THREE ESSAYS IN APPLIED MICROECONOMICS

By

Zijian Qi

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Economics – Doctor of Philosophy

2022

ABSTRACT

THREE ESSAYS IN APPLIED MICROECONOMICS

By

Zijian Qi

This dissertation has three chapters, each concentrating on a distinct aspect of information asymmetry. Each chapter approaches information asymmetry from a unique perspective: the first chapter explores a scenario with both hidden information and hidden action. The second chapter discusses another type of information asymmetry related to population uncertainty. Finally, the third chapter focuses on a natural result of information asymmetry — discrimination.

Chapter one explores a situation in which managers rely on their subordinates for local information that aids decision-making but cannot commit to a decision rule. When the firm and the workers have conflicting interests on how such information gets used, incentives for effort and information elicitation become intertwined. We explore how one may solve this incentive problem through job design—the choice between "individual assignment" where all tasks in a given job are assigned to the same worker, and "team assignment" where the tasks are split among a group. Team assignment facilitates information elicitation but suffers from "diseconomies of scope" in incentive provision. This trade-off drives the optimal job design, and it is shaped by two key parameters — the workers' ex-ante likelihood of being informed and the noise in the performance measure that is used to reward the worker. The individual assignment is optimal when the performance measure is well-aligned, but the team is optimal when the measure is noisy, and the workers are highly likely to be informed about the local conditions.

In chapter two, I study a contest with population uncertainty in which the value of the prize depends on the number of participants. There is friction between a contestant's perspective and an outsider's perspective regarding the number of contestants. This discrepancy drives the main result: under the assumption that the expected value of the prize is the same across all environments, if the value of the prize increases in the number of players, the players exert more effort; whereas, if the value of the prize declines in the number of players, the players exert less effort.

In the third chapter, I focus on discriminating as a consequence of information asymmetry. I construct a two-stage assimilation model to analyze the discrimination level in groups with different discount factors. I have three main results: First, there always exists an equilibrium for any discount factors and minority group size; the equilibrium will have an on-path action profile with a cutoff rule; second, as group size increases, both discrimination level and the ability cutoff will increase; third, when discount factors vary across different regimes, the effect is not monotonic.

To MY PARENTS

ACKNOWLEDGMENTS

I am overcome with appreciation and humility to express my gratitude to everyone who has helped me finish this dissertation. First, I would like to thank Dr. Jon Eguia, my graduate adviser, for his assistance throughout my studies. I would not be the same man I am today without his guidance, advice, and support. In addition, I thank Dr. Arijit Mukherjee for his patience with my adaptation, his innovative learning-by-doing practice, and his service on my committee. I would also like to express my gratitude to Dr. Jay Pil Choi and Dr. Julian Guo, my committee members, for their guidance and inspiration. Their constructive feedback and comments significantly improve the quality of this dissertation.

I am grateful to Dr. Luís Vasconcelos, who inspired me with his perseverance and determination during our collaboration. My gratitude also extends to the whole academic faculty and administrative staff of the Department of Economics at Michigan State University. I want to thank all of my friends at Michigan State University for their unwavering support and friendship. Thank you so much for the wonderful time we shared.

I acknowledge Dr. Yiran Fan. We have known each other for 25 years and have witnessed the development of each other. Yiran was always devoted and tenacious in making the world a better place, and his true and pure nature never ceases to inspire me. His philosophical insights and words of wisdom will enlighten me for the rest of my life.

My family has always been my greatest source of encouragement and motivation. My parents have done so much for me, from instilling reasoning and skepticism to caring for me when I am injured. They have supplied me with all the capabilities and resources I need to succeed for many years, and I will be eternally thankful for their love and support.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

Information asymmetry shatters the hall of welfare economics erected by neoclassical economic models. In a world with complete information (together with other assumptions), welfare economics's first and second fundamental theorems ensure a well-behaved economy, and Adam Smith's "invisible hand" property is satisfied in all markets. In such a Utopian society, government regulation is straightforward since there is no way to make Pareto improvements, leaving transfers as the only option for the government. However, reality no longer resembles such a beautiful world when information asymmetry occurs. Information asymmetry has been studied extensively over the last century, and it is still an active topic in economics.

Two traditional topics are well-studied on asymmetric information: adverse selection and moral hazard. Starting from the seminal work by Akerlof (1970) and Spence (1973), a large amount of literature has explored the topic of adverse selection. Adverse selection occurs when one party has hidden information. In other words, the type of player is private information. In some cases, the social planner could restore market efficiency through signaling or screening, but it cannot be guaranteed. In many scenarios, the social planner can only achieve the second-best. On the other hand, a moral hazard problem occurs when there is hidden action. Grossman and Hart (1983) devised a principal-agent model, which can be used to explain the moral hazard problem. In some scenarios, a carefully crafted contract could alleviate the moral hazard issue, although the first best cannot always be achieved. In the chapter "Optimal Job Design and Information Elicitation", we look at a situation where there is both hidden information and hidden action.

Levitt and Snyder (1997) are one of the first to analyze the interaction between hidden information (information elicitation) and hidden action (effort provision). The incentive problem for hidden information or hidden action is straightforward, but things become more complicated when the incentives become entwined. As a result, when the agents may be privately informed about projects' viability, the optimal contract is driven by the trade-off between efficiency in decision making and effort incentives. In order to induce both effort and truthful reporting of "bad news" (i.e., information that lowers the likelihood of the project's success), the optimal contract calls for an inefficiently "lenient" continuation policy where some projects with negative expected value are allowed to continue.

Our model examines a similar scenario in which incentives for effort provision and information elicitation are interwoven, but it focuses on a fundamental aspect of the organization structure — job design. Job design refers to how to group different tasks into jobs that may be assigned to workers. "individual assignment" and "team assignment" are the two most natural forms of job design. Under "individual assignment", all tasks associated with a specific production process are assigned to the same worker who is exclusively responsible for his job output. Alternatively, under "team assignment", different tasks in the production process are assigned to different workers who are jointly responsible for their work performance.

Our analysis highlights that team assignment facilitates information elicitation, whereas individual assignment facilitates the provision of effort incentives. Under individual assignment, an agent can fully control the project's outcome, as he can control the effort levels and the information of the project. Team assignment helps mitigate such incentive problems as the agent does not fully control the outcome of a project. For example, his attempt to conceal information may falter if his teammate happens to provide the same information to the principal, and he influences the effort only in a part of the project. Thus, team assignment facilitates information elicitation. However, team assignment suffers from "diseconomies of scope" because the principal must reward agents separately for distinct tasks to induce effort on a single project. Such diseconomies of scope might stifle motivation. On the other hand, individual assignment simply requires the principal to pay one reward for a single project. Thus, individual assignment facilitates effort provision.

Finally, we show that the optimal job design is driven by two salient informational frictions: the "availability" of agents' information and the "noise" in the agents' performance measure. Team assignment is strictly optimal when the agents are highly likely to observe the state, but there is a significant misalignment between the performance measure and the project output. In contrast, when the extent of misalignment is relatively small, individual assignment is strictly optimal regardless of the agents' likelihood of being informed about the state.

The scenario in which both hidden action and hidden information are present is the subject of the preceding analysis. In the chapter "Contests with Valuation Associated with Population Uncertainty", I explore information asymmetry in a another perspective. In a standard contest, a fixed number of players exert efforts to compete for a prize. If the players have different potential types and the type is private knowledge, the contest becomes one with hidden information. I investigate a different variant in which the number of players is random.

A contest with population uncertainty appears to be very similar to a contest with different types. To be more specific, a player in the former contest is uncertain about how many other players are, and a player in the latter contest is uncertain about who the other players are. A naïve intuition would lead one to believe that these two contests share the same feature of incomplete information in a general framework and that the solutions are fairly similar. This is incorrect since there is a significant distinction between "how many" and "who" in the contest setup. Once a player has entered the contest, the player's belief is updated, as "I am in the contest" already contains some information. As a result, a player's belief differs from an outsider's (game theorist's) without any further information. In the contest with different types, a player's belief of other players is a non-degenerate distribution and thus represents the uncertainty of "who"; whereas in the contest with population uncertainty, the uncertainty of "how many" contains not only a non-degenerate distribution as a belief but also the discrepancy between a player's belief and an outsider's belief.

Contests with population uncertainty are first studied by Myerson and Wärneryd (2006). In the paper, they set up a contest where the number of players is stochastic, and the value of the prize is fixed. They show that the total equilibrium expenditure is strictly lower in a contest with population uncertainty than in a contest without population uncertainty, even though the expected number of players is the same in both contests.

I extend the model such that the value of the prize could be dependent on the number of players. I then consider the following three scenarios:

(a) The value of the prize is constant;

(b) The value of the prize is increasing in the number of players;

(c) The value of the prize decreases in the number of players.

Under the assumption that the expected value of the prize is the same in (a), (b), and (c), I find that the effort level is high if the value of the prize is increasing in the number of players, and conversely, the effort level is low if the value of the prize is decreasing in the number of players. I also consider the following scenarios:

(d) the number of players is a constant, and the value of the prize is also a constant;

(e) the number of players is random, and the value of the prize is linear on the number of players with zero intercepts.

When the expected number of players and the expected value of the prize is the same in (d) and (e), I find the effort level is the same under (d) and (e).

The preceding chapters look at several types of information asymmetry. Then, in the chapter "Assimilation with Different Working Skill Acquisition", I study the potential outcome of information asymmetry in a real-world context. People are divided into separate groups, as they have diverse cultures. In addition, various groups may exhibit different traits due to information asymmetry. These traits could include how they discount the future, how they expose themselves to knowledge, and how they collaborate. The cultural barrier creates a chasm, and communication comes at a cost in the form of discrimination. I examine how the level of discrimination varies in different circumstances.

Eguia (2017) is one of the pioneering works on this topic. The paper presents a two-stage assimilation model: in the first stage, the majority group sets a level of discrimination, which is the barrier that people from the minority group must overcome if they want to assimilate; in the second stage, people from both the majority and minority groups choose their skill level, and people from the minority group can decide whether or not to assimilate. It concludes that the majority group utilizes discrimination as a screening technique and that only highly skilled minority members will assimilate. This screening equilibrium is optimal for the majority group because the persons who assimilated into the majority group are highly skilled and will generate positive peer effects. This paper restricts attention to circumstances where a minority group is at a disadvantage over the majority.

I investigate a situation in which a minority group has an advantage over the majority group. The minority has a higher discount factor and places a higher value on the future. I provide another two-stage assimilation model: in the first stage, people spend time learning working skills; in the second stage, the majority group establishes a level of discrimination, and minority group members can choose whether or not to assimilate. I show that an equilibrium exists for all discount factors and minority group size, and the equilibrium will have an on-path action profile with a cutoff rule. Also, when group size increases, both the discrimination level and the ability cutoff increase, but the effect is not monotonic when discount factors vary across various regimes.

This dissertation first investigates a scenario where there is hidden information and hidden action, then analyses the gap between beliefs in contests with population uncertainty, and lastly uses an assimilation model to study the potential outcome of information asymmetry.

CHAPTER 1

OPTIMAL JOB DESIGN AND INFORMATION ELICITATION*

1.1 Introduction

Managerial decision-making in a hierarchical organization often relies on local information that cannot be directly accessed by the headquarter but may be available to its lower-ranked employees. A host of key business decisions, such as launching new product lines, undertaking new business ventures, investments in new R&D initiatives, all require detailed information on customer preferences, profitability prospects, and technological capabilities that is more likely to be available to the junior workers who are more familiar with the local market conditions and the firm's production process. Effective decision-making, therefore, calls for timely provision of information that may be dispersed within an organization.

However, the firm and the workers may have conflicting interests on how information may be used, and when relaying local information to their manager, the workers may manipulate information to steer the firm's decision towards their own interests. A worker may deem his information "unfavorable" if the firm's expected action under such information could reduce the worker's future rents. Consequently, he may attempt to filter or conceal such information, particularly when the firm cannot commit on how the information may be used in its decision process. Such conflict of interest creates a complex incentive problem as the

*Disclaimer: This chapter was co-authored with Arijit Mukherjee (arijit@msu.edu) and Luís Vasconcelos (Luis.Vasconcelos@uts.edu.au). Both authors have approved that this work be included as a chapter in my dissertation.

incentives for effort and information elicitation get intricately entwined (Athey and Roberts, 2001).

Starting from the seminal work by Marschak (1955) and Marschak and Radner (1972) on team theory, a large literature has explored the limits on information provision in an organization and how these limits are influenced by the organization's structure (Aoki, 1986). However, this literature typically abstracts away from the problem of incentives as the employees' objective is assumed to be perfectly aligned with that of the employer. The goal of our paper is to explore how the problem of intertwined incentives for effort and information elicitation shapes a critical part of the organizational structure, namely, job design.

An essential problem in organizational design is how to group different tasks into jobs that may be assigned to the workers. An organization may typically choose between two natural designs: it may opt for "individual assignment" where all tasks associated with a specific production process are assigned to the same worker who remains solely accountable for his job output. Alternatively, it may choose "team assignment" where different tasks of the production process are assigned to different workers who are held jointly accountable for their job performance. When decision-relevant information is accessible only to the workers who are directly involved in the production process, the two job designs have distinct implications on how information may be dispersed within the organization. Under individual assignment, all information pertaining to a production process can be observed only by the worker who has been assigned to it, whereas multiple workers may access this information when they are working as a team.

The broad prevalence of individual and team assignments in project management structures has been well-documented in the management literature (Galbraith, 1971; Larson and

Gobeli, 1989; Hobday, 2000; Lechler and Dvir, 2010). Firms often adopt a "project-based" structure where a manager is assigned to oversee all aspects of a project, or a "functional" structure where projects are divided into segments and different segments are overseen by different managers. In exploring the relative merits of the two structures, this literature mostly focuses on the gains from task specialization vis-a-vis task coordination. We highlight that when the workers need to be incentivized for both effort provision and information elicitation, the choice between these two designs is shaped by a novel trade-off: Team assignment may facilitate information elicitation as no worker can fully control the information or project outcome, but it suffers from "diseconomies of scope" in incentive provision and may undermine workers' effort.

We explore this trade-off in a stylized model of job design in a principal-agent environment. In our setup a principal hires two agents to work on two projects. Each project has two tasks, and a project can either succeed or fail. The likelihood of success depends on the level of effort exerted in its tasks and the underlying "state of the world" that may be observable only to the agent(s) who are assigned to that project. At the beginning of the game, the principal chooses a job design: under individual assignment, each agent is responsible for a given project and is expected to exert effort in both tasks that are associated with it. Under team, an agent is assigned exactly one task from each of the two projects. While performing a task, an agent may observe the state of the world (pertaining to its associated project) with some probability and reports it to the principal. While the agent cannot misrepresent the state (i.e., observation on the state is "hard" information) he may conceal it by feigning ignorance. Up on receiving the agents' report on the state, the principal decides whether to continue or cancel a project. The project output is not verifiable, but the agents' effort in a project is reflected by a contractible but noisy performance measure.

Incentives are provided through a wage contract that ties an agent's pay to the principal's cancellation decision and the realization of the performance measures (if the project is continued). The misalignment between the performance measure and the project outcome gives rise to a conflict of interest between the principal and the agents. If the observed state does not bode well for the project's success but is unlikely to affect the performance measure (if the project is implemented), the agent may conceal his information to let the project proceed whereas the principal would have been better off by canceling it.

We show that the optimal job design is driven by two salient informational frictions: the "availability" of agents' information (i.e., the likelihood that the agent gets to observe the state while performing his assigned tasks) and the "noise" in the agents' performance measure (i.e., the extent of misalignment between the measure and the output). Team assignment is strictly optimal when the agents are highly likely to observe the state but there is significant misalignment between the performance measure and the project output. In contrast, when the extent of misalignment is relatively small, individual assignment is strictly optimal regardless of the agents' likelihood of being informed about the state.

The intuition for this result can be gleaned from the aforementioned trade-off between information elicitation and "diseconomies of scope" in incentive provision. Since the principal relies on the agents' report but cannot commit on how this information may be used, the agent may attempt to control the projects' outcome by manipulating his information and effort levels. Team assignment helps in mitigating such incentive problem as the agent does not fully control the outcome of a project. His attempt to conceal information may falter if his teammate happens to provide the same information to the principal, and he influences the effort only in a part of the project.

But incentive provision under team assignment suffers from “diseconomies of scope”: the principal needs to reward the two agents separately to induce effort on the two tasks that are associated with the project. And such diseconomies of scope can blunt incentives. As the principal cannot commit how she may use the agents’ report, in equilibrium, her continuation policy must be sequentially rational. If the principal proceeds with the project under a certain information, it must be that her expected payoff from proceeding with the project (conditional on the agents’ reports) is larger than what she might get from canceling it. These requirements put an upper bound on the amount of reward the principal pays to the agents when the project is successful (as per the performance measure). Under team, the total reward payout is larger and such bounds are harder to meet as the principal needs to pay the reward for success twice (paying each of the two agents separately) to elicit effort in both tasks. Consequently, strong incentives may be infeasible.

In contrast, under individual assignment, a single reward payment would have induced effort in both tasks, and such economies of scope in incentive provision makes it easier to provide strong incentives without violating the bounds on reward payments. However, under individual assignment, information elicitation becomes harder as the agent fully controls the outcome of a project through his report and effort.

Thus, between the two forms of job design, team assignment facilitates information elicitation whereas individual assignment facilitates the provision of effort incentives.

When the performance measure is considerably misaligned and can indicate success even when the project fails, the agents have strong incentives to conceal unfavorable information to let the project continue. This is when the team’s advantage in information elicitation is most useful: an agent's attempt to conceal information could be undone by his teammate, particularly when his teammate is very likely to have the same information. Due to such misalignment effort is also more sensitive to rewards and strong incentives may be feasible despite the scope diseconomies that arise under team. As a result, team assignment becomes optimal. In contrast, if the performance measure is relatively well-aligned with the project output, information elicitation is relatively easy as the agent has little to gain from concealing information from the principal. Thus, individual assignment becomes optimal—it allows the principal to exploit the economies of scope in incentive provision and offer strong incentives for effort without distorting the agent's reporting incentives.

Related literature: Our paper contributes to a growing literature on the interplay between incentives and communication of dispersed information within an organization. As mentioned earlier, the literature on team theory that followed from (Marschak and Radner, 1972) explores managerial decision-making when there are physical constraints on the flow of information (and the headquarters' ability to process it) but typically assumes that the workers are non-strategic in their communication (see, e.g., Cremer, 1980; Aoki, 1986; Geanakoplos and Milgrom, 1991; Bolton and Dewatripont, 1994). Several authors have subsequently analyzed strategic communication by privately informed workers and how it shapes the allocation of decision rights within organization (Dessein, 2002; Alonso, Dessein, and Matouschek, 2008; Rantakari, 2008). These papers focus on the tradeoff between the production efficiencies from coordination of actions and adaptation to local information but abstract away from the incentive problems in effort provision.

Levitt and Snyder (1997) is one of the first papers to analyze the interaction between the incentives for effort and truthful communication. They highlight a tradeoff between efficiency in decision making and effort incentives when the agents may be privately informed about their projects' viability. In order to induce both effort and truthful reporting of "bad news" (i.e., information that lowers the likelihood of the project's success), the optimal contract calls for an inefficiently "lenient" continuation policy where some projects with negative expected value are allowed to continue. But in their model the organizational structure is exogenously given; in contrast, we analyze how the interaction between the effort and reporting incentives drives the allocation of tasks within the organization.

Our paper complements the works by Athey and Roberts (2001), Friebel and Raith (2010), and Dessein, Garicano, and Gertner (2010), who explore organizational forms in the presence of the tradeoff between incentives for effort, communication, and efficient decision making. Athey and Roberts show that the tradeoff between effort incentives and efficient decision making can be mitigated by creating an organizational hierarchy by hiring a top-level manager who can obtain all information at a cost and coordinates the actions of her subordinates. However, they assume exogenous task allocation and do not allow for communication between agents. Strategic communication across organizational hierarchy plays a key role in Friebel and Raith (2010). They analyze the optimal firm structure for allocation of resources across its different divisions where the divisional managers are privately informed about the best use of such resources. The firm can integrate the units under a CEO with authority on resource allocation for more efficient allocation of resources but must elicit truthful reporting from the divisional managers. The optimality of such integration decision is driven by a tradeoff between the benefit of more efficient resource allocation and the cost of a distortion in the effort incentives that may be necessary for information elicitation. A similar integration issue is studied by Dessein, Garicano, and Gertner (2010) where a firm decides on whether to organize into business units (i.e., divisions with considerably autonomy) or create functional units that centralizes certain tasks for all divisions. The functional unit manager can implement standardization to capture synergy benefits but inflicts a cost on business unit managers by impeding adaptation to local information. The organization responds to this tradeoff by creating an incentive conflict between the business and functional unit managers and it drives the optimal allocation of authority and tasks within the organization. However, none of these papers explore the role of job design in incentivizing truthful communication within organization which is a key focus of our analysis.

This article also relates to a few other strands in the organizational economics literature. There is a vast literature on incentives in teams (Groves, 1973; Holmström, 1982; Mookherjee, 1984; McAfee and McMillan, 1991; Che and Yoo, 2001; Marino and Zábojník, 2004; Kvaløy and Olsen, 2006; Rayo, 2007; Blanes i Vidal and Möller, 2016; Friebel, Heinz, Krueger, and Zubanov, 2017) that takes the team structure as given and analyze how the underlying production and information environment drive the optimal provision of effort incentives. A notable exception is Gromb and Martimort (2007) who consider a setup where the decision-maker relies on experts to gather and report multiple signals on a risky project's profitability. They analyze a case where the decision-maker can either ask a single expert to acquire all signals or employ multiple experts where each one is responsible for acquiring exactly one signal. While this setup bears some resemblance to our job design problem, Gromb and Martimort's model differs from ours along various key dimensions. In particular, in their setup the agents' effort is useful for information acquisition but not for the project's value, experts have "soft information" (hence, can lie in their report), and the focus of their analysis is on the optimal incentives for such delegated expertise when the contracting parties may collude among themselves.

Job design has also been explored by several scholars, primarily as a possible remedy for the multitasking problem (Holmström and Milgrom, 1991; Dewatripont, Jewitt, and Tirole, 2000; Besanko, Régibeau, and Rockett, 2005; Corts, 2007; Schöttner, 2008; Mukherjee and Vasconcelos, 2011; Ishihara, 2017, 2020). In contrast, we abstract away from the multitasking problem; in our setting the conflict of incentives for effort and information elicitation is the key driver of the optimal job design. Finally, our work is reminiscent of the literature on authority and delegation where the contracting parties may have misaligned preferences over the managerial actions (Aghion and Tirole, 1997; Dessein, 2002; Alonso and Matouschek, 2008; Alonso, Dessein, and Matouschek, 2008; Deimen and Szalay, 2019). In this literature, the misalignment is assumed to stem from exogenous bias in the agents' preferences that may distort the communication within organization. However, in our setup the agents' possible gains from information manipulation arises endogenously due to the moral hazard problem in the agent's effort provision and the firm's lack of commitment power over its continuation policy.

This paper is structured as follows. Section 1.2 presents our model. A benchmark case with public signal is analyzed in Section 1.3. The optimal contracts under individual and team assignment is characterized in Section 1.4. In Section 1.5 we present our main result on the optimal job design and explore its comparative statics. A final section, Section 1.6, discusses a few extensions of our model and presents a conclusion. All proofs are given in the Appendix.

1.2 Model

Players: A principal $\mathcal{P}$ (she) hires two agents (he), $\mathcal{A}_1$ and $\mathcal{A}_2$ to work on two risky projects, A and B, and concurrently gather information on the projects' financial viability. Below we index the agents by $i \in \{1, 2\}$ and the projects by $j \in \{A, B\}$.

Technology: The production technology is reminiscent of the canonical setup of Dewatripont et al. (2000). Each project $j \in \{A, B\}$ consists of two tasks: T_{j1} and T_{j2}. To fix ideas, one may consider a firm exploring the launch of a new product, and a successful launch requires effort on product development and marketing. For notational clarity, we may refer to task T_{jk} simply as task k, $k \in \{1, 2\}$.

Each agent can perform at most two tasks. At the beginning of the game, the principal commits to a task allocation or "job design". The principal can choose one of two options: (i) "individual assignment", where each worker is assigned to a different project, and he works on the two tasks that are associated with his project, and (ii) "team assignment", where each worker performs exactly one task from each of the two projects. Without loss of generality, we assume that under individual assignment, agent $\mathcal{A}_1$ works on project A (and performs tasks $\{T_{A1}, T_{A2}\}$), and agent $\mathcal{A}_2$ works on project B (and performs tasks $\{T_{B1}, T_{B2}\}$); whereas under team assignment, $\mathcal{A}_1$ performs the first task in both jobs, $\{T_{A1}, T_{B1}\}$, and $\mathcal{A}_2$ performs the second task, $\{T_{A2}, T_{B2}\}$.

Let $e_{jk} \in [0, 1/2]$ denote the effort exerted in task T_{jk} (i.e., task $k \in \{1, 2\}$ of project $j \in \{A, B\}$). Effort is private, and it costs the agent (who has been assigned to this task) $c(e_{jk}) = e_{jk}^2/2$.

The outcome of project j, $Y_j \in \{0, y\}$, can be either a "success" ($Y_j = y$) or a "failure"

($Y_j = 0$). The project's outcome depends on the effort exerted in each of its two tasks and on its underlying "state of the world", $\omega_j \in \{G, B\}$ that can either be "good" ($\omega_j = G$) or "bad" ($\omega_j = B$). The production function is given as (denote $\mathbf{e}_j := (e_{j1}, e_{j2})$):

$$\Pr\left(Y_j = y \mid \mathbf{e}_j; \omega_j\right) = \begin{cases} e_{j1} + e_{j2} & if\ \omega_j = G \\ 0 & if\ \omega_j = B \end{cases}.$$

In a "bad" state, the project always fails regardless of the agents' effort, and yields $Y_j = 0$. In a "good" state failure can be averted as $Y_j \in \{0, y\}$, and effort is productive as it increases the chance of obtaining a high output of $Y_j = y$.

The project outcome is not verifiable, but the agent's performance is reflected by a metric $M_j \in \{0, 1\}$ that can be verified. However, the metric M_j is a noisy measure of the project outcome as:

$$\Pr\left(M_j = 1 \mid \mathbf{e}_j; \omega_j\right) = \begin{cases} e_{j1} + e_{j2} & if\ \omega_j = G \\ \mu\left(e_{j1} + e_{j2}\right) & if\ \omega_j = B \end{cases},$$

and $\mu \in [0, 1)$. In the context of the product launch example, one may consider Y_j to be the product's long-term value to the firm whereas M_j is a measure of the product's profitability in the short run. The extent of misalignment between the metric and the project output is reflected by the parameter μ; for $\mu = 0$ the distributions of Y_j and M_j are identical, but for $\mu > 0$, the metric may reflect a "success" ($M_j = 1$) even in a bad state when the project fails with certainty. And at the extreme, when $\mu \to 1$, the metric no longer depends on the underlying state.

Information structure: At the beginning of the production process, the underlying state of a project, ω_j, is unknown to all players but players hold a common prior belief given as $\Pr(\omega_j = G) = \frac{1}{2}$, where ω_A and ω_B are statistically independent. But an agent, up on completing an assigned task T_{jk}, *privately* observes the state ω_j with probability $\alpha \in [0, 1)$. Thus, under individual assignment, the agent assigned to project $j \in \{A, B\}$ learns the underlying state ω_j with probability $1 - (1-\alpha)^2$. And, under team assignment, the probability at least one of the two agents assigned to project j learns the state ω_j is also $1 - (1-\alpha)^2$. Denote $\mathcal{A}_i$'s observation on the state ω_j as $x_i^j \in \{G, B, \emptyset\}$, where $x_i^j = \emptyset$ if $\mathcal{A}_i$ does not observe ω_j.

Reporting: The agents simultaneously report their information on the underlying states to the principal. The observation on the state is "hard information": an agent cannot misreport the state but can hide his observation by feigning ignorance. Under individual assignment, denote $\mathcal{A}_i$'s report as $r_i \in \{G, B, \emptyset\}$, where $r_i = \emptyset$ when the agent claims to have failed to observe the state associated with his project. And under team assignment, $\mathcal{A}_i$ reports $r_i = \left(r_i^A, r_i^B\right)$ where $r_i^j \in \{G, B, \emptyset\}$ is the report on state ω_j, $j \in \{A, B\}$. With a slight abuse of notation, we denote the collective report of the two agents on state ω_j as $r^j \in \{G, B, \emptyset\}$ (i.e., the information on ω_j that the principal obtains from the two reports).

Given the agents' reports, the principal decides whether to implement a project or to cancel it. The project outcome Y_j and the associated performance measure M_j are realized only if the project is implemented. If a project is canceled, the principal earns her outside option, as described later in this section. The agents' reports, like the project outcomes Y_j, are not verifiable.

Contract: As mentioned above, the principal commits to a job design $d \in \{\mathcal{I}, \mathcal{T}\}$ that specifies either individual assignment ($d = \mathcal{I}$) or team assignment ($d = \mathcal{T}$). As neither the projects' outcomes nor the agents' reports are verifiable, the principal cannot commit to a cancellation policy, and can only commit to a wage schedule that depends on (i) whether the project has been implemented, and (ii) in the event the project is implemented, on the realization of the associated performance measure $M_j \in \{0,1\}$. To streamline notations, we set $M_j = \emptyset$ if project j gets canceled. Thus, under individual assignment, agent $\mathcal{A}_1$'s contract is given by the wage schedule $w_1^I(M_A)$, $M_A \in \{0,1,\emptyset\}$ as he is only responsible for project A (similarly, $w_2^I(M_B)$ for agent $\mathcal{A}_2$), and under team assignment, by the pair of schedules $\left\{w_{1A}^T(M_A); w_{1B}^T(M_B)\right\}$ as he works on parts of both projects (similarly, $\left\{w_{2A}^T(M_A); w_{2B}^T(M_B)\right\}$ for agent $\mathcal{A}_2$). Denote the wage schedule for $\mathcal{A}_i$ under the job design $d \in \{\mathcal{I}, \mathcal{T}\}$ as $\mathcal{W}_i^d$.

We denote a contract as $\phi := \left\{d, \mathcal{W}_1^d, \mathcal{W}_2^d\right\}$, and let Φ be the set of all such contracts.

Time line: The time line of the game is summarized below:

- $\mathcal{P}$ chooses a job design $d \in \{\mathcal{I}, \mathcal{T}\}$, and publicly offers a wage schedule $\left\{\mathcal{W}_1^d, \mathcal{W}_2^d\right\}$.
- $\mathcal{A}_1$ and $\mathcal{A}_2$ (simultaneously) accept or reject the contract $\phi = \left\{d, \mathcal{W}_1^d, \mathcal{W}_2^d\right\}$. The game proceeds only if both accept.
- $\mathcal{A}_i$ exerts effort in the two tasks that have been assigned to him.
- $\mathcal{A}_i$ may observe the state(s) ω_j from his assigned tasks and reports to $\mathcal{P}$.
- $\mathcal{P}$ decides which project, if any, to cancel.

- The project outcomes, performance measures, and payoffs are realized; and the game ends.

PAYOFFS: With a slight abuse of notation, we set $Y_j = \underline{\pi}$ if project j gets canceled. (Recall that in this case we also set the performance metric $M_j = \emptyset$.) Under individual assignment the agents' ex-post payoffs are:

$$\begin{aligned} u_1^I &:= \; w_1^I(M_A) - c(e_{A1}) - c(e_{A2}), \\ u_2^I &:= \; w_2^I(M_B) - c(e_{B1}) - c(e_{B2}); \end{aligned}$$

and the principal's ex-post payoff is $\pi^I := \pi_A^I + \pi_B^I$ where

$$\pi_A^I := Y_A - w_1^I(M_A), \text{ and } \pi_B^I := Y_B - w_2^I(M_B).$$

Analogously, the payoffs under team assignment are given as

$$\begin{aligned} u_1^T &:= \; w_{1A}^T(M_A) + w_{1B}^T(M_B) - c(e_{A1}) - c(e_{B1}), \\ u_2^T &:= \; w_{2A}^T(M_A) + w_{2B}^T(M_B) - c(e_{A2}) - c(e_{B2}), \end{aligned}$$

and $\pi^T := \pi_A^T + \pi_B^T$ where

$$\pi_j^T = Y_j - \left(w_{1j}^T(M_j) + w_{2j}^T(M_j)\right).$$

All players are risk neutral. If the agents accept the contract offered by the principal, the ex-ante payoff of an agent $\mathcal{A}_i$ is given by his expected wage net of his cost of effort. And the ex-ante payoff of the principal is given by the expected output from the two projects (when implemented) net of the expected wage payment. If a project is canceled, the principal can undertake an "outside option" that yields a payoff of $\underline{\pi}$ (> 0). Note that the expectations over project outcome and performance metric must account for the agents' reporting strategy and the principal's cancellation strategy (as we will elaborate below). But in our discussion below we do not explicitly mention this dependence to economize on notations.

We assume that a priori the principal is indifferent between canceling a project and implementing it without seeking any information from the agents, which implies the following restriction on the parameters.

Assumption 1. $\underline{\pi} = \max_{e_{j1}, e_{j2}} \ \frac{1}{2}\left(e_{j1} + e_{j2}\right) y - c\left(e_{j1}\right) - c\left(e_{j2}\right) = \frac{1}{4} y^2.$

We also assume that the outside option of both agents is 0.

STRATEGIES AND EQUILIBRIUM CONCEPT: The strategy of the principal, $\sigma_{\mathcal{P}}$, has two components: (i) A contract $\phi \in \Phi$ offered at the beginning of the game that stipulates the job design $d \in \{\mathcal{I}, \mathcal{T}\}$, and the agents' wage schedules given the chosen design, $\mathcal{W}_1^d$ and $\mathcal{W}_2^d$. (ii) A *continuation policy*, $\mathcal{C}_j$, that stipulates the principal's continuation decision on project j, $j \in \{A, B\}$, as a function of the agents' reports r_1 and r_2. The strategy of the agent $\mathcal{A}_i$, $\sigma_{\mathcal{A}_i}$, has three components: (i) accept or reject the contract offered by the principal, (ii) an *effort policy* $\mathcal{E}_i$ that stipulates effort levels on the assigned tasks, and (iii) a *reporting policy* ρ_i that maps the agent's observed signals to his report r_i. We use *perfect Bayesian Equilibrium* (PBE) in pure strategies as a solution concept.

As the projects are independent and the players' payoffs are additively separable across projects, without loss of generality, we limit attention to the class of equilibria where players use symmetric strategies (i.e., $\mathcal{C}_A = \mathcal{C}_B$, $\rho_1 = \rho_2$, $w_1^I(M_A) = w_2^I(M_B)$, and $w_{iA}^T(M_A) = w_{iB}^T(M_B)$, $i = 1, 2$). We look for the PBE that yields the highest payoff to the principal in each of the two continuation games that follows from a given job design $d \in \{\mathcal{I}, \mathcal{T}\}$. The optimal job design d is the one that yields the highest payoff to the principal.

1.3 A Public Information Benchmark

We begin our analysis by considering a benchmark case where the agents' observations on the state(s) are *publicly verifiable* information. Thus, the principal does not need to elicit any information from the agents on the projects' viability, and she can also commit at the outset to a cancellation policy that depends on the observed state. This case serves as an useful benchmark for the exploration of the optimal job design in our model: it highlights how the principal's need for information elicitation and her lack of commitment power on continuation decisions drive the key trade-off between individual and team assignment. [1]

As in our main model, denote $x^j \in \{G, B, \emptyset\}$ as the information on the state ω_j observed by the agent(s) assigned to project j ($x^j = \emptyset$ if neither of the two agents observes ω_j), but now assume that x^j is publicly observed. Suppose that the principal opts for individual assignment ($d = \mathcal{I}$), commits to proceed with project j if and only if $x^j \in X_P^j \subseteq \{G, B, \emptyset\}$, and offers the agents a wage schedule $\left\{\mathcal{W}_1^{\mathcal{I}}, \mathcal{W}_2^{\mathcal{I}}\right\}$.

In the continuation game that follows, the agent $\mathcal{A}_i$'s expected payoff from exerting effort

[1] The class of wage contracts in this benchmark case is assumed to be the same as the one defined in the main model. Even though the wage payments could be tied to the agents' observed state (when the observations are publicly verifiable), as we will explain below, the principal does not benefit from doing so.

$\mathbf{e}'_j := \left(e'_{j1}, e'_{j2}\right)$ is:

$$U_i^I\left(\mathbf{e}'_j, X_P^j\right) := \Pr\left(x^j \in X_P^j\right) \sum_{M_j \in \{0,1\}} w_i^I\left(M_j\right) \left[\sum_{\omega_j \in \{G,B\}} \Pr\left(M_j \mid \mathbf{e}'_j, \omega_j\right) \Pr\left(\omega_j \mid x^j \in X_P^j\right)\right]$$
$$+ \Pr\left(x^j \notin X_P^j\right) w_i^I\left(\emptyset\right) - \sum_{k \in \{1,2\}} c\left(e'_{jk}\right). \tag{1.1}$$

That is, with probability $\Pr\left(x^j \in X_P^j\right)$ the project continues, and agent $\mathcal{A}_i$ earns his expected wage conditional on the event that the observation on the underlying state is in X_P^j. Otherwise, the project is canceled, and the agent earns his "cancellation wage" $w_i^I\left(\emptyset\right)$. Notice that the agent incurs the cost of his effort regardless of the principal's decision on the project's implementation.

If the effort profile $\mathbf{e}_j$ is supported in equilibrium, it must satisfy $\mathcal{A}_i$'s incentive compatibility constraint:

$$\mathbf{e}_j = \arg\max_{e'_{j1}, e'_{j2}} \; U_i^I\left(\mathbf{e}'_j, X_P^j\right) \quad \forall\, j, \tag{IC_I}$$

and his participation constraint:

$$U_i^I\left(\mathbf{e}_j, X_P^j\right) \geq 0. \tag{IR_I}$$

Also, the principal's expected payoff under the effort profiles $\{\mathbf{e}_A, \mathbf{e}_B\}$ is (recall that we set $Y_j = \underline{\pi}$ if project j gets canceled):

$$\Pi^I := \mathbb{E}\left[Y_A - w_1^I(M_A \mid \mathbf{e}_A,\; X_P^A\right] + \mathbb{E}\left[Y_B - w_2^I\left(M_B\right) \;\mid \mathbf{e}_B,\; X_P^B\right].$$

The optimal contract stipulates the wage schedule and continuation policy (given by the sets X_P^j) that maximize Π^I subject to (IR_I) and (IC_I).

Next, consider the case where the principal opts for team assignment ($d = \mathcal{T}$) and offers a wage schedule $\left\{\mathcal{W}_1^{\mathcal{T}}, \mathcal{W}_2^{\mathcal{T}}\right\}$. In the continuation game that follows, the agents' subsequent effort choices constitute a Nash Equilibrium. Thus, if the contract induces the agent $\mathcal{A}_i$ to exert an effort profit $\mathbf{e}_i := (e_{Ai}, e_{Bi})$, it must be a best response to the other agent $\mathcal{A}_{-i}$'s effort level $\mathbf{e}_{-i}$.

Analogous to $U_i^I\left(\mathbf{e}'_j, X_P^j\right)$, denote $\mathcal{A}_i$'s expected payoff under team assignment as $U_i^T\left(\mathbf{e}'_i, \mathbf{e}_{-i}, X_P^A, X_P^B\right)$. The agent's incentive compatibility constraint parallels its counterpart under individual assignment, and can be written as:

$$\mathbf{e}_i = \arg\max_{\mathbf{e}'_i} \; U_i^T\left(\mathbf{e}'_i, \mathbf{e}_{-i}, X_P^A, X_P^B\right) \quad \forall i. \tag{IC_T}$$

Also, $\mathcal{A}_i$'s participation constraint requires:

$$U_i^T\left(\mathbf{e}_i, \mathbf{e}_{-i}, X_P^A, X_P^B\right) \geq 0 \;\; \forall i. \tag{IR_T}$$

Thus, the optimal contract stipulates the wage scheme and continuation policy (given by the sets X_P^j) that maximize the principal's expected payoff

$$\Pi^T := \sum_{j \in \{A,B\}} \mathbb{E}\left[Y_j - \left(w_1^T(M_j) + w_2^T(M_j)\right) \mid \mathbf{e}_1, \mathbf{e}_2, X_P^A, X_P^B\right],$$

subject to (IR_T) and (IC_T).

Proposition 1. *Under both individual and team assignment, in the optimal contract the principal proceeds with project j if and only if the bad state is not observed (i.e., $x^j \in \{G, \emptyset\}$) and obtains a payoff*

$$S^* := \left(1 + \alpha - \frac{1}{2}\alpha^2\right) \underline{\pi}.$$

That is, in the benchmark case, job design does not affect the principal's payoff under the optimal contract.

The above finding shows that the choice of job design is irrelevant when the agents' information is public. Regardless of job design, the principal can always commit to the optimal continuation policy, and use the wage contract to induce first-best effort while extracting all surplus from the agent. Thus, the issue of job design becomes relevant only when the agents' observations on the projects' underlying state remain private (as the agents' reports are non-contractible, the principal can no longer commit to her continuation policy).

1.4 Optimal Contract

In this section we explore how the principal's need for information elicitation while being unable to commit to her continuation policy shapes the choice between team and individual assignment. In contrast to the benchmark case, when the agents are privately informed, the wage contract not only affects the agents' effort but it also interferes with their incentives to reveal information as well as the principal's incentive to continue with the project. The analysis below highlights how the optimal job design is driven by such intertwined incentives.

1.4.1 Optimal Contract under Individual Assignment

We begin our analysis with the case of individual assignment. That is, we assume that the principal chooses $d = \mathcal{I}$, and in the continuation game we solve for the PBE that yields the highest payoff to the principal. But before we present the formal analysis, it is instructive to describe our solution method. Since we are looking for symmetric equilibria, we only focus on agent $\mathcal{A}_1$ who performs all tasks that are associated with project A. Also, to streamline notations, we drop the agent and project indices.

Our goal is to find the PBE with the largest ex-ante payoff for the principal, and we proceed in two steps: First, we fix a reporting and continuation policy pair $(\rho, \mathcal{C})$, i.e., a "communication protocol," and search for the optimal wage contract $\mathcal{W}$ and effort policy $\mathcal{E}$ such that the tuple $(\mathcal{W}, \mathcal{E}, \rho, \mathcal{C})$ can be supported in a PBE. Next, we compare the payoffs of the principal obtained in the first step across all possible communication protocols.

Lemma 1. *Without loss of generality, we can restrict attention to the following two communication protocols: (i) if the state is observed to be G, report G, otherwise report $\emptyset$; proceed with the project if and only if $r = G$, and (ii) if the state is observed to be B, report B, otherwise report $\emptyset$; proceed with the project if and only if $r \neq B$.*

Lemma 1 implies that we only have to consider two classes of PBE: one where the project proceeds if and only if there is "good news" , i.e., the agent's observation is $x \in X_P = \{G\}$, and another where the project proceeds if and only if there is "no bad news", i.e., the agent's observation is $x \in X_P = \{G, \emptyset\}$. Thus, without loss of generality, the communication protocols that are relevant for our analysis can be summarized by the set

$X_P \in \{\{G\}, \{G, \emptyset\}\}$. Also, for brevity of notation, we can denote $w_1^I(0) =: w_F$ (wage when the performance metric indicates "failure"), $w_1^I(\emptyset) - w_1^I(0) =: \Delta_C$ (wage premium for cancellation), and $w_1^I(1) - w_1^I(0) =: \Delta_S$ (wage premium for success).

Given a wage contract $\{w_F, \Delta_C, \Delta_S\}$, effort levels e_1 and e_2, and X_P (i.e., the set of agent's observation under which the project proceeds), the firm's ex-ante payoff is:

$$\Pi^I := \Pr(x \in X_P)\left[\Pr(\omega = G \mid x \in X_P)(y - \Delta_S) + \Pr(\omega = B \mid x \in X_P)(-\mu\Delta_S)\right]\sum_k e_k$$
$$+ \Pr(x \notin X_P)\left[\underline{\pi} - \Delta_C\right] - w_F.$$

If the project proceeds, it yields a revenue ($y = Y$) only when the state is good, but the wage premium for success may be paid even if the state is bad (as the performance measure is not perfectly aligned with the project's outcome). And if the project is canceled, the principal gets her outside option and pays the wage premium for cancellation. The agent's ex-ante payoff can be written analogously as:

$$U^I := \Pr(x \in X_P)\left[\Pr(\omega = G \mid x \in X_P) + \mu \Pr(\omega = B \mid x \in X_P)\right]\Delta_S \sum_k e_k$$
$$+ \Pr(x \notin X_P)\Delta_C + w_F - \tfrac{1}{2}\sum_k e_k^2.$$

Now, if the tuple $(w_F, \Delta_C, \Delta_S; e_1, e_2; X_P)$ is supported as a PBE, the following constraints must be met. First, for each of the two communication protocols given in Lemma 1, the principal's decision must be sequentially rational. In other words, if the principal believes that the agent's signal x is in X_P (given the agent's report), it must be more profitable for her to proceed with the project than to cancel it. Similarly, if the principal believes that the agent's signal is not in X_P, it must be more profitable for her to cancel the project than to proceed with it. Therefore, the principal's incentive compatibility constraints require:

$$\left[\Pr\left(\omega = G \mid x \in X_P\right)\left(y - \Delta_S\right) - \mu \Pr\left(\omega = B \mid x \in X_P\right)\Delta_S\right]\sum_k e_k \geq \underline{\pi} - \Delta_C, \quad (IC_P^I\text{-}1)$$

and

$$\left[\Pr\left(\omega = G \mid x \notin X_P\right)\left(y - \Delta_S\right) - \mu \Pr\left(\omega = B \mid x \notin X_P\right)\Delta_S\right]\sum_k e_k \leq \underline{\pi} - \Delta_C. \quad (IC_P^I\text{-}2)$$

Next, we have the agent's participation constraint:

$$U^I \geq 0. \qquad (IR_I)$$

Finally, consider the agent's incentive compatibility constraint. Let $U\left(e_1', e_2'; \rho'\right)$ be the agent's payoff given his efforts e_1', e_2', and reporting policy ρ' (fixing the wage contract and the principal's continuation policy). The agent's on-path payoff U^I must be the largest payoff attainable for any feasible choice of effort profile and reporting policy. So, we require:

$$U^I = \max_{e_1', e_2', \rho'} U\left(e_1', e_2'; \rho'\right). \qquad (1.2)$$

Stipulating (1.2) is equivalent to imposing the following two constraints: First, a standard incentive compatibility constraint that requires the effort levels to be optimal for the agent given his equilibrium reporting strategy (as per the communication protocol $(\rho, \mathcal{C})$); i.e.,

$$(e_1, e_2) = \arg\max_{e_1', e_2'} U\left(e_1', e_2'; \rho\right). \tag{1.2a}$$

Second, the agent may not gain from a "double deviation" either where he simultaneously deviates on his effort levels and his reporting strategy. Now, given a communication protocol $(\rho, \mathcal{C})$, if the agent can profitably deviate to some other reporting policy ρ' it must be that his report changes the principal's decision on whether to proceed with the project (under the continuation policy $\mathcal{C}$). Consider the two communication protocols mentioned in Lemma 1. In the first one the associated reporting policy is to report $x = G$ truthfully and report $\emptyset$ if $x \in \{\emptyset, B\}$; in the second one the agent reports $x = B$ truthfully and reports $\emptyset$ if $x \in \{G, \emptyset\}$. So, in the first case the only relevant deviation for the agent is to conceal information when $x = G$, and in the second case it is to conceal the information when $x = B$. Thus, in both of these cases, it is sufficient to consider only one type of deviation: the agent reports $\emptyset$ regardless of his observation. We denote this reporting policy as $\rho_\emptyset$. Hence, we must have:

$$U^I \geq \max_{e_1', e_2'} U\left(e_1', e_2'; \rho_\emptyset\right). \tag{1.2b}$$

It is instructive to elaborate on the conditions (1.2a) and (1.2b) as they, along with the principal's incentive constraints, illustrate the key trade-offs associated with information elicitation.

Consider a communication protocol from those specified in Lemma 1, and suppose that the project proceeds if $x \in X_P$, ($X_P \in \{\{G\}, \{G, \emptyset\}\}$). Regarding condition (1.2a), it is routine to check that U is concave in effort for any wage contract and communication protocol, and hence, the condition can be replaced by its associated first-order condition:

$$e_i = \Pr(x \in X_P)\left[\Pr(\omega = G \mid x \in X_P) + \mu \Pr(\omega = B \mid x \in X_P)\right]\Delta_S. \qquad (IC_A^I\text{-}1)$$

The condition (1.2b), however, is slightly more intricate. In order to simplify this condition one needs to account for the fact that when the agent deviates from his equilibrium reporting policy ρ to $\rho_\emptyset$ (i.e., reports $\emptyset$ regardless of his observation), it affects the project's continuation probability. And in case the project continues, the likelihood of a state ω conditional on the project being continued is the same as its prior probability as the project would continue regardless of the agent's observed signal x.

Let $p_\emptyset^I$ be the probability that the project continues when the agent deviates to the reporting policy $\rho_\emptyset$ given the equilibrium communication protocol, i.e., $p_\emptyset^I = 1$ if $X_P = \{G, \emptyset\}$ and $p_\emptyset^I = 0$ if $X_P = \{G\}$. Also, for brevity of notation, denote $p^I := \Pr(x \in X_P)$, and let

$$\begin{aligned} P^I &:= \Pr(\omega = G \mid x \in X_P) + \mu \Pr(\omega = B \mid x \in X_P), \\ P_\emptyset^I &:= \Pr(\omega = G) + \mu \Pr(\omega = B). \end{aligned}$$

Now, off-path, the agent's payoff can be derived as:

$$\begin{aligned}\max_{e_1',e_2'} U\left(e_1',e_2';\rho_\emptyset\right) = & \max_{e_1',e_2'} \; p_\emptyset^I \left[\Pr\left(\omega=G\right)+\mu \Pr\left(\omega=B\right)\right]\Delta_S \sum_k e_k' - \frac{1}{2}\sum_k e_k'^2 \\ & + \left(1-p_\emptyset^I\right)\Delta_C + w_F \\ = & \left(p_\emptyset^I P_\emptyset^I \Delta_S\right)^2 + \left(1-p_\emptyset^I\right)\Delta_C + w_F.\end{aligned}$$

The agent's on-path payoff can be computed analogously, and (1.2b) simplifies to:

$$\left[\left(p^I P^I\right)^2 - \left(p_\emptyset^I P_\emptyset^I\right)^2\right]\Delta_S^2 \geq \left(p^I - p_\emptyset^I\right)\Delta_C. \tag{IC_A^I-2}$$

Thus, the optimal wage contract that supports a communication protocol given by $X_P \in \{\{G\},\{G,\emptyset\}\}$ solves the following program:

$$\mathcal{P}^I : \max_{\substack{w_F,\Delta_C,\Delta_S,\\ e_1,e_2}} \Pi^I \;\; s.t. \; (IR^I), \; \left(IC_P^I\text{-}1\right), \; \left(IC_P^I\text{-}2\right), \; \left(IC_A^I\text{-}1\right), \text{ and } \left(IC_A^I\text{-}2\right).$$

Lemma 2. *The program $\mathcal{P}^I$ always admits a solution for $X_P = \{G,\emptyset\}$, and admits a solution for $X_P = \{G\}$ if and only if α is sufficiently large.*

The PBE that yields the highest payoff to the principal (under individual assignment) induces the communication protocol (given by $X_P \in \{\{G\},\{G,\emptyset\}\}$) for which the value of the program $\mathcal{P}^I$ is the largest.

1.4.2 Optimal Contract under Team Assignment

The analysis of team assignment resembles our above discussion on individual assignment, but the two forms of job design differ in two key aspects: First, under team assignment each agent gets exactly one signal from each job. In particular, both agents may observe the underlying state associated with a job. Thus, an agent cannot fully control the flow of information about a given project as his attempt to hide information would fail if the other agent happens to reveal it. Second, for each of the two projects, both agents must be (individually) incentivized for information elicitation and effort provision. (In contrast, under individual assignment the principal has to incentivize only one agent for each project; the agent is responsible for both tasks associated with the project and observes both signals on the project's underlying state). As we will explain later, these two distinctions give rise to the key trade-off between ease of information elicitation and economies of scope in incentive provision that drives the optimal job design.

Now, consider the principal's optimal contracting problem. As mentioned in the previous section, since the production environment and the wage schemes are both additively separable across projects, without loss of generality, we may require $w_{iA}^{T}(M_A) = w_{iB}^{T}(M_B)$. Consequently, we can formulate the principal's optimal contracting problem as one where there is only one project (with two tasks) and the principal hires two agents: each agent performs exactly one of the two tasks and observes exactly one of the two signals on the project's state.

Analogous to the case of individual assignment, we seek to characterize the PBE of this continuation game with the largest ex-ante payoff for the principal. The analysis follows the same two-step process that we have described above: first, we fix a communication protocol and derive the optimal wage contract that supports this protocol in equilibrium; and next, we compare the principal's payoff across all possible communication protocols that could be sustained in equilibrium.

With a slight abuse of notation, we continue to denote the strategies of the players in this game by the tuple $(\mathcal{W}_i, \mathcal{E}_i, \rho_i, \mathcal{C})$, $i = 1, 2$. To streamline notation, we drop the project index and relabel $w_{ij}^T(0) =: w_{iF}$, $w_{ij}^T(\emptyset) - w_{ij}^T(0) =: \Delta_{iC}$ (wage premium for cancellation), and $w_{ij}^T(1) - w_{ij}^T(0) =: \Delta_{iS}$ (wage premium for success). Also, we denote the team's collective observation on the state as x^T, where

$$x^T := \begin{cases} G & if\ x_i = G \text{ for some } i \\ B & if\ x_i = B \text{ for some } i \\ \emptyset & if\ x_1 = x_2 = \emptyset \end{cases} .$$

As in the case of individual assignment, we can again limit attention to only two communication protocols as stated in the lemma below. (We omit the proof of this lemma as it follows the same argument as that of Lemma 1.)

Lemma 3. *Without loss of generality, we can restrict attention to the following two communication protocols: (i) reporting policy for agent $\mathcal{A}_i$ ($i = 1, 2$): if the state is observed to be G, report G, otherwise report $\emptyset$; principal proceeds with the project only if $r_i = G$ for some i, and (ii) reporting policy for agent $\mathcal{A}_i$ ($i = 1, 2$): if state is observed to be B, report B, otherwise report $\emptyset$; principal proceeds only if $r_i \neq B$ for all i.*

Thus, without loss of generality, as before, the communication protocols that are relevant for our analysis of team assignment can be summarized by the set $X_P \in \{\{G\}, \{G, \emptyset\}\}$. Given the wage contracts $\{w_{iF}, \Delta_{iC}, \Delta_{iS}\}$, $i = 1, 2$, effort levels e_1 and e_2, and X_P, it is routine to check that the firm's ex-ante payoff is:

$$
\begin{aligned}
\Pi^T \;:=\; & \Pr\left(x^T \in X_P\right) \times \\
& \left[\Pr\left(\omega = G \mid x^T \in X_P\right)\left(y - \sum_i \Delta_{iS}\right) + \Pr\left(\omega = B \mid x^T \in X_P\right)\left(-\mu \sum_i \Delta_{iS}\right)\right] \sum_k e_k \\
& \qquad + \Pr\left(x^T \notin X_P\right)\left[\underline{\pi} - \sum_i \Delta_{iC}\right] - \sum_i w_{iF}.
\end{aligned}
$$

The agent i's participation constraint requires:

$$
\begin{aligned}
U_i^T \;:=\; & \Pr\left(x^T \in X_P\right) \times \\
& \left[\Pr\left(\omega = G \mid x^T \in X_P\right) + \mu \Pr\left(\omega = B \mid x^T \in X_P\right)\right] \Delta_{iS} \sum_k e_k \\
& \qquad + \Pr\left(x^T \notin X_P\right) \Delta_{iC} + w_{iF} - \tfrac{1}{2} e_i^2 \quad \geq 0.
\end{aligned}
\tag{IR_i^T}
$$

The principal's incentive compatibility constraints ensure that is it optimal for the principal to proceed with the project if x^T is in X_P and to cancel it otherwise:

$$
\begin{aligned}
& \left[\Pr\left(\omega = G \mid x^T \in X_P\right)\left(y - \sum_i \Delta_{iS}\right)\right. \\
& \qquad \left. + \Pr\left(\omega = B \mid x^T \in X_P\right)\left(-\mu \sum_i \Delta_{iS}\right)\right] \sum_k e_k \geq \underline{\pi} - \sum_i \Delta_{iC},
\end{aligned}
\tag{IC_P^T-1}
$$

and

$$
\begin{aligned}
& \left[\Pr\left(\omega = G \mid x^T \notin X_P\right)\left(y - \sum_i \Delta_{iS}\right)\right. \\
& \qquad \left. + \Pr\left(\omega = B \mid x^T \notin X_P\right)\left(-\mu \sum_i \Delta_{iS}\right)\right] \sum_k e_k \leq \underline{\pi} - \sum_i \Delta_{iC},
\end{aligned}
\tag{IC_P^T-2}
$$

Notice that in contrast to its counterpart under individual assignment, the (IC_P) constraints highlight that the project's success and cancellation both would require the principal to pay the corresponding wage premium to both of the two agents. As we will see later, the need for such "double payment" captures diseconomies of scope in incentive provision under team assignment.

Finally, consider the agents' incentive compatibility constraints. As before, the constraint would require that neither of the two agents can gain by unilaterally deviating to a different effort choice and reporting policy. However, there is a salient distinction between the constraints under team and their counterpart under individual assignment. Under team assignment, an agent chooses the effort in only one of the two tasks, and reports only one of the two signals on the project's underlying state. Thus, an agent cannot fully influence the project's output and the associated performance measure, nor he can fully control the information on the underlying state that may be communicated to the principal.

Let $U_i\left(e_i, \rho_i; e_j, \rho_j\right)$ be the agent $\mathcal{A}_i$'s payoff given the two agents' efforts and reporting policies (fixing the wage contracts and the principal's continuation policy). The agent's on-path payoff U_i^T must be the largest payoff attainable for any feasible choice of effort profile and reporting policy (given the other agent's equilibrium effort and reporting policy). So, the constraint requires:

$$U_i^T = \max_{e_i', \rho_i'} \; U_i\left(e_i', \rho_i'; e_j, \rho_j\right). \tag{1.3}$$

As before, it is sufficient to consider only two types of deviation: (i) the agent follows his equilibrium reporting policy ρ_i but deviates on his effort level, (ii) the agent reports $\emptyset$ regardless of his observation, and chooses his effort level accordingly. Again, with a slight abuse of notation, we denote the latter reporting policy (given in (ii)) as $\rho_{\emptyset}$. Thus, the incentive compatibility constraint (1.3) for agent $\mathcal{A}_i$ ($i = 1, 2$) is equivalent to the following two conditions:

$$e_i = \arg\max_{e_i'} U_i\left(e_i', \rho_i; e_j, \rho_j\right) \tag{1.3a}$$

and

$$U_i^T \geq \max_{e_i'} \; U_i\left(e_i', \rho_\emptyset; e_j, \rho_j\right). \tag{1.3b}$$

Now, (1.3a) implies that e_i satisfies the following first-order condition ($i = 1, 2$):

$$e_i = \Pr\left(x^T \in X_P\right)\left[\Pr\left(\omega = G \mid x^T \in X_P\right) + \mu \Pr\left(\omega = B \mid x^T \in X_P\right)\right]\Delta_{iS}. \tag{$IC_{A_i}^T$-1}$$

Also, (1.3b) can be simplified in the same fashion in which we streamlined its counterpart under individual assignment. However, one needs to account for the fact that under team, an agent's attempt to conceal information may be undermined by the report of the other agent. In parallel to our analysis of individual assignment, let $p_\emptyset^T$ be the probability that the project continues when agent i deviates to the reporting policy $\rho_\emptyset$, given the equilibrium communication protocol. Also, denote $p^T := \Pr\left(x^T \in X_P\right)$, and

$$\begin{aligned} P^T &:= \Pr\left(\omega = G \mid x^T \in X_P\right) + \mu \Pr\left(\omega = B \mid x^T \in X_P\right), \\ P_\emptyset^T &:= \Pr\left(\omega = G \mid \rho_\emptyset, \rho_j, \mathcal{C}\right) + \mu \Pr\left(\omega = B \mid \rho_\emptyset, \rho_j, \mathcal{C}\right), \end{aligned}$$

where $\Pr\left(\omega \mid \rho_\emptyset, \rho_j, \mathcal{C}\right)$ denotes the probability of the state ω conditional on the event that the project proceeds under the communication protocol $\{\rho_\emptyset, \rho_j, \mathcal{C}\}$. Now, plugging in the agent's on- and off-path payoffs, condition (1.3b) can be stated as:

$$\begin{array}{c}\frac{1}{2}\left[\left(p^T P^T\right)^2 - \left(p_\emptyset^T P_\emptyset^T\right)^2\right]\Delta_{iS}^2 + \left[\left(p^T P^T\right)^2 - \left(p_\emptyset^T P_\emptyset^T\right)\left(p^T P^T\right)\right]\Delta_{iS}\Delta_{jS} \\ \geq \left(p^T - p_\emptyset^T\right)\Delta_{iC}.\end{array} \qquad (IC_{A_i}^T\text{-}2)$$

Thus, the optimal wage contract under team assignment that supports a communication protocol given by $X_P \in \{\{G\}, \{G, \emptyset\}\}$ solves the following program:

$$\mathcal{P}^T : \max_{\substack{\{w_{iF}, \Delta_{iC}, \Delta_{iS}\}_{i=1,2} \\ e_1, e_2}} \Pi^T \;\; s.t. \;\left(IR_i^T\right),\; \left(IC_P^T\text{-}1\right),\; \left(IC_P^T\text{-}2\right),\; \left(IC_{A_i}^T\text{-}1\right), \text{ and } \left(IC_{A_i}^T\text{-}2\right).$$

Lemma 4. *(i) The program $\mathcal{P}^T$ always admits a solution for $X_P = \{G, \emptyset\}$ and admits a solution for $X_P = \{G\}$ if and only if both α and μ are sufficiently large.*

(ii) If $\mathcal{P}^T$ admits a solution, it also admits a symmetric solution where $w_{1F} = w_{2F} = w_F$, $\Delta_{1S} = \Delta_{2S} = \Delta_S$ and $\Delta_{1C} = \Delta_{2C} = \Delta_C$.

The PBE that yields the highest payoff to the principal (under team assignment) induces the communication protocol (given by $X_P \in \{\{G\}, \{G, \emptyset\}\}$) for which the value of the program $\mathcal{P}^T$ is the largest.

1.5 Optimal Job Design

By comparing the principal's payoffs associated with the optimal contracts under team and individual accountability, we can now characterize the optimal job design.

Proposition 2. ***(Optimal job design)*** *There exist two thresholds μ_0 and μ_1 (given α), $\mu_0 < \mu_1$, such that:*

(i) if $\mu < \mu_0$, it is optimal to choose individual assignment where the agent reports B only if he observes the state to be B, and reports $\emptyset$ otherwise; the principal proceeds with the project only if the report is not B. The associated optimal contract is efficient and the principal's payoff is S^* (as defined in Proposition 1).

(ii) If $\mu > \mu_1$, it is optimal to choose team assignment where the agent reports B only if he observes the state to be B, and reports $\emptyset$ otherwise; the principal proceeds with the project only if no agent reports B. The associated optimal contract is efficient and the principal's payoff is S^*.

(iii) Otherwise, ($\mu_0 \leq \mu \leq \mu_1$) the principal is indifferent between team and individual assignments: both designs, along with the corresponding communication protocol as stated in parts (i) and (ii) above, yield the same payoff of S^* for the principal.

Moreover, the parameter thresholds μ_0 and μ_1 vary with α in the following manner.

Proposition 3. ***(Comparative statics)*** *The threshold μ_0 is increasing in α. Also, there exists a cutoff α^* such that $\mu_1 = 1$ for $\alpha \leq \alpha^*$ and μ_1 is decreasing in α for $\alpha \geq \alpha^*$.*

Propositions 2 and 3 (illustrated in Figure 1.1) show how the optimal job design is driven by the "availability" of the agents' signal (as captured by α) and the "alignment" of the performance measure with the project's output (as captured by μ). For low α (i.e., $\alpha \leq \alpha^*$), individual assignment is always optimal; for low μ (i.e., $\mu < \mu_0$) it strictly dominates team assignment but otherwise (i.e., $\mu \geq \mu_0$) both designs yield the same (optimal) payoff. In contrast, when α is large, team assignment is strictly optimal provided μ is large as well (i.e.,

$\mu > \mu_1$). However, as before, for moderate μ the two designs yield the same payoff, and for small μ individual assignment remains strictly optimal.

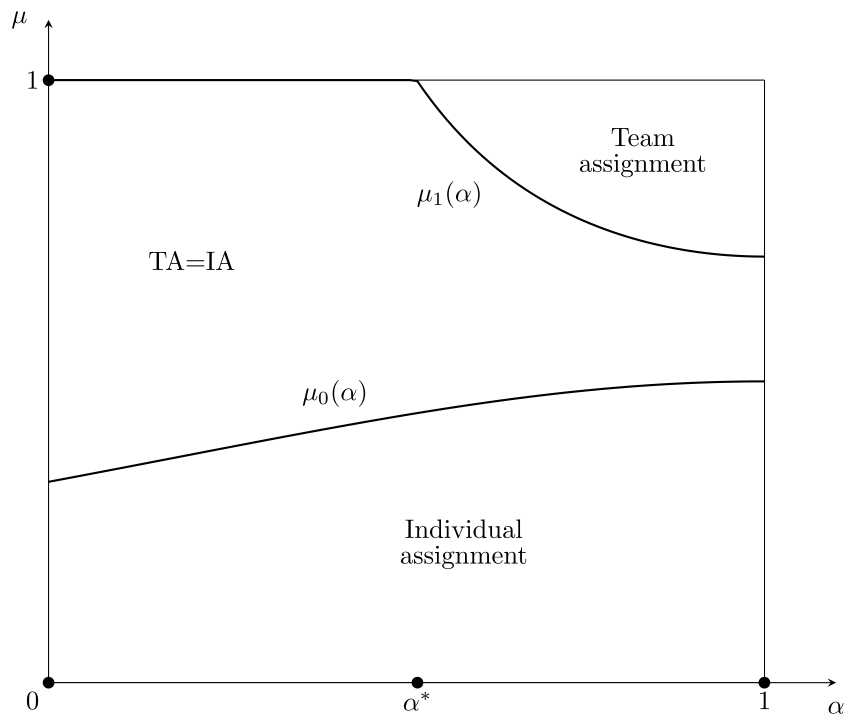


Figure 1.1: Optimal job design as a function of α and μ

To see the intuition behind the above result, recall that our setup highlights two key frictions. First, the principal lacks information on the project's viability and must elicit it from the agents. Second, even though the principal's continuation decision depends on the agents' information, she cannot commit to any continuation policy ex-ante. These two frictions give rise to a trade-off that drives the optimal job design: relative to individual assignment, team facilitates information elicitation but suffers from diseconomies of scope in incentive provision.

Team assignment helps in information elicitation as an agent cannot fully control the outcome of the project (and the performance measure). Even if the agent attempts to suppress information and adjust his effort (in his assigned task) accordingly, his gains from such deviations are muted by the fact that his teammate may still reveal the information to the principal. Also, the agent cannot control the level of effort on the task that is performed by his teammate. But, under individual assignment such a "double deviation", i.e., concurrent manipulation of reporting and effort, may be more profitable for the agent: he fully controls what the principal gets to learn about the project's underlying state and how much effort is exerted on both tasks that are associated with the project. In fact, he stands to profit from it when both α and μ are large.

When α is large, the agent's control over the project's continuation is more valuable as he is now more likely to observe the state and, under individual assignment, he can hide any unfavorable information. In particular, the agent would have a strong incentive to conceal the bad state (and let the project continue) if he expects to earn a large payoff even if the project fails. This is indeed the case when μ is large, i.e., the performance measure is significantly misaligned with the project's outcome: in a bad state, the measure is more likely to indicate success (given the effort levels) even though the project is sure to fail. Moreover, should the agent deviate on his reporting policy and hide the bad state, he may also exert more effort (vis-a-vis the on-path effort levels) so as to further increase his gains from deviation. Thus, when α and μ are both large, deterring the agent from double-deviation gets harder under individual assignment, and team's advantage over individual assignment in information elicitation becomes stronger. This is why team assignment dominates individual assignment when α and μ are high.

However, team assignment lacks economies of scope in incentive provision: in order to induce effort on both tasks associated with the project, the principal needs to incentivize the two agents separately. Notice that under individual assignment a single wage payment (w_F, w_S, or w_C based on the project's outcome) incentivizes the agent to exert efforts on all tasks. In contrast, in a team, each of the two agents are assigned to exactly one of the two tasks. Hence, if the principal were to induce the same level of effort in both tasks of the project her wage bill doubles ($2w_F$, $2w_S$, or $2w_C$).

Such *dis*economies of scope may be costly to the principal. As the principal lacks commitment power over the continuation policy, her (IC_P) constraints must hold. That is, for any given job design with communication protocol given by X_P, (i) the principal's expected payoff from proceeding when the agents' observation is in X_P must be larger than her payoff from canceling the project, and (ii) the payoff from canceling must be larger than her expected payoff from proceeding with the project if the agents' observation is not in X_P. Thus, any feasible contract must ensure that the principal earns more from proceeding when the signal is in X_P than when it is not. For example, $\left(IC_P^I\text{-}1\right)$ and $\left(IC_P^I\text{-}2\right)$ imply:

$$\begin{gathered}\left[\Pr\left(\omega=G\mid x\in X_P\right)\left(Y-\Delta_S\right)+\Pr\left(\omega=B\mid x\in X_P\right)\left(-\mu\Delta_S\right)\right]\sum_k e_k \geq \\ \left[\Pr\left(\omega=G\mid x\notin X_P\right)\left(Y-\Delta_S\right)+\Pr\left(\omega=B\mid x\notin X_P\right)\left(-\mu\Delta_S\right)\right]\sum_k e_k .\end{gathered}$$

This difference in earnings is given by the difference in the expected output of the project

$$\left[\left[\Pr\left(\omega=G\mid x\in X_P\right)-\Pr\left(\omega=G\mid x\notin X_P\right)\right]\sum_k e_k\right]Y,$$

and the difference in the expected wage payout

$$\left[(1-\mu)\left[\Pr\left(\omega=G \mid x\in X_P\right)-\Pr\left(\omega=G \mid x\notin X_P\right)\right]\sum_k e_k\right]\Delta_S.$$

Now, for any $X_P \in \{\{G\}, \{G, \emptyset\}\}$ the probabilities that the project and the performance measure indicate success (i.e., $y = Y$ and $M = 1$) are both larger (given the effort levels in the two tasks) when the agents' signal is in X_P than when it is not. Therefore, when the principal needs to pay the wage premium for success (Δ_S) twice in order to elicit the same amount of effort in both tasks—as is the case under team assignment—the difference in her expected wage payouts is larger. Consequently, the aforementioned feasibility constraint is harder to satisfy under team, and individual assignment becomes more favorable.

Also note that team's relative disadvantage (due to diseconomies of scope) becomes more acute when μ is small (i.e., the measure is well-aligned with the project's output). As the agent is unlikely to earn a reward for success when the state is bad, the wage premium for success needs to be sufficiently large so as to incentivize him to exert effort. And when the principal needs to pay such large premiums *twice*—as is the case under team assignment—her continuation policy is less likely to remain credible: proceeding with the project when the signal is in X_P may be less profitable than proceeding when it is not (i.e., (IC_P) gets violated). This explains why individual accountability dominates team when μ is low.

The above discussion may be summarized as follows: For low μ, provision of incentives under team assignment gets compromised due to acute diseconomies of scope, but incentives under individual assignment remain sharp as information elicitation is relatively easy ("double deviation" is less profitable as a successful performance is unlikely to arise when the state is bad). Thus, individual assignment strictly dominates team. However, for large

μ, diseconomies of scope does not distort incentive provision under team: as the required success premium is smaller, it may be feasible for the principal to pay it to both agents separately. Thus, both designs yield the same payoff as long as information elicitation does not distort incentives under individual assignment. But information elicitation gets harder under individual assignment if α is also large (along with μ), and team assignment becomes strictly optimal.

Notice that at the optimal job design diseconomies of scope does not distort incentives, and neither does the need for information elicitation. Therefore, the associated contract yields the efficient level of surplus as obtained in the public information benchmark (in Section 1.3). However, this observation critically hinges on our modeling assumption that the agents' observation on the state does not contain any noise (conditional on observing it in the first place). As we discuss in the next section, when the agent's signal is noisy, the optimal job design may entail inefficiencies both in the principal's continuation policy and in the agents' effort levels.

1.6 Discussion and Conclusion

While our model adopts a stylized information setup for analytical tractability, the key trade-off that we highlight here (between information elicitation and diseconomies of scope) may continue to shape the firm's job design decision in some related and more general settings. We consider two such extensions of our model. First, we relax the assumption that an informed agent observes the state without any noise, and assume that an agent's signal may be imprecise. Next, we relax the assumption that the observability of the underlying state of a project in each of its two tasks is statistically independent, and explore the case where they are mutually exclusive.

1.6.1 Imprecise Signals

In our model, the agent, conditional on observing the state, always observes it without any noise. While this assumption improves the analytical tractability of the model, it is conceivable that the agents may not be able to directly observe the state but only acquire an imprecise signal on the same. How would our characterization of the optimal job design change if the agents' information were noisy?

In order to explore this issue, we consider the following modification to our model: Suppose that the state $\omega_j \in \{G, B\}$ associated with the project j ($j \in \{A, B\}$) is never directly observed, but the agents' may observe a signal $\sigma_j \in \{G, B\}$ that is informative of ω_j. Let

$$\Pr(\omega_j = G \mid \sigma_j = G) = \Pr(\omega_j = B \mid \sigma_j = B) = \theta,$$

where $\theta \in (1/2, 1)$ reflects the precision of the signal. In parallel with the information structure of our model, we assume that the agent assigned in task T_{jk} privately observes σ_j with probability α. And with a slight abuse of notation, we denote the agent $\mathcal{A}_i$'s observation on the signal σ_j as $x_i^j \in \{G, B, \emptyset\}$, where $x_i^j = \emptyset$ if $\mathcal{A}_i$ does not observe σ_j in any of his assigned tasks. We keep all other aspects of our model unaltered. Notice that our main model corresponds to the case where $\theta = 1$.

Though a complete characterization of the optimal job design for this case is analytically intractable, the following proposition suggests that our main result is robust to a small noise in the agents' signal.

Proposition 4. *There exists a threshold $\theta^* < 1$ such that for $\theta > \theta^*$, the qualitative characterization of the optimal job design is the same as its counterpart in our main model (as given in Proposition 2), and the optimal contract is always efficient.*

However, if the agents' signal becomes sufficiently noisy (i.e., when θ is sufficiently low) our main result may no longer hold. Recall that under the optimal contract (in our main model), the project proceeds even when the agents fail to reveal their signal, i.e., the project continues unless the agent(s) report(s) a bad state. But when the agents' signal is sufficiently noisy, information elicitation gets harder. An agent now has a stronger incentive to hide a bad signal and let the project pass, since with some probability, a bad signal may still be associated with a good state.

This effect may introduce two sources of inefficiencies. First, the principal may reduce the effort incentives so as to mitigate the agent's incentive to hide a bad signal. (Recall that as the agent's effort increases, the performance measure is more likely to indicate success. Thus when the efforts are high, the agent has stronger incentive to continue the project under a bad signal.) Second, if such distortions to the effort level is too costly, the principal may also distort her continuation policy: the project may proceed only if the signal is good. And at the extreme, i.e., when θ is low enough, it is optimal for the principal to proceed with all projects without soliciting any information from the agents (or, equivalently, to settle for the outside option). These inefficiencies are illustrated in Figure 1.2 below that presents a numerical solution for the optimal job design problem.

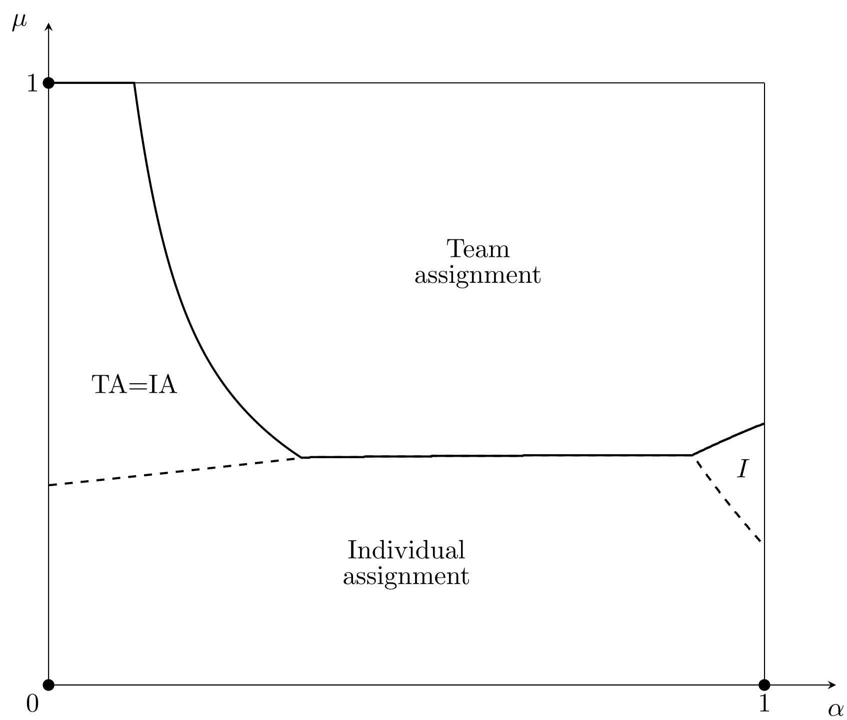


Figure 1.2: Optimal job design with imprecise signal ($\theta = 0.77$): In region I individual assignment is optimal but continuation decision is inefficient; project continues only if the report is good

1.6.2 Exclusive Signals

So far, we have assumed that the observability of the underlying state of a project in each of its two tasks is statistically independent. Such a setup may reflect a scenario where each task T_{jk} (of project j) gives access to a different (and independent) source of information, each of which may reveal the state ω_j with probability α. But it is conceivable that the informativeness of these sources may not be independent. In this subsection, we focus on one such scenario: sources being mutually exclusive in terms of their informativeness. An exploration of this case further illustrates how the agents' ability to control the outcome of a project through their efforts may affect the optimal job design.

To formalize this idea, we make the following modification to our model. We assume that exactly one of the two tasks associated with a given project may yield information about the project's underlying state. In particular, with probability 1/2, only task T_{j1} can yield information: the agent performing task T_{j1} observes the state with probability α, whereas the agent performing T_{j2} never observes it. And with probability 1/2, only T_{j2} is informative: the agent performing task T_{j1} never observes the state whereas the agent performing T_{j2} observes it with probability α. We keep all other aspects of the model unchanged.

Notice that in this setup, under individual assignment, the probability that an agent observes the state of his assigned project is α. And this is also the probability that under team accountability at least one of the two agents observes the state. However, in this setting team assignment appears to lose its advantage in information elicitation: as the observability of the state is mutually exclusive between tasks, should an agent observe an unfavorable information he can completely suppress it as his teammate would necessarily be uninformed.

One may anticipate that such complete control over the information on the state may make team suboptimal to individual assignment as team still continues to suffer from diseconomies of scope in incentive provision. However, this intuition is incomplete. Notice that an agent controls the outcome of a project in two ways: through his reporting on the state that affects the project's continuation probability, and also through his effort(s) that affect(s) the project's output and the performance metric (should the project proceed). When the signals are mutually exclusive, the advantage of team in muting the former channel is indeed diminished. However, team assignment may still help information elicitation as the agent cannot control the effort in all tasks that are associated with the project. Numerical result

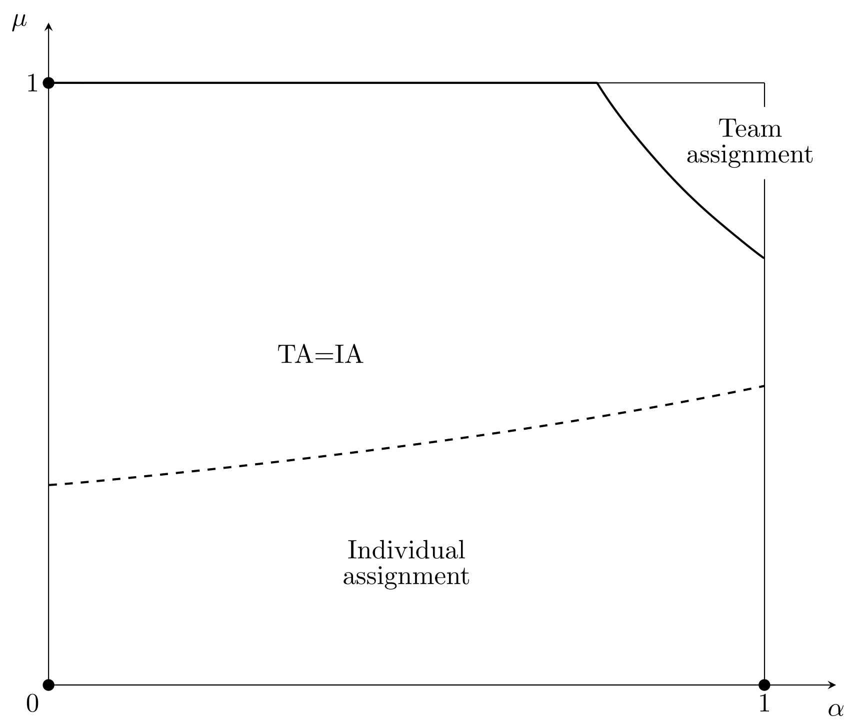


Figure 1.3: Optimal job design under mutually exclusive signals across tasks (within a project)

suggests (see Figure 1.3) that team's advantage in information elicitation remains sufficiently strong even under mutually exclusive signals and, as in our main model, it may still dominate individual assignment when both α and μ are sufficiently large.

1.6.3 Conclusion

When effective decision-making requires local information, the incentive structure in an organization must meet two goals at once: induce the workers to exert costly effort and truthfully report their information even if the information may be detrimental to their own interest. This article explores how job design—allocation of tasks among workers—interacts with such intertwined incentives. We argue that the optimal job design is shaped by a novel tradeoff between the ease of information elicitation and diseconomies of scope in incentive provision. And this tradeoff, in turn, is driven by the interplay between the "availability" of the workers' information and the "alignment" of their performance measure with the firm's objective. In particular, team assignment may be optimal when the performance measure is considerably misaligned, but the workers are highly likely to be informed about the local condition. Our findings suggest a novel explanation of why team can offer better incentive even when measures of individual performance remain available.

CHAPTER 2

CONTESTS WITH VALUATION ASSOCIATED WITH POPULATION UNCERTAINTY

2.1 Introduction

The term "contest" refers to a range of circumstances in which players exert efforts to surpass their opponents. Such circumstances include rent-seeking for rents allocated by policymakers, firms' advertising to compete for market shares, sports tournaments, patent races, and even military confrontation. Starting from the seminal work by Tullock (1980) on contest theory, substantial literature has investigated a range of applications using the Tullock contest success function. However, this literature typically assumes that both the number of players and the value of the prize is fixed.

The fact that these assumptions are overly strong is a significant problem. Players do not always know who their competitors are. In a rent-seeking situation, a firm typically lacks sufficient information about who and how powerful its competitors are; in a patent race, a firm lacks information about how many other firms are applying for the same patent when deciding how much R&D to invest. The players may have a list of potential competitors in a contest, but it can be tough to discern who is truly competing when they exert efforts. Games with population uncertainty can aid in the analysis of these situations.

A mathematical foundation for general games with population uncertainty is provided by Myerson (1998) and Milchtaich (2004), and a contest game suits it very well. Skaperdas (1996) axiomatizes the Tullock contest success function from several assumptions. One of the assumptions establishes sub-contests in which some players are excluded from the game. In an environment with a fixed number of participants, these sub-contests are manually constructed, in which only a subset of players participate in a hypothetical contest, and a sub-contest success function determines their chances of winning. In contrast, in an environment where the number of participants is random, a class of contest success functions for any number of players is well-defined. There is no need to manually construct hypothetical sub-contests. As a result, Skaperdas's assumption is more natural in contests with population uncertainty than in contests without.

Another common assumption in contest models is that the value of the prize remains constant regardless of the environment, i.e., it is unaffected by the number of contestants. This assumption is implicitly rooted in the classical contest models since the number of players is fixed, and it is explicitly stated in contest models with population uncertainty. This assumption becomes too strong in a variety of real-world scenarios.

For example, consider a scenario where firms compete in an R&D race for a new product, with the winner getting to launch the product first. If other firms can easily mimic the product, they will launch similar products after the winner debut the new product. In the real world, after the debut of the iPad from Apple, Samsung launched the Galaxy Tab, and Microsoft launched the Surface. In this scenario, the profit of successfully designing the new product decreases in the number of firms that participate in the R&D race. As a result, if the number of firms participating in the R&D race is big, the profit from launching the new product will be small because many firms will divide the market, and the market share of the winning firm will be small. On the other hand, if a firm is the only developer of a new product, it will become a market monopoly after launching it, and the profit will be significant. In this case, the value of the prize (profit of winning the R&D race) decreases in the number of contestants.

Cryptocurrency mining is another example. In most cases, the miner will invest resources to obtain a cryptocurrency. A large number of miners usually lead to a competitive contest, and the probability of a single miner getting a cryptocurrency is low. However, the large population of miners suggests that this cryptocurrency is popular among the general public, and its value will be substantial. In this case, the value of the prize (successfully mining for one unit of cryptocurrency) increases in the number of contestants.

As the assumption that the value of the prize is constant is too strong, I relax it and assume that the value of the prize is associated with the number of players. The value of the prize could be increasing/decreasing in the number of players, or it could be non-monotonic. The purpose of this study is to examine contests where the value of the prize is associated with population uncertainty and compare the effort put in under various situations.

In this paper, I first construct a contest with population uncertainty, and the value of the prize depends on the realization of the number of players. Then, I prove that the equilibrium exists and it is unique. I then consider the following three scenarios:

(a) The value of the prize is constant;

(b) The value of the prize increases as the number of players increases;

(c) The value of the prize decreases as the number of players increases.

The value of the prize is constant; The value of the prize increases as the number of players increases; The value of the prize decreases as the number of players increases. I assume that the expected value of the prize in these three scenarios is the same. When the value of the prize increases in the number of players, the effort level is high, while when the value of the prize decreases in the number of players, the effort level is low.

The driving force behind this result is the friction between the belief in the number of players from a player's perspective and the belief in the number of players from an outsider's perspective. From an outsider's perspective, the distribution of the number of players is just the prior distribution. However, from a player's perspective, "I am in the game" is informative, and the belief needs to be updated. As a result, when compared to the prior, the player's belief is skewed to the right. When the value of the prize increases in the number of players, as the player puts more weight on the event where the number of players is large, a player receives more rewards under the updated belief than under the prior, thus having more incentive to exert effort. Conversely, when the value of the prize decreases as the number of players increases, the logic is similar.

I then extend my analysis to the following scenarios: (i) the number of players is fixed, and the value of the prize is constant, (ii) the number of players is random, and the value of the prize is constant, and (iii) the number of players is random, and value of the prize is linear on the number of players. I also assume that these three contests have the same expected number of players and expected value of the prize. Myerson and Wärneryd (2006) show that expenditure under (ii) is smaller than (i). As mentioned above, the player's belief is skewed to the right compared with the prior. So, the competition is more "intense" in (ii) than in (i) from a player's perspective; thus, the effort level is lower in (ii). Further, I show that when the value of the prize is proportional to the number of players (linear with zero intercepts), then the effort level in (iii) is the same as in (i).

Related Literature: This paper contributes to the literature on games with population uncertainty. Population uncertainty arises when the assumption that the players' identities are common knowledge is relaxed, and the number of players is uncertain (or stochastic). Mcafee and Mcmillan (1987) are the first to investigate auction models with a stochastic number of players. They show that when bidders are risk-averse, auction with a stochastic number of players Pareto-dominates auctions that announces the number of players. Harstad, Kagel, and Levin (1990) and Levin and Ozdenoren (2004) concentrate on the revenue equivalence result when the number of bidders is uncertain in an auction. They show that the general results of revenue equivalence could be extended when the bidders are risk-neutral, but it breaks down when the bidders are an ambiguity aversion. Following that, several scholars examine bidder preference for auction forms (Matthews, 1987), endogenize entry decisions (Levin and Smith, 1994), and characterize information aggregation (Harstad, Pekeč, and Tsetlin, 2008) for auction with population uncertainty. These publications focus on population uncertainty in auctions, with no mention of other types of games.

Myerson (1998) provides formal definitions of games with population uncertainty, and Milchtaich (2004) proposes a more general mathematical framework. Myerson (1998) also points out that one particular game — Poisson game, has the following property: a player's environment (the number/type of players other than herself) is the same as an external game theorist's perception of the whole game. Poisson game is widely used in voting theories (Campbell, 1999; Myerson, 2000; Piketty, 2000; Myerson, 2002; Krishna and Morgan, 2012; Bouton and Castanheira, 2012; Bouton, 2013; Ekmekci and Lauermann, 2022), and game with population uncertainty are also studied in dynamic games (Satterthwaite and Shneyerov, 2007) and Bertrand competition (Ritzberger, 2009). This work differs from the previous studies in that it explores games with population uncertainty under an environment of the contest.

Tullock (1980) looks at the issue of competing rent-seekers who spend resources to sway policy outcomes. Tullock contest success function has a wide application: it may be used to describe the relationship between advertising expenditure and market shares (Schmalensee, 1976), to describe R&D contests (Fullerton and McAfee, 1999), and to describe the outcome of sports tournaments (Szymanski, 2003). Skaperdas (1996) axiomatizes the Tullock contest success function from several reasonable assumptions, thus giving strong support for its use in actual applications. Tullock contest variants have been investigated in subsequent research (Azmat and Möller, 2009; Münster, 2007, 2009; Wasser, 2013). These articles assume a fixed number of players, but I focus on a game with population uncertainty.

Our paper complements the works by Myerson and Wärneryd (2006), Münster (2006), and Lim and Matros (2009). Myerson and Wärneryd (2006) is one of the first papers to analyze the contest with population uncertainty. They set up a model where the number of players is stochastic and then show that total equilibrium expenditure is strictly lower in a contest with population uncertainty than in a contest without population uncertainty, even though the expected number of players is the same in both contests. Münster (2006) considers a rent-seeking model in which a group of potential players might be active or inactive. When the expected fraction of active players is low, a rise in the number of potential players boosts individual rent-seeking expenditure, which is the opposite of what happens in contests competitions without population uncertainty. Lim and Matros (2009) investigate a similar model where n potential players try to participate in a contest, and each player participates with probability p. They characterize the game's equilibrium and show that individual spending is single-peaked in p and the total spending is monotonically increasing in p and n. All three articles assume the prize has a fixed value, but I expand the analysis to a scenario in which the prize's value is contingent on the realization of the number of players.

This paper is structured as follows. The model is built up in Section 2.2. Section 2.3 examines different scenarios and summarizes the key findings. Section 2.4 calculates the magnitude of this effect through numerical examples. Section 2.5 provides two applications of the model. Section 2.6 concludes the paper. The Appendix contains all of the proofs.

2.2 Model Setup

Consider a contest with N identical risk-neutral players, where N is a random variable over $\mathbb{N} = \{1, 2, ...\}$. Let $\pi : \mathbb{N} \to [0, 1]$ be the prior probability distribution of N, so $\sum_{i=1}^{\infty} \pi(n) = 1$. Also, define μ as the expected number of players and assume it is finite, i.e. $\mu = \sum_{i=1}^{\infty} \pi(n)n < \infty$. If the support of π contains two or more elements, population uncertainty arises; otherwise the contest degenerates into one with fixed number of players. Players do not observe the realization of N, but the prior π is common knowledge.

Let $v : \mathbb{N} \to \mathbb{R}_+$, and players compete for a single reward of value $V = v(N)$. All players are identical and share the same valuation. V is also a random variable since N is a random variable. Also define η as the expected value of the reward and assume it is finite, i.e. $\eta = \sum_{i=1}^{\infty} \pi(n)v(n) < \infty$.

For any realization of n, denote $\mathbf{x}^n = (x_1, x_2, ..., x_n)$ be the efforts for the n players. The effort levels must be non-negative ($x_i \in [0, \infty)$, $\forall i$). The cost of effort is x_i itself for player i. Player i's winning probability is determined by a contest success function $p_i^n(\mathbf{x}^n)$. Similar to Skaperdas (1996) and Myerson and Wärneryd (2006), I assume $\{p_i^n\}$ satisfies the following assumptions:

(A1) $\forall n \in \mathbb{N}, \forall i \in \{1, ..., n\}, p_i^n \geq 0$; $\forall n \in \mathbb{N}$, $\sum_{i=1}^{n} p_i^n = 1$; if $x_i > 0$, $p_i^n(\mathbf{x}^n) > 0$.

(A2) $\forall n$, $\forall i$, p_i^n is increasing in x_i and decreasing in x_j for $j \neq i$.

(A3) Anonymity: for any $n \in \mathbb{N}$, for any permutation φ of $\{1, 2, ..., n\}$, we have

$$p_i^n(x_1, x_2, ..., x_n) = p_{\varphi(i)}^n(x_{\varphi(1)}, x_{\varphi(2)}, ..., x_{\varphi(n)})$$

(A4) Consistency: for any $i \leq m \leq n$, for any effort levels $(x_1, x_2, ..., x_n)$, we have:

$$p_i^m(x_1, x_2, ..., x_m) = \frac{p_i^n(x_1, x_2, ..., x_n)}{\sum_{j=1}^{m} p_j^n(x_1, x_2, ..., x_n)}$$

The following Lemma describes a contest success function.

Lemma 5. *A system of contest success functions $\{p_i^n\}$ that satisfied (A1)-(A4) must have the following form:*

$$p_i^n(x_1, x_2, ..., x_n) = \frac{f(x_i)}{\sum_{j=1}^{n} f(x_j)}$$

where $f(\cdot)$ is a positive increasing function.

I further assume that

(A5) $f(\cdot)$ is twice differentiable and concave.

Time line: The time line of the contest is summarized below:

- $N = n$ is realized according to distribution π.

- Without knowing the realization of N, each player i chooses an effort level x_i.

- The winner is chosen by the contest success function p_i^n, and payoffs accrue.

Equilibrium concept: As Myerson (1998) and Myerson and Wärneryd (2006) pointed out, the traditional concept of Nash equilibrium and its refinements do not apply to games with population uncertainty. In such games, a player can only be identified by her type, instead of her name. As a result, all players of the same type must share the same strategy. I further restrict attentions to pure strategies. Thus, in the game described above, a strategy x is an equilibrium if it satisfies the following:

- Belief $\tilde{\pi}$ satisfies Bayes' rule:
$$\tilde{\pi}(n) = \frac{\pi(n)n}{\sum_{n'=1}^{\infty} \pi(n')n'}$$

- The effort level maximizes the player's payoff, given other players play the equilibrium strategy:
$$\mathbb{E}[u_i(x_i, x_{-i}|\tilde{\pi})] \geq \mathbb{E}[u_i(x_i', x_{-i}|\tilde{\pi})] \quad \forall x_i'$$

The second condition is standard, and the first condition pins down a player's belief about the game she is in. In this game, players have only one type, so the belief is only about the number of players.

The equilibrium concept applied in this study expands upon the Bayesian Nash Equilibrium in the following sense: a) the player only takes other players' strategies, and her own belief into consideration, and b) higher-order beliefs will not affect the strategies. The equilibrium concept is equivalent to symmetric Bayesian Nash Equilibrium if π degenerates to a one-point distribution.

Proposition 5. *An equilibrium exists and is unique.*

Proposition 5 provides a foundation for the following analysis.

2.3 Analysis and Result

2.3.1 Value of Prize Is Monotonic

In this subsection, I focus on scenarios where the common-valued reward $v(n)$ is monotonic. To be more specific, consider the following three contests:

- Under contest C_1, $v(n) = v_1(n)$ where $v_1(n)$ is increasing in n, with $\mathbb{E}_\pi[v_1(n)] = v$.
- Under contest C_2, $v(n) = v_2(n)$ where $v_2(n)$ is decreasing in n, with $\mathbb{E}_\pi[v_2(n)] = v$.
- Under contest C_3, $v(n) = v$ where the value of the reward is independent to the number of players.

Proposition 6. *Let x_1^*, x_2^*, x_3^* be the equilibrium effort levels of the players in three contests respectively. Then $x_2^* \leq x_3^* \leq x_1^*$.*

The contest C_3 is a benchmark where the value of the prize is constant, and x_3^* is the effort level of each player. When the value of the prize increases in the number of players, the effort level is higher than the benchmark. In contrast, when the value of the prize decreases in the number of players, the effort level is lower than the benchmark.

To understand the intuition behind this effect, one needs to focus on the marginal benefit of exerting effort. When comparing contest C_1 and C_3, suppose the realization of the number of players is high. In that case, although the probability of getting the reward is low due to a large number of competitors, the marginal benefit of effort becomes high since the value of the prize is high. Similarly, if the realization of the number of players is low, the marginal benefit of effort becomes low. The updated belief $\tilde{\pi}$ puts more weight on the events where number of players is high. Thus, the former effect dominates the latter, so the marginal benefit of exerting effort is higher in contest C_1 than the benchmark.

To understand why the belief $\tilde{\pi}$ puts more weight on the events where the number of players is high, consider a simple example. Assume that the number of players is equally likely to be 1 or 3. The distribution would be $\pi(1) = \pi(3) = 0.5$ from the standpoint of nature (social planner/god mode). However, from the player's perspective, the belief would be $\tilde{\pi}(1) = 0.25$ and $\tilde{\pi}(3) = 0.75$. The player incorporates the information "I am in the game" into the belief. The player has a high probability of being chosen if the realized number of players is high, so the updated belief is skewed towards the right side.

The friction between the player's belief and the prior is found but not well studied in the economics literature. It is referred to as a "classroom size" problem by Mcafee and Mcmillan (1987), and Myerson (1998) shows that the expected number of players from a player's perspective is one more than the expected number of players from an outsider's perspective if and only if the distribution is Poisson. However, they do not focus on this friction in the above studies.

The difference between a player's belief and the prior may lead to other results. For example, if a player is competing in a market, population uncertainty drives the expected number of players in the player's perspective to be higher than the actual number. As a result, the player is always under the impression that the market is more competitive than it actually is. The population uncertainty may lead to excess competition in the market.

This friction also indicates that the number of players is special in the setup of a game. Suppose there is a game G1 with population uncertainty and a game G2 with uncertainty about the underlying state of the world. In the current game theory paradigm, if the player receives no information, the player's belief is the same as the prior in G2, but it differs from the prior in G1. It further suggests that the information analysis process differs depending on whether the uncertainty is about the population or the underlying state of the world. It is an open topic on why uncertainty about the population and uncertainty about the underlying state of the world are classified as different information categories. Otherwise, if one wants to treat these two uncertainties similarly, a more general framework for setting up a game may be needed.

2.3.2 Value of Prize Is Linear

In this subsection, I will focus on the scenarios where the value is linear in the number of players and compare it with contests with no population uncertainty. Consider the following three contests:

- Under contest C_4, the number of players is fixed at μ, and $v_4(n) = v$ where v is a constant.

- Under contest C_5, the number of players is a random number with density function π

where $\mathbb{E}_\pi[n] = \mu$, and $v_5(n) = v$ where v is a constant.

- Under contest C_6, the number of players is a random number with density function π where $\mathbb{E}_\pi[n] = \mu$, and $v_6(n) = a + bn$ where $\mathbb{E}_\pi[v_6(n)] = v$.

Proposition 7. *Let x_4^*, x_5^*, x_6^* be the equilibrium effort levels of the players in three contests respectively, and assume π is non-degenerate. Then:*

- *When $a > 0, b < 0$, $x_4^* > x_5^* > x_6^*$.*
- *When $a > 0, b = 0$, $x_4^* > x_5^* = x_6^*$.*
- *When $a > 0, b > 0$, $x_4^* > x_6^* > x_5^*$.*
- *When $a = 0, b > 0$, $x_4^* = x_6^* > x_5^*$.*
- *When $a < 0, b > 0$, $x_6^* > x_4^* > x_5^*$.*

Proposition 7 only focuses on contests where $v > 0$, since $x^* = 0$ when $v \leq 0$ (or players want to quit if there is an outside option). Myerson and Wärneryd (2006) show that $x_4^* > x_5^*$, which means the effort level is lower in the contests with population uncertainty than in contests without population uncertainty. My findings imply that when the reward value is in the form $v(n) = bn$, contests with population uncertainty will have the same amount of effort as contests without population uncertainty.

2.4 Numerical Example

In this section, I provide some numerically solved examples to measure the magnitude of this effect. First assume the contest success function $f(x) = x$. Then fix the expected number of players to μ, and let the distribution π be: $\pi(\mu-\tau)=\frac{1}{2}\alpha$, $\pi(\mu)=1-\alpha$, $\pi(\mu+\tau)=\frac{1}{2}\alpha$, and $\pi(n)=0$ for $n\notin\{\mu-\tau,\mu,\mu+\tau\}$. Both $\tau\in[0,\mu-1]$ and $\alpha\in[0,1]$ measures how diverse the distribution is, and the variance of distribution π is $\alpha\tau^2$. If $\alpha=0$ or $\tau=0$, the distribution degenerates to one point.

Now fix the expected value of the prize to 1 since it will not affect the relative magnitude of the effect. Assume the value of the prize is: $v(\mu-\tau)=1-\varepsilon$, $v(\mu)=1$, and $v(\mu+\tau)=1+\varepsilon$. $\varepsilon\in[-1,1]$ measures how the value of the prize changes according to the number of players. If $\varepsilon=0$, the value of the prize is constant; if $\varepsilon>0$, the value of the prize is increasing in n; and if $\varepsilon<0$, the value of the prize decreases in n.

I calculate the expected total efforts exerted by all players μx^*, as it is comparable across different μ.

$$\mu x^*=\frac{\mu-1}{\mu}-\frac{\tau^2}{(\mu-\tau)\mu(\mu+\tau)}\alpha+\frac{\tau}{(\mu-\tau)(\mu+\tau)}\varepsilon\alpha$$

The first part ($\frac{\mu-1}{\mu}$) is the solution to the classical contest problem where the number of players is fixed at μ. The second part ($-\frac{\tau^2}{(\mu-\tau)\mu(\mu+\tau)}\alpha$) represents the effect of introducing population uncertainty with a fixed value of prize, which is found by Myerson and Wärneryd (2006). The third part ($\frac{\tau}{(\mu-\tau)(\mu+\tau)}\varepsilon\alpha$) illustrates the effect of introducing the assumption that the value of the prize is associated with the number of players.

Here are several findings that can be derived from the formula:

- The effect is significant when μ is small.

  For example, $\mu=3$, $\tau=1$, and $\alpha=1$, associating the value of the prize with the number of players would increase (or decrease) the effort level by 20%.

$$\mu x^* = \begin{cases} 0.625 & \text{if} \quad \varepsilon = 0 \\ 0.75 & \text{if} \quad \varepsilon = 1 \\ 0.5 & \text{if} \quad \varepsilon = -1 \end{cases}$$

- The effect is small when μ is large.

  I give two examples here. For $\mu = 30$, $\tau = 1$, and $\alpha = 1$, the effort level fluctuate about 0.1% if the value of the prize is associated with he number of players.

$$\mu x^* \approx \begin{cases} 0.966630 & \text{if} \quad \varepsilon = 0 \\ 0.967742 & \text{if} \quad \varepsilon = 1 \\ 0.965517 & \text{if} \quad \varepsilon = -1 \end{cases}$$

  One may wonder how the effect changes if $\frac{\tau}{\mu}$ stay the same. The effect is still small when μ is large. Set $\mu = 30$, $\tau = 10$, and $\alpha = 1$, the effect account for about 1.3% of the effort:

$$\mu x^* \approx \begin{cases} 0.9625 & \text{if} \quad \varepsilon = 0 \\ 0.975 & \text{if} \quad \varepsilon = 1 \\ 0.95 & \text{if} \quad \varepsilon = -1 \end{cases}$$

When τ is fixed, $\frac{\tau}{(\mu-\tau)(\mu+\tau)}\varepsilon\alpha$ go to 0 at the speed of $\left(\frac{1}{\mu^2}\right)$. By setting $\frac{\tau}{\mu}$ to a constant, $\frac{\tau}{(\mu-\tau)(\mu+\tau)}\varepsilon\alpha$ go to 0 at the speed of $\left(\frac{1}{\mu}\right)$. In either case, the effect becomes small as μ increases.

The intuition behind this is the same as the main result. As the player's belief is skewed to the right compared with the prior, the relative magnitude of the difference between the player's belief and the prior becomes smaller as μ becomes larger. For $\mu = 3$, $\tau = 1$, and $\alpha = 1$, the expected number of players is 3 in an outsider's perspective and 3.33 in a player's perspective. This relative difference is quite large. If $\mu = 30$, $\tau = 1$, and $\alpha = 1$, the expected number of players would be 30 and 30.033 respectively, and the relative difference becomes small. Thus, the effect of associating the value of the prize to the number of players is large if μ is small, and vice versa.

2.5 Applications

2.5.1 Promoting Effort Levels

Consider the following scenario: an organization wants to elicit efforts from a potential group of people who may or may not be interested in such activities. Because the organization does not know who is interested in advance, the number of participants is uncertain. The effort does not directly benefit the player or generate little benefit compared to its cost, but the organization benefits from it. As a result, the organization must incentivize the players to put forth an effort. I also assume that the organization will be unable to provide individual incentives to the players. The reason could be that the effort level is difficult to contract or that the nature of the organization does not lend itself to issuing those incentives. I give some examples to illustrate this scenario.

Example 1: A firm wishes to promote effective communication skills among its departments. The ability to communicate will improve the department's overall efficiency but will not improve the efficiency of an individual worker. There are many workers, and each worker may be interested or disinterested in learning the communication skill. Hence, the number of workers willing to put forth the effort to learn is uncertain. Furthermore, it is difficult to provide individual incentives for workers to learn such skills. It is difficult to measure and contract a worker's effort level, and it is unreasonable to link a worker's wage to it. The company wants to encourage potential employees to learn as much as possible.

Example 2: An environmental organization seeks to encourage future farmers to adopt new environmentally friendly technology, and the potential farmers may be obstinate or open-minded. The proper application of technology will benefit both the environment and human welfare. On the other hand, farmers must learn how to use the technology and apply it to their own farms to reap the greatest environmental benefits. The goal of the organization would be to improve the environment; thus, the more farmers who learned, the more benefits the organization would receive. However, because learning is difficult to quantify, it is hard to provide direct subsidies. Therefore, the organization needs to find another way to incentivize the farmers.

One possible solution for the organization to address this issue described above would be to host a tournament/competition based on the amount of effort the player put in. The winner receives a monetary prize, and it functions similarly to an all-pay auction. The tournament/competition's details are not important in this paper, but I would assume that the process of generating the winner would satisfy $(A1)-(A4)$. As a result, the organization could avoid the cost of measuring each player's effort levels because selecting a winner requires less information than knowing each player's effort levels.

Now compare two scenarios in which the award is fixed versus variable. For simplicity, I'll assume that the prize is equal to the number of players multiplied by a constant (linear in the number of players). According to the main model's analysis, players with a variable award will exert more effort for the same expected monetary amount. Thus, the organization can elicit more efforts under the same expected payment.

Myerson and Wärneryd (2006) show that, given the same expected number of players, the effort level in the scenario with population uncertainty is less than the effort level in the scenario without population uncertainty for a fixed monetary award. In this paper, I demonstrate that the effort level could be partially restored by making the monetary award an increasing function of the number of players with the expected value of the prize stays the same.

2.5.2 Design Competition

Consider a design competition in which companies compete to design new products by investing in R&D. There are two kinds of products: those that are difficult to imitate and those that are easy to imitate. Because companies may have hidden developing initiatives that are only revealed after success, the number of companies that participate in the R&D of a certain product is random.

Consider a product that is hard to imitate, and assume that the new market has a unit demand of $p = 1-q$. Once the firm completes the product design, the firm that first designed it will have monopoly power on the market. The solution to the unit demand problem would be $p = q = \frac{1}{2}$, and the company's profit would be $\frac{1}{4}$. Because the product is difficult to imitate, profit will be zero for firms that are not the first to design it. As a result, the net payoff of designing a successful product is $\pi^d(n) = \frac{1}{4}$, and it is constant regardless of the number of competitors.

Now consider a product that is easy to imitate and has the same unit demand $p = 1 - q$. Because the product is easy to imitate, the firm that initially designs it will benefit from designing it for a short time, but not for a long time. Other businesses will follow suit after the debut of that new product. To simplify things, I assume that the company that successfully designed the product will be a Stackelberg leader for this product. In the n player Stackelberg game, the Stackelberg leader will receive $S(n) = \frac{1}{4n}$, while the followers will receive $F(n) = \frac{1}{4n^2}$. As a result, the net payoff of designing a successful product is $\pi^s(n) = S(n) - F(n) = \frac{n-1}{4n^2}$. The net payoff depends on the realized number of competitors, and it decreases as the number of competitors increases.

Since $\pi^s(n) = \frac{n-1}{4n^2} < \frac{1}{4} = \pi^d(n)$ for all $n > 1$ and $\lim_{n\to\infty} \pi^s(n) = 0$, it can be shown that $\mathbb{E}[\pi^s(n)] < \mathbb{E}[\pi^d(n)]$ for all non-degenerate distribution of n. That is, assuming the market has the same demand, the expected net payoff of designing a easy-to-imitate product is always less than the expected net payoff of designing a difficult-to-imitate product. As a result, the company will invest more resources to the research and development of difficult-to-imitate products.

Now, suppose that the product that is easy to imitate has a larger market. Assume the product that is difficult to imitate still has unit demand $p = 1 - q$, but the product that is easy to imitate has a demand of $p = b - q$ where $b > 1$. To make things comparable, I assume that $\mathbb{E}[\pi^s(n)] = \mathbb{E}[\pi^d(n)]$. That is, the expected net payoff of the two products is equal. A naïve intuition would suggest that the incentives to invest in both items are the same. However, based on the main model's analysis, the firm will continue to invest more in the product that is hard to imitate. This approach explains why firms invest more in hard-to-imitate products, even when the expected net payoff is the same. The very nature of the net payoff is decreasing in the number of realized competitors, leading to lower R&D investment in easy-to-imitate products.

2.6 Conclusion

In this paper, I explore contests with population uncertainty, in which the value of the prize depends on the number of players. When population uncertainty arises, the player's belief in the number of players is skewed to the right compared with the prior. This friction drives my results.

Assuming the expected number of players and the expected value of the prize stays the same, I compare the following three scenarios:

(a) the value of the prize is constant,

(b) the value of the prize increases as the number of players increases, and

(c) the value of the prize decreases as the number of players increases.

I find that the effort level is highest under (b) and lowest under (c).

I also compare the following three environments:

(i) the number of players is fixed, and the value of the prize is constant,

(ii) the number of players is random, the value of the prize is constant, and

(iii) the number of players is random, and the value of the prize is linear on the number of players.

I find that if the value of the prize is proportional to the number of players (linear with zero intercepts), the effort level is the same under (i) and (iii).

These analyses have many applications. One possible situation is that exerting efforts has certain positive externalities, but it is impossible to provide an incentive for each potential player individually. Then, a contest with escalating incentives as the number of players grows could be a viable option. Another situation would be a design competition. There are two new products to develop: H is hard to imitate, and E is easy to imitate. Even if E has a larger market and thus the expected profit of the two products is equal, the firm will nevertheless invest more in H because the net value of the product is declining as the number of competitors increases.

CHAPTER 3

ASSIMILATION WITH DIFFERENT WORKING SKILL ACQUISITION

3.1 Introduction

Discrimination between different groups is a widespread phenomenon around the world. For example, before WWII, the German government discriminated against Jews to the extreme, and the government wanted to eliminate them from the earth. However, during the first half of the 20th century, many European immigrants went to the US and experienced little discrimination. Therefore, the driving force behind discrimination in different countries is an exciting topic.

Recent literature focuses on the discrimination against people with low average working skill levels but seldom studies the discrimination against people with high average working skills. For example, the "Acting White" phenomenon is well studied by Eguia (2017) and Advani and Reich (2015). Both papers proposed a 2-stage game model: for the first paper, the agent of the advantaged group will choose a discrimination level in the first stage, and in the second stage, all agents choose a skill level, and agents from the disadvantaged group will choose their self-identity; for the second paper, in the first stage, agents from the minority group will choose their identity, and in the second stage, all agents will select their skill level. In both models, individuals from minority groups face a trade-off between cultural and economic incentives: assimilation will gain economic benefit, and non-assimilation will prevent the cost. These studies explained the discrimination against minorities with lower average working skill levels.

However, discrimination against minority with high working skill level do exist. For example, in the US, Asian Americans have a higher secondary school completion rate than white people (Espinosa, Turk, Taylor, and Chessman, 2019). The positive and negative dichotomy of Asian American stereotypes has been well documented (Fiske, Cuddy, Glick, and Xu, 2002; Gilbert, 1951; Ho and Jackson, 2001; Jackson, Hodge, Gerard, Ingram, Ervin, and Sheppard, 1996; Karlins, Coffman, and Walters, 1969; Katz and Braly, 1933). They are stereotyped as intelligent, diligent, hard-working, self-disciplined, good at math and sciences (implying competence), but quiet, shy, unpopular, reserved, traditional, and deriving less value on a leisurely life. With that said, Lai and Babcock (2013) studied how White male and female evaluators perceive an Asian American versus White job candidate on the dimensions of competence and social skills and how these perceptions affect evaluators' decisions in hiring and promotion. They found that female evaluators were less likely to select Asian than White candidates for positions involving social skills and were less likely to promote Asian than White candidates into these positions. These studies give us an example of discrimination against minorities with high working skill levels.

To understand the discrimination between different groups, the "self-identity" is an essential concept. Akerlof and Kranton (2000) pointed out the important relationship between self-identity and economic outcomes. The choice of self-identity affects the agent's utility function, so the choice changes the payoff of the agent herself and the payoff of other agents. Furthermore, the collective choice of self-identity may change the social norms, affecting identity-based preferences. It is important because discrimination against people is discrimination against race and discrimination against group choice. A person born in a family of a minority group can still choose the majority as "self-identity" and thus share the same culture with the majority group.

The empirical results prove that "self-identity" plays an essential role in the utility function. Benjamin, Choi, and Strickland (2010) conducted experiments to show that the social identity of an agent can affect her preference. The discount factor and propensity to save are affected by the choice of the majority group. The Asian American subjects exhibit more patient preferences when making their ethnicity salient. Similarly, black subjects with long-standing roots in the United States become more patient when their race becomes salient. There is also suggestive evidence that native blacks become more risk-averse and whites become more patient when their racial identity is salient.

To understand the discrimination among groups, one needs to figure out why there are differences between different groups' working skill levels and how "self-identity" affects utility. I model the difference in working skill level due to a difference in discount factors among groups. The "self-identity" will affect the utility function because there is a network effect within groups.

The discount factor is a generalized factor. There are many estimations about the discount factor, and the results are very different. The estimation conducted by Hausman (1979), Moore and Viscusi (1990) , Dreyfus and Viscusi (1995), Pender (1996), Coller and Williams (1999), Harrison, Lau, and Williams (2002) ranges from 0.53 to 0.99. By comparing Pender (1996) and Harrison, Lau, and Williams (2002), the first estimates the discount factor in India, and the second estimates the discount factor in Denmark. The first one gets a result between 0.59 to 0.79, and the second one gets a result of 0.78. It is clear that different groups have different discount factors.

I can explain the discount factor in my model in many ways. The high discount factor can be thought of as putting large weight into the future. It can also be explained as a longer expected life since the longer the life is, the more weight an agent will put in the future utility. It is also a factor of the cultural norm. For example, it has been shown before that Asian Americans place less value on leisure, so that the discount factor will be larger for Asian Americans.

There is evidence that networks play an essential role in the choice of assimilation. Verdier and Zenou (2017) studied the relationship between the social network and cultural assimilation. They show that agents in the center of the network have more incentive to assimilate than the agents in the marginal of the network. They also show that more people choose to assimilate with a denser network (interaction between agents is strong).

The utility of an agent depends on her group's average working skill and depends on how large the group is. The larger the group is, the more benefit that an agent can derive from being a member. Currarini, Jackson, and Pin (2009) find three significant results: first, larger groups (measured as a fraction of the population of their respective schools) form a greater fraction of their friendships with people of their same type; second, larger groups form significantly more friendships per capita, that is, members of a group that comprises a small minority in a school form roughly six friendships per capita, while members of groups that comprise large majorities (close to 100 percent of a school) form on average more than eight friendships; third, groups tend to form same-type friendships at rates that exceed the relative fractions in the population. These results give us a solid foundation that the utility function will depend on the group's size.

Our analysis proceeds as follows. Section 3.2 will setup the model with network effect and two groups have different discount factors. Section 3.3 solves the game. Section 3.4 solves the game for an explicit functional form. Section 3.5 will do the comparative status which explain the main result. Section 3.6 proposed some testable results. Section 3.7 discussed about some further extensions. Section 3.8 concludes the paper.

3.2 Model Setup

3.2.1 Players

Consider a society with a continuum of agents. Each agent is identified by her background and her ability. Assume that the set of possible backgrounds is $\{\mathcal{A}, \mathcal{I}\}$, where $\mathcal{A}$ represents the majority group, and $\mathcal{I}$ the minority group. Assume that the set of possible abilities is $[0, 1]$. Let $N = \{\mathcal{A}, \mathcal{I}\} \times [0, 1]$ denote the set of players. For each background $\mathcal{J} \in \{\mathcal{A}, \mathcal{I}\}$, let $N_{\mathcal{J}} = \mathcal{J} \times [0, 1]$ denote the subset of agents with background $\mathcal{J}$.

Assume the measure of N, denoted $m(N)$, is equal to 1, and that the distribution of agents is uniform over N. Also assume that $m(N_{\mathcal{I}}) = m$ ($m \in \left(0, \frac{1}{2}\right)$) and $m(N_{\mathcal{A}}) = 1 - m$. That is, the majority group has more population than the minority group. The distribution of ability conditional on background $\mathcal{J}$ is uniform over $[0, 1]$.

For any $i \in N$, let $\theta_i \in [0, 1]$ denote the ability of agent i. Individual ability is private information.

3.2.2 Lifetime of the Agent

The agent will have two stages. In stage 1, the agent will be considered young, and she will spend time learning working skills s_i and enjoy her leisure time. In stage 2, the agent will become an adult and choose between assimilation or not. The agent will then work, and payoff accrues.

Skill Level

Agents acquire working skills when they are young. I normalize the time agents have to 1 when they are young. For any agent $i \in N$, she can choose her leisure time l_i and spend the rest of her time $1 - l_i$ learning. Based on the ability θ_i, the working skill level agent i can get is $s_i = \theta_i(1 - l_i)$.

Discount Factor

Agents with different backgrounds have different time preferences. Assume that agents with background $\mathcal{A}$ have discount factor $\beta_{\mathcal{A}}$ and agents with background $\mathcal{I}$ have discount factor $\beta_{\mathcal{I}}$. I assume that $\beta_{\mathcal{A}} < \beta_{\mathcal{I}} < 1$ in this paper, that is, the minority group put more weight on the future utility.

3.2.3 Choice of Social Group

Assume that there are two self-identity groups, A and I, characterized by two sets of social norms and actions expected from their members. In-group networks are strong, and the networks across groups are very small.

In stage 2, assume that agents with background $\mathcal{A}$ will identify them self as A, $N_{\mathcal{A}} \subseteq A$. Assume that any agent with background $\mathcal{I}$ can choose to belong to social group I at no cost, or she can embrace the cultural norms of group A to then join A. Let $a_i \in \{0, 1\}$ be the choice of agent $i \in N_{\mathcal{I}}$. Let $a_i = 0$ denote that $i \in N_{\mathcal{I}}$ chooses to be part of group I and not to assimilate, and let $a_i = 1$ denote that agent $i \in N_{\mathcal{I}}$ chooses to adopt the majority cultural norms and to become a member of the majority group A. If $a_i = 1$, I say that

$i \in N_{\mathcal{I}}$ "assimilates."

3.2.4 The Cost of Assimilation

The cost of assimilation is d for agent i, where $d \in \mathbb{R}_+$ is the difficulty of assimilation to become a member of A. This difficulty of assimilation d is an endogenous, strategic variable. It can be interpreted as the level of discrimination: if agents with background $\mathcal{A}$ are welcoming to those who assimilate, d is small; if agents with background $\mathcal{A}$ are hostile, or if they give the cold shoulder to those who are trying to assimilate, then d is high.

The level of d is chosen endogenously in the model by an agent with background $\mathcal{A}$. In the setup, I assume that agents with background $\mathcal{A}$ collectively choose an agent $h \in N_{\mathcal{A}}$ as a representative. The agent h will then choose the discrimination level d. As shown in Section 3.3, all agents with background $\mathcal{A}$ will share the same optimal choice, so the mechanism of choosing the representative h will not affect the equilibrium.

3.2.5 Network Effect

Agents will benefit from the social group network effect. Agents in the same social group share the same behavior and culture so that they will be closely connected. With a larger group size, each member in the group will benefit more.

Mathematically, I use a function $f(m_J)$ to model the network effect. I will assume that $f(0) = 0, f'(.) \geq 0, f'(0) \leq 1, f''(.) \leq 0$. That is, the larger the group is, the greater the network effect. Also, the marginal benefit of the network effect is decreasing.

3.2.6 Timing of the Game

The timing of the game is as follows:

1. For any agent $i \in N$, i chooses the leisure time l_i when they are young and acquires working skill s_i accordingly. All agents will act simultaneously.

2.1 All agents become adults and observe the working skill s_i of other agents. Agents with background $\mathcal{A}$ choose a representative $h \in \mathcal{A}$ and h chooses the discrimination level d.

2.2 All agents observe d. Agents with background $\mathcal{I}$ make an assimilation choice a_i based on the information they have. All agents with background $\mathcal{I}$ will act simultaneously. Payoffs accrue.

3.2.7 Utility Functions

In stage 1, agent $i \in N$ will derive a utility level $log(l_i)$ for enjoying the leisure time $U_i^1(l_i) = Log(l_i)$.

In stage 2, agents become adults and start working. For each social group $J \in \{A, I\}$, let s_J be the average working skill of agents in J and m_J be the size of the group. Assume that an agent i with skill s_i in social group $J \in \{A, I\}$ with average skill s_J and size m_J derives a utility $f(m_J)s_J s_i$. In addition, agent i may experience costs of assimilation.

Let $U_i^2(d, a_i)$ denote the utility function of agent i in stage 2 as a function of the discrimination level d and the assimilation decisions a_i. I can fix $a_i = 0$ exogenously for any $i \in N_{\mathcal{A}}$, then the utility in stage 2 of an agent i in social group $J \in \{A, I\}$ can be written as:

$$U_i^2(d, a_i) = Log[f(m_J)s_J s_i - a_i d]$$

The agents are impatient and agents with different backgrounds have different discount factors. Let $\beta_{\mathcal{A}}$ denotes the discount factor for agents with background $\mathcal{A}$ and $\beta_{\mathcal{I}}$ denotes the discount factor for agents with background $\mathcal{I}$. I assume that $\beta_{\mathcal{A}} < \beta_{\mathcal{I}} < 1$.

Above all, the utility of an agent i with background $\mathcal{J} \in \{\mathcal{A}, \mathcal{I}\}$ in social group $J \in \{A, I\}$ can be written as:

$$U_i(l_i, d, a_i) = Log(l_i) + \beta_{\mathcal{J}} Log[f(m_J)s_J s_i - a_i d]$$

This completes the definition of game $\Gamma_{m,\beta_{\mathcal{A}},\beta_{\mathcal{I}}} = (N, S, U)$.

3.3 Solution to the Game

I will solve the game by backward induction. In stage 1, every agent i need to choose her leisure time l_i. In stage 2, the representative agent h with background $\mathcal{A}$ will choose the discrimination level d and every agent with background $\mathcal{I}$ will choose the assimilation action a_i.

Using backward induction, I will first characterize how agents make the assimilation decision in stage 2. We will then find the best choice of discrimination level d. After solving these, I will characterize the choice of leisure time for all agents.

3.3.1 Choice of Assimilation

For agents with background I, they make the assimilation choice simultaneously. With the proposition below, I can identify the structure of equilibria.

Proposition 8. *For any bounded measurable function s over N, for any discrimination level $d \in \mathbb{R}_+$, there exists $c \in (0, 1]$ and $p \in [0, 1]$ such that*

$$a_i(d, s_i) = \begin{cases} 1 & \textit{if } s_i > c \\ 0 & \textit{if } s_i < c \\ \begin{array}{c} 1 \textit{ with probability } p\textit{, and} \\ 0 \textit{ with probability } 1-p \end{array} & \textit{if } s_i = c \end{cases}$$

constitutes an equilibrium.

The proposition guarantees the existence of equilibrium but not uniqueness. In general, the uniqueness in stage 2 cannot be guaranteed since the distribution of working skills s over all agents N can be any function.

I focus on one specific functional form of the distribution of working skills s, which would be my on-path equilibrium result. The working skill distribution will take the form of

$$s_i = \begin{cases} \alpha_{\mathcal{A}}\theta_i & \text{if } i \in N_{\mathcal{A}}; \\ \alpha_{\mathcal{I}}\theta_i + s^0 & \text{if } \theta_i \geq \theta^0 \text{ and } i \in N_{\mathcal{I}}; \\ \alpha_{\mathcal{I}}\theta_i & \text{if } \theta_i < \theta^0 \text{ and } i \in N_{\mathcal{I}} \end{cases}$$

for some $\alpha_{\mathcal{A}}, \alpha_{\mathcal{I}} \in (0,1)$, $s^0 \in [0,1]$, $\theta^0 \in [0,1]$. Denote S be the set that contains all possible working skill distribution in this functional form.

Corollary 1. *For any working distribution $s \in S$, for any discrimination level $d \in \mathbb{R}_+$, there exists $c \in (0,1]$ such that*

$$a_i(d, s_i) = \begin{cases} 1 & \text{if } s_i \geq c; \\ 0 & \text{if } s_i < c \end{cases}$$

constitutes an equilibrium.

The uniqueness still cannot be guaranteed for this specific functional form since it would depend on the functional form of the network effect f. However, since there is no point mass in the working skill distribution, the equilibrium structure can be pinned down to the form above. Furthermore, the cutoff strategy simplifies my analysis because the representative of $N_{\mathcal{A}}$ can indirectly choose the cutoff θ^c by directly choosing the discrimination level d.

Denote $C(s, d)$ be a correspondence, such that for every element $c \in C(s, d)$, the action profile $a_i(d, s_i)$ with cutoff point c constitutes an equilibrium for working skill distribution $s \in S$ and discrimination level $d \in \mathbb{R}_+$.

3.3.2 Choice of Discrimination Level

The choice of discrimination level d is determined by the representative agent h with background $\mathcal{A}$. The representative agent faces the problem:

$$\max_{d \in [0,\infty)} f(m_A) s_A s_h$$

For agent h, s_h is fixed so the maximization problem would be the same as:

$$\max_{d \in [0,\infty)} f(m_A) s_A$$

For any agent $i \in N_{\mathcal{A}}$, the utility maximization problem will be the same. That is, all agents with background $\mathcal{A}$ share one preference profile. I assume that the representative h is randomly chosen from all agents with background $\mathcal{A}$. The choice of d will not be affected by choice of representative h. Thus, the mechanism of choosing h will not affect the equilibrium, as I discussed before.

The following proposition proves the existence of the d^* for any working skill distribution $s \in S$.

Proposition 9. *For any working skill distribution $s \in S$. There is a discrimination level d^* along with a cutoff $c^* \in C(s, d^*)$ such that, the representative h choose discrimination level*

d^* and agents with background $\mathcal{I}$ choose action profile

$$a_i(d^*, s_i) = \begin{cases} 1 & \text{if } s_i \geq c^*; \\ \\ 0 & \text{if } s_i < c^* \end{cases}$$

constitutes an equilibrium.

3.3.3 Choice of Working Skill

When choosing working skills, I assume that agents are sequentially rational and update their beliefs according to the Bayes' rule. I can characterize the equilibrium as follow:

Proposition 10. *At any equilibrium, $s_i = \frac{\beta_{\mathcal{J}}}{1+\beta_{\mathcal{J}}}\theta_i$ for any agent i in group J with $a_i = 0$ and $s_i = \frac{\beta_{\mathcal{I}}}{1+\beta_{\mathcal{I}}}\theta_i + s^*$ for any agent with $a_i = 1$, where $s^* = \frac{d^*}{(1+\beta_{\mathcal{I}})f(m_A)s_A}$ is a constant.*

By the proposition, the on-path choice of working skill will be in the set S at any equilibrium. Thus validating my definition of S. On the other hand, I cannot expand S to the set of any function on N since I need measurable functions to calculate the average working skill level.

Proposition 11. *There exists an equilibrium for the game $\Gamma_{m,\beta_{\mathcal{A}},\beta_{\mathcal{I}}}$.*

The uniqueness of equilibrium cannot be guaranteed, and it will highly depend on the functional form of f. Therefore, in the next section, I will solve the game with a specific f.

3.4 Numerical Example

I assume that $f(m) = m - \frac{1}{2}m^2$ as an explicit function of f. We can easily check that this function satisfies the assumptions above. We will solve the equilibrium for this specific case.

Proposition 12. *For $f(m) = m - \frac{1}{2}m^2$ and any $\beta_{\mathcal{A}}, \beta_{\mathcal{I}} \in (0,1)$, $m \in (0, \frac{1}{2})$, equilibrium exists. An equilibrium can be characterized by a pair (θ^*, s^*) where the on-path action profiles would be:*

$$s_i = \begin{cases} \frac{\beta_{\mathcal{A}}}{1+\beta_{\mathcal{A}}}\theta_i & \text{if } i \in N_{\mathcal{A}}; \\ \frac{\beta_{\mathcal{I}}}{1+\beta_{\mathcal{I}}}\theta_i + s^* & \text{if } \theta_i \geq \theta^* \text{ and } i \in N_{\mathcal{I}}; \\ \frac{\beta_{\mathcal{I}}}{1+\beta_{\mathcal{I}}}\theta_i & \text{if } \theta_i < \theta^* \text{ and } i \in N_{\mathcal{I}}. \end{cases}$$

$$d^* = \frac{1}{4}(1+m\theta^*)\left[\frac{\beta_{\mathcal{A}}}{(1+\beta_{\mathcal{A}})}(1-m) + \frac{\beta_{\mathcal{I}}}{(1+\beta_{\mathcal{I}})}m(1-(\theta^*)^2) + 2ms^*(1-\theta^*)\right](1+\beta_{\mathcal{I}})s^*.$$

$$a_i(s_i) = \begin{cases} 1 & \text{if } s_i \geq \frac{\beta_{\mathcal{I}}}{1+\beta_{\mathcal{I}}}\theta^* + s^*; \\ 0 & \text{if } s_i < \frac{\beta_{\mathcal{I}}}{1+\beta_{\mathcal{I}}}\theta^*. \end{cases}$$

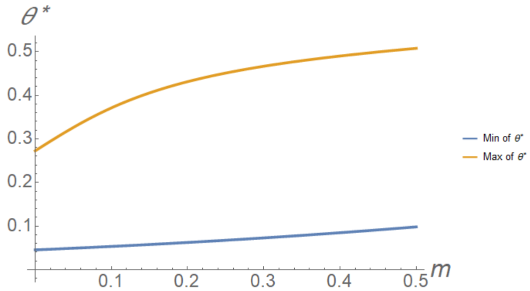
(θ^*, s^*) can be calculated by the following inequalities:

$$\begin{cases} 2P & \leq & (1+m\theta^*)^2(\gamma_{\mathcal{I}}\theta^* + s^*) \\ 2P & \geq & (1+m\theta^*)^2\gamma_{\mathcal{I}}\theta^* \\ (1+\beta_{\mathcal{I}})s^*P & \leq & (P-Q)(\gamma_{\mathcal{I}}\theta^* + s^*) \\ (1+\beta_{\mathcal{I}})s^*P & \geq & (P-Q)\gamma_{\mathcal{I}}\theta^* \end{cases}$$

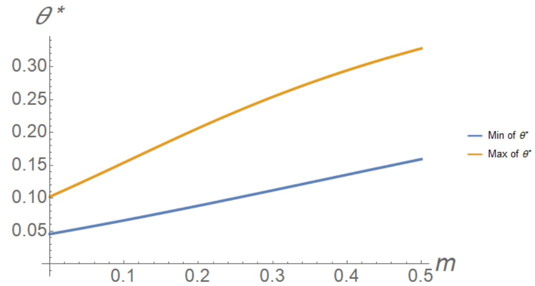
where

$$\begin{aligned} P &= \tfrac{1}{4}(1+m\theta^*)\left[\tfrac{\beta_{\mathcal{A}}}{(1+\beta_{\mathcal{A}})}(1-m) + \tfrac{\beta_{\mathcal{I}}}{(1+\beta_{\mathcal{I}})}m(1-(\theta^*)^2) + 2ms^*(1-\theta^*)\right] \\ Q &= \tfrac{\beta_{\mathcal{I}}}{4(1+\beta_{\mathcal{I}})}(2-m\theta^*)m(\theta^*)^2 \end{aligned}$$

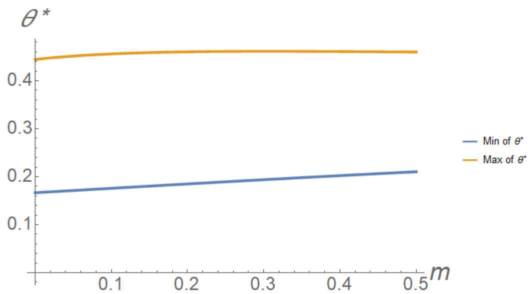
The proposition characterizes the equilibrium and provides a way to calculate it.



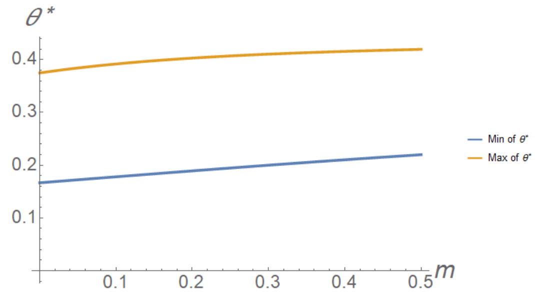
(a) Ability Cutoff ($\beta_{\mathcal{A}} = 0.1, \beta_{\mathcal{I}} = 0.2$)



(b) Ability Cutoff ($\beta_{\mathcal{A}} = 0.1, \beta_{\mathcal{I}} = 0.8$)



(c) Ability Cutoff ($\beta_{\mathcal{A}} = 0.5, \beta_{\mathcal{I}} = 0.6$)



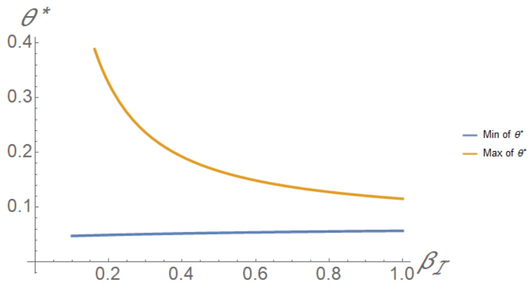
(d) Ability Cutoff ($\beta_{\mathcal{A}} = 0.5, \beta_{\mathcal{I}} = 0.8$)

Figure 3.1: The range of ability cutoff in different group sizes (m).

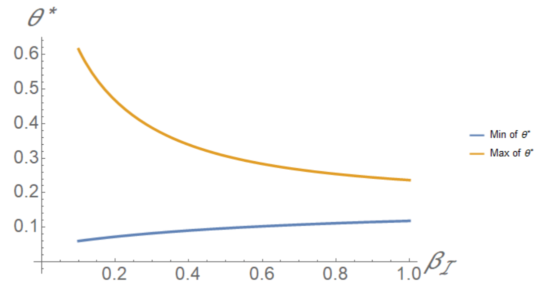
3.5 Comparative Status

Based on the equilibrium calculated in section 3.4, now I will talk about some comparative status of the equilibrium. The equilibrium is not unique with an explicit functional form of the network effect. For every pair of parameters $(\beta_{\mathcal{A}}, \beta_{\mathcal{I}}, m)$, I calculated $(d^{min}, d^{max}, \theta^{min}, \theta^{max})$. Note that there may exist an equilibrium in some conditions where two groups are separated, and no assimilation will happen. This equilibrium is not the main focus of this paper, so the discussion below will not consider this equilibrium.

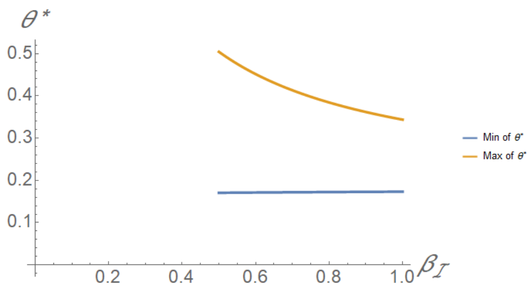
I calculated the minimum and maximum of both ability cutoff (θ^*) and the discrimination level (d^*) for some parameters. For every pair of parameters $(\beta_{\mathcal{A}}, \beta_{\mathcal{I}}, m)$, every discrimination level $d \in [d^{min}, d^{max}]$ can be achieved by some equilibrium. Similarly, every ability cutoff



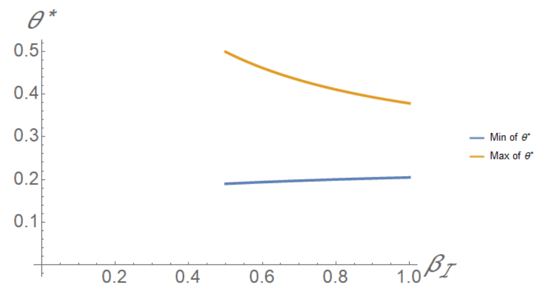
(a) Ability Cutoff ($\beta_{\mathcal{A}} = 0.1, m = 0.05$)



(b) Ability Cutoff ($\beta_{\mathcal{A}} = 0.1, m = 0.3$)



(c) Ability Cutoff ($\beta_{\mathcal{A}} = 0.5, m = 0.05$)



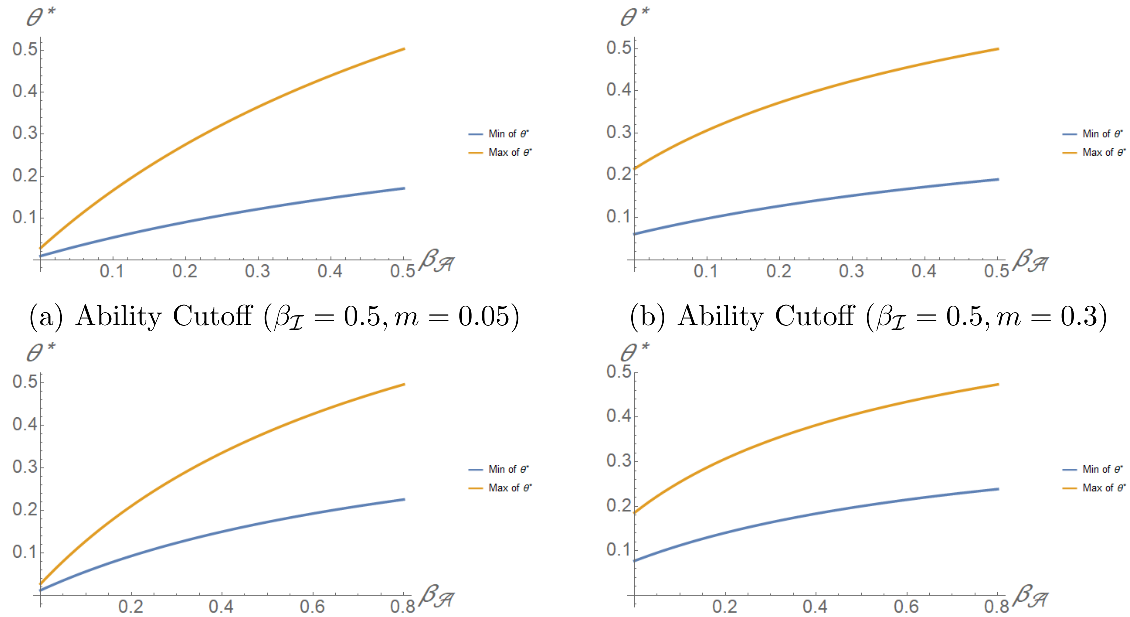
(d) Ability Cutoff ($\beta_{\mathcal{A}} = 0.5, m = 0.3$)

Figure 3.2: The range of ability cutoff in different discount factors ($\beta_{\mathcal{I}}$).

$\theta \in [\theta^{min}, \theta^{max}]$ can be achieved by some equilibrium, but (d^{min}, θ^{max}) or (d^{max}, θ^{min}) may not be achieved by some equilibrium. That is, the area in $\mathbb{R}^2$ that the pair (d, θ) can be achieved is not be a rectangle.

3.5.1 Group Size

First, I will talk about the effect of group size. By comparing Figure 3.1a to Figure 3.1d, in general, the ability cutoff θ^* is increasing as the group size become larger. When $\beta_{\mathcal{A}}$ is small, the increase of ability cutoff is very significant on both θ^{min} and θ^{max}, and when it is large, θ^{min} and θ^{max} is still increasing, but the slope is much smaller. When $\beta_{\mathcal{A}}$ is small, according to proposition 12, the average working skill level of agents with background $\mathcal{A}$ is low. If the group size of the minority group is small, the majority group would like almost all minority people to assimilate since the minority have a higher working skill level, so the ability cutoff is very small. However, if the group size of the minority group becomes larger,



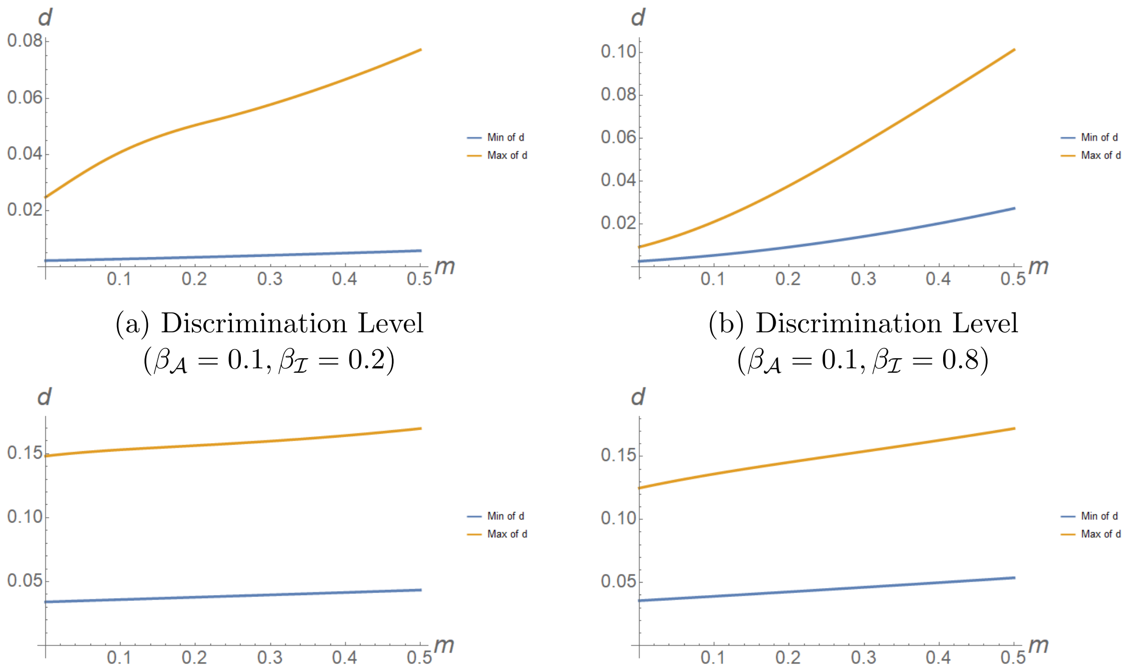
(a) Ability Cutoff ($\beta_{\mathcal{I}} = 0.5, m = 0.05$) (b) Ability Cutoff ($\beta_{\mathcal{I}} = 0.5, m = 0.3$)

(c) Ability Cutoff ($\beta_{\mathcal{I}} = 0.8, c = 0.05$) (d) Ability Cutoff ($\beta_{\mathcal{I}} = 0.8, m = 0.3$)

Figure 3.3: The range of ability cutoff in different discount factors ($\beta_{\mathcal{A}}$).

the assimilated minority will increase the average working skill level of the group A, so the ability cutoff becomes larger. The increase of ability cutoff is significant when $\beta_{\mathcal{A}}$ is small since the increase of average working skill of group A is very large due to assimilation. The increase of average working skill of group A is small when $\beta_{\mathcal{A}}$ is large, and in that case, both θ^{min} and θ^{max} increase slowly when the group size of minority (m) increases.

Then I focus on the discrimination level. By comparing Figure 3.4a to Figure 3.4d, in general both d^{min} and d^{max} increases as the group size (m) increase. When $\beta_{\mathcal{A}}$ is small, d^{max} increased significantly and when $\beta_{\mathcal{A}}$ is large, d^{max} increased slowly. This effect may have a similar reason as explained above. When the ability cutoff is small (in general), the discrimination level will be small; when the ability cutoff is large, the discrimination level will be large accordingly. The increasing speed of d^{max} is similar to the increasing speed of θ^*. Another result would be that the increase of d^{min} is significant only in Figure 3.4b where



(a) Discrimination Level ($\beta_{\mathcal{A}} = 0.1, \beta_{\mathcal{I}} = 0.2$)

(b) Discrimination Level ($\beta_{\mathcal{A}} = 0.1, \beta_{\mathcal{I}} = 0.8$)

(c) Discrimination Level ($\beta_{\mathcal{A}} = 0.5, \beta_{\mathcal{I}} = 0.6$)

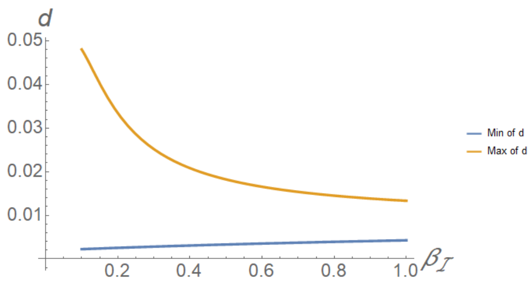
(d) Discrimination Level ($\beta_{\mathcal{A}} = 0.5, \beta_{\mathcal{I}} = 0.8$)

Figure 3.4: The range of discrimination in different group sizes (m).

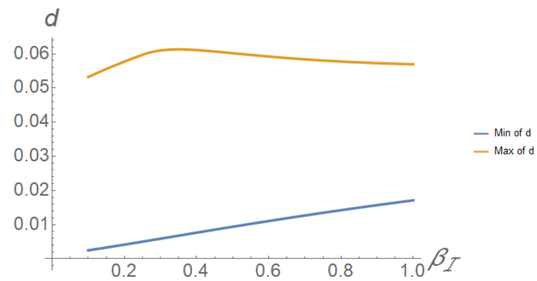
the difference between the discount factors is very large.

3.5.2 Different Discount Factors

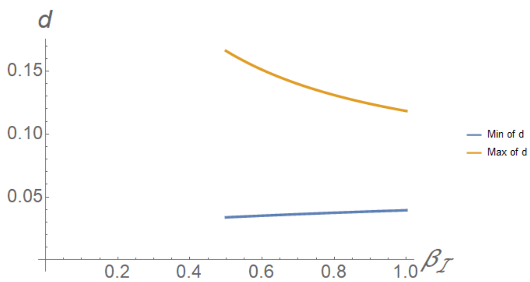
Different discount factors will affect the ability cutoff and the discrimination level differently. By comparing Figure 3.2a to Figure 3.3d, I can find several results. When $\beta_{\mathcal{I}}$ is fixed, the increasing of $\beta_{\mathcal{A}}$ will result in an increase on both θ^{min} and θ^{max}. On the other hand, when $\beta_{\mathcal{A}}$ is fixed, the increase of $\beta_{\mathcal{I}}$ will lead to a result where θ^{max} decrease θ^{min} increase. This result is very interesting. Intuitively, the ability cutoff will become smaller when the difference between discount factors becomes larger. In proposition 5, I can see that the on-path action profile is



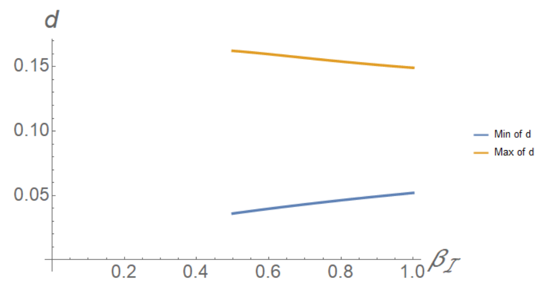
(a) Discrimination Level
($\beta_{\mathcal{A}} = 0.1, m = 0.05$)



(b) Discrimination Level
($\beta_{\mathcal{A}} = 0.1, m = 0.3$)



(c) Discrimination Level
($\beta_{\mathcal{A}} = 0.5, m = 0.05$)



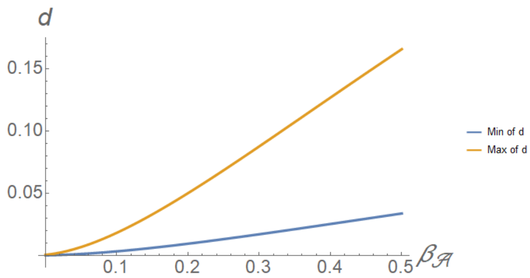
(d) Discrimination Level
($\beta_{\mathcal{A}} = 0.5, m = 0.3$)

Figure 3.5: The range of discrimination in different discount factors ($\beta_{\mathcal{I}}$).

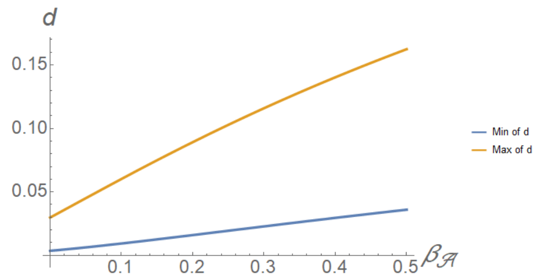
$$s_i = \begin{cases} \frac{\beta_{\mathcal{A}}}{1+\beta_{\mathcal{A}}}\theta_i & \text{if } i \in N_{\mathcal{A}}; \\ \frac{\beta_{\mathcal{I}}}{1+\beta_{\mathcal{I}}}\theta_i + s^* & \text{if } \theta_i \geq \theta^* \text{ and } i \in N_{\mathcal{I}}; \\ \frac{\beta_{\mathcal{I}}}{1+\beta_{\mathcal{I}}}\theta_i & \text{if } \theta_i < \theta^* \text{ and } i \in N_{\mathcal{I}} \end{cases}$$

There is a discontinuity at $\theta = \theta^*$. Agents with ability above the cutoff will exert extra effort to acquire an extra working skill s^*. In this way, even if the difference between the discount factor is very small, the discontinuity s^* will provide some extra working skill level so that the cutoff could be small. When $\beta_{\mathcal{A}}$ is fixed and $\beta_{\mathcal{I}}$ increased, the effect of s^* will dominate so the θ^{min} is increasing as $\beta_{\mathcal{I}}$ increasing. When $\beta_{\mathcal{I}}$ is fixed and $\beta_{\mathcal{A}}$ increases, the effect of the difference of discount factors will dominate, so θ^{min} increased as $\beta_{\mathcal{A}}$ increase.

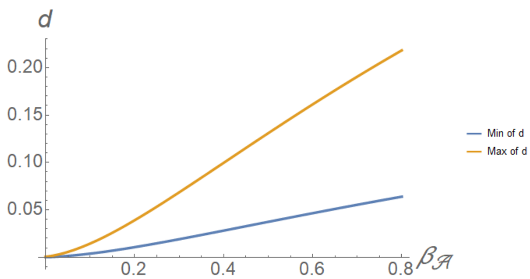
Now I will focus on the discrimination level. By comparing Figure 3.5a to Figure 3.6d,



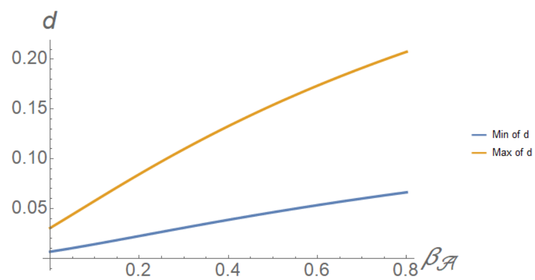
(a) Discrimination Level
($\beta_{\mathcal{I}} = 0.5, m = 0.05$)



(b) Discrimination Level
($\beta_{\mathcal{I}} = 0.5, m = 0.3$)



(c) Discrimination Level
($\beta_{\mathcal{I}} = 0.8, c = 0.05$)



(d) Ability Cutoff ($\beta_{\mathcal{I}} = 0.8, m = 0.3$)

Figure 3.6: The range of discrimination in different discount factors ($\beta_{\mathcal{A}}$).

I can find similar result as last paragraph. When fixed $\beta_{\mathcal{I}}$, both d^{min} and d^{max} increases as $\beta_{\mathcal{A}}$ increase. This is when the effect of difference between discount factors dominates. When $\beta_{\mathcal{A}}$ is fixed, in general, when $\beta_{\mathcal{I}}$ increases, d^{min} will increase and d^{max} will decrease. The increasing of d^{min} is because the effect of s^* (discontinuity of working skill level) dominates. There is another interesting result, that is when $\beta_{\mathcal{A}} = 0.1$, $m = 3$, d^{max} will first increase then decrease. This may be the effect of the combination of two effects.

3.6 Testable Results

3.6.1 Group Size Change

As explained in the last section, when the group size becomes large, the discrimination level will be higher than when the group size is small. The migration process may serve as empirical data of this change. When migration just started, the population of the minority group in a community was very small. Therefore, the discrimination against them should be small. As more and more minority people migrate to the community, the discrimination level should be larger than before. The discrimination level could be captured by the number of conflicts between majority and minority or the number of marriages between majority and minority groups.

3.6.2 Discount Factor Change

People who have different discount factors may experience different discrimination levels. For example, Jewish people are a minority group in many countries, and they share the same culture. Assume they are the minority group and they share the same $\beta_{\mathcal{I}}$ around the world. By comparing the discrimination level of Jews worldwide, I should see high discrimination levels in countries with high discount factors (a culture that puts more weight on future utility).

3.7 Further Discussion

As discussed before, there always exist an equilibrium that two groups remain separated. It is easily to check that when $f(1-m)\frac{\beta_{\mathcal{A}}}{1+\beta_{\mathcal{A}}} \leq f(m)\frac{\beta_{\mathcal{I}}}{1+\beta_{\mathcal{I}}}$, the separation equilibrium exists. The existence of this separation equilibrium will provide more interesting results for the model.

Another extension would be the study of the evolution of the group size. With assimilation, group size will change according to time. Different groups will have different growth rates, and the speed of assimilation will depend on the difference between discount factors.

A third extension would use a more general functional form of the network effect. Again, the equilibrium will generally exist, but the change in discrimination level and ability cutoff will vary across different functional forms.

3.8 Conclusion

In this paper, I construct a 2-stage game model to explain the difference in discrimination levels across different scenarios. There are several main results. First, there exists an equilibrium for any discount factors and minority group size, and the equilibrium will have an on-path action profile with a cutoff rule. Second, as group size increases, both the discrimination level and the ability cutoff will increase. Third, when discount factors vary across different regimes, there are two effects that drive the discrimination level and ability cutoff in opposite directions. When $\beta_{\mathcal{I}}$ is fixed, the larger the difference between discount factors, the larger the discrimination level and ability cutoff. When $\beta_{\mathcal{A}}$ is fixed, the two effects are mixed, and there are no general results for the discrimination level and ability cutoff.

APPENDICES

Appendix A: Proofs for Chapter 1

This article contains the proofs omitted in the text.

Proof of Proposition 1. Under individual assignment, the optimal contracting program is given as:

$$\max\ \Pi^I := \mathbb{E}\left[Y_A - w_1^I\left(M_A\right)\ \mid \mathbf{e}_A,\ X_P^A\right] + \mathbb{E}\left[Y_B - w_2^I\left(M_B\right)\ \mid \mathbf{e}_B,\ X_P^B\right]$$

$$s.t.$$

$$\mathbf{e}_j = \arg\max\nolimits_{e'_{j1},e'_{j2}}\ U_i^I\left(\mathbf{e}'_j, X_P^j\right)\quad \forall\ j \tag{IC_I}$$

$$U_i^I\left(\mathbf{e}_j, X_P^j\right) \geq 0 \tag{IR_I}$$

By standard argument, (IR_I) must bind, and any effort profile can be implemented (i.e., made to satisfy the (IC_I)) by choosing the wage schedules $w_1^I\left(M_A\right)$ and $w_2^I\left(M_B\right)$ appropriately. Thus, the program boils down to:

$$\max_{\mathbf{e}_A,\mathbf{e}_B} \sum_{j\in\{A,B\}} \left[\mathbb{E}\left[Y_j \mid \mathbf{e}_j, X_P^j\right] - \frac{1}{2}e_{j1}^2 - \frac{1}{2}e_{j2}^2\right].$$

Denote $\pi(X_P^j) := \max_{\mathbf{e}_j}\ \mathbb{E}\left[Y_j \mid \mathbf{e}_j, X_P^j\right] - \frac{1}{2}e_{j1}^2 - \frac{1}{2}e_{j2}^2$, and it is routine to check:

$$\pi(X_P^j) := \begin{cases} (1+(2\alpha-\alpha^2)(2\alpha-\alpha^2-\frac{1}{2}))\underline{\pi} & if\ X_P^j = \{g\} \\ (1+\alpha-\frac{1}{2}\alpha^2)\underline{\pi} & if\ X_P^j = \{g,\emptyset\} \\ \underline{\pi} & if\ X_P^j = \{g,\emptyset,b\} \end{cases}.$$

Comparing the values, we obtain that the optimal $X_P^j = \{g,\emptyset\}$ for all j. That is, under individual assignment, in the optimal contract the principal proceeds with project j if and only if the bad state is not observed, and obtains payoff $S^{*}=\left(1+\alpha-\frac{1}{2}\alpha^{2}\right)\underline{\pi}$.

Similarly, under team assignment, the optimal contracting program is give as:

$$\max\ \Pi^{T}:=\sum\nolimits_{j\in\{A,B\}}\mathbb{E}\left[Y_{j}-\left(w_{1}^{T}\left(M_{j}\right)+w_{2}^{T}\left(M_{j}\right)\right)\ \mid \mathbf{e}_{1},\mathbf{e}_{2},X_{P}^{A},X_{P}^{B}\right]$$

$$s.t.$$

$$\mathbf{e}_{i}=\arg\max_{\mathbf{e}_{i}'}\ U_{i}^{T}\left(\mathbf{e}_{i}',\mathbf{e}_{-i},X_{P}^{A},X_{P}^{B}\right)\quad\forall i \qquad (IC_{T})$$

$$U_{i}^{T}\left(\mathbf{e}_{i},\mathbf{e}_{-i},X_{P}^{A},X_{P}^{B}\right)\geq 0\ \ \forall i \qquad (IR_{T})$$

As in the case of individual assignment, we can plug (IR_{T}) in the objective function and ignore the (IC_{T}); the program boils down to:

$$\max_{e_{A1},e_{A2};\ e_{B1},e_{B2}}\ \sum_{j\in\{A,B\}}\mathbb{E}\left[Y_{j}-\frac{1}{2}e_{j1}^{2}-\frac{1}{2}e_{j2}^{2}\mid e_{j1},e_{j2},X_{P}^{j}\right].$$

Thus, given X_{P}^{A} and X_{P}^{B}, the principal's payoff is exactly the same as that in the case of individual assignment, and so claim follows.

Q.E.D.

Proof of Lemma 1. Note that there are four possible reporting policies: for each $x\in\{G,B\}$, $r=\rho(x)=x$ or $\emptyset$ (and $\rho(\emptyset)=\emptyset$); and eight possible continuationpolicies: for each $r\in\{G,\emptyset,B\}$, $\mathcal{C}(r)=$ *cancel* or *proceed.*

Step 1. *Without loss of generality we can consider only two continuation policies.* Trivially, $\mathcal{C}(r)=$ *cancel* $\forall r$ yields a payoff of $\underline{\pi}$ (principal's outside option), and $\mathcal{C}(r)=$ *proceed*

$\forall r$ also yields $\underline{\pi}$ (by Assumption 1). Also, as under any reporting policy,

$$\Pr\left(\omega = G \mid r = G\right) \geq \Pr\left(\omega = G \mid r = \emptyset\right) \geq \Pr\left(\omega = G \mid r = B\right),$$

in equilibrium, if $\mathcal{C}\left(B\right) = \textit{proceed}$ then $\mathcal{C}\left(r\right) = \textit{proceed}$ for all r, and if $\mathcal{C}\left(\emptyset\right) = \textit{proceed}$ then $\mathcal{C}\left(G\right) = \textit{proceed}$. Thus, without loss of generality, we can focus on equilibria that supports only one of the following two continuation policies: (i) $\mathcal{C}\left(r\right) = \textit{proceed}$ only if $r = G$, and (ii) $\mathcal{C}\left(r\right) = \textit{proceed}$ only if $r \in \{G, \emptyset\}$.

Step 2. *For each of the two continuation policies stated in Step 1, only one reporting policy may be played in equilibrium.*

Step 2a. Suppose, in the optimal contract, the principal's continuation policy (i), i.e., $\mathcal{C}\left(r\right) = \textit{proceed}$ if and only if $r = G$. The two reporting policies of the agent where $\rho\left(G\right) = \emptyset$ (and $\rho\left(B\right) = B$ or $\emptyset$) yield the same payoff as the project gets cancelled under both policies. Also, the two reporting policies, $\rho\left(G\right) = G$ and $\rho\left(B\right) = B$ or $\emptyset$, yield the same payoff. But the policy $\rho\left(G\right) = G$ and $\rho\left(x\right) = \emptyset$ if $x \in \{\emptyset, B\}$ relaxes the principal's incentive constraints relative to the policy $\rho\left(x\right) = x$ for all $x \in \{G, B\}$ as

$$\Pr\left(\omega = G \mid x = \emptyset\right) \geq \Pr\left(\omega = G \mid x \in \{\emptyset, B\}\right) \geq \Pr\left(\omega = G \mid x = B\right).$$

Thus, if in the optimal contract continuation policy (i) is played, then without loss of generality, we assume that the associated reporting policy is $\rho\left(G\right) = G$ and $\rho\left(x\right) = \emptyset$ if $x \in \{\emptyset, B\}$.

Step 2b. Now suppose in the optimal contract continuation policy (ii), i.e., $\mathcal{C}\left(r\right) =$

proceed if and only if $r \in \{G, \emptyset\}$, is played. The two reporting policies of the agent where $\rho(B) = \emptyset$ (and $\rho(G) = G$ or $\emptyset$) yield the same payoff as the project always proceeds. Also, the two reporting policies, $\rho(B) = B$ and $\rho(G) = G$ or $\emptyset$, yield the same payoff. But the policy $\rho(B) = B$ and $\rho(x) = \emptyset$ if $x \in \{G, \emptyset\}$ relaxes the incentive constraints relative to the policy $\rho(x) = x$ for all $x \in \{G, B\}$. Thus, if in the optimal contract continuation policy (ii) is played, then without loss of generality, we assume that the associated reporting strategy is $\rho(G) = G$ and $\rho(x) = \emptyset$ if $x \in \{\emptyset, B\}$.

Together, the observations in Steps 1 and 2 imply that, without loss of generality, we can limit attention to two communication protocols: (i) If the state is observed to be G, report G, otherwise report $\emptyset$; proceed with the project if and only if $r = G$. (ii) If the state is observed to be B, report B, otherwise report $\emptyset$; proceed with the project if and only if $r \neq B$.

Q.E.D.

Proof of Lemma 2. For brevity, we rewrite the objective function and all constraints by using the notations p^I, $p^I_\emptyset$, P^I and $P^I_\emptyset$ (as defined in Section 1.4.1), and the program $\mathcal{P}^I$ boils down to:

$$\max_{\substack{w_F,\Delta_C,\Delta_S,\\ e_1,e_2}} \quad \Pi^I := p^I[\left[\Pr(\omega = G \mid x \in X_P)y - P^I\Delta_S\right]\sum_k e_k$$

$$+(1-p^I)[\underline{\pi} - \Delta_C] - w_F$$

$$s.t.$$

$$p^I P^I \Delta_S \sum_k e_k + (1-p^I)\Delta_C + w_F - \tfrac{1}{2}\sum_k e_k^2 \geq 0 \qquad (IR^I)$$

$$[\Pr(\omega = G \mid x \in X_P)y - P^I\Delta_S]\sum_k e_k \geq \underline{\pi} - \Delta_C \qquad (IC_P^I\text{-}1)$$

$$[\Pr(\omega = G \mid x \notin X_P)y - P_C^I\Delta_S]\sum_k e_k \leq \underline{\pi} - \Delta_C \qquad (IC_P^I\text{-}2)$$

$$e_k = p^I P^I \Delta_S, \quad k = 1,2 \qquad (IC_A^I\text{-}1)$$

$$\left[\left(p^I P^I\right)^2 - \left(p_\emptyset^I P_\emptyset^I\right)^2\right]\Delta_S^2 \geq \left(p^I - p_\emptyset^I\right)\Delta_C. \qquad (IC_A^I\text{-}2)$$

By standard argument, we claim that $\left(IR^I\right)$ binds. Using (IR) and $(IC_A^I\text{-}1)$ we can eliminate w_F and e_is and the program can be further simplifies to:

$$\max_{\Delta_C,\Delta_S} \quad 2\left(p^I\right)^2 P^I \Pr(\omega = G \mid x \in X_P)y\Delta_S + (1-p^I)\underline{\pi} - \left(p^I P^I \Delta_S\right)^2$$

$$s.t.$$

$$[\Pr(\omega = G \mid x \in X_P)y - P^I\Delta_S]\left(2p^I P^I \Delta_S\right) \geq \underline{\pi} - \Delta_C \qquad (IC_P^I\text{-}1)$$

$$[\Pr(\omega = G \mid x \notin X_P)y - P_C^I\Delta_S]\left(2p^I P^I \Delta_S\right) \leq \underline{\pi} - \Delta_C \qquad (IC_P^I\text{-}2)$$

$$\left[\left(p^I P^I\right)^2 - \left(p_\emptyset^I P_\emptyset^I\right)^2\right]\Delta_S^2 \geq \left(p^I - p_\emptyset^I\right)\Delta_C \qquad (IC_A^I\text{-}2)$$

Case 1: $X_P = \{G, \emptyset\}$. Here $p^I_\emptyset = 1$, and the program becomes:

$$\mathcal{P}^I_{\{g,\emptyset\}}: \begin{cases} \max_{\Delta_C,\Delta_S} \quad 2\left(p^I\right)^2 P^I \Pr(\omega = G|x \in X_P)y\Delta_S + (1-p^I)\underline{\pi} - \left(p^I P^I \Delta_S\right)^2 \\ s.t. \\ \Delta_C \geq l_P := \underline{\pi} - [\Pr(\omega = G|x \in X_P)y - P^I\Delta_S]\left(2p^I P^I \Delta_S\right) & (IC^I_P\text{-}1) \\ \Delta_C \leq u_P := \underline{\pi} - [\Pr(\omega = G|x \notin X_P)y - P^I_C\Delta_S]\left(2p^I P^I \Delta_S\right) & (IC^I_P\text{-}2) \\ \Delta_C \geq l_A := \left[\left(P^I_\emptyset\right)^2 - \left(p^I P^I\right)^2\right]\frac{\Delta_S^2}{1-p^I} & (IC^I_A\text{-}2) \end{cases}$$

As Δ_C does not appear in the objective function, we can rewrite the program as:

$$\begin{cases} \max_{\Delta_S} \quad 2\left(p^I\right)^2 P^I \Pr(\omega = G \mid x \in X_P)y\Delta_S + (1-p^I)\underline{\pi} - \left(p^I P^I \Delta_S\right)^2 \\ s.t. \\ u_P \geq l_P \Leftrightarrow \underline{\pi} \geq \left[2p^I \Pr(\omega = G \mid x \notin X_P)P^I y\right]\Delta_S - \left[2p^I P^I P^I_C - \frac{\left(P^I_\emptyset\right)^2 - \left(p^I P^I\right)^2}{1-p^I}\right]\Delta_S^2 \\ u_P \geq l_A \Leftrightarrow \Delta_S \leq \frac{y}{1-\mu} \end{cases}.$$

By routine calculation one obtains $\Pr(\omega = G|x \in X_P) = 1/(2-\alpha')$, $\Pr(\omega = G|x \notin X_P) = 0$, and

$$p^I = 1 - \frac{1}{2}\alpha', \;\; P^I = \frac{1}{2-\alpha'} + \mu\left(1 - \frac{1}{2-\alpha'}\right), \;\; p^I_C = \mu,$$

where $\alpha' := 1 - (1-\alpha)^2$. Also, to streamline notation, without loss of generality, we set $y = 1$ (thus, by Assumption 1, $\underline{\pi} = \frac{1}{4}$). So, the program $\mathcal{P}^I_{\{g,\emptyset\}}$ boils down to:

$$\begin{cases} \max_{\Delta_S} \quad -\left[\frac{1}{2}(1+\mu(1-\alpha'))\Delta_S - \frac{1}{2}\right]^2 + \frac{1}{4} + \frac{1}{8}\alpha' \\ s.t. \\ \frac{1}{4} \geq \frac{1}{2}\left(\alpha'\Delta_S\right)^2 \;\; and \;\; \frac{1}{1-\mu} \geq \Delta_S \end{cases}.$$

Notice that the objective function is strictly concave with peak at $\frac{1}{1+\mu(1-\alpha')}$ and the feasible set is always non-empty. Thus the solution always exists and is given as:

$$\Delta_S^* = \begin{cases} \frac{1}{1+\mu(1-\alpha')} & if\ \frac{\alpha'\mu^2}{(1+\mu(1-\alpha'))^2} \leq \frac{1}{2} \\ \frac{1}{\mu\sqrt{2\alpha'}} & otherwise \end{cases}.$$

The associated value is:

$$V^I_{\{g,\emptyset\}} = \begin{cases} \frac{1}{4} + \frac{1}{8}\alpha' & if\ \frac{\alpha'\mu^2}{(1+\mu(1-\alpha'))^2} \leq \frac{1}{2} \\ \frac{1}{4} + \frac{1}{8}\alpha' - \frac{1}{2}\left[\frac{1}{\mu\sqrt{2\alpha'}}(1+\mu(1-\alpha')) - 1\right]^2 & otherwise \end{cases}. \tag{1}$$

Case 2: $X_P = \{G\}$. Here $p^I_\emptyset = 0$, and the program becomes:

$$\mathcal{P}^I_{\{g\}} : \begin{cases} \max_{\Delta_C, \Delta_S} \quad 2\left(p^I\right)^2 P^I \Pr(\omega = G \mid x \in X_P) y \Delta_S + (1-p^I)\underline{\pi} - \left(p^I P^I \Delta_S\right)^2 \\ s.t. \\ \Delta_C \geq l_P := \underline{\pi} - [\Pr(\omega = G \mid x \in X_P) y - P^I \Delta_S]\left(2p^I P^I \Delta_S\right) & (IC^I_P\text{-}1) \\ \Delta_C \leq u_P := \underline{\pi} - [\Pr(\omega = G \mid x \notin X_P) y - P^I_C \Delta_S]\left(2p^I P^I \Delta_S\right) & (IC^I_P\text{-}2) \\ \Delta_C \leq u_A := p^I \left(P^I\right)^2 \Delta_S^2 & (IC^I_A\text{-}2) \end{cases}.$$

As in Case 1, Δ_C does not enter into the objective function, and we can further simplify the program as:

$$\begin{cases} \max\limits_{\Delta_S} \quad 2\left(p^I\right)^2 P^I \Pr(\omega = G \mid x \in X_P) y \Delta_S + (1-p^I)\underline{\pi} - \left(p^I P^I \Delta_S\right)^2 \\ s.t. \\ l_P \leq u_A \Leftrightarrow \underline{\pi} \leq \left[2p^I \Pr(\omega = G \mid x \in X_P) P^I y\right] \Delta_S - \left[p^I \left(P^I\right)^2\right] \Delta_S^2 \\ l_P \leq u_P \Leftrightarrow \Delta_S \leq \frac{y}{1-\mu} \end{cases},$$

and plugging the values for the probablities and setting $y = 1$ we obtain:

$$\begin{cases} \max\limits_{\Delta_S} \quad \frac{1}{2}\alpha'^2 \Delta_S (1-\Delta_S) + \frac{1}{4}(1-\frac{1}{2}\alpha') \\ s.t. \\ \alpha' \Delta_S \left(1 - \frac{1}{2}\Delta_S\right) \leq \frac{1}{4} \textit{ and } \Delta_S \leq \frac{1}{1-\mu} \end{cases}.$$

The feasible set is non-empty if and only if $\alpha' \geq 1/2$ (equivalently, $\alpha \geq 1 - 1/\sqrt{2}$), and the objective function is concave with peak at 1. Thus, the solution of the program and the value would be:

$$\Delta_S^* = 1 \text{ and } V^I_{\{g\}} = \frac{1}{4} + \frac{1}{4}\alpha'\left(\alpha' - \frac{1}{2}\right) \text{ if } \alpha' \geq \frac{1}{2} \tag{2}$$

and no solution otherwise.

Q.E.D.

Proof of Lemma 4. The proof is similar to that of Lemma 2. For brevity, we rewrite the objective function and all constraints by using the notations p^T, $p^T_\emptyset$, P^T and $P^T_\emptyset$ (as defined in Section 1.4.2), and the program $\mathcal{P}^T$ boils down to:

$$\max_{\substack{\Delta_{iC},\Delta_{iS}\\ w_{iF},e_i}} \ \Pi^T = p^T\left[\Pr(\omega = G \mid x^T \in P)y - P^T\sum_i \Delta_{iS}\right]\sum_k e_k$$
$$+(1-p^T)\left[\underline{\pi} - \sum_i \Delta_{iC}\right] - \sum_i w_{iF}$$

$$s.t. \ \ \forall i \in \{1,2\}$$

$$p^T P^T \Delta_{iS}\sum_k e_k + (1-p^T)\Delta_{iC} + w_{iF} - \tfrac{1}{2}e_i^2 \geq 0 \qquad (IR_i^T)$$

$$\left[\Pr(\omega = G \mid x^T \in X_P)y - P^T\sum_i \Delta_{iS}\right]\sum_k e_k \geq \underline{\pi} - \sum_i \Delta_{iC} \qquad (IC_P^T\text{-}1)$$

$$\left[\Pr(\omega = G \mid x^T \notin X_P)y - P_C^T\sum_i \Delta_{iS}\right]\sum_k e_k \leq \underline{\pi} - \sum_i \Delta_{iC} \qquad (IC_P^T\text{-}2)$$

$$e_i = p^T P^T \Delta_{iS} \qquad (IC_{A_i}^T\text{-}1)$$

$$\tfrac{1}{2}\left[\left(p^T P^T\right)^2 - \left(p_\emptyset^T P_\emptyset^T\right)^2\right]\Delta_{iS}^2 + \left[\left(p^T P^T\right)^2 - \left(p_\emptyset^T P_\emptyset^T\right)\left(p^T P^T\right)\right]\Delta_{iS}\Delta_{jS}$$
$$\geq (p^T - p_\emptyset^T)\Delta_{iC} \qquad (IC_{A_i}^T\text{-}2)$$

We can eliminate w_{iF} and e_is using (IR_i^T) (that must bind) and $(IC_{A_i}^T\text{-}1)$, and the program further simplifies to:

$$\max_{\Delta_{iC},\Delta_{iS}} \quad \left(p^T\right)^2 P^T \Pr(\omega = G \mid x^T \in X_P) y \sum_i \Delta_{iS} + (1-p^T)\underline{\pi} - \tfrac{1}{2}\left(p^T P^T\right)^2 \sum_i \Delta_{iS}^2$$

$$s.t.$$

$$\left[\Pr(\omega = G \mid x^T \in X_P) y - P^T \sum_i \Delta_{iS}\right] p^T P^T \sum_i \Delta_{iS} \geq \underline{\pi} - \sum_i \Delta_{iC} \tag{IC_P^T-1}$$

$$\left[\Pr(\omega = G \mid x^T \notin X_P) y - P_C^T \sum_i \Delta_{iS}\right] p^T P^T \sum_i \Delta_{iS} \leq \underline{\pi} - \sum_i \Delta_{iC} \tag{IC_P^T-2}$$

$$\begin{aligned} &\tfrac{1}{2}\left[\left(p^T P^T\right)^2 - \left(p_\emptyset^T P_\emptyset^T\right)^2\right](\Delta_{1S})^2 + \left[\left(p^T P^T\right)^2 - p_\emptyset^T P_\emptyset^T p^T P^T\right]\Delta_{1S}\Delta_{2S} \\ &\qquad \geq (p^T - p_\emptyset^T)\Delta_{1C} \end{aligned} \tag{$IC_{A_1}^T$-2}$$

$$\begin{aligned} &\tfrac{1}{2}\left[\left(p^T P^T\right)^2 - \left(p_\emptyset^T P_\emptyset^T\right)^2\right](\Delta_{2S})^2 + \left[\left(p^T P^T\right)^2 - p_\emptyset^T P_\emptyset^T p^T P^T\right]\Delta_{1S}\Delta_{2S} \\ &\qquad \geq (p^T - p_\emptyset^T)\Delta_{2C} \end{aligned} \tag{$IC_{A_2}^T$-2}$$

Part (i). We now prove that if $\mathcal{P}^T$ admits a solution, it also admits a symmetric solution where $\Delta_{1S} = \Delta_{2S} = \Delta_S$ and $\Delta_{1C} = \Delta_{2C} = \Delta_C$. The proof is given in the following five steps.

Step 1: Suppose $\Delta^* := (\Delta_{1S}^*, \Delta_{2S}^*, \Delta_{1C}^*, \Delta_{2C}^*)$ is a solution to $\mathcal{P}_T$. If $\Delta_{1S}^* = \Delta_{2S}^* = \Delta_G^*$ (say), we argue that there also exists a symmetric solution $(\Delta_S^*, \Delta_S^*, \Delta_C^*, \Delta_C^*)$ where

$$\Delta_C^* = \frac{1}{2}\sum_i \Delta_{iC}^*.$$

To see this, notice that under Δ^*, $(IC^T_{A_i}\text{-}2)$s imply:

$$\begin{aligned}\left[\tfrac{3}{2}\left(p^T P^T\right)^2 - \tfrac{1}{2}\left(p^T_\emptyset P^T_\emptyset\right)^2 - p^T_\emptyset P^T_\emptyset p^T P^T\right](\Delta^*_S)^2 &\geq \max\{(p^T - P^T_\emptyset)\Delta^*_{1C}, (p^T - P^T_\emptyset)\Delta^*_{2C}\} \\ &\geq \tfrac{1}{2}(p^T - P^T_\emptyset)(\Delta^*_{1C} + \Delta^*_{2C}) \\ &= (p^T - P^T_\emptyset)\Delta^*_C.\end{aligned}$$

Thus, $(\Delta^*_S, \Delta^*_S, \Delta^*_C, \Delta^*_C)$ is also a solution as it satisfies $(IC^T_{A_i}\text{-}2)$ and does not affect $(IC^T_P\text{-}1)$ and $(IC^T_P\text{-}2)$.

Step 2: Denote

$$\begin{aligned}\Pi^T(\Delta_{1S}, \Delta_{2S}) \;&:= \left(p^T\right)^2 P^T \Pr(\omega = G|x^T \in X_P)y \sum_i \Delta_{iS} + (1 - p^T)\underline{\pi} \\ &\quad -\tfrac{1}{2}\left(p^T P^T\right)^2 \sum_i \Delta^2_{iS}.\end{aligned}$$

Suppose $\Delta^* := (\Delta^*_{1S}, \Delta^*_{2S}, \Delta^*_{1C}, \Delta^*_{2C})$ is a solution to $\mathcal{P}_T$ but $\Delta^*_{1S} \neq \Delta^*_{2S}$. Without loss of generality, assume $\Delta^*_{1S} > \Delta^*_{2S}$. We argue that then Δ^* cannot be a solution. In particular, there exists $\varepsilon > 0$ and cancellation premiums Δ'_{iC}s such that $(\Delta^*_{1S} - \varepsilon, \Delta^*_{2S} + \varepsilon, \Delta'_{1C}, \Delta'_{2C})$ is feasible and

$$\Pi^T(\Delta^*_{1S} - \varepsilon, \Delta^*_{2S} + \varepsilon) > \Pi^T(\Delta^*_{1S}, \Delta^*_{2S}).$$

Observe that $\Pi_T(\Delta_{1S}, \Delta_{2S})$ is symmetric and concave in $(\Delta_{1S}, \Delta_{2S})$ with peak at

$$\Delta_{1S} = \Delta_{2S} = \frac{y}{P^T}\Pr\left(\omega = G|x^T \in X_P\right).$$

Also, the following holds: take any $(\Delta_{1S}, \Delta_{2S})$ such that $\Delta_{1S} \neq \Delta_{2S}$, $\Delta_{1S} > \Delta_{2S}$, say. Then, there exists $\varepsilon > 0$ such that

$$\Pi^T \left(\Delta_{1S} - \varepsilon, \Delta'_{2S} + \varepsilon\right) > \Pi^T \left(\Delta_{1G}, \Delta_{2G}\right).$$

So, we only need to show that there exists an $\varepsilon > 0$, and $\Delta'_{1C}, \Delta'_{2C}$ values such that $(\Delta^*_{1S} - \varepsilon, \Delta^*_{2S} + \varepsilon, \Delta'_{1C}, \Delta'_{2C})$ is feasible. In order to prove this claim, it is worthwhile to first establish a few properties of the ($IC^T_{A_i}$-2) constraints, as given in the next step.

Step 3: Denote

$$L_i(\Delta_{1S}, \Delta_{2S}) := A\left(\Delta_{iS}\right)^2 + B\Delta_{1S}\Delta_{2S},$$

where

$$A := \frac{\left(p^T P^T\right)^2 - \left(p^T_\emptyset P^T_\emptyset\right)^2}{2\left(p^T - p^T_\emptyset\right)} \text{ and } B := \frac{\left(p^T P^T - p^T_\emptyset P^T_\emptyset\right) p^T P^T}{p^T - p^T_\emptyset}.$$

Note that the ($IC^T_{A_i}$-2) constraints can be written as:

$$L_i\left(\Delta_{1S}, \Delta_{2S}\right) \geq \Delta_{iC} \text{ if } p^T - p^T_\emptyset > 0, \text{ and } L_i\left(\Delta_{1S}, \Delta_{2S}\right) \leq \Delta_{iC} \text{ otherwise.}$$

Also,

$$(B - A) = \frac{\left(p^T P^T - p^T_\emptyset P^T_\emptyset\right)^2}{2\left(p^T - p^T_\emptyset\right)},$$

and hence,

$$\operatorname{sign}(B - A) = \operatorname{sign}\left(p^T - p^T_\emptyset\right)$$

It is routine to check that for $X_P = \{g\}$, $p^T - p^T_\emptyset > 0$ and $p^T P^T - p^T_\emptyset P^T_\emptyset > 0$, whereas for

$X_P = \{g, \emptyset\}$, $p^T - p^T_\emptyset < 0$ and $p^T P^T - p^T_\emptyset P^T_\emptyset < 0$. Thus,

$$A > 0, B > 0.$$

In the next two steps, we consider the two cases $p^T - p^T_\emptyset > 0$ and < 0, and show that the claim in Step 2 above holds in both cases.

Step 4: Suppose $p^T - p^T_\emptyset > 0$. So, ($IC^T_{A_i}$-2)s are given as:

$$L_i(\Delta_{1S}, \Delta_{2S}) \geq \Delta_{iC}.$$

There are three possibilities:

Case 1: Both ($IC^T_{A_i}$-2)s are slack at $(\Delta^*_{1S}, \Delta^*_{2S}, \Delta^*_{1C}, \Delta^*_{2C})$. Consider the solution

$$(\Delta^*_{1S} - \varepsilon, \Delta^*_{2S} + \varepsilon, \Delta^*_{1C}, \Delta^*_{2C})$$

where $\varepsilon > 0$. This solution leaves $\left(IC^T_P\right)$s unaffected, for sufficiently small ε, both ($IC^T_{A_i}$-2)s remain slack, and yields a higher value of Π^T(from Step 2).

Case 2: Exactly one of the two ($IC^T_{A_i}$-2)s is slack at $(\Delta^*_{1S}, \Delta^*_{2S}, \Delta^*_{1C}, \Delta^*_{2C})$. Suppose only ($IC^T_{A_1}$-2) is slack, say, (hence, (IC_{A_2}-2) is binding). Set

$$\Delta'_{1C} = \Delta^*_{1C} + \delta,\ \Delta'_{2C} = \Delta^*_{2C} - \delta$$

where $\delta > 0$. For δ sufficiently small, at $(\Delta^*_{1S}, \Delta^*_{2S}, \Delta'_{1C}, \Delta'_{2C})$, both ($IC^T_{A_i}$-2) become slack and $\left(IC_P^T\right)$s are unaffected, and hence, it is feasible. But then, as argued in Case 1, the solution $(\Delta_{1S}^* - \varepsilon, \Delta_{2S}^* + \varepsilon, \Delta_{1C}', \Delta_{2C}')$ is also feasible for $\varepsilon > 0$ sufficiently small, and attains a higher value of Π^T.

Case 3: Both $(IC_{A_i}^T$-2)s are binding at $(\Delta_{1S}^*, \Delta_{2S}^*, \Delta_{1C}^*, \Delta_{2C}^*)$. Consider changing $(\Delta_{1S}^*, \Delta_{2S}^*)$ to $(\Delta_{1S}^* - \varepsilon, \Delta_{2S}^* + \varepsilon)$. The left-hand side of $\left(IC_{A_i}^T\text{-}2\right)$ changes by

$$\delta_i := L_i\left(\Delta_{1S}^* - \varepsilon, \Delta_{2S}^* + \varepsilon\right) - L_i\left(\Delta_{1S}^*, \Delta_{2S}^*\right)$$

where

$$\begin{aligned}
\delta_1 &= -\varepsilon\left(2A\left(\Delta_{1S}^* - \frac{1}{2}\varepsilon\right) - B\left(\Delta_{1S}^* - \left(\Delta_{2S}^* + \varepsilon\right)\right)\right), \\
\delta_2 &= \varepsilon\left(2A\left(\Delta_{2S}^* + \frac{1}{2}\varepsilon\right) + B\left(\Delta_{1S}^* - \left(\Delta_{2S}^* + \varepsilon\right)\right)\right).
\end{aligned}$$

Note that by $A > 0$, $B > 0$ and ε small enough, $\delta_2 > 0$.

So, if $\delta_1 > 0$, the perturbation relaxes both $\left(IC_{A_i}^T\text{-}2\right)$s and by argument given in Case 1, $(\Delta_{1S}^* - \varepsilon, \Delta_{2S}^* + \varepsilon, \Delta_{1C}^*, \Delta_{2C}^*)$ is an improvement.

If $\delta_1 < 0$, $\left(IC_{A_1}^T\text{-}2\right)$ is now violated, but $(IC_{A_2}\text{-}2)$ has become slack. Also note that by $B - A > 0$,

$$\delta_2 + \delta_1 = 2\varepsilon\left(B - A\right)\left(\Delta_{1S}^* - \left(\Delta_{2S}^* + \varepsilon\right)\right) > 0.$$

Now, set

$$\Delta_{1C}' = \Delta_{1C}^* + \delta_1,\ \ \Delta_{2C}' = \Delta_{2C}^* - \delta_1.$$

Note that

$$L_1\left(\Delta_{1S}^* - \varepsilon, \Delta_{2S}^* + \varepsilon\right) = L_1\left(\Delta_{1S}^*, \Delta_{2S}^*\right) + \delta_1 = \Delta_{1C}^* + \delta_1 = \Delta_{1C}',$$

and

$$\begin{aligned} L_2\left(\Delta_{1S}^* - \varepsilon, \Delta_{2S}^* + \varepsilon\right) &= L_2\left(\Delta_{1S}^*, \Delta_{2S}^*\right) + \delta_2 \\ &= L_2\left(\Delta_{1S}^*, \Delta_{2S}^*\right) - \delta_1 + (\delta_2 + \delta_1) \\ &> L_2\left(\Delta_{1S}^*, \Delta_{2S}^*\right) - \delta_1 \\ &= \Delta_{2C}^* - \delta_1 = \Delta_{2C}'. \end{aligned}$$

Hence, $(\Delta_{1S}^* - \varepsilon, \Delta_{2S}^* + \varepsilon, \Delta_{1C}', \Delta_{2C}')$ is feasible (note that $\left(IC_P^T\right)$s are unaltered by construction), and for $\varepsilon > 0$ sufficiently small, attains a higher value of Π^T.

Step 5: Suppose $p^T - p_\emptyset^T < 0$. Thus, ($IC_{A_i}^T$-2)s are

$$L_i\left(\Delta_{1S}, \Delta_{2S}\right) \leq \Delta_{iC}.$$

As before, there are three possibilities:

Case 1: Both ($IC_{A_i}^T$-2)s are slack at $(\Delta_{1S}^*, \Delta_{2S}^*, \Delta_{1C}^*, \Delta_{2C}^*)$. By argument in case 1 in Step 4, this solution can be improved up on.

Case 2: Exactly one of the two ($IC_{A_i}^T$-2)s is slack at $(\Delta_{1S}^*, \Delta_{2S}^*, \Delta_{1C}^*, \Delta_{2C}^*)$. Suppose only ($IC_{A_1}^T$-2) is slack, say, (hence, ($IC_{A_2}^T$-2) is binding). Set

$$\Delta_{1C}' = \Delta_{1C}^* - \delta, \ \Delta_{2C}' = \Delta_{2C}^* + \delta$$

where $\delta > 0$. As in case 2 in Step 4, the solution $(\Delta_{1S}^* - \varepsilon, \Delta_{2S}^* + \varepsilon, \Delta_{1C}', \Delta_{2C}')$ is also feasible for $\varepsilon > 0$ sufficiently small, and attains a higher value of Π^T.

Case 3: Both $(IC^T_{A_i}\text{-}2)$s are binding at $(\Delta^*_{1S}, \Delta^*_{2S}, \Delta^*_{1C}, \Delta^*_{2C})$. Consider changing $(\Delta^*_{1S}, \Delta^*_{2S})$ to $(\Delta^*_{1S} - \varepsilon, \Delta^*_{2S} + \varepsilon)$. As in case 3 in Step 4, the left-hand side of $\left(IC^T_{A_i}\text{-}2\right)$ changes by

$$\delta_i := L_i\left(\Delta^*_{1S} - \varepsilon, \Delta^*_{2S} + \varepsilon\right) - L_i\left(\Delta^*_{1S}, \Delta^*_{2S}\right)$$

where $\delta_2 > 0$ and

$$\delta_2 + \delta_1 = 2\varepsilon\left(B - A\right)\left(\Delta^*_{1S} - \left(\Delta^*_{2S} + \varepsilon\right)\right) < 0.$$

for ε small enough.

So, if $\delta_1 > 0$, the perturbation relaxes both $\left(IC^T_{A_i}\text{-}2\right)$s and by argument given in Case 1, $(\Delta^*_{1S} - \varepsilon, \Delta^*_{2S} + \varepsilon, \Delta^*_{1C}, \Delta^*_{2C})$ is an improvement.

If $\delta_1 < 0$, $\left(IC^T_{A_2}\text{-}2\right)$ is now violated, but $\left(IC^T_{A_1}\text{-}2\right)$ has become slack. Now, set

$$\Delta'_{1C} = \Delta^*_{1C} - \delta_2,\ \ \Delta'_{2C} = \Delta^*_{2C} + \delta_2.$$

Note that

$$\begin{aligned} L_1\left(\Delta^*_{1S} - \varepsilon, \Delta^*_{2S} + \varepsilon\right) \quad &= L_1\left(\Delta^*_{1S}, \Delta^*_{2S}\right) + \delta_1 \\ &= L_1\left(\Delta^*_{1S}, \Delta^*_{2S}\right) - \delta_2 + \left(\delta_2 + \delta_1\right) \\ &< L_1\left(\Delta^*_{1S}, \Delta^*_{2S}\right) - \delta_2 \\ &= \Delta^*_{1C} - \delta_2 = \Delta'_{1C}. \end{aligned}$$

and

$$L_2\left(\Delta^*_{1S} - \varepsilon, \Delta^*_{2S} + \varepsilon\right) = L_2\left(\Delta^*_{1S}, \Delta^*_{2S}\right) + \delta_2 = \Delta^*_{2C} + \delta_2 = \Delta'_{2C},$$

Hence, $(\Delta^*_{1S} - \varepsilon, \Delta^*_{2S} + \varepsilon, \Delta'_{1C}, \Delta'_{2C})$ is feasible (note that $\left(IC^T_P\right)$s are unaltered by construction), and for $\varepsilon > 0$ sufficiently small, attains a higher value of Π^T.

Combining all cases stated above, we obtain that without loss of generality, we can focus on the solution where $\Delta_{1S} = \Delta_{2S} = \Delta_S$, $\Delta_{1C} = \Delta_{2C} = \Delta_C$. And from (IR^T_i), we obtain that under such a solution, we must have $w_{1F} = w_{2F} = w_F$. This observation completes the proof of part (i) of this lemma.

Part (ii). Since we focus on $\Delta_{1S} = \Delta_{2S} = \Delta_S$ and $\Delta_{1C} = \Delta_{2C} = \Delta_C$, the program can be simplified as:

$$\mathcal{P}^T \begin{cases} \max\limits_{\Delta_C, \Delta_S} \quad 2\left(p^T\right)^2 P^T \Pr(\omega = G \mid x^T \in X_P) y \Delta_S + (1 - p^T)\underline{\pi} - \left(p^T P^T\right)^2 \Delta_S^2 \\ \\ s.t. \quad \left[\frac{3}{2}\left(p^T P^T\right)^2 - \frac{1}{2}\left(p^T_\emptyset P^T_\emptyset\right)^2 - p^T_\emptyset P^T_\emptyset p^T P^T\right] (\Delta_S)^2 \geq (p^T - p^T_\emptyset)\Delta_C \quad (IC^T_A\text{-}2) \\ \quad 2\left[\Pr(\omega = G \mid x^T \in X_P) y - 2P^T \Delta_S\right] p^T P^T \Delta_S \geq \underline{\pi} - 2\Delta_C \quad (IC^T_P\text{-}1) \\ \quad 2\left[\Pr(\omega = G \mid x^T \notin X_P) y - 2P^T_C \Delta_S\right] p^T P^T \Delta_S \leq \underline{\pi} - 2\Delta_C \quad (IC^T_P\text{-}2) \end{cases}$$

As in the case of individual assignment, we have two cases: $X_P = \{g, \emptyset\}$ and $X_P = \{g\}$.

Case 1: $X_P = \{G, \emptyset\}$. Here, $p^T - p^T_\emptyset < 0$; so we have:

$$\mathcal{P}^T_{\{g,\emptyset\}} \begin{cases} \max\limits_{\Delta_C, \Delta_S} \quad 2\left(p^T\right)^2 P^T \Pr(\omega = G \mid x^T \in X_P) y \Delta_S + (1 - p^T)\underline{\pi} - \left(p^T P^T\right)^2 \Delta_S^2 \\ \\ s.t. \quad \Delta_C \geq l_A := \left[\frac{3}{2}\left(p^T P^T\right)^2 - \frac{1}{2}\left(p^T_\emptyset P^T_\emptyset\right)^2 - p^T_\emptyset P^T_\emptyset p^T P^T\right] \frac{\Delta_S^2}{p^T - p^T_\emptyset} \quad (IC^T_A\text{-}2) \\ \quad \Delta_C \geq l_P := \frac{1}{2}\underline{\pi} - \left[\Pr(\omega = G \mid x^T \in X_P) y - 2P^T \Delta_S\right] p^T P^T \Delta_S \quad (IC^T_P\text{-}1) \\ \quad \Delta_C \leq u_P := \frac{1}{2}\underline{\pi} - \left[\Pr(\omega = G \mid x^T \notin X_P) y - 2P^T_C \Delta_S\right] p^T P^T \Delta_S \quad (IC^T_P\text{-}2) \end{cases} .$$

Notice that Δ_C is not in the objective function, we can further simplify the program as:

$$\begin{cases} \max_{\Delta_S} \;\; 2\left(p^T\right)^2 P^T \Pr(\omega = G \mid x^T \in X_P) y \Delta_S + (1-p^T)\underline{\pi} - \left(p^T P^T\right)^2 \Delta_S^2 \\ s.t. \\ u_P \geq l_A \Leftrightarrow \frac{1}{2}\underline{\pi} \geq \left[\Pr(\omega = G \mid x^T \notin X_P) y - 2P_C^T \Delta_S\right] p^T P^T \Delta_S \\ \quad + \frac{\Delta_S^2}{p^T - p_\emptyset^T}\left[\frac{3}{2}\left(p^T P^T\right)^2 - \frac{1}{2}\left(p_\emptyset^T P_\emptyset^T\right)^2 - p_\emptyset^T P_\emptyset^T p^T P^T\right] \\ \\ u_P \geq l_P \Leftrightarrow \Delta_S \leq \frac{y}{2(1-\mu)}. \end{cases}$$

By routine calculation, one obtains $\Pr(\omega = G | x^T \in X_P) = \frac{1}{2-\alpha'}$, $\Pr(\omega = G | x^T \notin X_P) = 0$, and

$$p^T = 1 - \tfrac{1}{2}\alpha'; \quad P^T = \mu + (1-\mu)\tfrac{1}{2-\alpha'}; \quad p_C^T = \mu;$$

$$p_\emptyset^T = 1 - \tfrac{1}{2}\alpha; \quad P_\emptyset^T = \mu + (1-\mu)\tfrac{1}{2-\alpha},$$

where $\alpha' := 1 - (1-\alpha)^2$. Also, as in the proof of Lemma 2, to streamline notation, without loss of generality, we set $y = 1$ (and hence, $\underline{\pi} = 1/4$). Plugging the values, the program becomes:

$$\begin{cases} \max_{\Delta_S} \;\; -\frac{1}{4}\left(1 + \mu\left(1-\alpha'\right)\right)^2 \left(\Delta_S - \frac{1}{1+\mu(1-\alpha')}\right)^2 + \frac{1}{4}(1 + \frac{1}{2}\alpha') \\ \\ s.t. \qquad \Delta_S^2 \leq \frac{1}{2\mu^2\alpha(1-\alpha)} \; and \; \Delta_S \leq \frac{1}{2(1-\mu)} \end{cases}$$

The solution is given as:

$$\Delta_S^* = \begin{cases} \frac{1}{1+\mu(1-\alpha)^2} & if \; \frac{1}{1+\mu(1-\alpha)^2} \leq \frac{1}{2(1-\mu)} \\ \frac{1}{2(1-\mu)} & otherwise \end{cases},$$

and the associated value is:

$$V^T_{\{g,\emptyset\}} = \begin{cases} \frac{1}{4} + \frac{1}{8}\alpha(2-\alpha) & \textit{if } \frac{1}{1+\mu(1-\alpha)^2} \leq \frac{1}{2(1-\mu)} \\ \frac{1+\mu(1-\alpha)^2}{4(1-\mu)}\left[1 - \frac{1+\mu(1-\alpha)^2}{4(1-\mu)}\right] + \frac{1}{8}\alpha(2-\alpha) & \textit{otherwise} \end{cases} \tag{3}$$

Case 2: $X_P = \{G\}$. Here, $p^T - p^T_\emptyset > 0$, so we have:

$$\mathcal{P}^T_{\{g\}} \begin{cases} \max_{\Delta_C, \Delta_S} \; 2\left(p^T\right)^2 P^T \Pr(\omega = G \mid x^T \in X_P) y \Delta_S + (1-p^T)\underline{\pi} - \left(p^T P^T\right)^2 \Delta_S^2 \\ \\ s.t. \qquad \Delta_C \leq u_A := \left[\frac{3}{2}\left(p^T P^T\right)^2 - \frac{1}{2}\left(p^T_\emptyset P^T_\emptyset\right)^2 - p^T_\emptyset P^T_\emptyset p^T P^T\right] \frac{\Delta_S^2}{p^T - p^T_\emptyset} \qquad (IC^T_A\text{-}2) \\ \qquad \Delta_C \geq l_P := \frac{1}{2}\underline{\pi} - \left[\Pr(\omega = G \mid x^T \in X_P) y - 2P^T\Delta_S\right] p^T P^T \Delta_S \qquad (IC^T_P\text{-}1) \\ \qquad \Delta_C \leq u_P := \frac{1}{2}\underline{\pi} - \left[\Pr(\omega = G \mid x^T \notin X_P) y - 2P^T_C\Delta_S\right] p^T P^T \Delta_S \qquad (IC^T_P\text{-}2) \end{cases} .$$

As Δ_C does not appear in the objective function, we can replace the constraints by requiring $l_P \leq u_A$ and $l_P \leq u_P$, and the program simplifies to:

$$\begin{cases} \max_{\Delta_S} \; 2\left(p^T\right)^2 P^T \Pr(\omega = G \mid x^T \in X_P) y \Delta_S + (1-p^T)\underline{\pi} - \left(p^T P^T\right)^2 (\Delta_S)^2 \\ s.t. \\ \frac{1}{2}\underline{\pi} \leq \left[\Pr(\omega = G \mid x^T \in X_P) y - 2P^T\Delta_S\right] p^T P^T \Delta_S \\ \quad + \left[\frac{3}{2}\left(p^T P^T\right)^2 - \frac{1}{2}\left(p^T_\emptyset P^T_\emptyset\right)^2 - p^T_\emptyset P^T_\emptyset p^T P^T\right] \frac{\Delta_S^2}{p^T - p^T_\emptyset} \\ \Delta_S \leq \frac{y}{2(1-\mu)} \end{cases} .$$

Plugging the values for the probabilities (and parameters), we obtain:

$$\begin{cases} \max_{\Delta_S} \ \Pi^T_{\{g\}}(\Delta_S) := -\frac{1}{4}\left(\alpha'(\Delta_S - 1)\right)^2 + \frac{1}{4}\left(1 - \alpha'\left(\frac{1}{2} - \alpha'\right)\right) \\ s.t. \\ \alpha(2-\alpha)\Delta_S - \frac{1}{2}\alpha(1-\alpha)\Delta_S^2 \geq \frac{1}{4} \\ \Delta_S \leq \frac{1}{2(1-\mu)} \end{cases}$$

Let $\hat{\alpha} := 0.12445$ and $K(\alpha) := \frac{1}{1-\alpha}\left(2 - \alpha - \sqrt{(2-\alpha)^2 - \frac{1-\alpha}{2\alpha}}\right)$. It is routine to check that the program does not admit a solution if $\alpha < \hat{\alpha}$ or $K(\alpha) > 1/2\,(1-\mu)$. Otherwise, the solution is as follows:

$$\Delta_S^* = \begin{cases} 1 & if\ \alpha \geq \hat{\alpha}\ and\ K(\alpha) \leq 1 \leq \frac{1}{2(1-\mu)} \\ \frac{1}{2(1-\mu)} & if\ \alpha \geq \hat{\alpha}\ and\ K(\alpha) \leq \frac{1}{2(1-\mu)} < 1 \\ \tilde{\alpha} & if\ \alpha \geq \hat{\alpha}\ and\ 1 < K(\alpha) \leq \frac{1}{2(1-\mu)} \end{cases},$$

and the associated value function is

$$V^T_{\{g\}} = \begin{cases} \Pi^T_{\{g\}}(1) & if\ \alpha \geq \hat{\alpha}\ and\ K(\alpha) \leq 1 \leq \frac{1}{2(1-\mu)} \\ \Pi^T_{\{g\}}\left(\frac{1}{2(1-\mu)}\right) & if\ \alpha \geq \hat{\alpha}\ and\ K(\alpha) \leq \frac{1}{2(1-\mu)} < 1 \\ \Pi^T_{\{g\}}(\tilde{\alpha}) & if\ \alpha \geq \hat{\alpha}\ and\ 1 < K(\alpha) \leq \frac{1}{2(1-\mu)} \end{cases} \tag{4}$$

Thus, we conclude that the program $\mathcal{P}^T$ always admits a solution for $X_P = \{g, \emptyset\}$ and admits a solution for $X_P = \{g\}$ if and only if α and μ are sufficiently large.

Q.E.D.

Proof of Proposition 2. **Step 1.** Notice that program $\mathcal{P}^I_{\{g,\emptyset\}}$ and $\mathcal{P}^T_{\{g,\emptyset\}}$ have the objective function. Denote the the unconstrained maximum of that objective function as

$$V_{\{g,\emptyset\}} = \frac{1}{4} + \frac{1}{8}\alpha'.$$

Similarly, $\mathcal{P}^I_{\{g\}}$ and $\mathcal{P}^T_{\{g\}}$ have the same objective function, and we denote the unconstrained maximum as

$$V_{\{g\}} = \frac{1}{4}\left[1 - \alpha'\left(\frac{1}{2} - \alpha'\right)\right].$$

Since unconstrained maximum must be (weakly) larger than the value under a constrained maximization, we have

$$V^I_{\{g,\emptyset\}} \leq V_{\{g,\emptyset\}}, V^T_{\{g,\emptyset\}} \leq V_{\{g,\emptyset\}}, V^I_{\{g\}} \leq V_{\{g\}} \text{ and } V^T_{\{g\}} \leq V_{\{g\}}.$$

Further, we notice that $V_{\{g,\emptyset\}} - V_{\{g\}} = \frac{1}{4}\alpha'(1-\alpha') \geq 0$, so we have

$$V_{\{g\}} \leq V_{\{g,\emptyset\}}$$

and equality holds if and only if $\alpha' = 0$ or 1.

Step 2. Recall that the solutions for the programs $\mathcal{P}^I_{\{g,\emptyset\}}$ and $\mathcal{P}^T_{\{g,\emptyset\}}$ (see (1) and (3) ; we maintain $y = 1$ to streamline notation) stipulate

$$V^I_{\{g,\emptyset\}} = \frac{1}{4}\left(1 + \frac{1}{2}\alpha'\right) = S^*\left(= \frac{1}{4}\left(1 + \alpha - \frac{1}{2}\alpha^2\right)\right)$$

when $\frac{\alpha'\mu^2}{(1+\mu(1-\alpha'))^2} \leq \frac{1}{2}$, and

$$V^T_{\{g,\emptyset\}} = S^*$$

when $\frac{1}{1+\mu(1-\alpha)^2} \leq \frac{1}{2(1-\mu)}$.

Let μ_0 be the solution to the equation

$$\frac{1}{1+\mu(1-\alpha)^2} = \frac{1}{2(1-\mu)};$$

that is,

$$\mu_0 = \frac{1}{2+(1-\alpha)^2} = \frac{1}{3-\alpha'}. \tag{5}$$

Note that, for $\mu \in [0, \mu_0)$, $\frac{1}{1+\mu(1-\alpha)^2} > \frac{1}{2(1-\mu)}$; and for $\mu \in [\mu_0, 1)$, $\frac{1}{1+\mu(1-\alpha)^2} \leq \frac{1}{2(1-\mu)}$.

Next, define μ_1 as follows:

$$\mu_1 = \begin{cases} 1 & \textit{if } \frac{\alpha'\mu^2}{(1+\mu(1-\alpha'))^2} < \frac{1}{2} \ \forall \mu \in [0,1] \\ \mu^*(\alpha') & \textit{otherwise} \end{cases}, \tag{6}$$

where

$$\mu^*(\alpha') = \frac{1-\alpha'+\sqrt{2\alpha'}}{2\alpha'-(1-\alpha')^2}$$

is the unique solution to

$$\frac{\alpha'\mu^2}{(1+\mu(1-\alpha'))^2} = \frac{1}{2}$$

in $[0, 1]$.

Note that $\frac{\alpha'\mu^2}{(1+\mu(1-\alpha'))^2} \leq \frac{1}{2}$ for $\mu \in [0, \mu_1]$ and $\frac{\alpha'\mu^2}{(1+\mu(1-\alpha'))^2} > \frac{1}{2}$ for $\mu \in (\mu_1, 1]$.

Step 3. Notice that $\mu_0 < \mu_1 \ \forall \alpha \in [0,1]$ as using (5), one obtains

$$\frac{\alpha' \mu_0^2}{\left(1 + \mu_0 \left(1 - \alpha'\right)\right)^2} = \frac{\alpha'}{4 \left(2 - \alpha'\right)^2} < \frac{1}{2}.$$

Combining above observations we obtain: (i) if $\mu < \mu_0$, $S^* = V^I_{\{g,\emptyset\}} > \max\{V^T_{\{g,\emptyset\}}, V^I_{\{g\}}, V^T_{\{g\}}\}$; that is, individual assignment with $X_P = \{G, \emptyset\}$ is optimal; (ii) if $\mu > \mu_1$, $S^* = V^T_{\{g,\emptyset\}} > \max\{V^I_{\{g,\emptyset\}}, V^I_{\{g\}}, V^T_{\{g\}}\}$; that is, team assignment with $X_P = \{G, \emptyset\}$ is optimal; (iii) if $\mu_0 \leq \mu \leq \mu_1$, $S^* = V^I_{\{g,\emptyset\}} = V^T_{\{g,\emptyset\}} > \max\{V^I_{\{g\}}, V^T_{\{g\}}\}$; that is, both team and individual assignment with $X_P = \{G, \emptyset\}$ are optimal.

Q.E.D.

Proof of Proposition 3. **Step 1.** From (5) it directly follows that μ_0 is increasing in α.

Step 2. Now, consider the definition for μ_1 as given in (6). Note that when $\alpha' < \frac{1}{2}$, $\frac{\alpha' \mu^2}{(1+\mu(1-\alpha'))^2} \leq \alpha' \mu^2 \leq \alpha' < \frac{1}{2}$; so, $\mu_1 = 1$. And for $\alpha' \geq \frac{1}{2}$, we have

$$\mu_1 = \min\left\{1, \mu^*\left(\alpha'\right)\right\}.$$

Note that

$$\frac{d}{d\alpha'} \mu^*\left(\alpha'\right) = -\frac{\left(1 - \frac{1}{\sqrt{2\alpha'}}\right)\left(2\alpha' - \left(1 - \alpha'\right)^2\right) + 2\left(2 - \alpha'\right)\left(1 - \alpha' + \sqrt{2\alpha'}\right)}{\left(2\alpha' - \left(1 - \alpha'\right)^2\right)^2}.$$

For $\alpha' \in [\frac{1}{2}, 1)$ it is routine to check that $1 - \frac{1}{\sqrt{2\alpha'}} \geq 0$ and all other three terms in the numerator are strictly positive (denominator is positive by virtue of being a sqaure term).

So, $\frac{d}{d\alpha'}\mu^*(\alpha') < 0$. Hence, $\mu^*(\alpha')$ is also strictly decreasing in α when $\alpha \in \left[1 - 1/\sqrt{2}, 1\right]$ (recall $\alpha' := 1 - (1-\alpha)^2$).

Step 3. Finally, note that when $\alpha = 1 - \frac{1}{\sqrt{2}}$, $\mu^*(\alpha') = 2$; and when $\alpha = 1$, $\mu^*(\alpha') = \frac{1}{\sqrt{2}}$. As $\mu^*(\alpha')$ is decreasing in α, by Intermediate Value Theorem, there exists an α^* such that $\mu^*(\alpha^*) = 1$. Also, when $\alpha < \alpha^*$, $\mu^*(\alpha') > 1$; when $\alpha > \alpha^*$, $\mu^*(\alpha') < 1$.

Thus, for $1 - \frac{1}{\sqrt{2}} \leq \alpha \leq \alpha^*$, $\mu_1 = \min\{1, \mu^*(\alpha')\} = 1$ and for $\alpha \geq \alpha^*$, μ_1 is decreasing in α.

Q.E.D.

Proof of Proposition 4. **Step 1:** Since Lemma 1 and 3 hold for any $\theta \in (\frac{1}{2}, 1)$ (note that the proofs of these lemmas presented above do not rely on any specific value of θ), we may continue to limit attention to the set of four programs $\mathcal{P}^I_{\{g,\emptyset\}}$, $\mathcal{P}^I_{\{g\}}$, $\mathcal{P}^T_{\{g,\emptyset\}}$, and $\mathcal{P}^T_{\{g\}}$ as defined in the proofs of Lemma 2 and 4. In this step, we compute the unconstrained maximum of these four programs. That is, for $\mathcal{P}^I_{X_P}$, $X_P \in \{\{g,\emptyset\},\{g\}\}$, we solve for

$$\overline{V}^d_{X_P} := \max_{\Delta_S} \; 2\left(p^I\right)^2 P^I \Pr(\omega = G \mid x \in X_P) y\Delta_S + (1-p^I)\underline{\pi} - \left(p^I P^I\right)^2 \Delta_S^2,$$

and for $\mathcal{P}^T_{X_P}$, $X_P \in \{\{g,\emptyset\},\{g\}\}$, we solve for

$$\overline{V}^T_{X_P} := \max_{\Delta_S} \; 2\left(p^T\right)^2 P^T \Pr(\omega = G \mid x^T \in X_P) y\Delta_S + (1-p^T)\underline{\pi} - \left(p^T P^T\right)^2 \Delta_S^2.$$

Plugging in the values for all the probabilities, and solving for the optimization problem (notice that all objective functions are quadratic in Δ_S; hence solution exists and is unique) we obtain (recall that $\alpha' = 1 - (1-\alpha)^2$):

$$\overline{V}^I_{\{g,\emptyset\}} = \overline{V}^T_{\{g,\emptyset\}} = \frac{1}{4}\left[(1-\alpha'(1-\theta))^2 + \frac{1}{2}\alpha'\right] =: \overline{V},$$

and

$$\overline{V}^I_{\{g\}} = \overline{V}^T_{\{g\}} = \frac{1}{4}\left[(\alpha'\theta)^2 + (1-\frac{1}{2}\alpha')\right].$$

Note that

$$\overline{V} > \overline{V}^I_{\{g\}} = \overline{V}^T_{\{g\}}.$$

In what follows, we focus our attention on programs $\mathcal{P}^I_{\{g,\emptyset\}}$ and $\mathcal{P}^T_{\{g,\emptyset\}}$, as we show that for any given set of parameters, at least one of them achieves the value $\overline{V}$.

Step 2: We show that for θ sufficiently large, there exists a cutoff $\mu_0(\alpha;\theta)$ such that $V^T_{\{g,\emptyset\}} < \overline{V}$ if $\mu < \mu_0(\alpha;\theta)$; and $V^T_{\{g,\emptyset\}} = \overline{V}$ otherwise. Plugging the values of the probabilities, the program $\mathcal{P}^T_{\{g,\emptyset\}}$ can be written as:

$$\mathcal{P}^T_{\{g,\emptyset\}} : \left\{ \begin{array}{lr} \max\limits_{\Delta_S} \quad \overline{V} - \frac{1}{4}\left[(1-\alpha'(1-\theta)+\mu(1-\alpha'\theta))\,\Delta_S - (1-\alpha'(1-\theta))\right]^2 & \\ s.t. & \\ [1-\alpha'(1-\theta)+\mu(1-\alpha'\theta)](1-\theta)\Delta_S + \frac{1}{2}\alpha(1-\alpha)(1-\theta\,(1-\mu))^2\Delta_S^2 \le \frac{1}{4} & \left(C_1^T\right) \\ \Delta_S \le \frac{1}{2(1-\mu)} & \left(C_2^T\right) \end{array} \right. .$$

The objective function achieves its peak at $\Delta_S^* = \frac{1-\alpha'(1-\theta)}{1-\alpha'(1-\theta)+\mu(1-\alpha'\theta)}$. If Δ_S^* is feasible under constraints $\left(C_1^T\right)$ and $\left(C_2^T\right)$, $V^T_{\{g,\emptyset\}} = \overline{V}$; and $V^T_{\{g,\emptyset\}} < \overline{V}$ otherwise. Next, we analyze conditions under which this solution may be feasible.

Plugging Δ_S^* into $\left(C_2^T\right)$ and simplifying, we have:

$$\mu \geq \frac{1-\alpha'(1-\theta)}{3-\alpha'\theta-2\alpha'(1-\theta)} =: \mu_0(\alpha;\theta).$$

Thus, Δ_S^* satisfies constraint $\left(C_2^T\right)$ if and only if $\mu \geq \mu_0(\alpha;\theta)$.

We also claim that Δ_S^* satisfies $\left(C_1^T\right)$ if $\theta > 0.85$. To see this, plug Δ_S^* into the left-hand side of $\left(C_1^T\right)$, and we obtain:

$$\begin{aligned}
&(1-\theta)(1-\alpha'(1-\theta)) + \tfrac{1}{2}\alpha(1-\alpha)(1-\alpha'(1-\theta))^2 \left[\tfrac{1-\theta(1-\mu)}{1-\alpha'(1-\theta)+\mu(1-\alpha'\theta)}\right]^2 \\
&\leq (1-\theta)(1-\alpha'(1-\theta)) + \tfrac{1}{2}\alpha(1-\alpha)(1-\alpha'(1-\theta))^2 \left(\tfrac{1}{2-\alpha'}\right)^2 \\
&\leq (1-\theta) + \tfrac{\alpha(1-\alpha)}{2(2-\alpha')^2} \\
&\leq \tfrac{1}{4}.
\end{aligned}$$

The first inequality follows as the expression is increasing in μ, the second one follows as $1-\alpha'(1-\theta) \in [0,1]$, and the final one holds since $\frac{\alpha(1-\alpha)}{2(2-\alpha')^2} < 0.1$ (for $\alpha \in [0,1]$).

Hence, for $\theta > 0.85$, $V_{\{g,\emptyset\}}^T < \overline{V}$ when $\mu < \mu_0(\alpha;\theta)$, and $V_{\{g,\emptyset\}}^T = \overline{V}$ otherwise.

Step 3: We show that for θ sufficiently large, there exists a cutoff $\mu_1(\alpha;\theta)$ such that $V_{\{g,\emptyset\}}^T = \overline{V}$ when $\mu \leq \mu_1(\alpha;\theta)$ and $V_{\{g,\emptyset\}}^T < \overline{V}$ otherwise. The proof is analogous to the one given in Step 2 above.

Plugging the values of the probabilities, the program $\mathcal{P}^I_{\{g,\emptyset\}}$ can be written as:

$$\mathcal{P}^I_{\{g,\emptyset\}}:\begin{cases}\max\limits_{\Delta_S} \quad \overline{V}-\frac{1}{4}\left[(1-\alpha'(1-\theta)+\mu(1-\alpha'\theta))\,\Delta_S-(1-\alpha'(1-\theta))\right]^2 \\ \qquad\qquad\qquad s.t. \\ (1-\theta)[1+\mu-\alpha'(1-\theta+\mu\theta)]\Delta_S+\frac{1}{2}\alpha'\left(1-\theta\left(1-\mu\right)\right)^2\Delta_S^2\leq\frac{1}{4} \quad \left(C_1^I\right) \\ \qquad\qquad \Delta_S\leq\frac{1}{1-\mu} \qquad\qquad\qquad\qquad \left(C_2^I\right)\end{cases}.$$

The objective function achieves its peak at $\Delta_S^*=\frac{1-\alpha'(1-\theta)}{1-\alpha'(1-\theta)+\mu(1-\alpha'\mu)}$. If Δ_S^* is feasible under constraints $\left(C_1^I\right)$ and $\left(C_2^I\right)$, $V^I_{\{g,\emptyset\}}=\overline{V}$; and $V^I_{\{g,\emptyset\}}<\overline{V}$ otherwise. Next, we analyze conditions under which this solution may be feasible.

It is routinely to check that Δ_S^* is always feasible under $\left(C_2^I\right)$:

$$\Delta_S^*=\frac{1-\alpha'(1-\theta)}{1-\alpha'(1-\theta)+\mu(1-\alpha'\theta)}\leq 1\leq\frac{1}{1-\mu}.$$

Now, plugging Δ_S^* in the left-hand side of $\left(C_1^I\right)$ we get:

$$L(\mu;\alpha,\theta):=(1-\theta)(1-\alpha'(1-\theta))+\frac{1}{2}\alpha'\left(1-\alpha'(1-\theta)\right)^2\left(\frac{1-\theta+\mu\theta}{1-\alpha'(1-\theta)+\mu(1-\alpha'\theta)}\right)^2.$$

Note that $L(\mu;\alpha,\theta)$ is increasing in $\mu\in[0,1]$, so it achieves its maximum at $\mu=1$, where:

$$L(1;\alpha,\theta)=(1-\theta)(1-\alpha'(1-\theta))+\frac{1}{2}\alpha'\left(\frac{1-\alpha'(1-\theta)}{2-\alpha'}\right)^2.$$

Now, if $\theta > 0.85$, we have $L(1;0,\theta) = 1-\theta < \frac{1}{4}$ and $L(1;1,\theta) = \theta - \frac{1}{2}\theta^2 > \frac{1}{4}$; also

$$\frac{d}{d\alpha}L(1;\alpha,\theta) = \frac{2-2\alpha}{(2-\alpha')^3}\left[R_1+R_2+R_3+R_4\right],$$

where

$$R_1 = \tfrac{1}{2}(1-\theta)^2\left(2\alpha'+\alpha'^3\right), \quad R_2 = 3(1-\theta)^2(\alpha'-\alpha'^2),$$
$$R_3 = 8\left(\tfrac{3}{4}-\theta\right)^2\alpha', \qquad R_4 = 1-8(1-\theta)^2.$$

As $R_i \geq 0$ for $i = 1,...,4$, we have $\frac{d}{d\alpha}L(1;\alpha,\theta) \geq 0$. So by Intermediate Value Theorem, there exists a unique $\alpha^*(\theta) \in (0,1)$ such that $L(1;\alpha^*(\theta),\theta) = \frac{1}{4}$.

Next, define $\mu_1(\alpha;\theta)$ as follows: for $\alpha \leq \alpha^*(\theta)$, let $\mu_1(\alpha;\theta) = 1$; and for $\alpha > \alpha^*(\theta)$, let $\mu_1(\alpha;\theta)$ be the solution to $L(\mu;\alpha,\theta) = \frac{1}{4}$. That is:

$$\mu_1(\alpha;\theta) := \begin{cases} 1 & if\ \alpha \leq \alpha^*(\theta) \\ \frac{(1-\alpha'(1-\theta))\left(\sqrt{K}-(1-\theta)\sqrt{\alpha'}\right)}{(1-\alpha'(1-\theta))\theta\sqrt{\alpha'}-(1-\alpha'\theta)\sqrt{K}} & otherwise \end{cases},$$

where $K := \frac{1}{2} - 2(1-\theta)(1-\alpha'(1-\theta))$.

Notice that when $\alpha \leq \alpha^*(\theta)$, for all $\mu \leq 1 = \mu_1(\alpha;\theta)$, $L(\mu;\alpha,\theta) \leq \frac{1}{4}$, i.e., Δ_S^* satisfies $\left(C_1^I\right)$; when $\alpha > \alpha^*(\theta)$, for all $\mu \leq \mu_1(\alpha;\theta)$, $L(\mu;\alpha,\theta) \leq \frac{1}{4}$, i.e., Δ_S^* satisfies $\left(C_1^I\right)$, and for all $\mu > \mu_1(\alpha;\theta)$, $L(\mu;\alpha,\theta) > \frac{1}{4}$, i.e., Δ_S^* always violate $\left(C_1^I\right)$. As Δ_S^* always satisfies $\left(C_2^I\right)$ we conclude: for $\theta > 0.85$, $V_{\{g,\emptyset\}}^T = \overline{V}$ when $\mu \leq \mu_1(\alpha;\theta)$ and $V_{\{g,\emptyset\}}^T < \overline{V}$ otherwise.

Step 4: Define θ^* as the largest solution in $[0,1]$ to the equation $\mu_0(1;\theta) = \mu_1(1;\theta)$; i.e.,

$$\theta^* := \frac{1}{2}\left(1+\frac{1}{\sqrt{2}}\right).$$

As $\theta^* > 0.85$, the definition of μ_0 and μ_1 are valid for $\theta > \theta^*$.

Step 5: Note that $\mu_0(\alpha;\theta)$ is increasing in both α and θ for $\theta \in (\theta^*, 1]$:

$$\frac{d}{d\alpha}\mu_0(\alpha;\theta) = \frac{(2\theta-1)(2-2\alpha)}{\left[3-\alpha'\theta-2\alpha'\left(1-\theta\right)\right]^2} \geq 0,$$

and

$$\frac{d}{d\theta}\mu_0(\alpha;\theta) = \frac{\alpha'(2-\alpha')}{\left[3-\alpha'\theta-2\alpha'\left(1-\theta\right)\right]^2} \geq 0.$$

Step 6: Next, we claim that $\mu_1(\alpha;\theta)$ is decreasing in α and increasing in θ for $\theta \in (\theta^*, 1]$.

Recall that for $\alpha \leq \alpha^*(\theta)$, $\mu_1(\alpha;\theta) = 1$; for $\alpha > \alpha^*(\theta)$, taking the derivative of $\mu_1(\alpha;\theta)$ with respect to α we obtain:

$$\frac{d}{d\alpha}\mu_1(\alpha;\theta) = -\left(S_1S_2 + S_3S_4\right)S_5,$$

where

$$\begin{aligned}
S_1 &:= (1-\theta)\left[(1-\alpha'(1-\theta))\theta\sqrt{\alpha'} - (1-\alpha'\theta)\sqrt{K}\right], \\
S_2 &:= \tfrac{1}{2\sqrt{K}}\left(1-6(1-\theta)(1-\alpha'(1-\theta))\right) + \tfrac{1}{2\sqrt{\alpha'}}\left(1-3\alpha'(1-\theta)\right), \\
S_3 &:= (1-\alpha'(1-\theta))\left[\sqrt{K} - (1-\theta)\sqrt{\alpha'}\right], \\
S_4 &:= \tfrac{1}{2\sqrt{\alpha'}}\theta(1-3\alpha'(1-\theta)) + \tfrac{1}{2\sqrt{K}}\left(\theta + 2\theta^2 + 6\alpha'^2 - 2\right), \\
S_5 &:= (2-2\alpha)\,/\left[(1-\alpha'(1-\theta))\theta\sqrt{\alpha'} - (1-\alpha'\theta)\sqrt{K}\right]^2.
\end{aligned}$$

It is routine to check that $S_i \geq 0$ for all $i = 1, ..., 5$. Hence, $\frac{d}{d\alpha}\mu_1(\alpha;\theta) \leq 0$.

Next, consider the derivative of μ_1 with respect to θ:

$$\frac{d}{d\theta}\mu_1(\alpha;\theta) = \frac{\frac{\sqrt{\alpha'}}{\sqrt{K}}T_1 + T_2}{\left[1 - \alpha'(1-\theta)\theta\sqrt{\alpha'} - (1-\alpha'\theta)\sqrt{K}\right]^2},$$

where

$$\begin{aligned} T_1 :=&\ \tfrac{5}{2}(1-\theta) + \tfrac{1}{2}(-19+33\theta-13\theta^2)\alpha' + (1-\theta)\left(11-17\theta+4\theta^2\right)\alpha'^2 - 4(1-\theta^3)\alpha'^3, \\ T_2 :=&\ -\tfrac{1}{2}\alpha'(2-\alpha') + (1-\alpha'(1-\theta))\left[2-(3-2\theta)\alpha' + (1-\theta)(8\theta-3)\alpha'^2\right]. \end{aligned}$$

Below, we show that $T_1 > 0$ and $T_2 \geq 0$ that implies $\mu_1(\alpha;\theta)$ is increasing in θ.

Step 6a: To show $T_1 > 0$, we consider two cases: $A > 0$ and $A \leq 0$, where $A := -19+33\theta-13\theta^2$.

Case 1: When $A > 0$, we have $11-17\theta+4\theta^2 < 0$ and $13-17\theta+4\theta^2 \geq 0$. Now,

$$\begin{aligned} T_1 \geq&\ \tfrac{5}{2}(1-\theta) + \tfrac{1}{2}(-19+33\theta-13\theta^2)\alpha' + (1-\theta)\left(11-17\theta+4\theta^2\right) - 4(1-\theta^3) \\ =&\ \tfrac{1}{2}(-19+33\theta-13\theta^2)\alpha' + (1-\theta)\left(13-17\theta+4\theta^2\right) + (1-\theta)\left[\tfrac{1}{2} - 4(1-\theta)^2\right] \\ >&\ 0. \end{aligned}$$

Case 2: When $A \leq 0$ and $\theta > \theta^* > 0.85$, we have $11-17\theta+4\theta^2 < 0$. Now,

$$\begin{aligned} T_1 \geq&\ \tfrac{5}{2}(1-\theta) + \tfrac{1}{2}(-19+33\theta-13\theta^2) + (1-\theta)(11-17\theta+4\theta^2) - 4\left(1-\theta^3\right) \\ =&\ \tfrac{1}{2}\theta(5\theta-4) > 0. \end{aligned}$$

Step 6b: To show $T_2 \geq 0$, we first define

$$T_3 \ : \ = 2 - (3 - 2\theta)\alpha' + (1 - \theta)(8\theta - 3)\alpha'^2,$$

$$T_4 \ : \ = 2(1 - \theta)(2 - \alpha') + \alpha'(1 - \alpha'(1 - \theta)).$$

Now,

$$\begin{aligned} T_2 = & \ -\tfrac{1}{2}\alpha'(2 - \alpha') + (1 - \alpha'(1 - \theta))T_3 \\ \geq & \ -\tfrac{1}{2}\alpha'(2 - \alpha') + (1 - \alpha'(1 - \theta))T_4 \\ \geq & \ -\tfrac{1}{2}\alpha'(2 - \alpha') + \tfrac{1}{2}(2 - \alpha')^2 \\ = & \ (2 - \alpha')(1 - \alpha') \geq 0. \end{aligned}$$

The first inequality follows as $T_3 \geq T_4$ (routine to check). The second inequality follows from the fact that as we have $\alpha > \alpha^*(\theta)$, we have $L(1; \alpha, \theta) > \frac{1}{4}$. And,

$$L(1; \alpha, \theta) > \frac{1}{4} \Leftrightarrow (1 - \alpha'(1 - \theta))T_4 > \frac{1}{2}(2 - \alpha')^2.$$

Step 7: It is routine to check $\mu_1(1; \theta) > \mu_0(1; \theta)$. So, for any $\theta \in (\theta^*, 1]$, $\mu_1(\alpha; \theta) > \mu_0(\alpha; \theta)$ for all $\alpha \in [0, 1]$ (as μ_0 is strictly increasing, and μ_1 is decreasing in α). Thus, from Step 2 and 3, we find for $\mu < \mu_0(\alpha; \theta)$, $V^I_{\{g,\emptyset\}} = \overline{V} > \max\left\{V^I_{\{g\}}, V^T_{\{g,\emptyset\}}, V^T_{\{g\}}\right\}$; for $\mu > \mu_1(\alpha; \theta)$, $V^T_{\{g,\emptyset\}} = \overline{V} > \max\left\{V^I_{\{g\}}, V^I_{\{g,\emptyset\}}, V^T_{\{g\}}\right\}$; otherwise, $V^I_{\{g,\emptyset\}} = V^T_{\{g,\emptyset\}} = \overline{V}$. Thus, the characterization of optimal job design is qualitatively identical to that in Proposition 2.

Q.E.D.

Appendix B: Proofs for Chapter 2

Proof of Lemma 5. Skaperdas (1996) proved the following: in a contest with n players (fixed number), a contest success function satisfies (B1)-(B5) if and only if it satisfies (B6).

(B1) $\sum_{i=1}^{n} p_i(\mathbf{x}) = 1$ and $p_i(\mathbf{x}) \geq 0$ for all $i \in \{1, ..., n\}$ and all $\mathbf{x}$; if $x_i > 0$, then $p_i(\mathbf{x}) > 0$.

(B2) For all $i \in \{1, ..., n\}$, $p_i(\mathbf{x})$ is increasing in y_i and decreasing in y_j for all $j \neq i$.

(B3) For any permutation φ of $\{1, ..., n\}$ (i.e., a bijection $\varphi : \{1, ..., n\} \to \{1, ..., n\}$) we have

$$p_i(\mathbf{x}) = p_{\varphi(i)}(x_{\varphi(1)}, ..., x_{\varphi(n)}), \forall i \in \{1, ..., n\}.$$

(B4) Denote p_i^m the probability of winning in the subcontest where the players are in the subset M. Consistency requires

$$p_i^m(x_1, ..., x_n) = \frac{p_i(x_1, ..., x_n)}{\sum\limits_{j \in M} p_j(x_1, .., x_n)} \forall i \in M, \forall M \subseteq \{1, ..., n\}$$

(B5) p_i^m is independent of the x_is of the players not included in the subset M.

(B6) $p_i(\mathbf{x}) = \frac{f(x_i)}{\sum_{j=1}^{n} f(x_j)}$ for all $i \in \{1, ..., n\}$ and f is unique up to positive multiplicative transformations.

Let $\{p_i^n\}$ be a system of contest success functions that satisfies (A1)-(A4). Consider $n = 9$, $\{p_i^9\}$ satisfies (B1)-(B3), and $\{p_i^n\}_{n=1}^{8}$ satisfies (B4). (B5) is trivially satisfied as the definition of $\{p_i^n\}_{n=1}^{8}$ only contains the efforts of players in the game. Thus, there exists a $f(.)$ such that $p_i^n(\mathbf{x}) = \frac{f(x_i)}{\sum_{j=1}^{n} f(x_j)}, \forall n \leq 9$.

The next step is to prove p_i^n shares the same form for all $n > 9$. Suppose (by contradiction) that there exist a $k > 9, i \in \{1, ..., k\}$ such that $p_i^k(\mathbf{x}) \neq \frac{f(x_i)}{\sum_{j=1}^{k} f(x_j)}$. By similar arguments, $\{p_i^n\}_{n=1}^k$ satisfies (B1)-(B5), so there exists a $g(.)$ such that $p_i^n(\mathbf{x}) = \frac{g(x_i)}{\sum_{j=1}^{n} g(x_j)}, \forall n \leq k$. Thus, for $n \leq 9$, $p_i^n(\mathbf{x}) = \frac{f(x_i)}{\sum_{j=1}^{n} f(x_j)} = \frac{g(x_i)}{\sum_{j=1}^{n} g(x_j)}$. As Skaperdas (1996) have proved, $f(.)$ is unique up to positive multiplicative transformations, so $g(x_i) = \beta f(x_i)$ where $\beta > 0$. Plug it back into p_i^k:

$$p_i^k(\mathbf{x}) = \frac{g(x_i)}{\sum_{j=1}^{k} g(x_j)} = \frac{\beta f(x_i)}{\sum_{j=1}^{k} \beta f(x_j)} = \frac{f(x_i)}{\sum_{j=1}^{k} f(x_j)}$$

Contradiction! Thus, $p_i^n(\mathbf{x}) = \frac{f(x_i)}{\sum_{j=1}^{n} f(x_j)}, \forall n$.

Q.E.D.

Proof of Proposition 5. For player i, the maximization program can be written as:

$$\max_{x_i} \sum_{n=1}^{\infty} \tilde{\pi}(n) p_i^n(x_i, x_{-i}) v(n) - x_i.$$

Suppose the equilibrium effort level is x^*, and plug in the contest success function $p_i^n(\mathbf{x}) = \frac{f(x_i)}{\sum_{j=1}^{n} f(x_j)}$, the program becomes:

$$\max_{x_i} \sum_{n=1}^{\infty} \tilde{\pi}(n) \frac{f(x_i)}{f(x_i) + (n-1) f(x^*)} v(n) - x_i.$$

Taking derivative with respect to x_i:

$$\sum_{n=1}^{\infty} \tilde{\pi}(n) v(n) \frac{f'(x_i)(n-1) f(x^*)}{(f(x_i) + (n-1) f(x^*))^2} = 1.$$

In equilibrium, it must be $x_i = x^*$, so it requires:

$$\frac{f(x^*)}{f'(x^*)} = \sum_{n=1}^{\infty} \tilde{\pi}(n)v(n)\frac{n-1}{n^2}.$$

$f(.)$ is positive, concave and increasing in x , so $\frac{f(x)}{f'(x)}$ is positive and increasing in x. Thus, one of the two cases must be true:

- There is a unique x^* such that $\frac{f(x^*)}{f'(x^*)} = \sum_{n=1}^{\infty} \tilde{\pi}(n)v(n)\frac{n-1}{n^2}$.
- $\forall x \in [0, \infty)$, $\frac{f(x)}{f'(x)} > \sum_{n=1}^{\infty} \tilde{\pi}(n)v(n)\frac{n-1}{n^2}$

In the first case, the equilibrium exists (effort level is x^*), which is unique. In the second case, the effort level $x^* = 0$ is the only equilibrium, so it exists, and it is unique.

Q.E.D.

Lemma 6. *Consider two distributions, F and G, where F is a discreet distribution over $\mathbb{N} = \{1, 2, ..., n, ...\}$ with density function $f(n) = \pi(n)$ and G is a discreet distribution with density function $g(n) = \frac{\frac{1}{n}\pi(n)}{\sum_i \frac{1}{i}\pi(i)}$.*

F and G are valid distributions, and F has first-order stochastic dominance over G.

Proof of Lemma 6. First we verify that both F and G are valid distributions:

$$\sum_{n=1}^{\infty} \pi(n) = 1$$

$$\sum_{n=1}^{\infty} \frac{1}{n}\pi(n) \leq \sum_{n=1}^{\infty} \pi(n) = 1$$

$$\sum_{n=1}^{\infty} \frac{\frac{1}{n}\pi(n)}{\sum_i \frac{1}{i}\pi(i)} = 1$$

For any integer $j > 0$

$$
\begin{aligned}
G(j) - F(j) &= \tfrac{1}{\sum_i \frac{1}{i}\pi(i)} \sum_{n=1}^{j} \tfrac{1}{n}\pi(n) - \sum_{n=1}^{j} \pi(n) \\
&= \tfrac{1}{\sum_i \frac{1}{i}\pi(i)} \left[\sum_{n=1}^{j} \tfrac{1}{n}\pi(n) - \sum_{n=1}^{j} \pi(n) \textstyle\sum_i \frac{1}{i}\pi(i) \right] \\
&= \tfrac{1}{\sum_i \frac{1}{i}\pi(i)} \left[\sum_{n=1}^{j} \tfrac{1}{n}\pi(n) \sum_{n=1}^{\infty} \pi(n) - \sum_{n=1}^{\infty} \tfrac{1}{n}\pi(n) \sum_{n=1}^{j} \pi(n) \right] \\
&= \tfrac{1}{\sum_i \frac{1}{i}\pi(i)} \left[\sum_{n=1}^{j} \tfrac{1}{n}\pi(n) \sum_{n=j+1}^{\infty} \pi(n) - \sum_{n=j+1}^{\infty} \tfrac{1}{n}\pi(n) \sum_{n=1}^{j} \pi(n) \right] \\
&\geq \tfrac{1}{\sum_i \frac{1}{i}\pi(i)} \left[\sum_{n=1}^{j} \tfrac{1}{j}\pi(n) \sum_{n=j+1}^{\infty} \pi(n) - \sum_{n=j+1}^{\infty} \tfrac{1}{j+1}\pi(n) \sum_{n=1}^{j} \pi(n) \right] \\
&= \tfrac{1}{\sum_i \frac{1}{i}\pi(i)} \sum_{n=1}^{j} \pi(n) \sum_{n=j+1}^{\infty} \pi(n) \left(\tfrac{1}{j} - \tfrac{1}{j+1} \right) \\
&\geq 0
\end{aligned}
$$

Thus, we have $G(j) \geq F(j)$ for all j, so F first-order stochastic dominate G.

Q.E.D.

Proof of Proposition 6. Given equilibrium, we have

$$
\sum_{n=1}^{\infty} \tilde{\pi}(n) v(n) \frac{n-1}{n^2} = \frac{f(x^*)}{f'(x^*)}
$$

Under contest C_1, $v_1(n)$ is increasing in n and under contest C_3, v_3 is a constant. Once we plug in the probability where $\tilde{\pi}(n) = \frac{n\pi(n)}{\sum_i i\pi(i)}$, we have

$$
\frac{f(x^*)}{f'(x^*)} = \frac{1}{\sum_i i\pi(i)} \sum_{n=1}^{\infty} v(n)\pi(n) \frac{n-1}{n}
$$

To facilitate our proof, we set a benchmark value

$$B = \left[1 - \sum_{n=1}^{\infty} \frac{1}{n}\pi(n)\right] \sum_{n=1}^{\infty} \pi(n)v(n).$$

Since we have $\mathbb{E}[v_1(n)] = \mathbb{E}[v_2(n)] = v_3$, B is a constant under all contests. Thus, we have

$$\begin{aligned}
\sum_{n=1}^{\infty} v_3(n)\pi(n)\tfrac{n-1}{n} - B &= \sum_{n=1}^{\infty} v_3(n)\pi(n)[1-\tfrac{1}{n}] - \left[1 - \sum_{n=1}^{\infty} \tfrac{1}{n}\pi(n)\right] \sum_{n=1}^{\infty} \pi(n)v_3(n) \\
&= \sum_{n=1}^{\infty} \tfrac{1}{n}\pi(n) \sum_{n=1}^{\infty} \pi(n)v_3(n) - \sum_{n=1}^{\infty} v_3(n)\pi(n)\tfrac{1}{n} \\
&= v_3 \left[\sum_{n=1}^{\infty} \tfrac{1}{n}\pi(n) \sum_{n=1}^{\infty} \pi(n) - \sum_{n=1}^{\infty} \tfrac{1}{n}\pi(n)\right] \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\sum_{n=1}^{\infty} v_1(n)\pi(n)\tfrac{n-1}{n} - B &= \sum_{n=1}^{\infty} v_1(n)\pi(n)[1-\tfrac{1}{n}] - \left[1 - \sum_{n=1}^{\infty} \tfrac{1}{n}\pi(n)\right] \sum_{n=1}^{\infty} \pi(n)v_1(n) \\
&= \sum_{n=1}^{\infty} \tfrac{1}{n}\pi(n) \sum_{n=1}^{\infty} \pi(n)v_1(n) - \sum_{n=1}^{\infty} v_1(n)\pi(n)\tfrac{1}{n} \\
&= \sum_{n=1}^{\infty} \tfrac{1}{n}\pi(n)\mathbb{E}[v_1(n)] - \sum_{n=1}^{\infty} v_1(n)\pi(n)\tfrac{1}{n} \\
&= \sum_{n=1}^{\infty} \tfrac{1}{n}\pi(n) \left[\mathbb{E}[v_1(n)] - v_1(n)\right] \\
&= \left(\sum_{i=1}^{\infty} \tfrac{1}{i}\pi(i)\right) \sum_{n=1}^{\infty} \frac{\frac{1}{n}\pi(n)}{\sum_{i=1}^{\infty} \frac{1}{i}\pi(i)} \left[\mathbb{E}[v_1(n)] - v_1(n)\right] \\
&= \left(\sum_{i=1}^{\infty} \tfrac{1}{i}\pi(i)\right) \int_0^{\infty} \left[\mathbb{E}[v_1(n)] - v_1(n)\right] dG \\
&\geq \left(\sum_{i=1}^{\infty} \tfrac{1}{i}\pi(i)\right) \int_0^{\infty} \left[\mathbb{E}[v_1(n)] - v_1(n)\right] dF \\
&= \left(\sum_{i=1}^{\infty} \tfrac{1}{i}\pi(i)\right) \left[\mathbb{E}[v_1(n)] - \int_0^{\infty} v_1(n) dF\right] \\
&= 0
\end{aligned}$$

The inequality is true because F FOSD G. Similarly,

$$\begin{aligned}
\sum_{n=1}^{\infty} v_2(n)\pi(n)\tfrac{n-1}{n} - B &= \sum_{n=1}^{\infty} \tfrac{1}{n}\pi(n)\left[\mathbb{E}[v_2(n)] - v_2(n)\right] \\
&= \left(\sum_{i=1}^{\infty} \tfrac{1}{i}\pi(i)\right) \int_0^{\infty} \left[\mathbb{E}[v_2(n)] - v_2(n)\right] dG \\
&\leq \left(\sum_{i=1}^{\infty} \tfrac{1}{i}\pi(i)\right) \int_0^{\infty} \left[\mathbb{E}[v_2(n)] - v_2(n)\right] dF \\
&= 0
\end{aligned}$$

Thus, we have

$$\sum_{n=1}^{\infty} v_2(n)\pi(n)\frac{n-1}{n} \leq \sum_{n=1}^{\infty} v_3(n)\pi(n)\frac{n-1}{n} \leq \sum_{n=1}^{\infty} v_1(n)\pi(n)\frac{n-1}{n}$$

It is the same as

$$\frac{f(x_2^*)}{f'(x_2^*)} \leq \frac{f(x_3^*)}{f'(x_3^*)} \leq \frac{f(x_1^*)}{f'(x_1^*)}$$

We know that $\frac{f(x)}{f'(x)}$ is increasing in x, thus we have:

$$x_2^* \leq x_3^* \leq x_1^*.$$

Q.E.D.

Proof of Proposition 7. Myerson and Wärneryd (2006) showed that $x_4^* > x_5^*$. Proposition 6 showed that:

- when $b > 0$, $x_5^* < x_6^*$.
- when $b = 0$, $x_5^* = x_6^*$.

- when $b < 0$, $x_5^* > x_6^*$.

Now I want to compare x_4^* and x_6^*. x_4^* and x_6^* can be pinned down by:

$$\frac{f(x_4^*)}{f'(x_4^*)} = \frac{v}{\mu}\frac{\mu - 1}{\mu}$$

$$\frac{f(x_6^*)}{f'(x_6^*)} = \frac{1}{\mu}\sum_{n=1}^{\infty}(a + bn)\pi(n)\frac{n-1}{n}$$

Thus, we have

$$\begin{aligned}
\tfrac{f(x_4^*)}{f'(x_4^*)} - \tfrac{f(x_6^*)}{f'(x_6^*)} &= \tfrac{v}{\mu}\tfrac{\mu-1}{\mu} - \tfrac{1}{\mu}\textstyle\sum_{n=1}^{\infty}(a + bn)\pi(n)\tfrac{n-1}{n} \\
&= \tfrac{1}{\mu}\left[\textstyle\sum_{n=1}^{\infty}(a + bn)\pi(n)\tfrac{1}{n} - \tfrac{v}{\mu}\right] \\
&= \tfrac{1}{\mu}\left[\textstyle\sum_{n=1}^{\infty} a\pi(n)\tfrac{1}{n} + b - \tfrac{a+\mu b}{\mu}\right] \\
&= \tfrac{a}{\mu}\left[\textstyle\sum_{n=1}^{\infty} \pi(n)\tfrac{1}{n} - \tfrac{1}{\mu}\right] \\
&= \tfrac{a}{\mu}\left[\mathbb{E}_\pi[\tfrac{1}{n}] - \tfrac{1}{\mathbb{E}_\pi[n]}\right]
\end{aligned}$$

$\mathbb{E}_\pi[\frac{1}{n}] - \frac{1}{\mathbb{E}_\pi[n]} > 0$ since $\frac{1}{n}$ is a convex function, and $\frac{f(x_4^*)}{f'(x_4^*)} > \frac{f(x_6^*)}{f'(x_6^*)} \Leftrightarrow x_4^* > x_6^*$ since $f(.)$ is increasing and concave. Thus, the comparison between x_4^* and x_6^* depends on a:

- When $a > 0$, $x_4^* > x_6^*$.
- When $a = 0$, $x_4^* = x_6^*$.
- When $a < 0$, $x_4^* < x_6^*$.

The proposition focuses on scenarios with $\mu > 0$, and there are only five possible combinations: (i) $a > 0, b < 0$, (ii) $a > 0, b = 0$, (iii) $a > 0, b > 0$, (iv) $a = 0, b > 0$, and (v) $a < 0, b > 0$. The proposition is an immediate result of the above analysis.

Q.E.D.

Appendix C: Proofs for Chapter 3

Proof of Proposition 8. For any equilibrium, the assimilation choice must be a cutoff strategy. To see this, for any given agent i with skill level s_i and background $\mathcal{I}$, the payoff is

$$u_i = \begin{cases} f(m_A)s_A s_i - d & \text{if } a_i = 1 \\ f(m_I)s_I s_i & \text{if } a_i = 0 \end{cases}$$

As each agent has infinitesimal mass, the choice of one agent will not affect m_A, m_I, s_A and s_I. Thus, one of the three cases must be true:

- $f(m_A)s_A s_i - d \leq f(m_I)s_I s_i$ for all $s_i \in [0, 1]$. In this case, no one will assimilate.
- $f(m_A)s_A s_i - d \geq f(m_I)s_I s_i$ for all $s_i \in [0, 1]$. In this case, all agents will assimilate.
- $f(m_A)s_A s_i - d = f(m_I)s_I s_i$ for some $s_i = c \in [0, 1]$. In this case, agents with $s_i > c$ will assimilate, agents with $s_i < c$ will not assimilate, and agents with $s_i = c$ is indifferent between assimilate and not assimilate.

In either case, agents' action could be summarized as a cutoff strategy:

$$a_i = \begin{cases} 1 & \text{if } s_i > c \\ 0 & \text{if } s_i < c \\ 1 \text{ with proability } p & \text{if } s_i = c \end{cases}$$

The following steps prove the existence of the equilibrium.

First, denote $F(s)$ be the cumulative distribution function of skill levels of agents with background $\mathcal{I}$, and $G(s)$ be the cumulative distribution function of skill levels of agents with background $\mathcal{A}$. Denote $s_{\mathcal{I}}$ be the average skill level of agents with background $\mathcal{I}$, and $s_{\mathcal{A}}$ be the average skill level of agents with background $\mathcal{A}$:

$$s_{\mathcal{I}} = \int_0^1 sdF(s) \text{ and } s_{\mathcal{A}} = \int_0^1 sdG(s).$$

The cutoff strategy can be described as: a proportion (τ) of agents with highest skill levels will choose to assimilate. Thus, for any $\tau \in [0,1]$, denote $c(\tau) = \min_s\{F(s) \geq \tau\}$, and $p(\tau) = f(c(\tau)) - \tau$. Since F is increasing and right-continuous, the definition is valid. Thus, the cutoff strategy can be fully characterized by a parameter $\tau \in [0,1]$.

Consider three situations: (i) $\tau = 1$ is an equilibrium, (ii) $\tau = 0$ is an equilibrium, and (iii) $\tau \in (0,1)$ is an equilibrium.

Case (i) If $\tau = 1$ is an equilibrium, that means all agents will assimilate. This equilibrium exists if $d = 0$.

Case (ii) If $\tau = 0$ is an equilibrium, that means no agent will assimilate. This equilibrium exists if $f(1-m)s_{\mathcal{A}} - d \leq f(m)s_{\mathcal{I}}$.

Case (iii) If $\tau \in (0,1)$ is an equilibrium, then agent with skill level c is indifferent between assimilation and not assimilation. Thus, the following equation must hold:

$$f(1-m+m\tau)s_A c(\tau) - d = f(m(1-\tau))s_I c(\tau)$$

where $s_A = \frac{(1-m)s_A + m\int_{c(\tau)}^1 sdF(s)}{1-m+m\tau}$ and $s_I = \frac{\int_0^{c(\tau)} sdF(s)}{1-\tau}$. The above equation is the same as

$$[f(1-m+m\tau)s_A - f(m(1-\tau))s_I]c(\tau) = d$$

If $d > 0$ and $f(1-m)s_{\mathcal{A}} - f(m)s_{\mathcal{I}} > d$, then the left-hand side of the equation is larger than d if $\tau = 0$, and it is 0 if $\tau = 1$. Further, it is continuous in τ, so by Intermediate Value Theorem, there exists a $\tau \in (0,1)$ such that the equation holds. Thus, if $d > 0$ and $f(1-m)s_{\mathcal{A}} - f(m)s_{\mathcal{I}} > d$, an equilibrium exists.

Above all, for any bounded measurable function s over N and any discrimination level $d \in \mathbb{R}_+$, an equilibrium always exists.

Q.E.D.

Proof of Corollary 1. The Corollary is an immediate result from Proposition 8, there is no mass point in the distribution of skill levels.

Q.E.D.

Proof of Proposition 9. For any agent with background $\mathcal{A}$, the utility maximization program

$$\max_d f(m_A)s_A \quad s.t. \quad c \in C(s,d).$$

Since the working skill distribution $s \in S$, choosing c is equivalent to choosing a cutoff θ^c such that agents with $\theta \geq \theta^c$ will assimilate and $\theta < \theta^c$ will not assimilate. θ^c can be supported as an equilibrium as long as $d = [f(m_A)s_A - f(m_I)s_I]\theta^c \geq 0$. It is evident that m_A is decreasing in θ^c, m_I is increasing in θ^c, and s_I is increasing in θ^c. For s_A, it is decreasing when $\theta^c < \tilde{\theta}$ and increasing when $\theta^c > \tilde{\theta}$, where $\tilde{\theta}$ is unique and can be calculated by $s_A = s_i(\tilde{\theta})$.

If $\theta^c \leq \tilde{\theta}$, then $s_A > s_I$ and $m_A > m_I$, so any θ^c can be supported as an equilibrium by some d.

If $\theta^c > \tilde{\theta}$, $[f(m_A)s_A - f(m_I)s_I]$ is decreasing in θ^c, and it is continuous. Denote $\bar{\theta} = 1$ if $[f(m_A)s_A - f(m_I)s_I] > 0$ when $\theta^c = 1$; otherwise, denote $\bar{\theta}$ be the solution to the equation $[f(m_A)s_A - f(m_I)s_I] = 0$. $\bar{\theta}$ is unique as $[f(m_A)s_A - f(m_I)s_I]$ is monotonic on $[\tilde{\theta}, 1]$.

Thus, choosing $d \in [0, \infty)$ is equivalent to choosing $\theta^c \in [0, \bar{\theta}]$. As a result, I can write the maximization program of h as:

$$\max_{\theta^c \in [0,\bar{\theta}]} f(m_A)s_A$$

Since $f(m_A)s_A$ is continuous in θ^c, and $[0, \bar{\theta}]$ is compact. $f(m_A)s_A$ achieves maximum on $[0, \bar{\theta}]$. Denote the maximizer as θ^*. The corresponding d^* can be calculated by $d^* = [f(m_A^*)s_A^* - f(m_I^*)s_I^*]\theta^*$.

The following steps show that $\theta = 0$ and $\theta = 1$ cannot be the maximum, so the maximum is achieved on $(0, 1)$.

First m_A is decreasing in θ^c, since high θ^c means less agents will assimilate. Also, s_A is decreasing in θ^c at $\theta^c = 1$, as agents with highest skill levels assimilate will increase the average skill level of group A. Thus, $f(m_A)s_A$ is decreasing at $\theta^c = 1$, so it does not achieve maximum at $\theta = 1$.

Then I want to show $f(m_A)s_A$ does not achieve maximum at $\theta^c = 0$. Calculate $f(m_A)s_A$

as a function of θ^c:

$$f(m_A)s_A = \frac{f(1-m\theta^c)}{1-m\theta^c}\left[(1-m)s_A + m\int_{\theta^c}^{1} s(t)dt\right]$$

Denote $H(\theta^c) = \frac{f(1-m\theta^c)}{1-m\theta^c}$. Taking derivative with respect to θ^c:

$$\begin{aligned} H'(\theta^c) &= \tfrac{f'(1-m\theta^c)(-m)(1-m\theta^c)-f(1-m\theta^c)(-m)}{(1-m\theta^c)^2} \\ &= \tfrac{-m}{(1-m\theta^c)^2}\left[f'(1-m\theta^c)(1-m\theta^c) - f(1-m\theta^c)\right] \end{aligned}$$

$f'(1) < f(1) \Rightarrow H'(0) > 0$, so $H(\theta^c)$ is increasing in θ^c at $\theta^c = 0$. Also, $\left[(1-m)s_A + m\int_{\theta^c}^{1} s(t)dt\right]$ is increasing in θ^c, so $f(m_A)s_A$ does not achieve maximum at $\theta^c = 0$.

Q.E.D.

Proof of Proposition 10 . For agent with background $\mathcal{A}$, the utility maximization program becomes:

$$\max_{l_i\in[0,1]} \log(l_i) + \beta_{\mathcal{A}}\log(f(m_A)s_A s_i) \quad s.t. \quad s_i = \theta(1-l_i).$$

Substitute l_i with s_i would result in

$$\max \log(1-\frac{s_i}{\theta}) + \beta_{\mathcal{A}}\log(f(m_A)s_A s_i).$$

FOC:

$$\frac{1}{1-\frac{s_i}{\theta}}\left(-\frac{1}{\theta}\right) + \frac{\beta_{\mathcal{A}}}{s_i} = 0.$$

After simplification:

$$s_i^* = \frac{\beta_{\mathcal{A}}}{1+\beta_{\mathcal{A}}}\theta.$$

Similarly, for agent with background $\mathcal{I}$ and $a_i = 0$, $s_i^* = \frac{\beta_{\mathcal{I}}}{1+\beta_{\mathcal{I}}}\theta$.

For agent with background $\mathcal{I}$ and $a_i = 1$, the program becomes:

$$\max_{l_i \in [0,1]} \log(l_i) + \beta_{\mathcal{I}} \log(f(m_A)s_A s_i - d) \quad s.t. \quad s_i = \theta(1 - l_i).$$

FOC:

$$\frac{1}{1-\frac{s_i}{\theta}}\left(-\frac{1}{\theta}\right) + \beta_{\mathcal{I}}\frac{f(m_A)s_A}{f(m_A)s_A s_i - d} = 0.$$

After simplification:

$$s_i^* = \frac{\beta_{\mathcal{I}}}{1+\beta_{\mathcal{I}}}\theta + s^*$$

where $s^* = \frac{d^*}{(1+\beta_{\mathcal{I}})f(m_A)s_A}$.

Q.E.D.

Proof of Proposition 11. Denote $\gamma_{\mathcal{A}} = \frac{\beta_{\mathcal{A}}}{1+\beta_{\mathcal{A}}}$ and $\gamma_{\mathcal{I}} = \frac{\beta_{\mathcal{I}}}{1+\beta_{\mathcal{I}}}$. According to Proposition 8, 9, and 10, if the equilibrium exists, it must have the following form: agents will assimilate iff $\theta \geq \theta^*$. The on path strategies are:

$$s_i = \begin{cases} \gamma_{\mathcal{A}}\theta_i & \text{if agent has background } \mathcal{A} \\ \\ \gamma_{\mathcal{I}}\theta_i + s^* & \text{if agent has background } \mathcal{I} \text{ and } \theta_i \geq \theta^* \\ \\ \gamma_{\mathcal{I}}\theta_i & \text{if agent has background } \mathcal{I} \text{ and } \theta_i < \theta^* \end{cases}$$

For agents with background $\mathcal{I}$, the assimilation choice

$$a_i = \begin{cases} 1 & \text{if } s_i \geq c^* \\ \\ 0 & \text{if } s_i < c^* \end{cases}$$

and the following conditions must hold:

$$\begin{cases} d^* &=& [f(m_A)s_A - f(m_I)s_I]c^* \\ \\ \theta^* &=& \arg\max_{\theta} f(m_A)s_A \\ \\ c^* &\in& [\gamma_{\mathcal{I}}\theta^*, \gamma_{\mathcal{I}}\theta^* + s^*] \\ \\ s^* &=& \frac{d^*}{(1+\beta_{\mathcal{I}})f(m_A)s_A} \end{cases}$$

Simplify these conditions by substituting c^* and d^*:

$$\begin{cases} \theta^* &=& \arg\max_{\theta} f(m_A)s_A \\ \\ \frac{(1+\beta_{\mathcal{I}})f(m_A)s_A s^*}{f(m_A)s_A - f(m_I)s_I} &\in& [\gamma_{\mathcal{I}}\theta^*, \gamma_{\mathcal{I}}\theta^* + s^*] \end{cases}$$

I first focus on the first condition. Given the skill acquisition strategies above, the objective function could be written as:

$$f(m_A)s_A = \begin{cases} \frac{f(1-m\theta)}{1-m\theta}\left[\frac{1}{2}(1-m)\gamma_{\mathcal{A}} + \frac{1}{2}m\gamma_{\mathcal{I}}(1-\theta^2) + m(1-\theta)s^*\right] & \text{if } \theta \geq \theta^* \\ \frac{f(1-m\theta)}{1-m\theta}\left[\frac{1}{2}(1-m)\gamma_{\mathcal{A}} + \frac{1}{2}m\gamma_{\mathcal{I}}(1-\theta^2) + m(1-\theta^*)s^*\right] & \text{if } \theta < \theta^* \end{cases}$$

As $f(m_A)s_A$ achieves maximum at θ^*, it is equivalent to $f(m_A)s_A$ achieves maximum at θ^* on both $[0, \theta^*]$ and $[\theta^*, 1]$. Thus, I can take derivative with respect to θ on two intervals separately. $\theta^* = \arg\max_{\theta} f(m_A)s_A$ is equivalent to:

$$\begin{cases} [f(1-m\theta^*) - (1-m\theta^*)f'(1-m\theta^*)]s_A & \leq & f(1-m\theta^*)[\gamma_{\mathcal{I}}\theta^* + s^*] \\ [f(1-m\theta^*) - (1-m\theta^*)f'(1-m\theta^*)]s_A & \geq & f(1-m\theta^*)\gamma_{\mathcal{I}}\theta^* \end{cases}$$

Now combine it with the second condition. The equilibrium exists as long as there exists (θ^*, s^*) that satisfies the condition below:

$$\begin{cases} [f(1-m\theta^*) - (1-m\theta^*)f'(1-m\theta^*)]s_A & \leq & f(1-m\theta^*)[\gamma_{\mathcal{I}}\theta^* + s^*] \\ [f(1-m\theta^*) - (1-m\theta^*)f'(1-m\theta^*)]s_A & \geq & f(1-m\theta^*)\gamma_{\mathcal{I}}\theta^* \\ (1+\beta_{\mathcal{I}})s^* f(m_A)s_A & \leq & [f(m_A)s_A - f(m_I)s_I][\gamma_{\mathcal{I}}\theta^* + s^*] \\ (1+\beta_{\mathcal{I}})s^* f(m_A)s_A & \geq & [f(m_A)s_A - f(m_I)s_I]\gamma_{\mathcal{I}}\theta^* \end{cases} \tag{7}$$

I then show the existence of (θ^*, s^*) that satisfies condition (7), thus finish the proof.

First, I find (θ^*, s^*) that satisfies the following equations:

$$\begin{cases} s^* &= \frac{1}{1+\beta_{\mathcal{I}}}\gamma_{\mathcal{I}}\theta^* \\ [f(1-m\theta^*) - (1-m\theta^*)f'(1-m\theta^*)]s_A &= \frac{2+\beta_{\mathcal{I}}}{1+\beta_{\mathcal{I}}} f(1-m\theta^*)\gamma_{\mathcal{I}}\theta^* \end{cases} \tag{8}$$

I want to show the definition of this (θ^*, s^*) is valid. I substitute s^* and reduce condition (8) to:

$$\frac{\frac{1}{1-m\theta^*}[f(1-m\theta^*)-(1-m\theta^*)f'(1-m\theta^*)]}{\left[(1-m)\frac{1}{2}\gamma_{\mathcal{A}}+m\frac{1}{2}\gamma_{\mathcal{I}}(1-(\theta^*)^2)+m(1-\theta^*)\frac{1}{1+\beta_{\mathcal{I}}}\gamma_{\mathcal{I}}\theta^*\right]} = \frac{2+\beta_{\mathcal{I}}}{1+\beta_{\mathcal{I}}} f(1-m\theta^*)\gamma_{\mathcal{I}}\theta^*$$

The only variable in this equation is θ^*, and the rest parameters are all exogenous given. The LHS and RHS are all continuous in θ^*. The solution of this equation is guaranteed by Intermediate Value Theorem, as $LFS > RHS$ if $\theta^* = 0$, and $LFS < RHS$ if $\theta^* = 1$ (note $\gamma_{\mathcal{A}} < \gamma_{\mathcal{I}}$). Thus, I proved that there always exists (θ^*, s^*) that satisfies the condition (8)

Then I want to show that if (θ^*, s^*) satisfies condition (8), it must also satisfy condition (7). For the four inequalities in condition (7):

- The first inequality holds with equality. It is a rearrangement of condition (8).

- The second inequality holds, since the first inequality holds with equality, and $s^* > 0$.

- The third inequality holds, as:

$$\begin{aligned}
\frac{s_A}{s_I} &= \frac{2s_A}{\gamma_{\mathcal{I}}\theta^*} \\
&= (2+\beta_{\mathcal{I}})\frac{2}{1+\beta_{\mathcal{I}}}\frac{f(m_A)}{f(m_A)-m_A f'(m_A)} \\
&> (2+\beta_{\mathcal{I}})
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \qquad \frac{f(m_A)s_A}{f(m_I)s_I} &\geq 2+\beta_{\mathcal{I}} \\
\Leftrightarrow \qquad f(m_A)s_A &\leq \frac{2+\beta_{\mathcal{I}}}{1+\beta_{\mathcal{I}}}[f(m_A)s_A - f(m_I)s_I] \\
\Leftrightarrow \quad (1+\beta_{\mathcal{I}})s^* f(m_A)s_A &\leq [f(m_A)s_A - f(m_I)s_I][\gamma_{\mathcal{I}}\theta^* + s^*]
\end{aligned}$$

- The forth inequality holds, as:

$$(1+\beta_{\mathcal{I}})s^* f(m_A)s_A \;=\; f(m_A)s_A\gamma_{\mathcal{I}}\theta^* \;\geq\; [f(m_A)s_A - f(m_I)s_I]\gamma_{\mathcal{I}}\theta^*$$

Q.E.D.

Proof of Proposition 12. For $f(m) = \frac{1}{2} - \frac{1}{2}(1-m)^2 = m - \frac{1}{2}m^2$, I can write $f(1-m\theta) = \frac{1}{2}(1-m\theta)(1+m\theta)$. The condition (7) becomes:

$$\begin{cases}
(1-m\theta^*)s_A \leq (1+m\theta^*)(\gamma_{\mathcal{I}}\theta^* + s^*) \\
(1-m\theta^*)s_A \geq (1+m\theta^*)\gamma_{\mathcal{I}}\theta^* \\
(1+\beta_{\mathcal{I}})s^* f(m_A)s_A \leq [f(m_A)s_A - f(m_I)s_I][\gamma_{\mathcal{I}}\theta^* + s^*] \\
(1+\beta_{\mathcal{I}})s^* f(m_A)s_A \geq [f(m_A)s_A - f(m_I)s_I]\gamma_{\mathcal{I}}\theta^*
\end{cases}$$

Let $P = f(m_A)s_A, Q = f(m_I)s_I$, then $s_A = \frac{2P}{(1-m\theta^*)(1+m\theta^*)}$. Substituting $f(m_A), s_A, f(m_I), s_I$ in the conditions:

$$\begin{cases} 2P & \leq & (1+m\theta^*)^2(\gamma_{\mathcal{I}}\theta^* + s^*) \\ 2P & \geq & (1+m\theta^*)^2\gamma_{\mathcal{I}}\theta^* \\ (1+\beta_{\mathcal{I}})s^*P & \leq & (P-Q)(\gamma_{\mathcal{I}}\theta^* + s^*) \\ (1+\beta_{\mathcal{I}})s^*P & \geq & (P-Q)\gamma_{\mathcal{I}}\theta^* \end{cases}$$

The existence of (θ^*, s^*) has been proved in Proposition 11. When (θ^*, s^*) satisfies the condition above, the equilibrium exists and the on-path strategies are described below. Thus, I finish the proof.

$$s_i = \begin{cases} \gamma_{\mathcal{A}}\theta_i & \text{if } i \in N_{\mathcal{A}}; \\ \gamma_{\mathcal{I}}\theta_i + s^* & \text{if } \theta_i \geq \theta^* \text{ and } i \in N_{\mathcal{I}}; \\ \gamma_{\mathcal{I}}\theta_i & \text{if } \theta_i < \theta^* \text{ and } i \in N_{\mathcal{I}}. \end{cases}$$

$$d^* = (1+\beta_{\mathcal{I}})f(m_A)s_As^*.$$

$$a_i(s_i) = \begin{cases} 1 & \text{if } s_i \geq \gamma_{\mathcal{I}}\theta^* + s^*; \\ 0 & \text{if } s_i < \gamma_{\mathcal{I}}\theta^*. \end{cases}$$

Q.E.D.

BIBLIOGRAPHY

BIBLIOGRAPHY

Advani, Arun and Bryony Reich (2015), "Melting pot or salad bowl: The formation of heterogeneous communities." Technical report, IFS Working Papers.

Aghion, Philippe and Jean Tirole (1997), "Formal and Real Authority in Organizations." *Journal of Political Economy*, 105, 1–29.

Akerlof, George A. (1970), "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics*, 84, 488–500.

Akerlof, George A. and Rachel E. Kranton (2000), "Economics and identity." *The Quarterly Journal of Economics*, 115, 715–753.

Alonso, Ricardo, Wouter Dessein, and Niko Matouschek (2008), "When Does Coordination Require Centralization?" *American Economic Review*, 98, 145–179.

Alonso, Ricardo and Niko Matouschek (2008), "Optimal Delegation." *Review of Economic Studies*, 75, 259–293.

Aoki, Masahiko (1986), "Horizontal vs. Vertical Information Structure of the Firm." *The American Economic Review*, 971–983.

Athey, Susan and John Roberts (2001), "Organizational design: Decision rights and incentive contracts." *American Economic Review*, 91, 200–205.

Azmat, Ghazala and Marc Möller (2009), "Competition among contests." *The RAND Journal of Economics*, 40, 743–768.

Benjamin, Daniel J., James J. Choi, and A. J. Strickland (2010), "Social identity and preferences." *The American Economic Review*, 100, 1913–1928.

Besanko, David, Pierre Régibeau, and Katharine E. Rockett (2005), "A Multi-Task Principal-Agent Approach to Organizational Form." *Journal of Industrial Economics*, 53, 437–467.

Blanes i Vidal, Jordi and Marc Möller (2016), "Project selection and execution in teams." *The RAND Journal of Economics*, 47, 166–185.

Bolton, P. and M. Dewatripont (1994), "The Firm as a Communication Network." *The Quarterly Journal of Economics*, 109, 809–839.

Bouton, Laurent (2013), "A Theory of Strategic Voting in Runoff Elections." *American Economic Review*, 103, 1248–1288.

Bouton, Laurent and Micael Castanheira (2012), "One Person, Many Votes: Divided Majority and Information Aggregation." *Econometrica*, 80, 43–87.

Campbell, Colin M. (1999), "Large Electorates and Decisive Minorities." *Journal of Political Economy*, 107, 1199–1217.

Che, Yeon-Koo and Seung-Weon Yoo (2001), "Optimal Incentives for Teams." *The American Economic Review*, 525–541.

Coller, Maribeth and Melonie B. Williams (1999), "Eliciting individual discount rates." *Experimental Economics*, 2, 107–127.

Corts, Kenneth S. (2007), "Teams versus individual accountability: Solving multitask problems through job design." *The RAND Journal of Economics*, 467–479.

Cremer, Jacques (1980), "A Partial Theory of the Optimal Organization of a Bureaucracy." *The Bell Journal of Economics*, 11, 683–693.

Currarini, Sergio, Matthew O. Jackson, and Paolo Pin (2009), "An economic model of friendship: Homophily, minorities, and segregation." *Econometrica*, 77, 1003–1045.

Deimen, Inga and Dezsö Szalay (2019), "Delegated Expertise, Authority, and Communication." *American Economic Review*, 109, 1349–1374.

Dessein, Wouter (2002), "Authority and Communication in Organizations." *Review of Economic Studies*, 69, 811–838.

Dessein, Wouter, Luis Garicano, and Robert Gertner (2010), "Organizing for Synergies." *American Economic Journal: Microeconomics*, 2, 77–114.

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole (2000), "Multitask agency problems: Focus and task clustering." *European Economic Review*, 44, 869–877.

Dreyfus, Mark K. and W. K. Viscusi (1995), "Rates of time preference and consumer valuations of automobile safety and fuel efficiency." *The Journal of Law and Economics*, 38, 79–105.

Eguia, Jon X. (2017), "Discrimination and assimilation at school." *Journal of Public Economics*, 156, 48 – 58.

Ekmekci, Mehmet and Stephan Lauermann (2022), "Information aggregation in Poisson elections." *Theoretical Economics*, 17, 1–23.

Espinosa, Lorelle L, Jonathan M Turk, Morgan Taylor, and Hollie M Chessman (2019), "Race and ethnicity in higher education: A status report."

Fiske, Susan T., Amy J. C. Cuddy, Peter Glick, and Jun Xu (2002), "A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition." *Journal of Personality and Social Psychology*, 82, 878–902.

Friebel, Guido, Matthias Heinz, Miriam Krueger, and Nikolay Zubanov (2017), "Team Incentives and Performance: Evidence from a Retail Chain." *American Economic Review*, 107, 2168–2203.

Friebel, Guido and Michael Raith (2010), "Resource Allocation and Organizational Form." *American Economic Journal: Microeconomics*, 2, 1–33.

Fullerton, Richard L and R Preston McAfee (1999), "Auctioning Entry into Tournaments." *journal of political economy*, 107, 573–605.

Galbraith, Jay R. (1971), "Matrix organization designs How to combine functional and project forms." *Business Horizons*, 14, 29–40.

Geanakoplos, John and Paul Milgrom (1991), "A theory of hierarchies based on limited managerial attention." *Journal of the Japanese and International Economies*, 5, 205–225.

Gilbert, G. M. (1951), "Stereotype persistence and change among college students." *The Journal of Abnormal and Social Psychology*, 46, 245–254.

Gromb, Denis and David Martimort (2007), "Collusion and the organization of delegated expertise." *Journal of Economic Theory*, 137, 271–299.

Grossman, Sanford J and Oliver D Hart (1983), "An analysis of the principal-agent problem." *Econometrica*, 51, 7–45.

Groves, Theodore (1973), "Incentives in Teams." *Econometrica*, 41, 617–631.

Harrison, G.W, M.I Lau, and M.B Williams (2002), "Estimating individual discount rates in denmark: A field experiment." *The American Economic Review*, 92, 1606–1617.

Harstad, Ronald M., John H. Kagel, and Dan Levin (1990), "Equilibrium bid functions for auctions with an uncertain number of bidders." *Economics Letters*, 33, 35–40.

Harstad, Ronald M., Aleksandar Saša Pekeč, and Ilia Tsetlin (2008), "Information aggregation in auctions with an unknown number of bidders." *Games and Economic Behavior*, 62, 476–508.

Hausman, Jerry A. (1979), "Individual discount rates and the purchase and utilization of energy-using durables." *The Bell Journal of Economics*, 10, 33–54.

Ho, Colin and Jay W. Jackson (2001), "Attitudes toward asian americans: theory and measurement." *Journal of Applied Social Psychology*, 31, 1553.

Hobday, Mike (2000), "The project-based organisation: an ideal form for managing complex products and systems?" *Research Policy*, 29, 871–893.

Holmström, Bengt (1982), "Moral Hazard in Teams." *The Bell Journal of Economics*, 13, 324–340.

Holmström, Bengt and Paul Milgrom (1991), "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization*, 24–52.

Ishihara, Akifumi (2017), "Relational contracting and endogenous formation of teamwork." *The RAND Journal of Economics*, 48, 335–357.

Ishihara, Akifumi (2020), "On Multitasking and Job Design in Relational Contracts." *The Journal of Industrial Economics*, 68, 693–736.

Jackson, Linda A., Carole N. Hodge, Donna A. Gerard, Julie M. Ingram, Kelly S. Ervin, and Lori A. Sheppard (1996), "Cognition, affect, and behavior in the prediction of group attitudes." *Personality and Social Psychology Bulletin*, 22, 306–316.

Karlins, Marvin, Thomas L. Coffman, and Gary Walters (1969), "On the fading of social stereotypes: Studies in three generations of college students." *Journal of Personality and Social Psychology*, 13, 1–16.

Katz, D. and K. Braly (1933), "Racial stereotypes of one hundred college students." *The Journal of Abnormal and Social Psychology*, 28, 280–290.

Krishna, Vijay and John Morgan (2012), "Voluntary voting: Costs and benefits." *Journal of Economic Theory*, 147, 2083–2123.

Kvaløy, Ola and Trond E. Olsen (2006), "Team Incentives in Relational Employment Contracts." *Journal of Labor Economics*, 24, 139–169.

Lai, Lei and Linda C. Babcock (2013), "Asian americans and workplace discrimination: The interplay between sex of evaluators and the perception of social skills: Asian americans and workplace discrimination." *Journal of Organizational Behavior*, 34, 310–326.

Larson, Erik W. and David H. Gobeli (1989), "Significance of project management structure on development success." *IEEE Transactions on Engineering Management*, 36, 119–125.

Lechler, Thomas G. and Dov Dvir (2010), "An Alternative Taxonomy of Project Management Structures: Linking Project Management Structures and Project Success." *IEEE Transactions on Engineering Management*, 57, 198–210.

Levin, Dan and Emre Ozdenoren (2004), "Auctions with uncertain numbers of bidders." *Journal of Economic Theory*, 229–251.

Levin, Dan and James L Smith (1994), "Equilibrium in Auctions with Entry." *American Economic Review*, 84, 585–599.

Levitt, Steven D. and Christopher M. Snyder (1997), "Is no News Bad News? Information Transmission and the Role of "Early Warning" in the Principal-Agent Model." *The RAND Journal of Economics*, 28, 644–661.

Lim, Wooyoung and Alexander Matros (2009), "Contests with a stochastic number of players." *Games and Economic Behavior*, 67, 584–597.

Marino, Anthony M. and Ján Zábojník (2004), "Internal Competition for Corporate Resources and Incentives in Teams." *The RAND Journal of Economics*, 35, 710–727.

Marschak, Jacob (1955), "Elements for a Theory of Teams." *Management Science*, 127–137.

Marschak, Jacob and Roy Radner (1972), *Economic Theory of Teams*. Yale University Press, New Haven and London.

Matthews, Steven (1987), "Comparing Auctions for Risk Averse Buyers: A Buyer's Point of View." *Econometrica*, 55, 633–646.

Mcafee, R. Preston and John Mcmillan (1987), "Auctions with a Stochastic Number of Bidders." *Journal of Economic Theory*, 43, 1–19.

McAfee, R. Preston and John McMillan (1991), "Optimal Contracts for Teams." *International Economic Review*, 32, 561–577.

Milchtaich, Igal (2004), "Random-player games." *Games and Economic Behavior*, 47, 353–388.

Mookherjee, Dilip (1984), "Optimal Incentive Schemes with Many Agents." *The Review of Economic Studies*, 51, 433–446.

Moore, Michael J. and W. K. Viscusi (1990), "Models for estimating discount rates for long-term health risks using labor market data." *Journal of Risk and Uncertainty*, 3, 381–401.

Mukherjee, Arijit and Luis Vasconcelos (2011), "Optimal job design in the presence of implicit contracts." *The RAND Journal of Economics*, 42, 44–69.

Myerson, Roger B. (1998), "Population uncertainty and Poisson games." *International Journal of Game Theory*, 27, 375–392.

Myerson, Roger B. (2000), "Large poisson games." *Journal of Economic Theory*, 94, 7–45.

Myerson, Roger B (2002), "Comparison of Scoring Rules in Poisson Voting Games." *Journal of Economic Theory*, 103, 219–251.

Myerson, Roger B. and Karl Wärneryd (2006), "Population uncertainty in contest." *Economic Theory*, 469–474.

Münster, Johannes (2006), "Contests with an unknown number of contestants." *Public Choice*, 129, 353–368.

Münster, Johannes (2007), "Contests with investment." *Managerial and Decision Economics*, 28, 849–862.

Münster, Johannes (2009), "Repeated Contests with Asymmetric Information." *Journal of Public Economic Theory*, 11, 89–118.

Pender, John L. (1996), "Discount rates and credit markets: Theory and evidence from rural india." *Journal of Development Economics*, 50, 257–296.

Piketty, Thomas (2000), "Voting as Communicating." *Review of Economic Studies*, 67, 169–191.

Rantakari, Heikki (2008), "Governing Adaptation." *Review of Economic Studies*, 75, 1257–1285.

Rayo, Luis (2007), "Relational Incentives and Moral Hazard in Teams." *Review of Economic Studies*, 937–963.

Ritzberger, Klaus (2009), "Price competition with population uncertainty." *Mathematical Social Sciences*, 58, 145–157.

Satterthwaite, Mark and Artyom Shneyerov (2007), "Dynamic Matching, Two-Sided Incomplete Information, and Participation Costs: Existence and Convergence to Perfect Competition." *Econometrica*, 75, 155–200.

Schmalensee, Richard (1976), “A Model of Promotional Competition in Oligopoly.” *The Review of Economic Studies*, 43, 493–507.

Schöttner, Anja (2008), “Relational Contracts, Multitasking, and Job Design.” *Journal of Law, Economics, and Organization*, 24, 138–162.

Skaperdas, Stergios (1996), “Contest success functions.” *Economic Theory*, 7, 283–290.

Spence, Michael (1973), “Job Market Signaling.” *The Quarterly Journal of Economics*, 87, 355–374.

Szymanski, Stefan (2003), “The Economic Design of Sporting Contests.” *Journal of Economic Literature*, 16, 1137–1187.

Tullock, Gordon (1980), “Efficient rent seeking.” In *Toward a theory of the rent-seeking society*, 97–112, College Station: Texas A & M University.

Verdier, Thierry and Yves Zenou (2017), “The role of social networks in cultural assimilation.” *Journal of Urban Economics*, 97, 15–39.

Wasser, Cédric (2013), “Incomplete information in rent-seeking contests.” *Economic Theory*, 53, 239–268.