \in \Theta$ is interpreted and defined as her type. She chooses a candidate who gives her a higher utility.

We consider two cases: i) the candidates receive utility only from winning and try to maximize their winning probability, and ii) where candidates maximize the share of votes. We assume that the historical platform of each candidate is strictly in the interior of Θ. All information is public, and both acquiring information and casting a vote are costless. Hence, there is no abstention and no "swing voter's curse" (Feddersen and Pesendorfer, 1996).

2.2.2 Voter's Utility

The platform chosen by candidate j, w_j, affects the utility of voters in two ways. First, as is standard, the closer is candidate j's platform w_j to a voter's ideal platform θ, the higher is the voter's utility from her political preference. We define $v(w_j, \theta)$ to represent voter's utility from the closeness of platform, termed *policy preference*, which is single-peaked and symmetric at $w_j = \theta$.

Second, voters form an evaluation of candidate j's non-policy attribute, proxied as his *sincerity* $s_j(w_j; h_j)$, represents the voter's utility from candidate j's sincerity, which is a function of his current platform w_j and his historical platform h_j. Note that s_j is independent of θ: all types of voter have an identical assessment about candidate j's sincerity.

In addition to sincerity, there is another kind of non-policy attribute, termed as *valence* which is not affected by political baggage. Voters' evaluation of candidate j's valence z_j is, therefore, the same regardless of his platform choice and given as a constant.

There exists a common utility shock ε_j to all voters about candidate j. This shock reflects news about the candidate that may emerge between the time the platform is chosen and the election day. This commonness implies that the realization of randomness, ε_i and ε_e, respectively, is perfectly correlated for all voters.

Note that the magnitude of the difference in candidates' utility has *no* effect on voters' decision; only the sign of the difference, or who gives the higher utility, matters. In other words, the voting behavior follows the *Downsian* approach.

A voter's utility from candidate j, if elected, is defined by the sum of these elements as

$$u_j(w_j, \theta, \varepsilon_j) = v\left(w_j, \theta\right) + s_j(w_j; h_j) + z_j + \varepsilon_j \tag{2.1}$$

As Feddersen and Pesendorfer (1997) note, a candidate's policy platform can be considered as a private value to voters while their evaluation of his non-policy attributes like sincerity and valence as common values.

A candidate's platform choice w_j is assumed to be commonly known to all candidates and voters. There is no uncertainty for a voter's policy preference on candidates. On the other hand, the common random shock ε_j is realized between the time of platform choice and the actual votes. This uncertainty creates the positive chance to win for a lesser regarded candidate, and removes the instable, knife-edge characteristic of the canonical model's prediction.

We make the following assumptions.

A.1 Each ε_j ($j \in \{i, e\}$) is independently drawn from an identical, well-defined, and zero-mean distribution. Further, ε_j can be sufficiently high or low to lead to a positive chance of winning at any platform.

Given assumption A.1, even a less favored candidate has a positive probability of winning. This feature greatly reduces the set of equilibria.

We have a standard regularity assumption for $v(\cdot)$ as follows:

A.2 $v(w_j, \theta) = v(|w_j - \theta|)$ is continuous, strictly concave, and twice differentiable.

We assume the following properties of sincerity $s_j(w_j; h_j)$.

A.3 i) Voters' evaluation of sincerity $s_j(w_j; h_j) = s_j(|w_j - h_j|)$ is a differentiable function with $s'_j < 0$ and $s'_j = 0$ at $w_j = h_j$.

ii) At any given distance $d > 0$ from a candidate's historical platform, or $d = |w_e - h_e| = |w_i - h_i|$, $s''_i < s''_e$ and, consequently, $0 > s'_e > s'_i$.

iii) $s_e(h_i; h_e) + z_e < s_i(h_i; h_i) + z_i$, and $s_e(h_e; h_e) + z_e > s_i(h_e; h_i) + z_i$.

iv) At any w_j and θ, $|s''_j| < |v''_j|$ is always satisfied.

In A.3, i) says that the wider gets the distance between h_j and w_j, the lower is the voters' evaluation of candidate j's sincerity, $s_j(\cdot)$. ii) states that the incumbent's baggage leads to a larger marginal cost for changing his platform. In other words, *the incumbent has greater baggage.*[8] It is also a mathematical definition of greater

[8] Intuitively, once the incumbent has built his reputation at a certain historical platform, his political history is less useful to evaluate him if he moves farther from this platform. On the other hand, the entrant may be relatively free from this constraint due to lack of (or shorter) political history, or the challenging party may choose a politician who fits better on a certain platform out of its candidate pool.

baggage in our model. iii) says that if both candidates choose their platforms at candidate j's historical platform h_j, candidate j has strictly greater perceived sincerity than his opponent. Moreover, i) and iii) lead that within the range of $\boldsymbol{\Theta}$, there exists a unique platform $\overline{w}$ which satisfies $s_e(\overline{w}; h_e) = s_i(\overline{w}; h_i)$. Finally, iv) states that the change in policy preference is always steeper than that in sincerity. Note that the differences between an incumbent and an entrant exist only in ii) and nowhere else. For this reason, any property for one candidate also holds for the other as long as ii) is not required.

Finally, without loss of generality, we assume that the incumbent's historical platform, h_i, is on the left of the median voter platform m, and the entrant's, h_e, is on the right of m. That is, $h_i < m < h_e$.

2.2.3 Platform Choice and the Envelope of Expected Utility

With the Downsian approach, if one candidate successfully attract voters of type θ, at least the whole voters on one side would also choose that candidate. This property simplifies a candidate's electoral objective as attracting a specific type of voters (e.g. the median voters). We will investigate and use it further in the next chapter.

A voter of type θ has *expected* utility U_j from choosing candidate j as

$$U_j(\theta, w_j) = E[u(w_j, \theta, \varepsilon_j)] = v(w_j, \theta) + s_j(w_j; h_j) + z_j \tag{2.2}$$

With A.2 and A.3, we expect that $U_j(\theta, w_j)$ is continuous, differentiable, single-peaked and, therefore, the first order condition $\frac{\partial U_j(w_j,\theta)}{\partial w_j} = 0$ must be satisfied at $w_j = w_j^\theta$, where w_j^θ is defined as $w_j^\theta \in \arg\max_{w_j \in \boldsymbol{\Theta}} U_j(\theta, w_j)$. In other words, candidate j maximizes the expected utility for the voters of type θ when he chooses w_j^θ as his platform. The following proposition states that this is indeed true and this platform w_j^θ is not θ.

Proposition 2.1. *Candidate j's ($j \in \{i, e\}$) most appealing platform to the voter of type θ, w_j^θ, is (i) unique and falls between h_j and θ unless $\theta = h_j$, and (ii) a monotone increasing and continuous function of θ.*

Proposition 2.1 explains the reason why a candidate in general *does not* choose his platform $w_j = \theta$ to give the highest expected utility to the voters of type θ. If a voter has ideal point $\theta \neq h_j$, then candidate j's platform most appealing to that voter of type θ will *never* be θ since a marginal departure from θ will cause no first order loss in v but will cause a first order gain in z. It is a crucial property that leads to the non-convergence of equilibrium platform choices. The continuity of U_j implies that any w_i can be a platform that maximize a type θ except for the small fringe regions around the bounds of policy space Θ, which do not need considering.

Proposition 2.1 also implies that $U_j(\theta, w_j^\theta)$ is a continuous and differentiable function of $\theta \in \Theta$. It is actually the envelope of the expected utility for candidate j. For simplicity, we let $\overline{U}_j(\theta) := U_j(\theta, w_j^\theta)$ and call $\overline{U}_j(\theta)$ as candidate j's *envelope*. Figure 2.1 graphically describes the expected utilities and envelopes. Envelope determines which candidate is expected to win the majority of votes in the current election and the boundary of the voters' type a candidate is expected to attract. We investigate further in the next chapter.

A candidate's envelope specifies the range of types that a candidate cannot be expected to capture *ex ante*. For example, given the entrant's platform w_e, if $U_e(\theta, w_e)$ is higher than $\overline{U}_i(\theta)$ at a platform θ, the voters of type θ would not expected to prefer the incumbent regardless of his platform choice w_i.

Lemma 2.1. *Candidate j's envelope, $\overline{U}_j(\theta)$, decreases as θ moves away from j's historical platform h_j.*

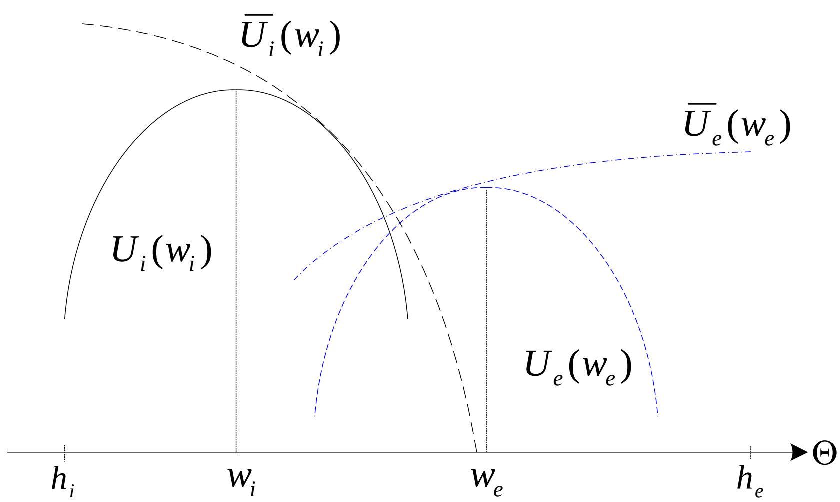

Figure 2.1: Expected utility and envelope

Finally, *Favor* in an election is defined to be used as follows in our model. As m moves closer to candidate j's historical platform, the election becomes more favorable for j.

2.3 Elections

In an election, each candidate chooses a platform to maximize their probability of winning or expected share of votes, depending on his electoral objective. Without loss of generality, we assume that the incumbent's historical platform is on the left of the entrant's, or $h_i < h_e$ and the median is in between.

We define the difference between the incumbent's utility and the entrant's utility for voters of type θ, $b(\theta; w_i, w_e)$, as

$$\begin{aligned} b(\theta; w_i, w_e) &= u(\theta, w_i, \varepsilon_i) - u(\theta, w_e, \varepsilon_e) \\ &= [v(|w_i - \theta|) + s_i(w_i; h_i) + z_i + \varepsilon_i] - [v(|w_e - \theta|) + s_e(w_e; h_e) + z_e + \varepsilon_e] \end{aligned}$$

and the difference between expected utilities as

$$\begin{aligned} B(\theta; w_i, w_e) &= U_i(\theta, w_i) - U_e(\theta, w_e) = E[b(\theta; w_i, w_e)] \\ &= [v(|w_i - \theta|) + s_i(w_i; h_i) + z_i] - [v(|w_e - \theta|) + s_e(w_e; h_e) + z_e] \end{aligned}$$

As well as the difference in candidates' utility, $b(\theta)$ can also be interpreted as the voters of type θ's inclination for the incumbent. With Downsian approach, voters of type θ would vote for the incumbent if $b(\theta) > 0$ and for the entrant if $b(\theta) < 0$. With probabilistic voting approach (Enelow and Hinich, 1981; Hinich, 1977), higher $b(\theta)$ implies that an individual voter have higher probability of choosing the incumbent.

2.3.1 Platform Choices under Winning-Probability Maximization

Each candidate's objective is to maximize the probability of winning the majority of votes. Remember that each candidate's platform choice decision should be made before the external common random utility shocks ε_i and ε_j, respectively, are realized.

With Downsian approach, if one candidate can successfully attract the median voters, the rest of the voters on one side also choose that candidate. Therefore, given the entrant's platform w_e, the incumbent's objective is represented as $\max_{w_i} \Pr\left(b(m; w_i, w_e) \geq 0\right)$, and, similarly, given w_i, the entrant's objective is represented as $\min_{w_e} \Pr\left(b(m; w_i, w_e) \geq 0\right)$. Each candidate tries to attract the median voters by maximizing the median voters' expected utilities from choosing the incumbent, $U_i\left(m, w_i\right)$, and the entrant, $U_e\left(m, w_e\right)$, respectively.

The following proposition specifies the equilibrium platform choices of both candidates and their property.

Proposition 2.2. *Under the objective of winning-probability maximization, it is a strictly dominant strategy for candidate j ($j \in \{i, e\}$) chooses the platform which maximizes $U_j(m, w_j)$ from choosing him. Further, candidate j's valence has no effect on his equilibrium platform choice.*

Because the first derivative, not the level, of candidate j's sincerity affects his choice, any non-policy value added, either positive or negative, to his sincerity does not change his equilibrium platform choice as long as the non-policy value is not affected by w_j. This result shows a good contrast with the previous literature about the correlation between valence and policy extremism. The higher level of sincerity, however, *does* lead to the greater chance of winning and the margin of victory in our model.

Note that each candidate's platform choice, w_i^m and w_e^m, is not affected by the opponent's platform. w_i^m and w_e^m are dominant platforms for each candidate who wants to maximize his winning probability. Therefore, whether it is a simultaneous move or a sequential move game, regardless who moves first, a candidate's equilibrium platform choice is always the same. Figure 2.2 describes this equilibrium. The incumbent and the entrant's equilibrium platform choices, w_i^m and w_e^m satisfy $w_i^m < m < w_e^m$.

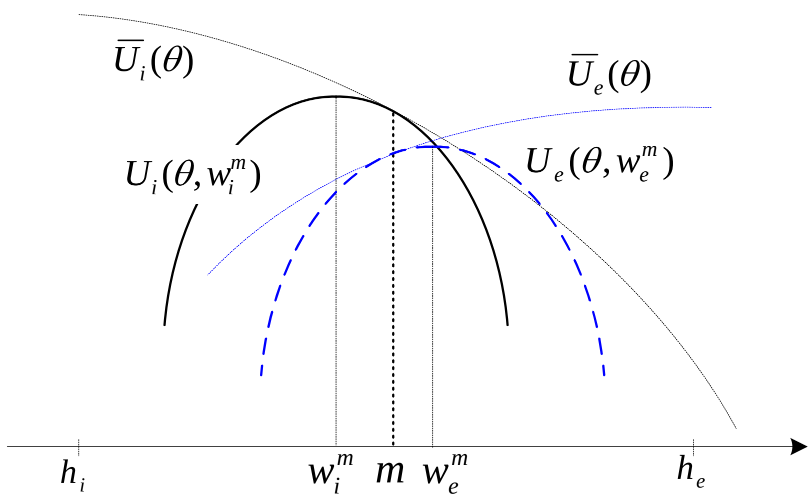

Figure 2.2: Equilibrium platform choice when candidates maximize their winning probability.

Proposition 2.3 states the comparative statics of policy extremism, which is defined as the distance from the median voter platform to a candidate's platform choice, and policy rigidity in equilibrium under winning-probability maximization.

Proposition 2.3. *In equilibrium under winning-probability maximization, we have the following properties: i) greater baggage leads to greater policy extremism, ii) a more favorable election leads to less policy extremism, and iii) greater baggage leads to greater policy rigidity.*

Note that Proposition 2.3 does *not* predict that the candidate with the greater baggage always chooses the more extreme platform than his opponent. For example, if the median voters are located sufficiently close to the historical platform of the candidate with the greater baggage, his platform choice can be less extreme than his opponent's.

2.3.2 Platform Choices under Expected-Share Maximization

Candidates now try to maximize their expected share of votes because a larger share of votes can lead to a bigger political power for the winning candidate, or the political institution actually adopts a proportional system. We assume that candidates are risk-neutral about the share of votes. Therefore, each candidate's platform choice is based only on the voters' expected utilities from choosing him.

Unlike the electoral competition between candidates maximizing their winning-probability, the game structure, simultaneous or sequential move game, now may change the equilibrium platform choices because there is no dominant strategy when candidates try to maximize their expected share. In the next chapter, however, we show that the equilibrium platform choices coincide for both sequential and simultaneous move games if the election is tight.

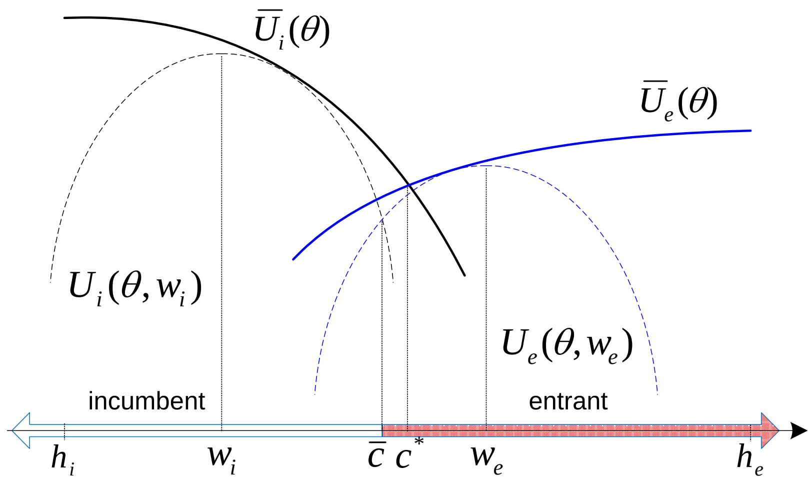

Figure 2.3: Expected utility, envelope, threshold platform and border platform

Figure 2.3 graphically describes the utilities, envelopes and some important platforms to understand equilibrium under expected share maximization. c^* is the platform at which both envelopes $\overline{U}_i$ and $\overline{U}_e$ intersect each other. A unique intersection of the envelopes exists only if A.3 i), ii) and iii) are satisfied. We define the intersection of both envelopes, c^* in Figure 2.3, as *border platform*. border platform c^* can be interpreted as the platform beyond which a candidate does not expect to attract the voters. w_i^* is the incumbent's platform where the voters at the border platform have the maximum utility from the incumbent and w_e^* is the entrant's platform where they have the maximum utility from the entrant. In other words, $w_i^* \in \arg\max U_i(c^*, w_i)$ and $w_e^* \in \arg\max U_e(c^*, w_e)$.

We define *threshold platform* $c(w_i, w_e)$ ($c \in \mathbf{\Theta}$) to be a platform where a voter at

c becomes a threshold voter, one who is indifferent between the incumbent and the entrant. In other words, we have $u(w_i, c, \varepsilon_i) = u(w_e, c, \varepsilon_e)$ and, therefore, $b(c; w_i, w_e) = 0$. We can also define expected threshold platform $\overline{c}$ in the same approach: $U_i(\overline{c}, w_i) = U_e(\overline{c}, w_e)$ and $B(\overline{c}; w_i, w_e) = 0$. Obviously, c (and $\overline{c}$) is a function of both candidates' platform choice w_i and w_e. For the later part of this paper, we use threshold platform to indicate expected threshold platform.

The following lemma shows the existence and uniqueness of this threshold platform.

Lemma 2.2. *If both candidates' platforms are not the same, there is a unique threshold platform* $\overline{c}(w_i, w_e)$.

Note the difference between the (expected) threshold platform $\overline{c}(w_i, w_e)$ and border platform c^*. $\overline{c}$ is the intersection of both candidates' expected and changes with the platform choices and a function of both candidates' platform choices. On the other hand, c^* is the intersection of the *envelopes* of their expected utilities and does not change with the platform choices. While the voters of type $\overline{c}$, or threshold voters, swing their votes, these votes are *not pivotal* in general.

Without loss of generality, we investigate the case of the entrant as a first mover. With Downsian model, a candidate's objective is now choosing a platform that send the expected threshold platform as far as possible. In other words, a candidate's objective is to maximize, or minimize the threshold platform $\overline{c}$.

From A.3.iii) we can derive a platform $\overline{w}$ which satisfies $s_i(\overline{w}; h_i) + z_i = s_e(\overline{w}; h_e)+$

z_e. It is obvious that $B(\theta, \overline{w}, \overline{w}) = 0$ at any θ in this case. $\overline{w}$ can be interpreted as the platform beyond which a candidate's sincerity becomes lower enough that he cannot be expect to attract any type of voters as a first mover. For example, if the entrant's platform choice w_e is left of $\overline{w}$, or $w_e < \overline{w}$, the incumbent can expect to attract all types of voters by setting his platform at the entry platform (i.e. $w_i = w_e$). The value of $B(\theta; w_i, w_e)$ is positive for all θ because the incumbent's initial sincerity at w_e is higher than the entrant's, or $r_i(w_i = w_e) > r_e(w_e)$ for any $w_e \leq \overline{w}$. Therefore, no equilibrium platform choice by the entrant includes the platform on the left of $\overline{w}$. The same logic is applied to the incumbent, too.

This situation can be compared to the Bertrand competition with asymmetric cost structure; the incumbent can be considered as a lower cost seller. Therefore, with Downsian approach, the incumbent's best response is to set his platform at the entrant's entry platform and receives all of the votes, as the lower cost seller slightly undercuts the opponent's price and gets all of the customers. It is obvious that the first-mover entrant should not choose his entry platform at $w_e \leq \overline{w}$ in equilibrium.

Unlike Bertrand competition, however, there exists no pure strategy equilibrium at $w_i = w_e = \overline{w}$. No tie breaking-rule can lead $(\overline{w}, \overline{w})$ to a pure-strategy equilibrium. We assume that in this case, the tie-breaking rule gives the second mover all of the votes to preserve continuity of best responses in a sequential move game.

Lemma 2.3. *There is no equilibrium in which $w_i = w_e = \overline{w}$: no tie-breaking rule that leads $w_i = w_e = \overline{w}$ to a pure-strategy Nash equilibrium.*

In the case of $\overline{w} \leq w_e$, on the other hand, it is obviously impossible for the incumbent to give higher expected utility to all types of voters. In other words, there exists some θ at which $U_i(\theta, w_i) < \overline{U}_j(\theta)$ is satisfied.

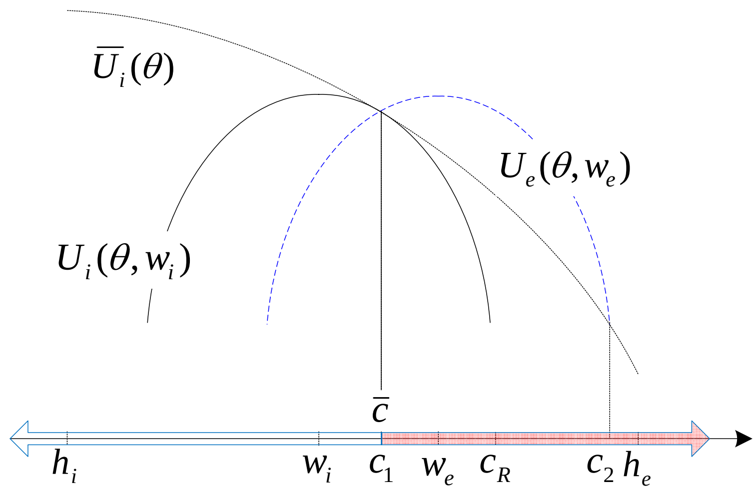

Figure 2.4: Incumbent's best response when candidates maximize their expected share. $(F(c_1) \geq 1 - F(c_2))$

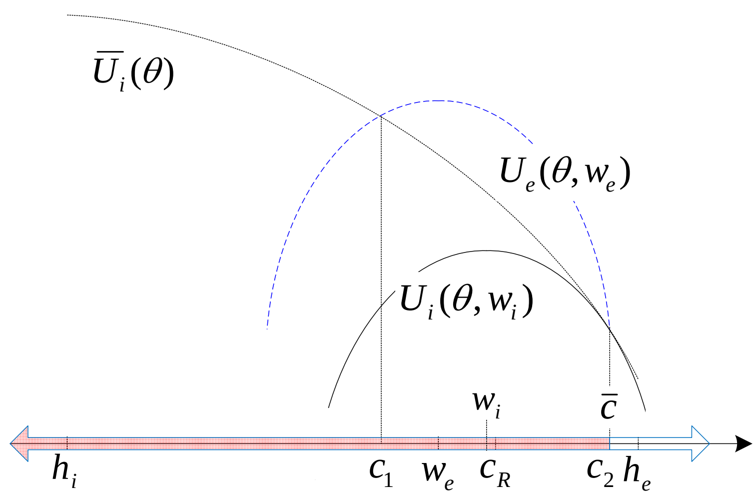

Figure 2.5: Incumbent's best response when candidates maximize their expected share. $(F(c_1) < 1 - F(c_2))$

From our Downsian assumption, we expect that the incumbent, who wants to maximize the threshold platform, also wants to maximize the expected utility for the voters of threshold platform, and this platform is expected to satisfy the first order condition $\frac{dU_i(\overline{c}, w_i)}{dw_i} = 0$. Figure 2.4 graphically describes this situation. If a threshold platform $\overline{c}$ is located as $w_i \leq \overline{c} \leq w_e$, the incumbent's share is $F(\overline{c})$ and the entrant's is $1 - F(\overline{c})$. If $w_e \leq w_i \leq \overline{c}$, the incumbent's share is $1 - F(\overline{c})$ and the entrant's share is $F(\overline{c})$. Given w_e, the incumbent's objective is given as

$$\max_{w_i}[\max\{F(\overline{c}\,(w_i, w_e)), 1 - F(\overline{c}\,(w_i, w_e))\}] \tag{2.3}$$

Figure 2.4 hints that there exist two platforms c_1 and c_2 ($c_1 < c_2$, $c_1, c_2 \in \Theta$) where $U_e(c_1, w_e) = \overline{U}_i(c_1)$ and $U_e(c_2, w_e) = \overline{U}_i(c_2)$ as long as $\overline{w} < w_e$. Note that c_1 and c_2 are solely determined by the entrant's choice w_e and can represented as a function of w_e, $c_1(w_e)$ and $c_2(w_e)$, respectively. Figure 2.4 also hints that the threshold platform $\overline{c}$ be either c_1 or c_2 when entrant moves first. The following lemma gives a formal proof for this claim.

Lemma 2.4. *Suppose $w_i < \overline{w} < w_e$. There exists two platforms c_1 and c_2 ($c_1 < c_2$) where $U_e(c_1, w_e) = \overline{U}_i(c_1)$ and $U_e(c_2, w_e) = \overline{U}_i(c_2)$, and the incumbent's best response leads the threshold platform $\overline{c}(w_i, w_e)$ to be either on c_1 or c_2.*

Figure 2.4 and Figure 2.5 describe the incumbent's best response. Given the entrant's platform choice w_e and corresponding $c_1(w_e)$ and $c_2(w_e)$, the incumbent's platform choice is now simply either $w_i^{c_1}$ or $w_i^{c_2}$. Let $c_R := \frac{c_1 + c_2}{2}$. Because of symmetry of the voter distribution, $F(c_1) > 1 - F(c_2)$ can be satisfied if and only if $m < c_R$.

Therefore, if $m < c_R$, the incumbent chooses $w_i^{c_1}$ which satisfies $\frac{dU_i(w_i|c_1)}{dw_i} = 0$. w_i is on the left of w_e in this case. Otherwise, he will choose $w_i^{c_2}$, which is on the right of w_e.

We now investigate the most important and unique property regarding a candidate's best response. The incumbent's best response is *non-monotone* even when the location of median m still lets the incumbent take the share of votes on the left of the median. The incumbent's best response is strategic substitute when the entrant's platform is located between $\overline{w} < w_e < w_e^*$ while it is strategic complement when $w_e^* < w_e$. Therefore, w_e^* is the entrant's platform where the incumbent's best response changes from strategic substitute to strategic complement. The following lemma shows that, without loss of generality, a candidate's best response is non-monotone with his opponent's platform.

Lemma 2.5. *When $m \leq c_R$, if $\overline{w} < w_e < w_e^*$, the incumbent's best response decreases as w_e increases, and it increases when $w_e^* \leq w_e$.*

Figure 2.6 shows an example of an incumbent's best response when the entrant platform is between $\overline{w}$ and w_e^*.

We now investigate the entrant's equilibrium platform choice as the first mover. If $m \leq c_R$, then $F(c_1) \geq 1 - F(c_2)$ and the incumbent will choose $w_i^{c_1}$ which satisfies $\frac{dU_i(w_i|\overline{c})}{dw_i} = 0$, and is on the left of w_e. Otherwise, he will choose $w_i^{c_2}$ which is on the right of the entrant.

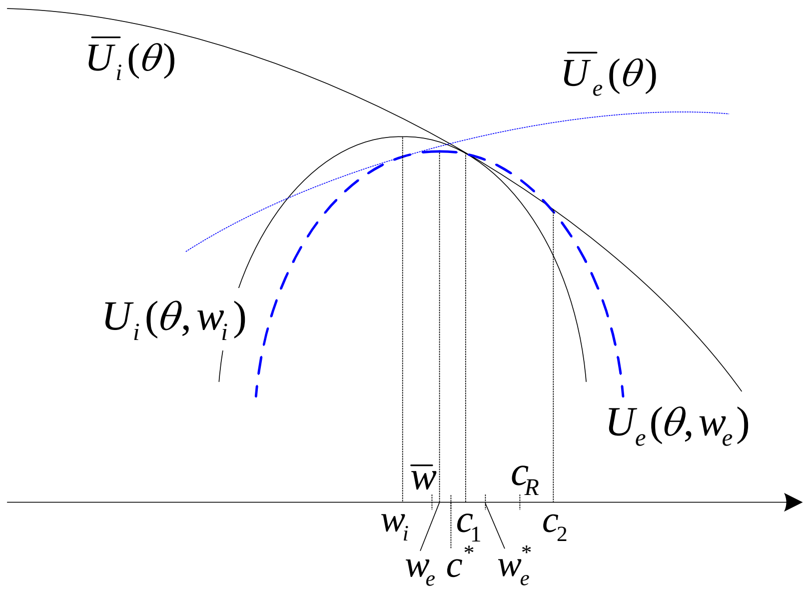

Figure 2.6: Incumbent's best response when $\overline{w} < w_e < w_e^*$.

Expecting the incumbent's best response, the entrant's objective is given as

$$\min_{w_e}[\max_{w_i(w_e)}\{F(\overline{c}\,(w_i, w_e)), 1 - F(\overline{c}\,(w_i, w_e))\}] \tag{2.4}$$

When the entrant's platform choice $w_e = w_e^*$, we know that $c_1 = c^*$. Let $c_R^* = \frac{c^*+c_2^*}{2}$ ($c^* < c_2^*$) where c_2^* and c_R^* are the corresponding c_2 and c_R when $w_e = w_e^*$. As the first mover, when $c_R^* \leq m$, the entrant chooses his platform further from his historical platform than his choice under winning-probability maximization. On the other hand, when $m \leq c_R^*$, he does not follow the median any more. The following proposition formally states the entrant's platform choice as the first mover.

Proposition 2.4. *Consider the game when the entrant moves first. If $m \leq c_R^*$, the entrant sets his platform at w_e^* and the incumbent follows at w_i^*. If $c_R^* < m$, the entrant sets his platform w_e so that $c_R(w_e) = \frac{c_1(w_e)+c_2(w_e)}{2} = m$, and the incumbent* *chooses either* $w_i^{c_1}$ *or* $w_i^{c_2}$ *in equilibrium.*

The same logic is applied when the incumbent moves first. With abuse of notation, $c_1(w_i)$ and $c_2(w_i)$ ($c_1 < c_2$) now represent the intersections of $U_i(\theta, w_i)$ and $\overline{U}_e$.

Let $c_L(w_i) := \frac{c_1(w_i)+c_2(w_i)}{2}$ and let $c_L^* := \frac{c_1^*+c^*}{2}$ where $c_1^* = c_1(w_i^*)$ and $c^* = c_2^* = c_2(w_i^*)$. Again, if w_i is located between w_i^* and $\overline{w}$, the entrant's best response *increases* as w_i decreases.

Proposition 2.5. *Consider the game when the entrant moves first. If* $c_L^* \leq m$*, the platform choices are* $w_i = w_i^*$ *and* $w_e = w_e^*$ *in equilibrium. If* $m < c_L^*$*, the incumbent chooses his platform* w_i *so that* $c_L = \frac{c_1(w_i)+c_2(w_i)}{2} = m$*, and the entrant chooses either* $w_e^{c_1}$ *or* $w_e^{c_2}$ *in equilibrium.*

Proof. It is just the mirror image of Proposition 2.4. □

The location of median voter platform can give an implication regarding how competitive an election is as well as whether a candidate is favored or not. If the median voter platform is closer to one candidate's historical platform, he is expected to be more strongly preferred among the majority of voters in equilibrium while the other voters who prefer his opponent only slightly more to him. Therefore, the candidate can expect the greater chance of victory or the larger share of votes. On the other hand, when the median voter location gets closer to the border platform, neither candidate can expect the greater chance of winning or larger share of votes as before. The electoral race in this case would become more competitive, or tight. We call an election is *one-sided* when $m \leq c_L^*$ or $c_R^* \leq m$: the former case is *incumbent-sided* and the latter case is *entrant-sided*. When $c_L^* \leq m \leq c_R^*$, we call this election *tight*.

In a one-sided election between candidates maximizing expected share, the seek-and-hide behavior reappears; a favored candidate wants to hide from the opponent while the opponent wants to manage a close, but not too far, distance. Note that it does not exclude the existence of a mixed-strategy equilibrium.[9]

Lemma 2.6. *In a one-sided election, there is no pure-strategy equilibrium in a simultaneous move game.*

In a tight election, however, a pure-strategy equilibrium exists for a simultaneous choice game between expected share maximizing candidates. Proposition gives a formal proof about the emergence of complete insensitivity of the equilibrium platform choices with respect to the median voter platform change in a tight election between candidates maximizing their expected share. Further, each candidate's equilibrium platform choice is the same regardless of the game structure or small change of the median as long as the election remains tight. It provides alternative explanations to the emergence of policy rigidity as a result of political competition, without assuming influence from special interest politics.[10]

[9] Theorem 5 in Dasgupta and Maskin (1986) suggests that there exists a mixed-strategy equilibrium. The formal proof and analysis of its property, however, is not provided here.

[10] Alesina and Holden (2008) provide an interesting prediction regarding the influence of special interest politics on policy choice. The compromise between contributions from special interest groups and proximity to the median voter platform leads to the emergence of platform *ambiguity*: announcement of a range of platforms instead of a specific one.

Proposition 2.6. *In a tight election between candidates maximizing their expected share, whether an election is a simultaneous or sequential move game, and regardless of the median voter location, there exists a unique pure-strategy equilibrium platform choice* (w_i^*, w_e^*).

Figure 2.7 graphically describe this equilibrium. Lemma 2.5 tells us that this equilibrium cannot be derived from a supermodular structure, which guarantees the existence of a Nash equilibrium (Milgrom and Roberts, 1990). While a candidate's policy platform is a strategic component, his perceived sincerity is not. Thus, if both candidates' gain from the best response for the strategic component can be lower than the loss of his non-strategic component value, candidates' best responses may not be monotone.

The non-monotonicity of best response in our model, however, actually helps the coincidence of the equilibria for both simultaneous and sequential choice games. The equilibrium platforms (w_i^*, w_e^*) are actually where both candidates' best response functions turn from strategic substitutes to strategic complement and the order of platform choice does not alter the equilibrium outcome for a sequential move game and, therefore, a simultaneous move game.[11]

The complete insensitivity in platform choices does not appear when candidates try to maximize their winning probability because, in that case, wooing the median

[11] The non-monotonicity of best responses is not a necessary condition for this coincidence of equilibrium outcomes. There can be pathological cases where the coincidence happens with monotone best responses if the payoff structures are different among players.

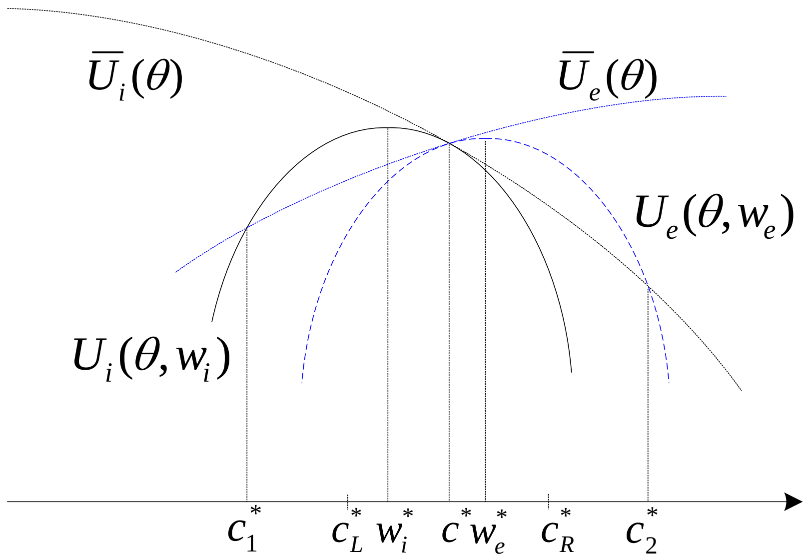

Figure 2.7: Equilibrium in a tight electoral competition between expected share maximizing candidates

voters is the dominant strategy regardless of the opponent's platform. It does not mean, however, that the insensitivity necessarily appears under expected share maximization, nor a change of the voter distribution does not matter. When an election is one-sided under expected share maximization and the favored candidate is a first mover in a sequential choice game, no insensitivity happens, either. This is because the first mover needs to set his platform so that $c_L(w_i)$ (or $c_R(w_e)$) $= m$. Voters' policy preferences distribution just affects which side of the opponent, left or right, a candidate's platform should be chosen at for his best response.

2.4 Probability of Winning Vs. Expected Share of Votes

As we have seen in the previous chapter, candidates' electoral objectives, maximizing probability of winning or expected share of votes, affect the degree and extent of policy extremism and policy rigidity. In this chapter, we focus on the remaining issues about the influence of electoral objectives on the equilibrium platform choices.

We refer that the incumbent is favored if $m < c^*$, and the entrant is favored if $c^* < m$, where m is the median voter platform. In other words, when both candidates choose the most appealing platform to the median voters, their expected utility from choosing the favored candidate is greater than choosing the other.

Proposition 2.7. *In any equilibrium under the expected-share maximization, the favored candidate's platform choice under winning-probability maximization is closer to his historical platform than his choice under expected share maximization while the unfavored candidate's platform choice shows the opposite characteristic.*

Proposition tells us that both equilibrium platform choices are closer to a favored candidate's historical platform when candidates try to maximize their chance of winning. Considering that the benefit of reaching more voters is important under expected-share maximization, compared to winning-probability maximization, this difference is in accordance with our intuition about the effect of electoral objectives. Proposition , therefore, leads to the following conclusion about equilibrium outcome fluctuation under different electoral objectives.

Corollary 2.1. *Equilibrium platform choices fluctuate greater with respect to voters' policy preferences change when candidates rather seek to maximize their probability of winning than their expected share of votes.*

We now look at the policy rigidity and extremism in a tight election. Proposition obviously states that policy rigidity is more salient when candidate try to maximize their expected share of votes. Policy extremism, on the other hand, is a little more complicated.

Proposition still implies that the unfavored candidate's platform choice is always more extreme under expected share maximization. The favored candidate's equilibrium platform, however, may sometimes pass the median voter's platform toward the opponent's platform, as we have seen at Proposition . In this case, the distance between the equilibrium platform and the median voter's ideal may be farther than that under winning-probability maximization.[12]

Still, based on Proposition , we have the following conclusion about policy extremism under different electoral objectives.

Corollary 2.2. *When the median voters' ideal platform m is located between w_i^* and w_e^*, or $w_i^* \leq m \leq w_e^*$, the favored candidate shows greater policy extremism when candidates maximize their probability of winning than expected share of votes. On the other hand, an unfavored candidate shows the opposite characteristic.*

2.5 Summary and Discussion

We consider a model of spatial electoral competition between two candidates when voters care about each candidate's sincerity as well as his policy platform. In

[12] We argue that policy extremism should be a lesser issue when the favored candidate's equilibrium platform overshoots the median voters' ideal. Considering that the reason for the overshooting is to reach more voters, it seems to be a little ironical to call this platform choice more extreme.

equilibrium, a candidate's political baggage leads to stubbornness in his platform choices and the non-convergence of equilibrium platforms. More baggage leads to greater stubbornness.

If each candidate maximizes his probability of winning, there exists a pure-strategy equilibrium with differing platform choices. The entrant, the candidate with less baggage, chooses a platform closer to the median voter than the incumbent regardless of the game structure. The level of perceived sincerity does not affect the equilibrium platform choices. This result differs from the conclusions of previous work dealing with candidates' valence (Aragones and Palfrey, 2002; Groseclose, 2001), which concludes that the candidate with the higher valence chooses a relatively more moderate platform than one with the lower valence. Our result is derived from the incumbent's stronger need to keep his expected sincerity greater than the entrant.

If each candidate instead maximizes his expected share of total votes, there exists a pure-strategy Nash equilibrium with differing platform choices, provided the election is tight. Each candidate's best response then becomes completely insensitive to a small change in the median voters' policy preference. This property comes from the non-monotonicity of best responses. The equilibrium platform choices occur where candidates' best responses switch from being strategic substitutes to strategic complements. Moreover, these equilibrium choices (under either electoral objective) are the same for both simultaneous and sequential move games. If an election is not tight, we have pure-strategy equilibrium only for a sequential move game, and when an unfavored candidate moves first, the equilibrium platform choices happen to be the same as in a tight electoral competition.

We have different equilibrium outcomes under different electoral objectives. In particular, under the objective winning probability maximization, the platform choice of a favored candidate, who possesses the higher expected utility for the median voters, is more stubborn than under the objective of expected share maximization rule while an unfavored candidate's choice shows the opposite behavior. Thus, the equilibrium platform choices are closer to the favored candidate's historical platform under winning probability maximization, compared to expected share maximization. Our result contrasts with some previous work that takes a probabilistic voting approach, and predicts the equivalence of equilibrium platform choices under both objectives (Banks and Duggan, 2005; Duggan, 2000; Patty, 2002). The perfectly correlated random shock and the Downsian voting approach in our model violate one of the key assumptions in this previous work: the independence of utility shocks across voters.

Our model reconciles seemingly contradictory views about the effect of uncertainty about a candidate's platform. For example, Shepsle (1972) and Alesina and Cukierman (1990) suggest that a candidate has an incentive to make his platform ambiguous while Enelow and Hinich (1981) suggest the opposite. A candidate with more ambiguity in his platform choice may be viewed as less sincere by voters. However, if a candidate can have a smaller political baggage by making his platform ambiguous, he would have less cost for his future potential platform change. A candidate, therefore, has an incentive to make his platform choice more ambiguous if the political winds change frequently, and less ambiguous if the political surroundings are stable.

The expected share maximization case may also be interpreted in terms of spatial competition model in differentiated products (Calem and Rizzo, 1995; Liu et al., 2004) where the good has both a private value and a common value component. It can be applied to such questions as whether a new brand should be launched when consumer tastes change. Extending our model to multiple brands, if the median consumer is shifted too far from the incumbent's prior location, it would be better to launch another brand rather than to produce a different product under the same brand name. The rise and fall of Pontiac gives an example. A *New York Times* article describes this situation as the follows: "When you deviate too far from it, that's when you run into trouble as a brand and a company," and "More than any other G.M. brand, Pontiac stood for performance, speed and sex appeal. Its crosstown rivals followed with similar muscle cars, giving Detroit bragging rights over the cars that Japanese automakers were selling based on quality and reliability."[13]

There are many issues left for future research. First, it is not clear whether welfare is maximized under either of the candidates' electoral objectives. Welfare comparisons between candidates' different electoral objectives in our model are not straightforward, since voters' care about sincerity as well. Another question of inter-

[13]Micheline Maynard, "Its Muscle Car Glory Faded, Pontiac Shrivels Up", *New York Times*, Feb. 19, 2009

est is the effect of candidates' risk preference over the share of votes on equilibrium outcomes. While we conjecture that greater risk aversion leads to greater rigidity in platform choices, we have not established this result yet. Finally, empirical studies are needed to check whether real world election data support our predictions.

2.6 Appendix: Proofs

2.6.1 Propositions

Proposition 2.1

Proof. Without loss of generality, the incumbent's case when $h_i < \theta$ is shown.

i) Suppose $\theta > h_i$. If w_i is set optimally, it maximizes the expected utility of voters of type θ. The first order condition $\frac{\partial U_i}{\partial w_i} = \frac{\partial v(w_i,\theta)}{\partial w_i} + \frac{\partial s_i(w_i;h_i)}{\partial w_i} = 0$ needs to be satisfied. With A.2 and A.3, for any $\frac{\partial s_i}{\partial w_i} < 0$, $\frac{\partial v(w_i,\theta)}{\partial w_i} > 0$ must be determined to satisfy the first order condition. Note that $\frac{\partial v(w_i,\theta)}{\partial w_i} > 0$ implies that $h_i < w_i^\theta < \theta$.

The strict concavity of $v(\cdot)$, or $v'' < 0$, implies that for any given type θ and the value of $\frac{\partial s_i}{\partial w_i}$, there exists a unique platform w_i^θ which satisfies the first order condition $\frac{\partial v(w_i^\theta,\theta)}{\partial w_i} + \frac{\partial s_i(w^\theta;h_i)}{\partial w_i} = 0$. A.3 iv) guarantees the satisfaction of the second order condition.

ii) Given a type θ, there always exists unique w_i^θ. At $w_i = w_i^\theta < \theta$, if θ is increased by any positive ϵ, strict concavity of v leads to the lower value of $\frac{\partial v(w_i,\theta)}{\partial w_i}$ and $\frac{\partial v(w_i,\theta)}{\partial w_i} + \frac{\partial s_i}{\partial w_i}$ becomes less than zero at $\theta + \epsilon$.

If w_i moves to the left of w_i^θ, the value of $\frac{\partial v(w_i,\theta)}{\partial w_i}$ becomes lower and $\frac{\partial s_i}{\partial w_i}$ higher.

From A.3 iv), the decrease in $\frac{\partial v(w_i,\theta)}{\partial w_i}$ is faster than the increase in $\frac{\partial s_i}{\partial w_i}$, which means that the value of $\frac{\partial v(w_i,\theta)}{\partial w_i}+\frac{\partial s_i}{\partial w_i}$ decreases. There exists no positive δ where $w_i = w_i^\theta - \delta$ satisfies the first order condition.

If w_i moves to the right of w_i^θ, the value of $\frac{\partial v(w_i,\theta)}{\partial w_i}$ becomes higher and $\frac{\partial s_i}{\partial w_i}$ lower. Now, the increase in $\frac{\partial v(w_i,\theta)}{\partial w_i}$ is faster than the decrease in $\frac{\partial s_i}{\partial w_i}$, which means the value of $\frac{\partial v(w_i,\theta)}{\partial w_i} + \frac{\partial s_i}{\partial w_i}$ increases. Because $\frac{\partial v(w_i,\theta)}{\partial w_i}$ and $\frac{\partial s_i}{\partial w_i}$ are continuous and differentiable with w_i, we can find some positive δ where $w_i = w_i^\theta + \delta$ satisfies the first order condition for any positive ϵ. Therefore, w_i^θ is a continuous and increasing function of θ. □

Proposition 2.2

Proof. Both candidates' platform choices do not affect the common random shocks ε_i and ε_e. Thus, given the opponent's platform choice, maximizing or minimizing $\Pr\left(b(m; w_i, w_e) \geq 0\right)$ is equivalent to a maximizing or minimizing $B(m; w_i, w_e)$.

For the incumbent, $B(m; w_i, w_e)$ is maximum at $w_i = w_i^m$ regardless of w_e. Only the first order condition of the median voters' expected utility from choosing the incumbent affects his equilibrium platform choice. For the entrant, $B(m; w_i, w_e)$ is obviously minimum at $w_e = w_e^m$ regardless of w_i for the same reason.

Finally, w_i^m and w_e^m satisfy $w_i^m = \arg\max_{w_i} \Pr(b(m; w_i, w_e) + k \geq 0)$, and $w_e^m = \arg\min_{w_e} \Pr(b(m; w_i, w_e) + k \geq 0)$ for any fixed value of k because, like ε_i and ε_e, k is not what candidates can alter. This implies that only the first momentum of each candidate's sincerity affects the equilibrium platform choice; valence has no effect. □

Proposition 2.3

Proof. We first look at the incumbent's case. The first order condition is

$\frac{\partial v(|w_i - m|)}{\partial w_i} + \frac{\partial s_i(w_i; h_i)}{\partial w_i} = 0$

The greater baggage means the lower value of the first derivative of sincerity $\frac{\partial s_i}{\partial w_i}$, which is negative, than before, and the higher value of the first derivative of policy preference $\frac{\partial v(|w_i - m|)}{\partial w_i}$, which is positive. Therefore, $|w_i^m - m|$ increases: w_i^m moves leftward and the distance from the median m becomes wider.

Suppose that an election becomes more favorable to the incumbent, or m moves closer to h_i. We already know that the incumbent's best response is w_i^m. The first order condition is given as

$$\frac{\partial v(w_i^m, m)}{\partial w_i} + \frac{\partial s(w_i^m; h_i)}{\partial w_i} = 0$$

and it must be always satisfied for any m. Thus, we have the following equation

$$\frac{\partial}{\partial m}\left(\frac{\partial v(w_i^m, m)}{\partial w_i} + \frac{\partial s_i(w_i^m; h_i)}{\partial w_i}\right) = \frac{\partial^2 v}{\partial m \partial w_i} + \frac{\partial^2 v}{\partial w_i^2}\frac{\partial w_i^m}{\partial m} + \frac{\partial^2 s_i(w_i^m; h_i)}{\partial^2 w_i}\frac{\partial w_i^m}{\partial m} = 0$$

which leads to the following equation:

$$\frac{\partial w_i^m}{\partial m} = \frac{-\frac{\partial^2 v}{\partial m \partial w_i}}{\frac{\partial^2 v}{\partial w_i^2} + \frac{\partial^2 s_i(w_i^m; h_i)}{\partial^2 w_i}}$$

Because $\frac{\partial^2 v}{\partial m \partial w_i} > 0$, $\frac{\partial^2 v}{\partial w_i^2} < 0$, $\frac{\partial^2 s_i(w_i^m; h_i)}{\partial^2 w_i} < 0$, and $\left|\frac{\partial^2 v}{\partial m \partial w_i}\right| = \left|\frac{\partial^2 v}{\partial m \partial w_i}\right|$ we can conclude that $0 < \frac{\partial w_i^m}{\partial m} < 1$. Thus, when the median m moves marginally toward left, w_i^m also moves toward left, but less than the marginal shift of the median, which means less policy extremism.

Finally, the greater baggage means that $\frac{\partial^2 s(w_i^m;h_i)}{\partial^2 w_i}$ is lower. Thus, the value of $\frac{\partial w_i^m}{\partial m}$ is also lower, which implies greater policy rigidity. The entrant's case can be drawn in the same way. □

Proposition 2.4

Proof. Two possible threshold platforms c_1 and c_2 are decided after the entrant choose a platform w_e. If the following inequality $F(c_1(w_e)) \geq 1 - F(c_2(w_e))$ is satisfied in equilibrium, the incumbent's platform choice is $w_i^{c_1}$, which gives the highest expected utility to the voter of type c_1, and the entrant's platform choice minimizes this c_1. We already know that the minimum c_1 is c^*.

If $m \leq c_R^*$, the incumbent's platform choice is $w_i = w_i^{c_1}$ because the entrant cannot minimize c_1 less than c^* and the median is still on the left of c_R^*. The entrant's platform choice is w_e^* and the incumbent's best response is w_i^*.

If $m \geq c_R^*$, the entrant's objective (2.4) leads to $F(c_1) = 1 - F(c_2)$ to minimizes the incumbent's share of votes. With the symmetric voter distribution, $F(c_1) = 1 - F(c_2)$ leads to the distance from median to c_1 and c_2 must be equal. Therefore, the entrant sets his platform w_e so that $m = c_R$. □

Proposition 2.4

Proof. In a tight election, the median voter platform m is located between c_L^* and c_R^*, or $c_L^* \leq m \leq c_R^*$. From Proposition 2.4 and 2.5, we know that w_i^* and w_e^* are platform choices as the first mover, respectively, and that at $w_e = w_e^*$, the incumbent's best response is $w_i^{c_1} = w_i^*$, and that at $w_i = w_i^*$, the entrant's best response is $w_e^{c_2} = w_e^*$. Therefore, whoever moves first, (w_i^*, w_e^*) is the equilibrium in a sequential move game. In addition, (w_i^*, w_e^*) is also a Nash Equilibrium because both are the best responses for each other.

Now we want to show (w_i^*, w_e^*) is the only pure-strategy equilibrium. We know that platform $w_i > \overline{w}$ for the incumbent and $w_e < \overline{w}$ for the entrant are never chosen as a first mover and can be ignored. Lemma 2.3 also tells us that $w_i = w_e = \overline{w}$ cannot be a pure-strategy equilibrium platform. Suppose that $\overline{w} < w_e < w_e^*$. Corresponding $c_1(w_e) > c^*$ and $c_2(w_e)$ are determined. From Lemma 2.5, we know that $c_1(w_e)$ decreases by increasing w_e until $c_1(w_e) = c^*$. The increase of w_e also leads to the increase of $c_2(w_e)$ for any $\overline{w} < w_e$. Whether the incumbent choose $w_i^{c_1}$ or $w_i^{c_2}$, the incumbent has always incentive to increase w_e when $\overline{w} < w_e < w_e^*$.

Suppose now that $w_e^* < w_e$. Corresponding $c_1(w_e)$ is determined. From Lemma 2.4 and the location of the median $m \leq c_R^*$, we know that the incumbent's choice is $w_i^{c_1}$. The incumbent then wants to decrease w_e to give higher expected utility to the voters of c_1 until $c_1(w_e) = c^*$. Therefore, the entrant always has an incentive to deviate unless the incumbent's platform w_i is w_i^*. □

Proposition 2.4

Proof. We assume that the incumbent is the favored candidate. The following proof can be easily applied when the entrant is favored.

(Sequential move game. The incumbent moves first) If $m < c_L^*$, then the incumbent choose w_i which makes $c_L(w_i) = \frac{c_1(w_i)+c_2(w_i)}{2} = m$ $(c_1 < c_2)$. Let us call this w_i as w_i^1 and this c_L as c_L^m. It is obvious that $c_1(w_i^1) < w_i^1 < c_2(w_i^2)$. Because $\overline{U}_e(c_1) < \overline{U}_e(c_2)$, $v(w_i^1, c_1) < v(w_i^1, c_2)$ for any given $w_i < h_e$. This inequality implies that w_i^1 is on the right of c_L^m, or $c_L^m < w_i^1$ and, therefore, the on the right of m, too. The equilibrium choice w_i^1 is farther from the historical platform h_i than w_i^m, which is the incumbent's most appealing platform for the median voters, or $w_i^m < w_i^1$. On the other hand, the unfavored entrant's choice, assuming that he chooses to remain on the right side of the incumbent, is $w_e^{c_2}$ which is the most appealing platform for voters of type c_2 and obviously on the right of w_e^m, or $w_e^m < w_e^{c_2}$

If $c_L^* \leq m < c^*$, w_i^* is the incumbent's equilibrium platform, which is always on the right of w_i^m, or $w_i^m < w_i^*$. The unfavored entrant's choice w_e^* is obviously on the right of w_e^m, or $w_e^m < w_e^*$.

(Sequential move game. The incumbent moves second) The entrant chooses w_e^* in his first move and the incumbent's best response is w_i^* as long as $m \leq c_R^*$. The unfavored entrant's choice is always w_e^* as a first mover.

(Simultaneous move game in a tight election) The incumbent always chooses w_i^* and the entrant always chooses w_e^*. □

2.6.2 Lemmas

Lemma 2.1

Proof. Without loss of generality, the incumbent's case when $h_i < \theta$ is shown. For any θ' on the right of θ, or $\theta < \theta'$, the first order condition implies that the following inequalities $0 > \frac{\partial s_i(w_i^{\theta}, h_i)}{\partial w_i} > \frac{\partial s_i(w_i^{\theta'}, h_i)}{\partial w_i}$ and $0 < \frac{\partial v(w_i, \theta)(w_i^{\theta}, \theta)}{\partial w_i} < \frac{\partial v(w_i, \theta)(w_i^{\theta'}, \theta')}{\partial w_i}$ must be satisfied. In other words, the distance between w_i^{θ} and θ must be smaller than that between $w_i^{\theta'}$ and θ'.

Suppose $U_i(\theta', w_i^{\theta'}) \geq U_i(\theta, w_i^{\theta})$. Then, $v(w_i^{\theta'}, \theta') - v(w_j^{\theta}, \theta) \geq s_i(w_i^{\theta}; h_i) - s_i(w_i^{\theta'}; h_i)$ must be satisfied, which implies the distance between w_i^{θ} and θ must be greater than that between $w_i^{\theta'}$ and θ', which contradicts the condition to satisfy the first order condition. □

Lemma 2.2

Proof. For any θ, $\frac{\partial v(w_i, \theta)}{\partial \theta} < \frac{\partial v_e}{\partial \theta}$ because $w_i < w_e$, $\frac{\partial s_i}{\partial \theta} < 0$ and $\frac{\partial s_e}{\partial \theta} > 0$. Therefore, $\frac{\partial B}{\partial \theta} = \frac{\partial v(w_i, \theta)}{\partial \theta} - \frac{\partial v_e}{\partial \theta} + \frac{\partial s_i}{\partial \theta} - \frac{\partial s_e}{\partial \theta} < 0$ for any θ. In other words, $B(\theta; w_i, w_e)$ is strictly decreasing with θ. By A.3 ii) and iii), there exist one and only one $\theta = \overline{c}$ such that $B(\overline{c}; w_i, w_e) = 0$. The same idea can be applied to the case of $w_i > w_e$. □

Lemma 2.3

Proof. Assume that α $(0 < \alpha < 1)$ is the incumbent's share when $w_i = w_e = \overline{w}$. At $\overline{w}$, the incumbent maximizes the expected utility of voter of type θ_{w_i}, whose expected utility is maximal from the incumbent's platform w_i and the entrant does the utility of type θ_{w_e}, which is defined in the similar way, where $\theta_{w_e} < \overline{w} < \theta_{w_i}$.

From a marginal shift from $\overline{w}$, the incumbent can earn his share $F(\theta_{w_i})$, and, therefore, $\alpha \geq F(\theta_{w_i})$ for the incumbent to stay at $\overline{w}$. By the same logic, the entrant can earn his share $1 - F(\theta_{w_e})$ from a marginal shift from $\overline{w}$ and, therefore, $1 - \alpha \geq 1 - F(\theta_{w_e})$ for the entrant to stay at $\overline{w}$. These two conditions, however, contract each other for any value of α between zero and one. We cannot have any tie-breaking rule that prevents both candidates from unilateral deviation. $\square$

Lemma 2.4

Proof. We know that when $\overline{w} < w_e$, there exist θ which satisfies $U_e(\theta, w_e) > \overline{U}_i(\theta)$. The first derivatives are given as $\frac{\partial U_e}{\partial \theta} = \frac{\partial v(w_e,\theta)}{\partial \theta}$ and $\frac{\partial \overline{U}_i(\theta)}{\partial \theta} = \frac{\partial v(w_i,\theta)(w_i^\theta,\theta)}{\partial \theta} + \frac{\partial v(w_i,\theta)}{\partial w_i^\theta}\frac{dw_i^\theta}{d\theta} + \frac{\partial s_i}{\partial w_i^\theta}\frac{dw_i^\theta}{d\theta} = \frac{\partial v(w_i,\theta)(w_i^\theta,\theta)}{\partial \theta}$ because $\frac{\partial v(w_i,\theta)}{\partial w_i^\theta} + \frac{\partial s_i}{\partial w_i^\theta} = 0$ by the first order condition. Initially, at $\theta = w_e$, $\frac{\partial v(w_e,\theta)}{\partial \theta} = 0$ is obviously greater than $\frac{\partial v(w_i,\theta)(w_i^\theta,\theta)}{\partial \theta} < 0$, but $\frac{\partial v(w_e,\theta)}{\partial \theta}$ decreases faster than $\frac{\partial v(w_i,\theta)(w_i^\theta,\theta)}{\partial \theta}$ because $\frac{dw_i^\theta}{d\theta} > 0$. Therefore, there must be two platforms $\theta = c_1$ and c_2 which satisfy $U_e(\theta, w_e) = \overline{U}_i(\theta)$ when $\overline{w} < w_e$.

Next, we want to show that the threshold platform $\overline{c}(w_i, w_e)$ satisfies the first order condition with w_i and that the incumbent's optimal platform satisfies $\frac{dU_i(\overline{c},w_i)}{dw_i} = 0$. At any threshold platform $\overline{c}(w_i, w_e)$, $v(w_i, \overline{c}(w_i, w_e)) + s_i(w_i) = v\left(w_e, \overline{c}(w_i, w_e)\right) + s_e(w_e)$ is always satisfied. Therefore, the equation $\frac{\partial v(w_i,\overline{c}(w_i,w_e))}{\partial w_i} + \frac{\partial v(w_i,\overline{c}(w_i,w_e))}{\partial \overline{c}}\frac{\partial \overline{c}}{\partial w_i} + \frac{ds_i}{dw_i} =$

$\frac{\partial v(w_e,\overline{c}(w_i,w_e))}{\partial \overline{c}}\frac{\partial \overline{c}}{\partial w_i}$ must also be satisfied. $\frac{\partial \overline{c}}{\partial w_i}$ is represented as

$$\frac{\partial \overline{c}}{\partial w_i} = \frac{\frac{\partial v(w_i,\overline{c}(w_i,w_e))}{\partial w_i} + \frac{ds_i}{dw_i}}{\frac{\partial v(w_e,\overline{c}(w_i,w_e))}{\partial \overline{c}} - \frac{\partial v(w_i,\overline{c}(w_i,w_e))}{\partial \overline{c}}} = \frac{\frac{\partial U_i}{\partial w_i}}{\frac{\partial v(w_e,\overline{c}(w_i,w_e))}{\partial \overline{c}} - \frac{\partial v(w_i,\overline{c}(w_i,w_e))}{\partial \overline{c}}}$$

The denominator is always positive for a given $\overline{w} < w_e$. Therefore, from the property of U_i, we can conclude that $\overline{c}(w_i, w_e)$ is single-peaked with respect to w_i.

Suppose now that $\frac{d\overline{c}}{dw_i} = 0$ but $\frac{dU_i(\overline{c},w_i)}{dw_i} \neq 0$. If $\frac{dU_i(\overline{c},w_i)}{dw_i} > 0$, then, for $\frac{d\overline{c}}{dw_i} = 0$ to be satisfied, $\frac{dU_e(\overline{c},w_e)}{dw_i} > 0$ should be satisfied at the same time, which is impossible because $\frac{d\overline{c}}{dw_i} = 0$ should be satisfied at optimum, and $\frac{dw_e}{dw_i} = 0$. By the same logic, if $\frac{dU_i(\overline{c},w_i)}{dw_i} < 0$, then, for $\frac{d\overline{c}}{dw_i} = 0$ to be satisfied, $\frac{dU_e(\overline{c},w_e)}{dw_i} < 0$ at the same time, which is impossible for the same reason.

Finally, in a threshold platform, the expected utilities are the same, $U_i(\overline{c}, w_i) = U_e(\overline{c}, w_e)$, and $U_i(\overline{c}, w_i)$ is the maximum expected utility for threshold voters. In other words, the incumbent's platform choice leads to $U_i(\overline{c}, w_i) = \overline{U}_i(\overline{c}) = U_e(\overline{c}, w_e)$ Therefore, the threshold platform is either c_1 or c_2. □

Lemma 2.5

Proof. First, we want to recall that w_e^* is the platform where the incumbent maximizes his expected utility for voters at the border platform c^*. As long as $m < c_R$, $w_i^{c_1}$ is the best response for the incumbent.

If $\overline{w} < w_e < w_e^*$, the voters of type θ_{w_e}, whose expected utility is maximal from the entrant's platform w_e, is $\theta_{w_e} < c^* < c_1(w_e)$ because $U_e(\theta_{w_e}, w_e) = \overline{U}_e(\theta_{w_e}) < \overline{U}_i(\theta_{w_e})$ when $w_e < w^*$, and $U_e(c_1(w_e), w_e) = \overline{U}_i(c_1(w_e))$ must be satisfied at c_1.

A marginal increases of w_e implies the entrant gives the marginally greater expected utility to voters of type $\theta > \theta_{w_e}$. As we have seen in Lemma 2.4, the incumbent's platform choice $w_i = w_i^{c_1}$ needs to satisfy the condition $U_e(c_1, w_e) = \overline{U}_i(c_1) = U_i(c_1, w_i^{c_1})$. Therefore, c_1 and the corresponding incumbent's platform choice $w_i^{c_1}$ decreases as w_e increases.

If $w_e^* \leq w_e$, on the other hand, $c_1(w_e) < \theta_{w_e}$ because $U_e(\theta_{w_e}, w_e) = \overline{U}_e(\theta_{w_e}) > \overline{U}_i(\theta_{w_e})$, but $U_e(c_1(w_e), w_e) = \overline{U}_i(c_1(w_e))$ still needs to be satisfied at c_1. A marginal increase of w_e now marginally decreases the expected utility of the voters on the left of c_1. Therefore, c_1 and the corresponding $w_i^{c_1}$ increases as w_e increases. □

Lemma 2.6

Proof. Assume an incumbent is the favored candidate and he has to move first. Then, the incumbent has to choose his platform so that $c_L^* = m$. If an entrant's best response is $w_e^{c_1}$ (or $w_e^{c_2}$), the incumbent's share would increase by switching his platform to the left (or right). This logic can be applied when an entrant is the favored candidate and he has to move first. □

Chapter 3

Getting Advice with Optimal Contract

3.1 Introduction

This essay explores mechanisms through which a principal can best elicit information from multiple experts. In particular, we focus on a contractual situation, implicitly assuming the information the principal needs to gather is at least partially specific to him. Two important issues emerge as the principal makes contracts with multiple experts. Firstly, the principal should consider how to aggregate information from multiple sources. Secondly, she should determine the wage offer to each expert, which may depend on the information quality if available.

Efficient information aggregation crucially depends on whether the principal can screen the precision of information each expert possesses. To see this, consider an environment in which experts have heterogeneous private signal distributions and the signal is independently informative of the true state. The heterogeneity reflects different abilities in processing raw information, analytic technologies, and/or levels of 'animal spirits.'

Provided the principal knows the precision of each expert, she can easily aggregate information from multiple experts by Bayesian updating. Assuming the compensation scheme is designed so that experts are paid off according to the ex-post accuracy, and experts have no payoff other than the wage paid by the principal, each expert must submit the report on the true state at his posterior mean given his available information[1]

Without message pooling, the one-to-one relationship between posterior mean and private signal allows the principal to discern each expert's private information. It follows that the principal's best prediction of the true state is then the weighted average of the prior and all private signals, where the weights are determined by the precision of the signal.

When the precision of each expert's signal is unknown to the principal, however, information aggregation cannot be achieved with the simple compensation scheme described above. A posterior mean is no longer matched to the private signal in a one-to-one relationship, and the principal would not know the weight she should assign to the report of each expert. The compensation must be more sophisticated. It should be designed not only to induce the honest report of the true state but also to elicit the precision of expert's report.

Sorting through compensation helps the principal via another channel. When the information quality of each expert is heterogeneous and the reservation wage

[1]Experts' payoff other than the wage from the principal, including nonpecuniary or implicit compensation, may drive shaded or pooled messages on true state. For example, reputation concern induces experts to shade or pool messages in the model of (Ottaviani and Sørensen, 2006a,b).

depends on the quality, the principal also needs to decide the wage offer and it would be beneficial if the wage offer were contingent on the information quality.

We show, in an environment where the reservation wage is type dependent, that there exist payoff function(s) in which the true type revelation is implemented and the honest report on the true state is induced. In addition, the compensation scheme induces the first-best outcome in the sense that no information rent exceeding the reservation utility is paid in equilibrium. When the reservation utility schedule is convex in type, a simple linear payoff function with respect to the mean squared error of the report on true state achieves the first-best. In the case when the reservation utility schedule is concave, the optimal payoff function is more complicated but keeps the linearity in a certain form of performance measure.

The intuition behind this sorting mechanism is straightforward. In the optimal compensation scheme we propose, the principal asks each expert what his type is. The optimal contract is designed so that the penalty for the incorrect report is increasing in type announced. The less accurate expert then incurs more cost when he pretends to be a more accurate type, barring untruthful type revelation. Moreover, due to the cheap talk feature of the 'production' of advice, there is no intrinsic utility or cost for experts. This implies the virtual surplus is linear in the control variable of the principal, and the principal makes the information rent arbitrarily small up to the reservation utility to achieve sorting.

We then propose a game in which the principal achieves not only the efficient information aggregation but also the optimal employment. In the game, the principal announces the payoff function which depends on the type, the precision of the private signal announced by each expert, the report on true state, and the true state to be revealed ex-post. Experts from population then apply for the job (pre-screening stage.) Among those applicants, the principal decides which experts to hire (employment stage.) Compensation is paid after the true state is revealed.

This chapter is organized as follows: A brief literature survey follows the introduction in the next section. We describe the model and present the optimization problem in section 3.3. In section 3.4, we derive the optimal contract in which honest reporting and truthful type revelation are achieved and the participation constraint is binding. In section 3.5, we propose a game to achieve the optimal employment. Section 3.6 concludes and addresses issues for further research.

3.2 Related Literature

The sorting mechanism in the paper is an application of a screening problem under asymmetric information. For example, Maskin and Riley (1984) address the problem in the context of an optimal quantity discount by a monopolist. The main difference is that in professional advising, the information asymmetry occurs not only in the type of each agent but also in the true state which is realized ex-post. Indeed, the type, or the information quality of each expert is revealed ex-post through the realized true state and the forecast. The principal, therefore, needs to get messages from each agent about her type in addition to the forecast on the true state. In a sense, the model presented here is a hybrid model of screening and moral hazard because the latter message is often sent after the principal's employment decision is made.

Bhattacharya and Pfleiderer (1985) is more directly related to our work in the motivation and the model specification. They examine the compensation problem for risk-averse portfolio managers whose signal and signal distribution are both private information. They also derive the compensation scheme which achieves the first best outcome. It differs from ours in the objective function and the risk attitude. They assume the utility function of both principal and agents exhibits constant absolute risk aversion, which makes sense in the context of the delegation of portfolio management. With risk neutral agents, as in our model, the problem is not well defined since the portfolio choice position would be extreme. In this sense, the first main result of this paper is a risk neutral agent version of section 4 in Bhattacharya and Pfleiderer (1985). The second main theorem is new. While Bhattacharya and Pfleiderer (1985) derives the first best outcome under some regularity conditions on the reservation utility, we show it for quite general case by varying the performance measure.[2]

Crémer and McLean (1985, 1988) study mechanisms in which a principal, or a seller, extracts full surplus in the context of the independent value auction. In their

[2]Osband (1989) also studies the incentive provision problem for forecasters. The precision of each forecaster in his model depends on the effort level, so the focus is on the moral hazard problem, not on the hidden type problem as ours.

model, the valuations of the bidders are correlated and they know this fact. The seller then designs an auction mechanism in which payments depend on the types announced by bidders. Under some regularity conditions, the seller can induce each bidder to announce his type truthfully, which results in the full surplus extraction.

The types of experts in our model are also correlated, but they are *conditionally* independent. The true state itself, which is assumed to be verifiable ex-post, becomes a reference point that each expert's type is measured. Each expert, thus, is induced to announce his type truthfully without guessing other experts' type. This allows us to develop an *independent* compensation scheme that does not depend on the type announcements by other experts. Auctions with state-dependent payments are studied in Hansen (1985), but it deals with very special cases.

Recent literature on professional advisors is based on the cheap talk game model first introduced by Crawford and Sobel (1982). Departing from partisan bias exogenously given to the payoff functions, Scharfstein and Stein (1990) explores how reputation concerns affect the pattern of messages in equilibrium. They show the reputation concern drives experts to herd in a binary model. The model is generalized in Ottaviani and Sørensen (2006a,b).

The main difference between this paper and the previous literature on professional advising is twofold. First, we give the principal an active role in determining the compensation scheme. Secondly, our focus is on efficiency in information aggregation and employment, not on the strategic bias. To do so, we assume the principal has no private information, and experts are not concerned the reputation effect of the current report.

In our model, the information asymmetry is two dimensional: the signal and its distribution. Except for a few papers, most existing papers on professional advising assume experts share the common private signal distribution, and asymmetric information lies only in realized value of their private signal. Avery and Chevalier (1999), Levy (2004), and Ottaviani and Sørensen (2006a) consider heterogeneous private signal distribution but usually the uncertainty is assumed to be symmetric across the players in the model. Trueman (1994) and section 6 in Ottaviani and Sørensen (2006a) model asymmetric information on signal distribution. The information structure in this paper is mostly similar to Ottaviani and Sørensen (2006a). Battaglini (2002) explores a cheap talk game with multi dimensional uncertainty and multiple referrals, but his results are mainly derived from the orthogonality between uncertain variables, which is different from our setting.

3.3 Model

An uninformed principal tries to make his best prediction of the true state, for example the profitability of a project. To get better information, the principal wishes to hire privately informed agents, who are called 'experts' hereafter. Experts are heterogeneous in the precision of their private signal, which is labeled their *type*. The principal designs a game as follows.

The true state x is drawn from a normal distribution with mean μ_x and variance $1/\tau_x$, which is common knowledge. We assume the true state is verifiable and thus contractible. While the principal has no private information[3], expert $i \in I$ receives a conditionally independent private signal s_i, where I is the set of experts who are employed.[4] The distribution of s_i is assumed to follow a normal distribution with mean x and precision τ_i, or

$$s_i = x + \epsilon_i, \quad \epsilon_i \sim N\left(0, 1/\tau_i\right).$$

The precision, or type, of each expert τ_i is drawn from the population with distribution function F on the support of $[\underline{\tau}, \overline{\tau}] \subset \mathbb{R}^+$. We assume F is continuously differentiable so that the probability continuous density function f exists. The principal cannot discern the type of each agent, but each expert knows his own type. In the pre-screening stage, each expert is requested to submit a message on his own type t_i. Once hired, he has to submit a *report* on the true state, denoted by $r_i \in \mathbb{R}$ for expert i.[5]

The principal's objective is to maximize *revenue* less payoffs to employed experts. The revenue function R depends on the principal's prediction on the true state, denoted by $\widehat{x}$, and the true state. We assume the revenue is decreasing in the ex-

[3]The assumption of a fully uninformed principal, in addition to that of the payoff being conditioned on the true state, precludes the 'yesman effect' in Prendergast (1993).

[4]We fix the employment set of experts in this chapter, as though the employment decision is made before the contract and the information aggregation. However, the order may be reversed in order for the contract to be used as a pre-screening device. The whole recruiting, contracting, and information aggregation process is discussed in the later section.

[5]We follow the convention that each expert reports his best prediction, not directly revealing his private signal. However, reporting the prediction is equivalent to reporting the signal in equilibrium provided there is no message pooling, which is the case of this paper.

post error, $| \widehat{x} - x |$. For example, the revenue can be the negative mean-squared error where $R(x, \widehat{x}) \equiv -a(\widehat{x} - x)^2$ for a constant $a > 0$. In this case, the revenue function is a decreasing function of the mean squared error.

The only cost for the principal is the wage she pays to the experts, where the payoff function is denoted by $C(r_i, t_i, x)$. Note that the payoff does not depend on other experts' messages. In other words, we restrict the compensation to be independent, which implies that the principal cannot use group incentives to implement the information revelation and the performance must be evaluated through the absolute performance basis[6].

Experts are assumed to be risk neutral utility maximizers with the identical vNM utility function $u(c) = c$. We assume that the only benefit from information provision is the payoff from the principal. Each expert has a reservation utility which is type-dependent. Type-dependent reservation utility function $\underline{u}(\tau)$ is assumed to be strictly increasing and continuously differentiable. Increasing reservation utility is realistic when the private information is not fully relation-specific. The expert might use the private information outside of the principal-agent relationship to derive some personal benefits from it.[7]

[6]Relative performance evaluation has been an important issue in contract theory with multi-agent models. We exclude such evaluation on report for simplicity and tractability of the payoff function, since we are focusing on the screening procedure. Extant literature in contract theory find the merits of relative performance evaluation in that it reduces risk-sharing cost as to the common noise. See Holmstrom (1982). Another branch of literature regarding relative evaluation studies rank order compensation or tournament. While it has been shown that tournament scheme can provide approximately the same incentive for agents as the standard contractual form(See, for example, Green and Stokey (1983)), it is less susceptible to extreme output volatilities. Both benefits mentioned above are not relevant to the current model. Ottaviani and Sørensen (2006b) consider forecasting contest, an extreme case of compensation scheme based on relative performance evaluation, in the context of reputational cheap talk game, but their information structure is different from ours.

[7]See Jullien (2000) for examples of type-dependent reservation utility and the general solution in the context of

The Principal's action is denoted by $(\widehat{x}, C)$. The optimization problem is formally described as follows.

$$\max_{\widehat{x},C} E_x \left[R(\widehat{x}, x) - \sum_{i \in I} C(r_i(s_i), \tau_i, x) \right]$$

subject to the expert's problem

$$(r_i(s_i), \tau_i) \in \arg\max_{r,t} E_x \left[C(r, t, x) \mid s_i, \tau_i \right]$$

subject to the participation constraint

$$E_x \left[C(r_i(s_i), \tau_i, x) \mid s_i, \tau_i \right] \geq \underline{u}(\tau_i).$$

where $r_i(s_i)$ is agent i's true posterior mean after observing s_i.

In the next section, we begin our analysis with the pre-screening stage.

3.4 Compensation Scheme for Sorting

In this section, we aim at finding a compensation scheme which achieves the first-best. We ask whether there exists a payoff function $C(r_i, t_i, x)$ which induces the expert to report his posterior mean, implements him to message his own type, and further the expected payoff is just his reservation utility[8]. Formally, we want to find C satisfying

$$\left(\frac{s_i \tau_i + \mu_x \tau_x}{\tau_i + \tau_x}, \tau_i \right) \in \arg\max_{t,r} E_x \left[C(r, t, x) \mid s_i, \tau_i \right]$$

screening problem. Bhattacharya and Pfleiderer (1985) also assume the type dependent reservation utility.

[8]We assume the massage related to the true state is the posterior mean of each expert. However, it is equivalent to assume that the private signal itself is reported.

and

$$E_x\left[C\left(\frac{s_i\tau_i+\mu_x\tau_x}{\tau_i+\tau_x},\tau_i,x\right)\mid s_i,\tau_i\right]=\underline{u}(\tau_i).$$

Note that once the sorting and truthful reporting are implemented, the principal's optimal action is straightforward. Provided the revenue $R(\widehat{x},x)$ is decreasing in error $\mid \widehat{x}-x \mid$, for a fixed I, the best prediction on the true state, $\widehat{x}^*$, is the posterior mean given reports from $|I|$ experts, where $|A|$ is the number of elements in a set A. Formally, we have

$$\widehat{x}^*=\frac{\sum_{i\in I}r_i(\tau_i+\tau_x)-(|I|-1)\mu_x\tau_x}{\sum_{i\in I}\tau_i+\tau_x}.$$

For example, if the revenue is negative mean squared error, the resulting expected net gain is

$$E\left[R(\widehat{x}^*,x)\right]-\sum_{i\in I}\underline{u}(\tau_i)=-\frac{a}{\tau_x+\sum_{j\in J}\tau_j}-\sum_{j\in J}\underline{u}(\tau_j).$$

Our strategy to show existence is as follows. We first restrict to a subclass of payoff functions. We then solve the standard screening problem within the class and check whether the participation constraint is binding for all types of expert.

Proposition 1 is our first main result. It states that if the reservation utility function is non-convex, the first-best outcome is achieved through a payoff function which is linear in the mean squared error of the report. With the linearity restriction, the expert with precision τ_i should solve

$$(r_i,t_i)\in\arg\max_{t,r}E_x\left[-\alpha(t)(r-x)^2+\beta(t)\mid s_i,\tau_i\right] \tag{3.1}$$

We want to find $\alpha(t)$ and $\beta(t)$ such that the solution to (3.1) satisfies the conditions for both honest reporting and truthful type revealing, as well as the participation constraint. We must therefore have the following conditions:

- Incentive Compatibility for Honest Reporting (ICR)

$$\alpha(t) \geq 0. \tag{3.2}$$

- Incentive Compatibility for Truthful Type Revelation (ICT)

  Given (3.2), the expert with precision τ_i will solve the following problem:

$$\tau_i \in \arg\max_{t_i} -\frac{\alpha(t_i)}{\tau_x + \tau_i} + \beta(t_i). \tag{3.3}$$

- Participation Constraint (PC)

  Once (3.2) and (3.3) are satisfied, the participation constraint for type τ_i becomes

$$\max_{r_i, t_i} E_x\left[-\alpha(t_i)(r_i - x)^2 + \beta(t_i)\right] = -\frac{\alpha(\tau_i)}{\tau_x + \tau_i} + \beta(\tau_i) \geq \underline{u}(\tau_i). \tag{3.4}$$

Proposition 3.1. *Suppose the reservation utility is convex on the support of τ. Then, the first-best is strictly implemented through the payoff function within the class of linear functions in mean squared error. That is, it is the strict best response for each expert to message his own type and submit his posterior mean, and the payoff is only his reservation utility if the payoff function is designed to be*

$$C(r, t, x) = -\alpha(t)(r - x)^2 + \beta(t)$$

where

$$\alpha(t) = (\tau_x + t)^2 \underline{u}'(t)$$

and

$$\beta(t) = (\tau_x + t)\underline{u}'(t) + \underline{u}(t).$$

Proof. Let α and β be $\mathcal{C}^2$ function on $\mathbb{R}_{++}$.[9] Let $\alpha(t) \geq 0$ to satisfy (ICR). Define $\mathbf{C}(\tau, t)$ be the expected payoff when type τ expert announces that his type is t and he reports posterior mean honestly. Given (ICR), we have

$$\begin{aligned} \mathbf{C}(\tau, t) &= E_{x,s}\left[C(\frac{\tau s + \tau_x \mu_x}{\tau + \tau_x}, t, x) \mid \tau, s\right] \\ &= -\alpha(t) E_{x,s}\left(\frac{\tau s + \tau_x \mu_x}{\tau + \tau_x} - x\right)^2 + \beta(t) \\ &= \frac{-\alpha(t)}{\tau + \tau_x} + \beta(t) \end{aligned}$$

The first order condition for (ICT) is then

$$\forall \tau \in [\underline{\tau}, \overline{\tau}], \quad -\frac{\alpha'(\tau)}{\tau_x + \tau} + \beta'(\tau) = 0. \tag{3.5}$$

To see the second order condition given (IRC) and the first order condition of (ICT), consider the following formula.

$$\frac{\partial \mathbf{C}}{\partial t}(\tau, t) = -\alpha'(t) \cdot \frac{t - \tau}{(\tau_x + \tau)(\tau_x + t)}, \tag{3.6}$$

[9] It is required that $\alpha'(t)$ and $\beta'(t)$ are right continuous at 0.

which implies that $t = \tau$ is the global maximizer of $\mathfrak{C}(\tau, t)$ for all τ if and only if $\alpha(t)$ is nondecreasing. We will temporarily ignore (3.6) to solve for the optimal contract with (3.5), and then check whether the contract satisfies (3.6).

Let $c(\tau)$ be the utility of expert type τ at the optimum, so that $c(\tau) = \mathbf{C}(\tau, \tau) = -\alpha(\tau)/(\tau_x + \tau) + \beta(\tau)$. Note that from envelope theorem,

$$c'(\tau) = \frac{\alpha(\tau)}{(\tau_x + \tau)^2} \geq 0 \tag{3.7}$$

which implies that the experts with higher precision are paid more.

From (3.5), we have

$$\beta'(\tau) = \frac{\alpha'(\tau)}{\tau_x + \tau}$$

$$\Rightarrow$$

$$\beta(\tau) = \beta(\underline{\tau}) + \int_{\underline{\tau}}^{\tau} \frac{\alpha'(s)}{\tau_x + s} ds$$

$$= \beta(\underline{\tau}) - \frac{\alpha(\underline{\tau})}{\tau_x + \underline{\tau}} + \frac{\alpha(\tau)}{\tau_x + \tau} + \int_{\underline{\tau}}^{\tau} \frac{\alpha(s)}{(\tau_x + s)^2} ds \tag{3.8}$$

and

$$c(\tau) = -\frac{\alpha(\tau)}{\tau_x + \tau} + \beta(\tau) = \beta(\underline{\tau}) - \frac{\alpha(\underline{\tau})}{\tau_x + \underline{\tau}} + \int_{\underline{\tau}}^{\tau} \frac{\alpha(s)}{(\tau_x + s)^2} ds.$$

Given that the expert reports his posterior mean, the principal's problem is

$$\min_{\alpha(\tau)} E_\tau [c(\tau)] = \int_{\underline{\tau}}^{\overline{\tau}} \left(\beta(\underline{\tau}) - \frac{\alpha(\underline{\tau})}{\tau_x + \underline{\tau}} + \int_{\underline{\tau}}^{\tau} \frac{\alpha(s)}{(\tau_x + s)^2} ds \right) f(\tau) d\tau$$

$$= \int_{\underline{\tau}}^{\overline{\tau}} \left(\beta(\underline{\tau}) - \frac{\alpha(\underline{\tau})}{\tau_x + \underline{\tau}} + \frac{\alpha(\tau)}{(\tau_x + \tau)^2} \frac{1 - F(\tau)}{f(\tau)} \right) f(\tau) d\tau \tag{3.9}$$

subject to the participation constraint

$$c(\tau) = \beta(\underline{\tau}) - \frac{\alpha(\underline{\tau})}{\tau_x + \underline{\tau}} + \int_{\underline{\tau}}^{\tau} \frac{\alpha(s)}{(\tau_x + s)^2} ds \geq \underline{u}(\tau). \tag{3.10}$$

We can solve this problem through point-wise minimization. Since the formula in the bracket of (3.9), the so called *virtual cost*, is linear in α, the principal can let $\alpha(\tau)$ be the least possible cost $c(\tau)$ for all τ. That is, the participation constraint (3.10) should bind for all τ. Differentiation of the binding participation constraint gives

$$\alpha(t) = (\tau_x + t)^2 \underline{u}'(t)$$

and from (3.8),

$$\beta(t) = \beta(\underline{\tau}) - \frac{\alpha(\underline{\tau})}{\tau_x + \underline{\tau}} + (\tau_x + t)u'(t) + \underline{u}(t) - \underline{u}(\underline{\tau}).$$

Let $\beta(\underline{\tau}) = (\tau_x + \underline{\tau})u'(\underline{\tau}) + u(\underline{\tau})$ for the participation constraint of the lowest type to bind. We now solve for α and β to satisfy the first order condition of (ICT) and binding participation constraint (PC). Finally we need to check the second order condition, which is equivalent to α monotone nondecreasing. Since

$$\alpha'(t) = 2(\tau_x + t)u'(t) + (\tau_x + t)^2 u''(t),$$

the second order condition is satisfied provided u is non-concave. This completes the proof.[10] □

[10]The result still holds in the case of type-independent reservation utility. Suppose w is the constant reservation utility. From (3.7), the participation constraint is binding for the lowest type, i.e., $\beta(\underline{\tau}) - \alpha(\underline{\tau})/(\tau_x + \underline{\tau}) = w$.

Then, the principal should solve

$$\min_{\alpha(\tau)} E\left[c(\tau)\right] = \int_{\underline{\tau}}^{\overline{\tau}} \left(w + \frac{\alpha(s)}{(\tau_x + s)^2} \frac{1 - F(\tau)}{f(\tau)} \right) f(\tau) d\tau$$

Point-wise minimization gives $\alpha(\tau) = 0$ for all τ and consequently $\beta(t) = w$ for all τ. That is, the optimal contract indicates that the principal offers flat wage.

The problem in this case is that the honest reporting and the truthful type revelation are implemented only weakly: experts are indifferent between sending truthful messages and lying. However, the principal can achieve the first best with arbitrary small cost by setting $\alpha(t)$ to be increasing in t very slowly but still keeping $\beta(t) = w$.

The following examples show how the concave reservation utility function obstructs the truthful type revelation.

Example 3.1. Suppose $\tau_x = 1$ and $\underline{u}(\tau) = \tau$ for $\tau \geq 0$. From Proposition 3.1, we have $C(r,t,x) = -(1+t)^2(r-x)^2 + (1+2t)$, and since (ICR) is satisfied, $E_x[C(r,t,x)] = -(1+t)^2/(1+\tau) + (1+t)$. The first order condition gives $t = \tau$ and the second order condition is satisfied.

Suppose now $\tau_x = 1$ and $\underline{u}(\tau) = 1 - 1/(1+\tau)^2$. Note the concavity of the reservation utility function. If we construct the compensation function with α and β in Proposition 1, we have

$$E_x[C(r,t,x)] = -\frac{2}{(1+t)(1+\tau)} + 1 + \frac{1}{(1+t)^2}.$$

The first order condition still gives $t = \tau$, but the second derivative of the expected compensation is

$$\frac{\partial^2}{\partial t^2} E_x[C(r,t,x)] = -\frac{4}{(1+t)^3(1+\tau)} + \frac{6}{(1+t)^4}$$

which is positive at $t = \tau$, violating the second order condition. ■

The result of Proposition 3.1 holds only for non-concave reservation utility functions. When the reservation function is sufficiently concave, the compensation scheme derived from the first order condition becomes convex, barring the expert from revealing his own type to maximize compensation. To achieve the first best outcome with a concave reservation utility function, the principal needs to make the compensation function more concave in equilibrium. This can be done by restricting the expected payoff to be linear in a geometric power of variance. We state this result in Proposition 3.2.

Proposition 3.2. *Suppose $\underline{u}$ is a strictly increasing $\mathcal{C}^2$ function and the support of types is bounded, i.e. $\overline{\tau} < \infty$. Then, there exist $p \in \mathbb{N}$ such that the compensation scheme*

$$C(r, t, x) = -\alpha(t)(r - x)^{2p} + \beta(t)$$

achieves the first best outcome, where

$$\alpha(t) = \frac{2^p(p-1)!}{(2p)!}(\tau_x + t)^{p+1}\underline{u}'(t)$$

and

$$\beta(t) = \frac{(\tau_x + t)}{p}\underline{u}'(t) + \underline{u}(t).$$

Proof. Let α and β be $\mathcal{C}^2$ functions on $\mathbb{R}_{++}$. Let $\alpha(t) \geq 0$ to satisfy (ICR). Then, the expected payoff for type τ is

$$E_x\left[C(r, t, x)\right] = -\alpha(t)\mu_{2p}(\tau) + \beta(t)$$

where $\mu_{2p}(\tau)$ is the $(2p)$'th central moment. Under the Gaussian specification, we have

$$\mu_{2p}(\tau) = E(r - x)^{2p} = \frac{(2p)!}{2^p p!}\left(\frac{1}{\tau_x + \tau}\right)^p$$

The first order condition for (ICT) is

$$\forall \tau \in [\underline{\tau}, \overline{\tau}], \quad -\alpha'(\tau)\mu_{2p}(\tau) + \beta'(\tau) = 0 \tag{3.11}$$

Defining $\mathbf{C}(\tau, t)$ as in proposition 1, we have

$$\frac{\partial \mathbf{C}}{\partial t}(\tau, t) = -\alpha'(t) \cdot \frac{t-\tau}{(\tau_x + t)} \mu_{2p}(\tau),$$

which implies that $t = \tau$ is the global maximizer of $\mathbf{C}(\tau, t)$ if and only if $\alpha(t)$ is nondecreasing.

The expected payoff at the optimum, $c(\tau)$, is now $c(\tau) = \mathbf{C}(\tau, \tau) = -\alpha(\tau)\mu_{2p}(\tau) + \beta(\tau)$. From (3.11), we get

$$\begin{aligned} \beta(\tau) &= \beta(\underline{\tau}) + \int_{\underline{\tau}}^{\tau} \alpha'(s)\mu_{2p}(s)ds \\ &= \beta(\underline{\tau}) - \alpha(\underline{\tau})\mu_{2p}(\underline{\tau}) + \alpha(\tau)\mu_{2p}(\tau) - \int_{\underline{\tau}}^{\tau} \alpha(s)\mu'_{2p}(s)ds \qquad (3.12) \end{aligned}$$

and

$$c(\tau) = -\alpha(\tau)\mu_{2p}(\tau) + \beta(\tau) = \beta(\underline{\tau}) - \alpha(\underline{\tau})\mu_{2p}(\underline{\tau}) - \int_{\underline{\tau}}^{\tau} \alpha(s)\mu'_{2p}(s)ds.$$

The principal's problem is now

$$\min_{\alpha(\tau)} E_\tau\left[c(\tau)\right] = \int_{\underline{\tau}}^{\overline{\tau}} \left(\beta(\underline{\tau}) - \alpha(\underline{\tau})\mu'_{2p}(\underline{\tau}) - \alpha(\tau)\mu'_{2p}(\tau)\frac{1-F(\tau)}{f(\tau)}\right) f(\tau)d\tau. \qquad (3.13)$$

Note that the virtual cost in (3.13) still keeps the linearity in α, which implies the participation constraint should bind for all τ in the optimal contract, i.e.,

$$c(\tau) = \beta(\underline{\tau}) - \alpha(\underline{\tau})\mu_{2p}(\underline{\tau}) - \int_{\underline{\tau}}^{\tau} \alpha(s)\mu'_{2p}(s)ds = \underline{u}(\tau). \qquad (3.14)$$

Differentiating (3.14) with respect to τ, we get

$$\alpha(t) = -\frac{\underline{u}'(t)}{\mu'_{2p}(t)}$$

and from (3.12) and the appropriate boundary condition,

$$\beta(t) = -\frac{\mu_{2p}(t)}{\mu'_{2p}(t)}\underline{u}'(t) + \underline{u}(t).$$

The remaining part is to check the second order condition or monotonicity of α. Since $\mu'_{2p}(t) = -pL(\tau_x + t)^{-p-1} < 0$ and $\mu''_{2p}(t) = -p(-p-1)L(\tau_x + t)^{-p-2}$ where $L = (2p)!/(2^p p!)$, we have

$$\alpha'(t) = -\frac{\underline{u}''(t)\mu'_{2p}(t) - \underline{u}'(t)\mu''_{2p}(t)}{\left(\mu'_{2p}(t)\right)^2} > 0$$

or

$$\frac{\underline{u}''(t)}{\underline{u}'(t)} > \frac{\mu''_{2p}(t)}{\mu'_{2p}(t)} = \frac{-p-1}{\tau_x + t}. \tag{3.15}$$

Since t is defined on a compact set and $\underline{u}'$ and $\underline{u}''$ are continuous, the left side of (3.15) is bounded. Therefore, for large p, the inequality holds for all t in the support of τ. □

The logic of proposition 3.2 is as follows. To satisfy the second order condition, the sorting variable $\alpha(t)$ must be monotone increasing.[11] The $\alpha(t)$ derived from the first order condition is the product of $\underline{u}'(t)$ and a function of the announced posterior precision, which we call here $h(\tau_x+t)$. In the proposition, $h(\tau_x+t) = M(\tau_x+t)^{p+1}$ for a constant M. Though h turns out to be increasing and positive, α is not guaranteed to be monotone increasing for a concave $\underline{u}$. The principal, however, can take arbitrarily large p so that h increases fast enough to cover the effect of decreasing $\underline{u}'(t)$ so that

[11] This is indeed equivalent to the supermodularity of the objective function in (t, τ).

the product is monotone increasing. Indeed, for a given reservation utility, we can find p^* such that any payoff function with $p > p^*$ achieves the first best. We present an example.

Example 3.2. Let $\tau_x = 1$ and $\underline{u}(\tau) = 1 - 1/(1+\tau)^2$, as in the second case in example 1. Then, from proposition 2, the expected payoff given honest reporting on the true state is, in the optimal contract with $p = 3$,

$$E_x\left[C(r,t,x)\right] = -\alpha(t)\mu_{2p}(\tau) + \beta(t) = -\frac{2}{3}\frac{1+t}{(1+\tau)^3} + 1 - \frac{1}{3(1+t)^2}.$$

The first order condition gives

$$\frac{2}{3(1+\tau)^3} - \frac{2}{3(1+t)^2} = 0 \quad \Rightarrow \quad t = \tau$$

and the second order condition is satisfied since

$$\frac{\partial^2}{\partial t^2} E_x\left[C(r,t,x)\right] = -\frac{2}{(1+t)^4} < 0. \quad \blacksquare$$

Figure 1.1 shows how the power of the ex-post error affects the expected payoff. If the compensation is linear in mean squared error ($p = 1$), the truthful report $t = \tau = 1$ does not maximize the expected payoff. When the performance measure is more sensitive to the error, for example with $p = 3$ in this example, the truthful report becomes optimal for the expert.

The results in this section are interesting from two perspectives. First, sorting and honest reporting are implemented through a simple linear-form payoff function.

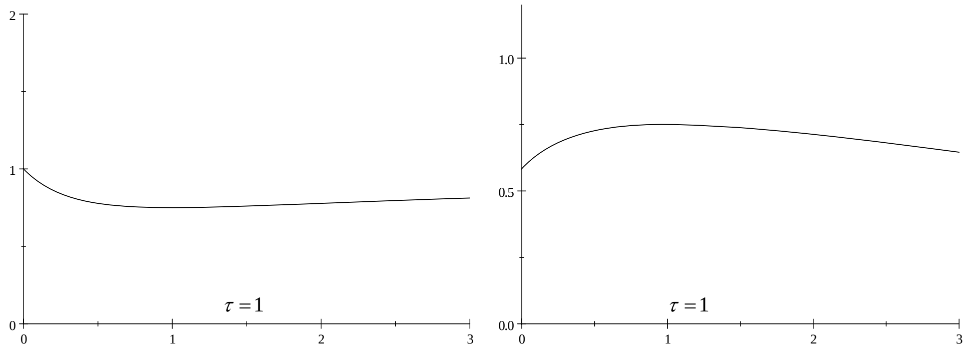

Figure 3.1: Expected Payoffs when $p = 1$ (right) and $p = 3$ (left)

This is because the linear payoff function under honest reporting satisfies the Spence-Mirrlees condition, or $-\alpha(t)/(\tau_x + \tau)$ is supermodular in α and τ. Moreover, it is supermodular in t and τ provided α is increasing in t. This simplifies the problem since the second order condition is equivalent to the monotonicity of α.

Another striking result is that the minimal information rent is paid in equilibrium. This is because, unlike the standard screening problem, the professional advising has a cheap talk feature in the sense that the sorting variable α does not affect the intrinsic cost or utility of the expert. This makes the virtual surplus (or virtual cost) linear in α. Therefore, the principal can fully control the payoff so the participation constraint is binding for all types of experts.

It is worthwhile noting that this mechanism is not a unique. One may design other mechanisms that achieve truthful type revelation and honest reporting. In addition, we should emphasize the compensation of each expert depends only on each expert's own report, not others'. This *independent* compensation scheme, in conjunction with the binding participate constraint, is beneficial to the principal because she can design the employment policy independent of the compensation scheme. We now turn our focus to the employment stage.

3.5 Optimal Employment

In the previous section, we showed that the principal can elicit each expert's precision and induce the honest reporting through a compensation scheme. Those features do not change when the principal wishes to hire more than one expert, since the optimal compensation scheme proposed is independent. Each expert would not care what types of experts he will co-work. Furthermore, since each expert will be paid at his reservation utility level, he would not concern about whether he will be hired or not. This implies that once pre-screening is done before the employment decision is made, the employment policy can be independent of the compensation scheme.

Specifically, consider the following game. At the beginning of the game, the true state is realized, but not revealed to anyone in the game. Then the pre-screening stage begins. The principal announces the compensation scheme, which is designed to screen the type of each applicant. Each expert, drawn from the population, applies for the job positions and send a message t on his own type. In the employment stage, the principal decides which applicants he will hire, based on the information he learns from the pre-screening stage. Once hired, each expert submits his report on the true state. Finally, the true state is revealed and payoffs are made according to the compensation scheme.

The screening through compensation simplifies the optimization problem. After the principal pre-screens experts, she knows the type of each applicant and how much she should pay if she hires some of them. Since experts hired are expected to submit honest reports on the true state, the objective function of the principal becomes a function of the precisions of employed experts less the sum of their reservation utilities. With the mean squared error specification of the revenue, for example, we have the optimization problem as follows:

$$\max_{S \subset I} \left(-\frac{a}{\tau_x + \sum_{j \in S} \tau_j} \right) - \sum_{j \in S} u(\tau_j) \tag{3.16}$$

where I is the set of all applicants. Now, the optimization problem becomes a combinatorial optimization, or a discrete portfolio problem, which is covered in Chapter 4 of this dissertation.

3.6 Conclusion and Discussion

This paper considers a principal who wishes to get advice from one or more experts. To aggregate information from possibly multiple sources and pay the least amount to each type of expert, the principal needs to design a mechanism which induces truthful type revelation and honest reporting of the true state. Under a Gaussian specification, it is shown that there exists a payoff function which achieves this first-best outcome. In the optimal contract proposed, the penalty for an incorrect report is increasing in type (precision) revealed by experts, preventing less precise experts from hiding behind more precise experts.

We derive the optimal compensation scheme in the class of linear functions in a specified performance measure. We show that if the reservation utility is convex, the first-best outcome is achieved with a payoff function linear in mean squared error. In the case with concave reservation utility, the performance measure should be more sensitive to the ex-post error, but still we can design the payoff function which is linear in the power of mean squared error.

In the paper, we assume each expert's gain from providing information depends only on the *current period* compensation paid by the principal, and his ability is elicited through it. However, from a dynamic perspective, the ability or precision of each agent may be evaluated by two parties: the decision maker (the principal) and the outside evaluator (the market). In this case, the gain from information provisions would come, at least partly, from future payoffs which depend on the reputation built today. Recent empirical studies show that career concerns matter in expert advising. Ottaviani and Sørensen (2006a,b) explore the theoretical approaches on this topic.

However, Ottaviani and Sørensen (2006a) assume the compensation is solely determined by reputation. In this sense, the approach of the paper is in the opposite direction to ours. A complete theory would consider the compensation determined by both factors: future payoff from reputation and current payoff from compensation. As two parties are involved in evaluation, there would be a conflict of interests between the decision maker and the evaluator. Since the report tends to be shaded or biased in the presence of reputation concerns, the principal's objective is to reduce such effect, without paying too much. We leave these topics for future research.

Chapter 4

Optimal Employment of Multiple Experts

4.1 Introduction

This chapter considers an employer's employment portfolio choice in which only employees' abilities, not efforts, affect the outcome, and employees are heterogeneous in their ability and reservation wage. This environment usually appears when a principal hires multiple experts for a single project.

Levitt (1995) shows that hiring multiple agents can increase the payoff when an employer cares only about the best outcome among agents. In other words, multiple samples from a distribution are better than a single sample. The school application problem studied by Chade and Smith (2006) falls into this category. As in their work, this chapter focuses on the optimal portfolio choice problem ignoring information asymmetry issues.

When an employer hires multiple experts, while an expert with higher ability may contribute more than those with lower ability, the level of an individual's marginal contribution often depends on the other experts' contributions. For example, a financial forecaster's forecasting may be more valuable when there is no other forecasting value available. Two pharmaceutical scientists with the same capability are more likely to come out with a successful medicine when they work together for the same research project than when one scientist works alone, but the probability of success may not be doubled up. In other words, the output produced by multiple experts can be less than the sum of their individual outputs; therefore, production is submodular.

An optimal employment portfolio choice is especially a big problem, even without information asymmetry, with a submodular production function and an arbitrary reservation wage schedule.[1] Further, it is almost impossible to characterize the property of the optimal portfolio in general. Indeed, it is widely known that the maximization of a general submodular set function with arbitrary price, or reservation utility, schedule is computationally intractable.

A naïve approach to this problem is to find a local maximum, or an agent with the highest marginal contribution, at each step and repeat it until there is no agent who makes a positive marginal contribution.[2] This kind of myopic decision approach, however, does not always lead to a global optimum. More often than not, the optimal employment portfolio does not necessarily include the most capable agent, and an agent may be omitted while others with lower and higher ability are hired.

Under some conditions, this myopic approach can still lead to an optimum.

[1]This kind of problem is known to be NP-hard: there is no algorithm for it to be solved in polynomial time (Cormen et al., 2001). It is possible, however, to find a maximum for a general *super*modular function in polynomial time, see Lovász (1983).

[2]This approach is also known as the greedy algorithm; see Cormen et al. (2001).

Murota and Shioura (2003) show that if the *single improvement* property (Gul and Stacchetti, 1999) is satisfied for an objective function, the following properties must hold: (i) a greedy algorithm leads to the global optimum, and (ii) local and global optima coincide. By definition, the single improvement property is satisfied if, for any sub-optimal employment portfolio, adding, removing or changing only one element leads to a higher payoff. A unit demand function, for example, satisfies the single improvement property.[3] Gul and Stacchetti (1999) show that for monotone objective functions, the single improvement property is equivalent to the gross substitutes condition of Kelso and Crawford (1982).

Unfortunately, the single improvement property is often too strong to be satisfied in many portfolio choice problems. Kelso and Crawford (1982) show that an objective function is submodular if it is non-decreasing and satisfies the gross substitutes property, but the opposite is not necessarily true. Many portfolio choice problems, including simultaneous search with a downward recursive payoff function (Chade and Smith, 2006) and hiring multiple experts (Chapter 3) do not satisfy the single improvement property even though the objective function is non-decreasing and submodular. In such problems, the naïve approach often fails to achieve the optimum.

In this essay, we focus on the employment portfolio decision when the single improvement property of the production function is violated, but a myopic employment

[3] Gul and Stacchetti (1999) give several classes of functions which satisfy the single improvement property. In addition, they give two operations that allow us to drive a new function satisfying the single improvement property from other functions satisfying it.

approach is still optimal. Assuming the agents' abilities and reservation wages are known to the employer, we show that when (i) an agent's reservation wage is a fraction of his production as an individual, and (ii) the production function is increasing, submodular and exhibits increasing differences in marginal production, the optimal employment portfolio simply includes all experts with higher capability than a cut-off level and excludes everyone else. We refer to this property as monotone employment. In this case, the employment decision is simply to sequentially choose the expert who offers the largest marginal improvement, until the marginal contribution turns negative.

This essay can be viewed both as a case of submodular function optimization when the single improvement condition is violated, as in Chade and Smith (2006), and as a justification for a monotone employment policy, which is often taken for granted.The rest of this chapter is organized as follows: We describe the model in the next section. In Section 4.3, we present the specifications under which the monotone employment result becomes optimal. Section 4.4 presents an example for a decision-maker who wishes to gather information from multiple professional forecasters. Section 4.5 concludes and relates the result with the literature on combinatorial optimization.

4.2 Model

There is a monopolistic employer with many potential employees, or applicants, each of whom is heterogeneous in ability. We define $S = \{\tau_1, \tau_2, \cdots, \tau_N\}$ as the set of all potential employees and $\tau_i \in S$ $(i = 1, \dots N)$ as the ability of each potential employee to the job position. S is a discrete, finite subset of $\mathbb{R}_+^N$. An employee with the ability τ has his (reservation) wage $w(\tau)$. The elements in S be ordered as $\{\tau_{i+1}\} \succcurlyeq \{\tau_i\}$ for $i = 1 \dots N-1$.

The production function, denoted by $f(\cdot)$, is a set function from 2^S to $\mathbb{R}_+$. The employees, once hired, jointly produce outputs. The production depends solely on the employees' ability. Without any employment, the employer cannot produce any output, or $f(\emptyset) = 0$. For any $A \in 2^S$, The principal's payoff can be represented as $f(A) - \sum_{a \in A} w(a)$. Her objective is, therefore, maximizing this value. For convenience, we sometimes use $w(A)$ to represent $\sum_{a \in A} w(a)$, and $f(A, \tau)$ to represent $f(A \cup \{\tau\})$.

We define the binary relation $\succcurlyeq$ as follows. For any two subsets A_i and A_j, $A_j \succcurlyeq A_i$ if and only if $f(A_j) \geq f(A_i)$. Because the value of f is a non-negative real number, every subset of S is completely ordered with respect to the binary relationship $\succcurlyeq$.

We assume the following regularity conditions for the set production function $f(X)$ $(X \subseteq S)$.

Assumption 4.1. *The set production function $f(\cdot)$ satisfies*

(i) Strictly monotonicity with the set inclusion ordering relation: If $A_1 \subseteq A_2$, then $f(A_1) \leq f(A_2)$, and if $A_1 \subset A_2$, then $f(A_1) < f(A_2)$.

(ii) Order preservation in marginal production: For any $A \subset S$ and τ_i, $\tau_j \notin A$,

$$f(\tau_i) \leq f(\tau_j) \Leftrightarrow f(A, \tau_i) \leq f(A, \tau_j).$$

(iii) Decreasing marginal production: If $f(\tau_i) \leq f(\tau_j)$, then for any $A \subseteq S \backslash \{\tau_i, \tau_j\}$,

$$f(A, \tau_j) - f(A, \tau_i) \leq f(\tau_j) - f(\tau_i).$$

A set function $f(\cdot)$ is *submodular* (Topkis, 1998) if for two sets A and B, $f(A \cup B) + f(A \cap B) \leq f(A) + f(B)$ and strictly submodular if $f(A \cup B) + f(A \cap B) < f(A) + f(B)$ when $A \neq B$.

Lemma 4.1. *$f(X)$ is submodular in X on S.*

Proof. Item (iii) in Assumption 4.1 implies that the production function has decreasing marginal returns. Lemma 1 in Gul and Stacchetti (1999) shows that when a set function $f : 2^S \to \mathbb{R}$ is monotone, it is submodular if and only if the set function has decreasing marginal returns. □

Finally, following Gul and Stacchetti (1999), we define that f satisfies the single improvement property if for any wage schedule w and subset $A \subseteq S$ such that $A \notin \arg\max_{J \subseteq S} f(J) - w(J)$, there exists $B \subseteq S$ such that (i) $f(A) - w(A) < f(B) - w(B)$, (ii) $\#(A \backslash B) \leq 1$ and (iii) $\#(B \backslash A) \leq 1$. Note that the single improvement property is a property of f only, not related to the wage schedule w. The single improvement property, which guarantees that a greedy algorithm reaches the global optimum, is too strong to be satisfied in many cases. Our specification on f do not necessarily satisfy the single improvement property, either. We give an example in subsection 4.4.1.

4.3 Optimal Monotone Employment

When the optimal employment portfolio simply includes all experts with higher capability than a cutoff level and excludes everyone else, we refer to it as *optimal monotone employment*. In this chapter, we propose a set of conditions which leads to the optimal monotone employment, through a simple approach called *marginal improvement algorithm* (Chade and Smith, 2006).

When a principal adopts a sequence of myopic decisions, she begins with the null employment set, searches the best expert and adds him to the employment set if he generates a positive profit. Otherwise, the optimal employment set is empty. She then searches for the best one among experts not employed yet given the current employment set. If this expert generates a positive marginal profit, he is added to the employment set. Otherwise, the algorithm stops. This procedure is repeated until the best remaining expert generate a negative marginal profit is negative. Following Chade and Smith (2006), we call this approach the marginal improvement algorithm. It can be classified as a greedy Algorithm (Cormen et al., 2001).[4] Note that the optimal monotone employment is not necessarily achieved even if the marginal improvement algorithm leads to an optimum.

We define the positive single crossing condition of the marginal production by a single expert and cutoff element.

[4]In a marginal improvement algorithm process, once an employment decision about an expert is made, it is not revisited afterward. Because the marginal contribution of an expert depends not just on his own ability but also on the abilities of the remaining employees, this once-and-for-all feature leaves the possibility of failing to reach the optimum in general.

Definition 4.1. (Positive Single Crossing Property) Positive single crossing is satisfied if, for any subset $A \subset S$, there exists at most one element $a^* \in \{S \backslash A\}$ such that for any $b \geq a^*$, the inequality

$$f(A,b) - \left(\sum_{a \in A} w(a) + w(b) \right) \geq 0$$

is satisfied, and for any $b \leq a^*$,

$$f(A,b) - \left(\sum_{a \in A} w(a) + w(b) \right) \leq 0$$

are satisfied. This a^* is called the *cutoff element* for set A.

We want to note that it is not the same as the single-crossing property defined in Milgrom and Shannon (1994). In Milgrom and Shannon's definition, if it is (strictly) preferable to have more of the second component, an element in this case, given a particular level for the first component, an employment subset in this case, then it would still be (strictly) preferable to have the greater second component given a greater level for the first component. Unlike Milgrom and Shannon (1994), our single-crossing property does not require preserving the preference between *any* two elements $a_1, a_2 \in S$ for different employment subsets. For example, in our model, given $a_1 \leq a_2$, $A \preccurlyeq B$, it is possible that $0 < f(A, a_1) - w(A) - w(a_1) < f(A, a_2) - w(A) - w(a_2)$, but $0 > f(B, a_1) - w(B) - w(a_1) > f(B, a_2) - w(B) - w(a_2)$, which is not consistent with Milgrom and Shannon's single-crossing property.

We assume another property for the production function.

Assumption 4.2. *f has increasing differences in marginal production. That is, for any $\tau, \tau', \tau'' \in \mathbb{R}_+$ ($\tau \leq \tau' \leq \tau''$) and $A \in 2^S \backslash \{\emptyset\}$, if $0 > f(A, \tau'') - 2f(A, \tau') + f(A, \tau)$, then, the following inequality*

$$0 > f(A, \tau'') - 2f(A, \tau') + f(A, \tau) > f(\tau'') - 2f(\tau') + f(\tau)$$

is satisfied.

The intuition behind the Assumption 4.2 is that if f were a real-valued function, the sign of the third derivative of f would be positive, or $f''' > 0$. In other words, f would show decreasing concavity. This condition is satisfied for both a decreasing absolute risk-aversion (DARA) and a constant absolute risk-averse (CARA) function.

We also assume that the reservation utility schedule $w(\tau)$ is exogenous to the model as follows.

Assumption 4.3. *$w(\tau_i)$ is proportional to the individual agent's single production, i.e., $w(\tau) = \beta f(\tau)$ where $0 < \beta < 1$ and ($A \subseteq S$).*

The following lemma states that the positive single crossing condition is satisfied if the submodular production function shows a decreasing curvature and the reservation utility is given as a fraction of single production.

Lemma 4.2. *For $f(\cdot)$, if assumption 4.1, 4.2 and 4.3 are satisfied, the positive single crossing property is satisfied.*

Proof. If the statement of the lemma is true, satisfying the inequality $f(A, \tau_{i^*}) - \beta \left(\sum_{a \in A} f(a) + f(\tau_{i^*}) \right) \geq 0$ must always imply $f(A, \tau_i) - \beta \left(\sum_{a \in A} f(a) + f(\tau_i) \right) \geq 0$

for any $i > i^*$ and a subset $A \subseteq S \backslash \{\tau_i, \tau_{i^*}\}$. This condition is satisfied if the following inequality

$$f(A, \tau_i) - f(A, \tau_{i^*}) > \beta f(\tau_i) - \beta f(\tau_{i^*}) \tag{4.1}$$

is satisfied for any $\tau_i > \tau_{i^*}$ and not otherwise. When $0 \geq f(A, \tau_{i+1}) - 2f(A, \tau_i) + f(A, \tau_{i^*})$, then, by Assumption 4.2, the decrease of marginal production from $f(A, \tau_i) - f(A, \tau_{i^*})$ to $f(A, \tau_{i+1}) - f(A, \tau_i)$ is slower than from $f(\tau_i) - f(\tau_{i^*})$ to $f(\tau_{i+1}) - f(\tau_i)$. Thus, the following inequality

$$f(A, \tau_{i+1}) - f(A, \tau_i) + f(A, \tau_i) - f(A, \tau_{i^*}) > \beta(f(\tau_{i+1}) - f(\tau_i) + f(\tau_i) - f(\tau_{i^*}))$$

which can be reduced as

$$f(A, \tau_{i+1}) - f(A, \tau_{i^*}) > \beta(f(\tau_{i+1}) - f(\tau_{i^*})) \tag{4.2}$$

must be satisfied as long as (4.1) is true. When $0 < f(A, \tau_{i+1}) - 2f(A, \tau_i) + f(A, \tau_{i^*})$, then we can find a sequence of numbers $(t_0 = \tau_{i^*}, t_1 = \tau_i, t_2, \ldots, t_{K-1}, t_K = \tau_{i+1})$ where $t_k \leq t_{k+1}$ $(k = 0, \ldots, K-1)$ so that $f(A, \tau_{i+1}) - f(A, \tau_{i^*})$ can be rewritten as

$$f(A, t_K) - f(A, t_{K-1}) + f(A, t_{K-1}) - f(A, t_{K-2}) + \ldots + f(A, t_1) - f(A, t_0)$$

and that $0 \geq f(A, t_{k+2}) - 2f(A, t_{k+1}) - f(A, t_k)$. Again, Assumption 4.2 and equation (4.2) are satisfied.

As long as we can find a $\tau \in S \backslash A$ which makes the lefthand side of (4.1) strictly positive for given β, we have one cutoff element which satisfies (4.1). Otherwise, there is no cutoff element.

Finally, we want to show that $f(A,\tau_i) - \beta\left(\sum_{a\in A} f(a) + f(\tau_i)\right) < 0$ for any $\tau_i < \tau_{i^*}$. Suppose that there exists such $\tau' \notin A$ that satisfies $\tau' \leq \tau_{i^*}$ and $f(A,\tau') - \beta\left(\sum_{a\in A} f(a) + f(\tau')\right) \geq 0$, and that there exists such τ'' that satisfies $\tau' \leq \tau'' \leq \tau_{i^*}$ and $f(A,\tau'') - \beta\left(\sum_{a\in A} f(a) + f(\tau'')\right) < 0$. Then, we have the following inequality $f(A,\tau'') - \beta f(\tau'') < f(A,\tau') - \beta f(\tau')$ which clearly violates (4.2), which must be satisfied because of the existence of τ', and, therefore, the assumption $\tau' \leq \tau''$. This completes the proof. □

It should be noticed that the existence of a cutoff element does not itself permit an immediate identification of the optimal set. Lemma 4.2, however, implies that the marginal contribution, as long as it stays positive, is always proportional to applicants' capability and that monotone employment can lead to the global optimum. To show optimal monotone employment achieved, we need the following lemma which states that the cutoff element is non-decreasing with respect to the set of previous employees.

Lemma 4.3. *Given the properties of the set production function and the reservation utility in Assumption 4.3, for any set $A, B \subseteq S$ such that $B \succcurlyeq A$ with cutoff elements a^* and b^* for A and B, respectively, it must be that $b^* \succcurlyeq a^*$ is satisfied.*

Proof. Based on property 3 in Assumption 4.1, if $B \succcurlyeq A$, then it is possible that

$$\begin{aligned} f(A,a^*) - f(A) &\geq \beta f(a^*), \text{ but} \\ f(B,a^*) - f(B) &< \beta f(a^*) \end{aligned}$$

On the other hand, if $f(B, b^*) - f(B) \geq \beta f(b^*)$, then $f(A, b^*) - f(A) \geq \beta f(b^*)$ must be satisfied. □

Combined together, Lemma 4.2 and Lemma 4.3 predict that monotone employment is optimal. Proposition 4.1 formally states it.

Proposition 4.1. *Under the given assumptions, the optimal employment portfolio is monotone. Further, it can be achieved by the marginal improvement algorithm.*

Proof. In a local optimum, there is no applicant unemployed whose capability is greater than the cutoff value and every employee makes a positive marginal contribution.

Suppose that there is an element a such that $a \succcurlyeq b$, but $a \notin A$ and $b \in A$ in a locally optimal employment set. Then, the following inequality

$$f(A \setminus \{b\}, a) - \beta f(a) < f(A \setminus \{b\}) < f(A) - \beta f(b)$$

must be satisfied and, therefore

$$f(A \setminus \{b\}, a) - f(A) < \beta f(a) - \beta f(b)$$

must be satisfied, too. Because b must be greater than the cutoff element of the previous employment set $A \setminus b$, by Lemma 4.2, the following inequality

$$f(A \setminus \{b\}, a) - f(A) \geq \beta f(a) - \beta f(b)$$

must be satisfied, which contradicts the previous inequality. Thus, there must be no element missed in a locally optimal employment set. Lemma 4.2 implies that a locally optimal employment set must comprise an interval from the highest type.

Now, suppose that there are two locally optimal employment sets E_1 and E_2 ($E_1 \subset E_2$). Then, there must be at least one element e such that $e \in E_2$, but $e \notin E_1$. Therefore, e must be smaller than the cutoff value for E_1 but greater than that for E_2. This result, however, violate Lemma 4.3. Thus, there must be at most one locally optimal employment set, which is eventually global optimal. □

4.4 Example: Hiring Multiple Professional Forecasters

This chapter considers, as an example, the optimal employment when a monopolistic information demander, a principal, wishes to gather information from multiple professional forecasters. We assume each forecaster has heterogeneous information quality, or precision of the signal, and his reservation wage depends on the quality. When the information quality is known to the principal, his objective is to select the set of forecasters which provides the best information.

A principal tries to make the best prediction of the true state, for example the profitability of a project. To get better information beyond common prior, the principal wishes to hire privately informed professional forecasters, called experts hereafter. Experts are heterogeneous in the precision of their private signals on the true state, which can be interpreted as their ability, is labeled as their *type*.

The true state x is drawn from a normal distribution with mean μ_x and variance

$1/\tau_x$, which is common knowledge. While the principal has no private information, expert i of type $\tau_i \in S$ receives a conditionally independent private signal s_i, where $S = \{\tau_1, \ldots, \tau_N\}$ is the set of experts who applied for the job positions. The distribution of s_i is assumed to follow a normal distribution with mean x and precision τ_i, or

$$s_i = x + \epsilon_i, \quad \epsilon_i \sim N\left(0, 1/\tau_i\right).$$

Once hired, each expert should submit a *report* on the true state, denoted by $r_i \in \mathbb{R}$ for expert i. We assume the compensation scheme is designed for each expert to report his posterior mean. In other words, the honest report is a priori assumed.[5] Since the principal knows the precision of each expert, the honest report on the true state is equivalent to the honest report on the signal received. Formally, we have the one-to-one relationship between the private signal and the honest report:

$$s_i = \frac{(\tau_x + \tau_i)r_i - \tau_x \mu_x}{\tau_i}$$

The principal's objective is to maximize *revenue* less payoffs to employed experts. The revenue function R depends on the principal's prediction on the true state, denoted by $\widehat{x}$, and the true state. We specify it as $R(x, \widehat{x}) \equiv \alpha - \gamma(\widehat{x} - x)^2$ for constants $\alpha > 0$ and $\gamma > 0$. In other words, the principal tries to minimize the mean squared error[6].

[5]We can actually design an optimal contract which can derive the first-best outcome for this case if the true state is verifiable ex post. See chapter 3 for more.

[6]The qualitative results of this chapter hold provided the principal minimizes any power function of ex-post error, $\mid \hat{x} - x \mid$. This is because the expectation of the power of error is a constant times the power of the posterior variance. One can easily transform the optimization problem into one with the mean squared error.

With known precisions and honest reports, the principal's best prediction is the posterior mean of the true state given all information available for her. If $J \subseteq S$ is the employment set, we have

$$\widehat{x} = \frac{\sum_{\tau_j \in J} r_j(\tau_j + \tau_x) - (|J| - 1)\mu_x \tau_x}{\sum_{\tau_j \in J} \tau_j + \tau_x},$$

This prediction is unbiased, so the expected mean squared error becomes the posterior variance, $\tau_x + \sum_{\tau_j \in J} \tau_j$. Principal's objective is to select the best subset $E \subseteq S$ which maximizes the revenue less the wage payments. The optimization problem is formally described as follows.

$$\text{(4.3)} \quad \max_{J \subseteq S} E_x \left[\alpha - \gamma(\widehat{x} - x)^2 - \sum_{\tau_j \in J} w(\tau_j) \right] = \left(\alpha - \frac{\gamma}{\tau_x + \sum_{\tau_j \in E} \tau_j} \right) - \sum_{\tau_j \in E} w(\tau_j)$$

In principle, the problem can be solved through computing values of the objective function over the power set of alternatives (in our setting over the power set of applicants). It belongs to the class of *combinatorial optimization problems*, which aims at finding the best subset from a finite set of alternatives.[7]

Before we begin the analysis, we define some functions for notational convenience. In this example, the production function f is defined as

$$f(J) \equiv g(\sum_{\tau_j \in J} \tau_j) = \alpha - \frac{\gamma}{\tau_x + \sum_{\tau_j \in J} \tau_j}, \qquad J \subseteq S.$$

If J is a singleton, $f(J)$ is called the *single production* function.

Let $T_J = \tau_x + \sum_{\tau_j \in J} \tau_j$, which is the precision of the employees in J. In the later part of this paper, we sometimes suppress the subscript J and use T to represent T_J.

[7]The utility maximization problem given a price vector in Gul and Stacchetti (1999) is also isomorphic to ours, where the utility and the price are analogue to the information gain and the reservation wage, respectively. In Gul and Stacchetti (1999), however, the SI condition is satisfied.

The *marginal production* function is derived from information contribution function. For prior precision T and an additional signal with precision τ, let

$$\triangle\,(T,\tau) \equiv g(T+\tau) - g(T) = \frac{\gamma\tau}{T(T+\tau)}.$$

Similarly, we define the *marginal loss* function when we remove a signal with precision τ from a set J with precision T.

$$\nabla(T,\tau) \equiv g(T) - g(T-\tau) = \frac{\gamma\tau}{T(T-\tau)}.$$

We have the following properties of the information contribution function.

Lemma 4.4. *The production function f satisfies all properties in Assumption 4.1 and Assumption 4.2.*

Proof. $g(\sum_{\tau_j\in J}\tau_j) = \alpha - \frac{\gamma}{\tau_x+\sum_{\tau_j\in J}\tau_j} = \alpha - \frac{\gamma}{T_J}$ is increasing concave with respect to T_J. By Topkis (1998), g satisfies the properties in Assumption 4.1. The third derivative of g is $\frac{6\gamma}{T_J^4}$, which is obviously greater than zero. Therefore, $g'' < 0$, which is equivalent to the difference in marginal production, increases as T_J increases. Assumption 4.2 is satisfied. □

4.4.1 Discussion on the Specification

In our model, the assumptions for the production function does not necessarily satisfy the single improvement property. f satisfies the single improvement property if, for *any* reservation wage schedule and an employment portfolio $A \subseteq S$, if A is not optimal, then there exists another subset $B \neq A$ such that $f(A) - w(A) < f(B) - w(B)$, $\#(A\backslash B) \leq 1$, and $\#(B\backslash A) \geq 1$.

The single improvement condition is too strong to be satisfied in many cases. We show an example that a production function which satisfies Assumptions 4.1 and 4.2 may violate the single improvement property (Example 4.1).

Instead of relaxing the single improvement property in the production function, we put a constraint on the reservation wage schedule to achieve the optimal monotone employment through the marginal improvement algorithm. The assumption on reservation wage is critical to achieve the optimal monotone employment. To motivate, we provide examples which show that (i) if the specification on the reservation wage schedule violates Assumption 4.3, the marginal improvement algorithm fails to achieve the optimum even if the production function satisfies Assumptions 4.1 and 4.2(Example 4.1), and (ii) the optimal employment portfolio is not monotone (Example 4.2).

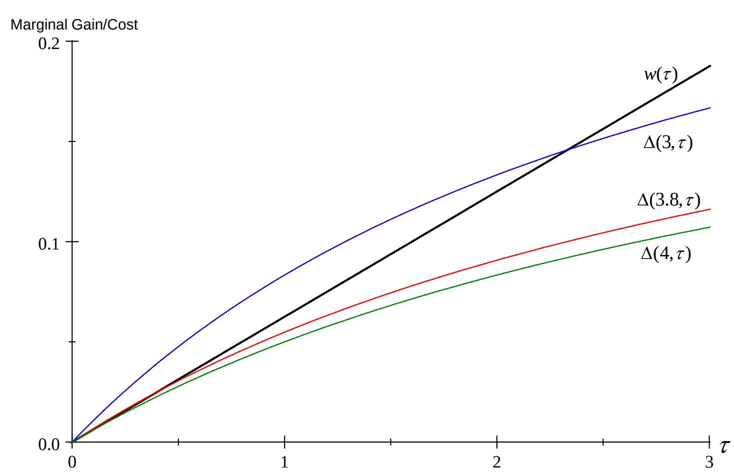

Figure 4.1: Marginal Information Contribution and Reservation Utility

Example 4.1. Figure 4.1 shows a linear reservation utility function $w(\tau) = \frac{1}{16}\tau$ and marginal information contribution functions when $\gamma = 1$. Let $\tau_x = 1$ and $\tau \in (0, 3]$. Note first that when the principal is to hire one expert in the beginning, the highest type is always preferred since $\triangle\ (T, \tau) - w(\tau) = \frac{\tau}{(1+\tau)} - \frac{1}{16}\tau$ is increasing when $\tau \leq 3$.

Suppose that the pool of applicants S has three potential employees $\{\tau_1, \tau_2, \tau_3\}$ and that from the pre-screening the principal knows their types are $\tau_1 = 1$, $\tau_2 = 2$, and $\tau_3 = 2.8$. If the principal chooses the marginal improvement algorithm, or finding the local maximum in each step, she will first hire $\tau_3 = 2.8$, who makes the biggest marginal production. Then, both τ_1 and τ_2 cannot create positive contributions and the hiring process stops. $T = 3.8$ and the principal's profit is $\alpha - \frac{1}{1+2.8} - \frac{2.8}{16} = \alpha - 0.43816$. On the other hand, if she starts from hiring $\tau_2 = 2$, she will be able to hire another agent, $\tau_1 = 1$. Then, her profit is $\alpha - \frac{1}{1+2+1} - \frac{1+2}{16} = \alpha - 0.4375$, which is bigger than $\alpha - 0.43816$ and, actually, the global optimum. The marginal improvement algorithm fails to achieve a global optimum. The marginal production functions for each case are demonstrated in Figure 4.1.

Adding, deleting, or substituting only one element from $J = \{\tau_3 = 2.8\}$ does not improve the principal's payoff. The optimal employment portfolio, however, is $E = \{\tau_1 = 1, \tau_2 = 2\}$. The single improvement property obviously fails here. ■

Example 4.2. Consider now three applicants of type $\tau_1 = 1$, $\tau_2 = 1.5$, and $\tau_3 = 2$. Under the marginal improvement algorithm, the principal first hires $\tau_3 = 2$, whose marginal profit is biggest. Then, the marginal production from hiring one more agent becomes $\Delta(1 + 2, \tau) = \frac{1}{1+2} - \frac{1}{1+2+\tau}$, and the marginal profit is $\Delta(1 + 2, \tau) - \frac{1+2+\tau}{16}$, which is positive if $\tau = 1$, but not if $\tau = 1.5$. The principal chooses to hire τ_1 and the employment portfolio is $E = \{\tau_1 = 1, \tau_3 = 2\}$, which happens to be optimal, but obviously not monotone. ■

The above example shows that the marginal improvement algorithm does not always lead to the global optimum, and the optimal employment set may not comprise an interval. The key point in this example is the quasi-concavity of the marginal profit. As the reservation utility function is close to linear, the marginal profit is always more concave than the reservation utility function. If the reservation utility is linear or convex, the marginal profit is always concave due to the submodularity of f. This implies that as the employment set is enlarged, or equivalently the information is cumulated, the lower type has a better chance of being hired than the higher type, though initially the higher type contributes more. This breaks down the monotonicity of the positive profit and finding a local maximum at each step, or marginal improvement algorithm, may not lead to the global optimum.

4.4.2 Properties of Optimal Employment Set

Even though it does not satisfy the single improvement property, the objective function of our model has a nice feature. Any set of experts can be characterized by a single real number, the sum of precisions of experts in the set. This allows us to transform the objective set function to a function on the two dimensional Euclidian space. This transformation allows us to solve the problem through a greedy algorithm, as will be discussed later.

Yet the optimal employment set is quite arbitrary since it depends heavily on the form of reservation utility function. We make a critical but reasonable assumption on the reservation utility in the next section: the reservation utility is proportional to the marginal single information contribution. Under this specification, the optimal employment set is shown to follow a cut-off property. We then discuss on general cases, providing an example of complicated optimal employment set. The comparison of our model and other combinatorial optimization problem is presented in the final subsection.

The result crucially depends on the quasi-convexity of the marginal profit, which is due to the fact that, roughly speaking, the marginal production is less convex than the reservation wage. Moreover, the marginal production becomes less concave as T gets large. This implies that even though the marginal production is more concave than reservation utility at the initial state (when $T = \tau_x$), it might become less concave when T approaches the optimal cut-off point. The specification of reservation utility in the previous subsection shows exactly this case. Initially, the curvature of the reservation utility is the same as that of the marginal information contribution. For $T > \tau_x$, however, the curvature of the former is always bigger than the latter.

We now apply Assumption 4.3 on the reservation wage schedule. It is proportional to the single production, or production by single agent.[8], that is, for $\kappa > 1$,

$$w(\tau) \equiv \frac{1}{\kappa} \triangle (\tau_x, \tau) = \frac{\gamma\tau}{\kappa\tau_x(\tau_x + \tau)}.$$

We also normalize $\gamma = \tau_x = 1$. The assumption is quite strict but reasonable. It says that if the expert utilizes his private information outside of the relationship with the principal, his gain is proportional to his single production under the principal. Behind the assumption we think potential principals in the market share the same information on each expert so that the gain inside of the market should be the same across principals. Since all employers do not know who is what type, the reservation utility cannot be type dependent. The only situation where an expert with higher ability gains more lies in the case when he uses the private information for his own gain.

To prove the key characterization of the optimal employment set, we need the following lemma:

Lemma 4.5. *The profit from adding τ from a set with collective precision T, $\triangle (T, \tau) - w(\tau)$ crosses zero on $\tau > 0$ at most once and from below. Likewise, the profit from dropping τ, $w(\tau) - \nabla(T, \tau)$ crosses zero at most once and from above.*

Proof. $\triangle (T, \tau) - w(\tau) > 0$ if and only if $T(T+\tau) - \kappa(1+\tau) < 0$. Since it is linear in τ, for some $\tau > 0$ to satisfy the equality we must have either $\kappa - T^2 > 0$ and $T - \kappa > 0$

[8] We strongly conjecture that the main result still hold if the reservation utility is a concave function of the single production. This is because, as will be clarified later, the main result depends on the fact that the marginal contribution function near the global optimum crosses the reservation utility only once and from below. The fact is still satisfied when the reservation utility is concave in the marginal single contribution.

or $\kappa - T^2 < 0$ and $T < \kappa$. However, the latter inequality cannot hold because $\kappa > 1$. Thus, we need to check only the case of $\sqrt{\kappa} < T < \kappa$. Then, $T(T+\tau) - \kappa(1+\tau)$ is a decreasing function of τ and crosses zero only once rom above. This implies that $\triangle\ (T,\tau) - w(\tau)$ crosses zero at most once from below.

For the dropping case, $w(\tau) - \nabla(T,\tau) > 0$ if and only if $(T+\kappa)\tau - T^2 + \kappa < 0$. This crosses zero at most once regardless of the value of T. Since it crosses from below, $w(\tau) - \nabla(T,\tau)$ crosses from above. □

The intuition of Lemma 4.5 is as follows. For the marginal information contribution function to cross the reservation utility function, T must be in an appropriate range. Since the marginal production becomes less concave as T increases, it is flat relative to the reservation wage function in the range of T.

We need an additional lemma to prove the main proposition.

Lemma 4.6. *Given $T^2 > \kappa$, if τ_2 satisfy $\triangle\ (T,\tau) - w(\tau) = 0$ and τ_1 satisfy $w(\tau) - \nabla(T,\tau) = 0$, then, $\tau_1 < \tau_2$.*

Proof. When $T^2 \leq \kappa$, there is no applicant who can make a positive profit for the principal. Then, there will no more employment. When $T^2 > \kappa$, the existences of τ_1 and τ_2 are immediate from the proof of Lemma 4.5. We have $(T-\kappa)\tau_2 + T^2 - \kappa = 0$ and $(T+\kappa)\tau_1 - T^2 + \kappa = 0$. But then, $(T+\kappa)\tau_2 - T^2 + \kappa > -(T-\kappa)\tau_2 = T^2 - \kappa$. Thus, $w(\tau_2) - \nabla(T,\tau_1) < 0$, which implies that $\tau_1 < \tau_2$. □

Proposition 4.2. *Under the given specification, there exists at most one cutoff ele-*

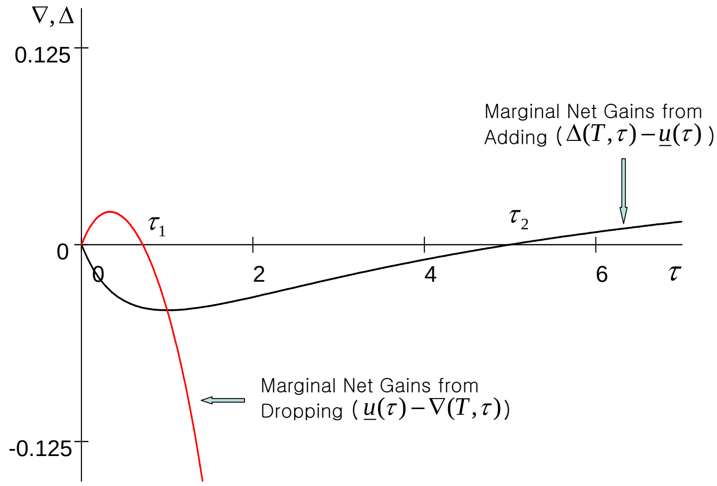

Figure 4.2: Marginal Contribution from Adding/Dropping an Expert with τ ($\kappa = 4$, $T = 3$)

ment $\tau^* \leq \tau_N$ such that the optimal employment portfolio includes any expert $\tau_i \geq \tau^*$ and can be achieved by the marginal improvement algorithm.

Proof. Let $E \subseteq S$ be the optimal set and let T_E be the associated collective precision. Note first that if $T_E^2 < \kappa$, E cannot be the optimum unless $E = S$ since adding any expert in $S \setminus E$ yields positive net gain. We only consider the case $T_E^2 > \kappa$.

Define τ_1 and τ_2 as in the proof of lemma 4.6. Suppose $e \in E$ is less than τ_1. Then, dropping it improves net gain, contradicting the optimality. Similarly, any $u \in S \setminus E$ cannot be bigger than τ_2. The only thing we need to check is the case in which there are i and j such that both are between τ_1 and τ_2, $\tau_i < \tau_j$, and $\tau_i \in E$ but $\tau_j \in S \setminus E$. Consider $E \setminus \{\tau_i, \tau_j\}$. The optimality implies that $\triangle\ (T_E - \tau_i, \tau_i) - w(\tau_i) > \triangle\ (T_E - \tau_i, \tau_j) - w(\tau_j)$. But then $\triangle\ (T - \tau_i, \tau) - w(\tau)$ crosses zero from above, which contradicts Lemma 4.6. (Refer to Figure 4.2) This completes the proof. □

4.5 Conclusion and Discussion

An employer may need to hire multiple experts for a single project such as financial forecasting, consulting, or pharmaceutical production even if there is little complementarity among the experts. If the employer cares about the maximal output across agents, hiring multiple agents can increase the expected value of the output. In other words, multiple samples from a distribution are better than a single sample.

There are two issues in finding an optimal employment portfolio in this environment. First, if the production function is submodular, it is computationally complex even when the experts' abilities are known (Lovász, 1983). Second, the optimal employment portfolio is not well-behaved.

If the gross substitutes condition (Kelso and Crawford, 1982) or single improvement property (Gul and Stacchetti, 1999) is satisfied for a set production function, the global maximum can be achieved by a greedy algorithm (Murota and Shioura, 2003). Unfortunately, the single improvement property is too restrictive in many cases.

Like Chade and Smith (2006), we first provide conditions under which, even though the single improvement property is not satisfied, the optimal employment portfolio is achieved through a simple myopic approach, the marginal improvement algorithm. On the other hand, while Chade and Smith (2006) focus on the specification of the production function and cost structure under which the marginal improvement algorithm leads to the global optimum, we propose conditions under which the optimal employment portfolio can be simply described by a cutoff element where all experts with greater ability than the cutoff are hired and the rest are not. We call this result optimal monotone employment. Such a property does not necessarily obtain in Chade and Smith (2006).

In our model, if an increasing, strictly submodular production function shows increasing differences in marginal production, and an agent's reservation wage is a fraction of his individual production, monotone employment is optimal and, therefore, can be achieved by the marginal improvement algorithm. We provide an example in which the single improvement property is violated, but monotone employment is nevertheless optimal.

One drawback of our result is that there is no guarantee that the core exists with our proposed conditions. Because the single improvement condition guarantees the existence of the core, a Walrasian equilibrium can be achieved in matching or auction applications. The possibility that the core may not exist limits the application of our model to other economic problems.

A few questions remain as future research topics. First, while we conjecture that our condition is weaker than the single improvement property, the exact relationship between these two conditions needs to be investigated further. Second, we believe that a positive concave transformation of the reservation wage schedule we use will lead to the same result. Finally, topological features of the production function and reservation wage schedule need to be explored.

Bibliography

Alesina, A. and A. Cukierman (1990). The politics of ambiguity. *Quarterly Journal of Economics 105*, 24–52.

Alesina, A. and R. Holden (2008). Ambiguity and extremism in elections. *Working Paper*.

Ansolabehere, S. and J. M. Snyder Jr. (2000). Valence politics and equilibrium in spatial election models. *Public Choice 103*(3-4), 327–336.

Aragones, E. and T. R. Palfrey (2002). Mixed equilibrium in a downsian model with a favored candidate. *Journal of Economic Theory 103*, 131–161.

Avery, C. N. and J. A. Chevalier (1999). Herding over the career. *Economics Letters 63*(3), 327–333.

Banks, J. and J. Duggan (2005). Probabilistic voting in the spatial model of elections: The theory of office-motivated candidates. In D. Austen-Smith and J. Duggan (Eds.), *Social Choice and Strategic Decisions: Essays in Hornor of Jeffrey S. Banks*. Springer Berlin Heidelberg.

Battaglini, M. (2002). Multiple referrals and multidimensional cheap talk. *Econometrica 70*(4), 1379–1401.

Berger, M. M., M. M. Munger, and R. F. Potthoff (2000). Expository note: The downsian model predicts divergence. *Journal of Theoretical Politics 12*(2), 228–240.

Bernhardt, D., O. Câmara, and F. Squintani (2008). Competence and ideology. *Working Paper*.

Bernhardt, M. D. and D. E. Ingberman (1985). Candidate reputations and the 'incumbency effect'. *Journal of Public Economics 27*, 47–67.

Bhattacharya, S. and P. Pfleiderer (1985). Delegated portfolio management. *Journal of Economic Theory 36*(1), 1–25.

Calem, P. S. and J. A. Rizzo (1995). Competition and specialization in the hospital industry: An application of the hotelling's location model. *Southern Economic Journal 61*, 1182–1198.

Callander, S. and S. Wilkie (2007). Lies, damned lines, and political campaigns. *Games and Economic Behavior 60*, 262–286.

Chade, H. and L. Smith (2006). Simultaneous search. *Econometrica 74*(5), 1293–1307.

Coate, S. and S. Morris (1999). Policy persistence. *American Economic Review 89*(5), 1327–1336.

Cormen, T. H., C. E. Leiserson, R. L. Rivest, and C. Stein (2001). *Introduction to Algorithms* (2nd ed.). The MIT Press.

Crawford, V. and J. Sobel (1982). Strategic information transmission. *Ecoometrica 50*(6), 1431–1451.

Crémer, J. and R. P. McLean (1985). Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica 53*(2), 345–361.

Crémer, J. and R. P. McLean (1988). Full extraction of the surplus in bayesian and dominant strategy auctions. *Econometrica 56*(6), 1247–1257.

Dasgupta, P. and E. Maskin (1986). The existence of equilibrium in discontinous game: Part i theory. *Review of Economic Studies 53*, 1–26.

DeBacker, J. (2008). Flip-flopping: Ideological adjustment costs in the united states senate. *Working paper*.

Downs, A. (1957). *An Economic Theory of Democracy*. New York: Harper and Row.

Duggan, J. (2000). Equilibrium equivalence under expected plurality and probability of winning maximization. *Working Paper*.

Enelow, J. and M. J. Hinich (1981). A new approach to voter uncertainty in the downsian spatial model. *American Journal of Political Science 25*, 483–493.

Eyster, E. and T. Kittsteiner (2007). Party platforms in electoral competition with heterogeneous constituencies. *Theoretical Economics 2*, 41–70.

Feddersen, T. and W. Pesendorfer (1996). The swing voter's curse. *American Economic Review 86*(3), 408–424.

Feddersen, T. and W. Pesendorfer (1997). Voting behavior and information aggregation in elections with private information. *Econometrica 65*, 1029–1058.

Fiorina, M. P. (1973). Electoral margins, constituency influence, and policy moderation: A critical assessment. *American Politics Quarterly 1*(4), 479–498.

Green, J. R. and N. L. Stokey (1983). A comparison of tournaments and contracts. *Journal of Political Economy 91*(3), 349–364.

Griffin, J. D. (2006). Electoral competition and democratic responsiveness: A defense of the marginality hypothesis. *Journal of Politics 68*(4), 909–919.

Groseclose, T. (2001). A model of candidate location when one candidate has a valence advantage. *American Journal of Political Science 45*, 862–886.

Grossman, G. M. and E. Helpman (1994). Protection for sale. *American Economic Review 84*(4), 833–850.

Gul, F. and E. Stacchetti (1999). Walrasian equilibrium with gross substitutes. *Journal of Economic Theory 87*, 95–124.

Hansen, R. G. (1985). Auctions with contingent payments,. *American Economic Review 75*(4), 862–865.

Hinich, M. J. (1977). Equilibrium in spatial voting: The median voter result is an artifact. *Journal of Economic Theory 16*, 208–219.

Holmstrom, B. (1982). Moral hazard in teams. *Bell Journal of Economics 13*(2), 324–340.

Jullien, B. (2000). Participation constraints in adverse selection models. *Journal of Economic Theory 92*, 1–47.

Kartik, N. and R. P. McAfee (2007). Signaling character in electoral competition. *American Economic Review 97*(3), 852–870.

Kelso, A. S. and V. P. Crawford (1982). Job matching, coalition formation,and gross substitutes. *Econometrica 50*(6), 1483–1504.

Levitt, S. D. (1995). Optimal incentive schemes when only the agents' "best" output matters to the principal. *Rand Journal of Economics 26*(4), 744–760.

Levy, G. (2004). Anti-herding and strategic consultation. *European Economic Review 48*(3), 503–525.

Liu, Y., D. S. Putler, and C. B. Weinberg (2004). Is having more channels really better? a model of competition among commercial television broadcasters. *Marketing Science 23*, 120–133.

Lizzeri, A. and N. Persico (2001). The provision of public goods under alternative electoral incentives. *American Economic Review 91*(1), 225–239.

Lovász, L. (1983). Submodular functions and convexity. In A. Bachem, M. Grotschel, and B. H. Körte (Eds.), *Mathematical Programming: The State of the Art*, pp. 235–257. Berlin: Springer-Verlag.

Maskin, E. and J. Riley (1984). Monopoly with incomplete information. *RAND Journal of Economics 15*(2), 171–196.

Milgrom, P. and J. Roberts (1990). Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica 58*, 1255–1277.

Milgrom, P. and C. Shannon (1994). Monotone comparative statics. *Econometrica 62*(1), 157–180.

Murota, K. and A. Shioura (2003). Quasi m-convex and l-convex functions - quasiconvexity in discrete optization. *Discrete Applied Mathematics 131*, 467–494.

Osband, K. (1989). Optimal forecasting incentives. *Journal of Political Economy 97*(5), 1091–1112.

Ottaviani, M. and P. N. Sørensen (2006a). Reputational cheap talk. *RAND Journal of Economics 37*, 155–175.

Ottaviani, M. and P. N. Sørensen (2006b). The strategy of professional forecasting. *Journal of Financial Economics 81*(2), 441–466.

Patty, J. W. (2002). Equivalence of objectives in two candidate elections. *Public Choice 112*, 151–166.

Persson, T., G. Roland, and G. Tabellini (2000). Comparative politics and public finance. *Journal of Political Economy 108*(6), 1121–1161.

Persson, T. and G. Tabellini (1999). The size and scope of government: Comparative politics with rational politicians. *European Economic Review 43*, 699–735.

Prendergast, C. (1993). A theory of "yes man ". *American Economic Review 83*(4), 757–770.

Scharfstein, D. S. and J. C. Stein (1990). Herd behavior and investment. *American Economic Review 80*(3), 465–479.

Shepsle, K. (1972). The strategy of ambiguity: Uncertainty and electoral competition. *American Political Science Review 66*, 551–558.

Topkis, D. M. (1998). *Supoermodularity and Complementarity.* Princeton University Press.

Trueman, B. (1994). Analyst forecasts and herding behavior. *Review of Financial Studies 7*(1), 97–124.

Wittman, D. (2007). Candidate quality, pressure group endorsements, and the nature of political advertising. *European Journal of Political Economy 23*(2), 360–378.