2017

Essays In Mechanism Design

Yunan Li
University of Pennsylvania, yunanli@sas.upenn.edu

Essays In Mechanism Design

Abstract
In this thesis, I study mechanism design problems in environments where the information necessary to make decisions is affected by the actions of principal or agents.

The first chapter considers the problem of a principal who must allocate a good among a finite number of agents, each of whom values the good. Each agent has private information about the principal's payoff if he receives the good. There are no monetary transfers. The principal can inspect agents' reports at a cost and punish them, but punishments are limited because verification is imperfect or information arrives only after the good has been allocated for a while. I characterize an optimal mechanism featuring two thresholds. Agents whose values are below the lower threshold and above the upper threshold are pooled, respectively. If the number of agents is small, then the pooling area at the top of value distribution disappears. If the number of agents is large, then the two pooling areas meet and the optimal mechanism can be implemented via a shortlisting procedure. The fact that the optimal mechanism depends on the number of agents implies that small and large organizations should behave differently.

The second chapter considers the problem of a principal who wishes to distribute an indivisible good to a population of budget-constrained agents. Both valuation and budget are an agent's private information. The principal can inspect an agent's budget through a costly verification process and punish an agent who makes a false statement. I characterize the direct surplus-maximizing mechanism. This direct mechanism can be implemented by a two-stage mechanism in which agents only report their budgets. Specifically, all agents report their budgets in the first stage. The principal then provides budget-dependent cash subsidies to agents and assigns the goods randomly (with uniform probability) at budget-dependent prices. In the second stage, a resale market opens, but is regulated with budget-dependent sales taxes. Agents who report low budgets receive more subsidies in their initial purchases (the first stage), face higher taxes in the resale market (the second stage) and are inspected randomly. This implementation exhibits some of the features of some welfare programs, such as Singapore's housing and development board.

The third chapter studies the design of ex-ante efficient mechanisms in situations where a single item is for sale, and agents have positively interdependent values and can covertly acquire information at a cost before participating in a mechanism. I find that when interdependency is low or the number of agents is large, the ex-post efficient mechanism is also ex-ante efficient. In cases of high interdependency or a small number of agents, ex-ante efficient mechanisms discourage agents from acquiring excessive information by introducing randomization to the ex-post efficient allocation rule in areas where the information's precision increases most rapidly.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Economics

First Advisor
Rakesh V. Vohra

Subject Categories
Economic Theory

ESSAYS IN MECHANISM DESIGN

Yunan Li

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2017

Supervisor of Dissertation

Rakesh V. Vohra, George A. Weiss and Lydia Bravo Weiss University Professor

Graduate Group Chairperson

Jesus Fernandez-Villaverde, Professor of Economics

Dissertation Committee

Steven A. Matthews, Professor of Economics

Mallesh M. Pai, Assistant Professor of Economics

ESSAYS IN MECHANISM DESIGN

This thesis is dedicated to my parents for their endless love, support and encouragement.

ACKNOWLEDGMENT

I would like to thank all the people who contributed in some way to the work described in this dissertation.

First and foremost, I would like to express my sincere gratitude to my adviser Professor Rakesh Vohra for his continuous encouragement, support and guidance, for his patience and immense knowledge. During my PhD study, he helped me to grow as a researcher by supporting my attendance at various conferences, engaging me in new ideas, and demanding a high quality of work in all my endeavors.

I am also deeply grateful to the rest of my dissertation committee, Professor Steven Matthews and Professor Mallesh Pai, for their encouragement, guidance and insightful comments. Without their help, this dissertation could not be completed.

I thank Professors Aislinn Bohren, David Dillenberger, Hanming Fang, Sangmok Lee, George Mailath, Andrew Postlewaite and Yuichi Yamamoto for their encouragement and support at various stages of my dissertation. They have been kind enough to extend their help whenever I approached them. I also thank Professors Harold Cole, Cecilia Fieler, Iourii Manovskii, Wojciech Olszewski, Guillermo Ordonez, Michael Richter, Andrew Shephard and Holger Sieg for their comments and suggestions to my dissertation and their assistance during my job market process.

Special thanks to my fellow students Daniel Hauser and Nick Janetos for the stimulating discussions and for all the fun we have had in our PhD studies. I would like to thank my fellow students and friends Qing Gong and Yu Liu for listening, offering me advice, and supporting me. They have made my PhD study more enjoyable and memorable. I would also like to thank my other fellow students Rodrigo Azuero, Yiran Chen, Mustafa Dogan, Selman Erol, Xiang Fang, Juan Hernandez, Ju Hu, Ami Ko, Yang Liu, Joonbae Lee, Rossa O'Keeffe-O'Donovan, Nishant Ravi, Daniel Wills, Xingtan Zhang and Junyuan Zou, and my friends Sungki Hong, Shaowei Ke, Mariann Ollar, Chang Sun, Yu Zhang and Chen Zhao for their time discussing about my research.

The Xinmei Zhang Fellowship, the Sydney Weintraub Memorial Fellowship and the SAS Dissertation Completion Fellowship at the University of Pennsylvania are gratefully acknowledged.

Last but not the least, I would like to thank my family: my parents and my husband for their endless love and support.

ABSTRACT

ESSAYS IN MECHANISM DESIGN

Yunan Li

Rakesh V. Vohra

In this thesis, I study mechanism design problems in environments where the information necessary to make decisions is affected by the actions of principal or agents.

The first chapter considers the problem of a principal who must allocate a good among a finite number of agents, each of whom values the good. Each agent has private information about the principal's payoff if he receives the good. There are no monetary transfers. The principal can inspect agents' reports at a cost and punish them, but punishments are limited because verification is imperfect or information arrives only after the good has been allocated for a while. I characterize an optimal mechanism featuring two thresholds. Agents whose values are below the lower threshold and above the upper threshold are pooled, respectively. If the number of agents is small, then the pooling area at the top of value distribution disappears. If the number of agents is large, then the two pooling areas meet and the optimal mechanism can be implemented via a shortlisting procedure. The fact that the optimal mechanism depends on the number of agents implies that small and large organizations should behave differently.

The second chapter considers the problem of a principal who wishes to distribute an indivisible good to a population of budget-constrained agents. Both valuation and budget are an agent's private information. The principal can inspect an agent's budget through a costly verification process and punish an agent who makes a false statement. I characterize the direct surplus-maximizing mechanism. This direct mechanism can be implemented by a two-stage mechanism in which agents only report their budgets. Specifically, all agents report their budgets in the first stage. The principal then provides budget-dependent cash subsidies to agents and assigns the goods randomly (with uniform probability) at budget-dependent prices. In the second stage, a resale market opens, but is regulated

with budget-dependent sales taxes. Agents who report low budgets receive more subsidies in their initial purchases (the first stage), face higher taxes in the resale market (the second stage) and are inspected randomly. This implementation exhibits some of the features of some welfare programs, such as Singapore's housing and development board.

The third chapter studies the design of ex ante efficient mechanisms in situations where a single item is for sale, and agents have positively interdependent values and can covertly acquire information at a cost before participating in a mechanism. I find that when interdependency is low or the number of agents is large, the ex post efficient mechanism is also ex ante efficient. In cases of high interdependency or a small number of agents, ex ante efficient mechanisms discourage agents from acquiring excessive information by introducing randomization to the ex post efficient allocation rule in areas where the information's precision increases most rapidly.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1 : INTRODUCTION

Mechanism design is “an analytical framework for thinking clearly and carefully about what exactly a given institution can achieve when the information necessary to make decisions is dispersed and privately held.”[1]. In many important applications, this information is *endogenously* influenced by actions of principal or agents. A useful example to illustrate this point is a start-up company that wants to bring a new product to the market and needs to attract funding. The start-up company can learn the worth of the product by developing a prototype and gathering information about target customers. A venture capital firm can also investigate the technical and the economic feasibility of the product, which is initially privately known by the start-up company. These issues do not fit into the standard mechanism design literature which largely focuses on environments in which the private information of agents is given exogenously and is non-verifiable. In this thesis, I study mechanism design problems in a richer information environment. Specifically, I explore the following two environments: one where principal can verify agents’ information; and one where agents can covertly acquire information.

1.1. Mechanism design with costly verification

The standard mechanism design literature on allocation problems has largely focused on the use of monetary transfers to induce truthful revelation. and has ignored the possibility of *principal verifying agents’ information*. In some cases, the principal can obtain information about agents at a cost. For example, the head of personnel for an organization can verify a job applicant’s claim or monitor his performance once he is hired. A venture capital firm can investigate the competing start-ups or audit the progress of a start-up once it is funded. Hence, I think it is important to consider this option.

Two recent papers, Ben-Porath et al. (2014) and Mylovanov and Zapechelnyuk (2014), have studied this problem by taking the opposite position from the standard model and by ruling out transfers but allowing costly inspection. They examined two extreme cases. In Ben-Porath et al. (2014), verification is costly but punishment is unlimited in the sense that an agent can be rejected and does

[1] Vohra (2011)

not receive the good. In Mylovanov and Zapechelnyuk (2014), verification is free but punishment can be limited because verification is imperfect or information arrives only after the agent has been hired for a while.

In Chapter 2, I consider a situation with both costly verification and limited punishment. I characterize an optimal mechanism which has two thresholds. Agents whose values are below the lower threshold and above the upper threshold respectively are pooled. If the number of agents is small, then the pooling area at the top of the value distribution disappears, as seen in Ben-Porath et al. (2014). In the case of intermediate and large numbers of agents, the optimal allocation rule also involves pooling at the top, as seen in Mylovanov and Zapechelnyuk (2014). If the number of agents is sufficiently large, then the two pooling areas meet and the optimal mechanism can be implemented via a shortlisting procedure.

In Chapter 3, I study the problem of a principal who wishes to distribute an indivisible good to a population of budget-constrained agents, such as public housing and social health care programs. Both valuation and budget are the private information of an agent, but the principal can inspect an agent's budget through a costly verification process. Indeed, in many public programs, applicants are subject to a set of eligibility conditions such as monthly income and family nucleus. Based on the literature which studies allocation problems among financially-constrained agents (such as Che et al. (2013a)), I consider mechanisms with monetary transfers and add the option of costly verification on budgets.

I characterize the (direct) surplus-maximizing mechanism and also provide an implementation via a two-stage mechanism. For tractability, I assume there are only two budget types: low and high. In the first stage, agents report their budgets and the principal allocates the goods randomly. Agents who report low budgets receive more cash and in-kind subsidies, and their reports are verified randomly. In the second stage, a resale market opens but is regulated. Agents who report low budgets face higher resale taxes, and their reports are verified randomly if they do not sell. This resembles the affordable housing program in Singapore, which imposes more restrictions on the resale of agents whose initial purchases are subsidized by the government.

A technical challenge of this chapter is that *we cannot anticipate a priori the set of binding incentive compatibility constraints*. This problem is also ubiquitous in multidimensional screening problems with only monetary transfers and mechanism design problems with costly verification. To overcome this difficulty, I develop a novel method which can potentially be used to solve other problems.

1.2. Mechanism design with information acquisition

In most of the literature on mechanism design, agents are assumed to have a given amount of information and their incentives to acquire more information are not modeled. However, in many practical settings, this assumption does not apply. For example, in auctions for offshore oil and gas leases in the U.S., companies collect information about the tracts offered for sale using seismic surveys before participating in the auctions. Another example is the sale of financial or business assets, in which buyers perform due diligence to investigate the quality and compatibility of the assets before submitting offers. Moreover, it is costly to acquire information. In the example of U.S. auctions for offshore oil and gas leases (see Haile et al. (2010)), it cost \$100, 000 to conduct a 50 square mile 3-D seismic survey in 2000. Similarly, in the sale of a business asset, the legal and accounting costs of performing due diligence often amount to millions of dollars (see Quint and Hendricks (2013)).

Earlier papers have studied agents' incentives to acquire information for *fixed* mechanisms. In a recent paper, Bergemann and Välimäki (2002) considered the socially optimal information acquisition in the context of general mechanism design. They focus on mechanisms that implement the ex-post efficient allocations given acquired private information, and find that ex ante efficient information acquisition can be achieved if agents have *independent private values*. However, if agents' values are *interdependent*, then ex-post efficient mechanisms will result in socially sub-optimal information acquisition. In other words, there is a conflict between the provision of ex ante efficient incentives to acquire information and the ex post efficient use of information.

What are ex-ante efficient mechanisms that balances the two trade-offs? In Chapter 4, I provide an answer to this question. Specifically, I study the design of ex-ante efficient mechanisms in the sale of a single object when agents' values are positively interdependent. I focus on symmetric mechanisms that treat all agents identically, and on symmetric equilibria in which agents acquire the same amount of information before participating in a mechanism. I find that, when interdependency is low or the number of agents is large, then the ex-post efficient mechanism is also ex-ante efficient. In cases of high interdependency or a small number of agents, then ex-ante efficient mechanisms discourage agents from acquiring excessive information by introducing randomization to the ex-post efficient allocation rule in areas where the information's precision increases most rapidly.

CHAPTER 2 : MECHANISM DESIGN WITH COSTLY VERIFICATION AND LIMITED PUNISHMENTS

2.1. Introduction

In many large organizations scarce resources must be allocated internally without the benefit of prices. Examples include the head of personnel for an organization choosing one of several applicants for a job, venture capital firms choosing which startup to fund and funding agencies allocating a grant among researchers. In these settings the principal must rely on verification of agents' claims, which can be costly. For example, the head of personnel can confirm a job applicant's past work experience, or monitor his performance once he is hired. A venture capital firm can investigate the competing startups, or audit the progress of a startup once it is funded. Furthermore, the principal can punish an agent if his claim is found to be false. For example, the head of personnel can reject an applicant, fire an employee or deny a promotion. Venture capitals and funding agencies can cut off funding.

Prior work has examined two extreme cases. In Ben-Porath et al. (2014), verification is costly, but punishment is unlimited in the sense that an agent can be rejected and does not receive the resource. In Mylovanov and Zapechelnyuk (2014), verification is free, but punishment is limited. In this paper, I consider a situation with both costly verification *and* limited punishment. I interpret verification as acquiring information (e.g., requesting documentation, interviewing an agent, or monitoring an agent at work), which could be costly. Moreover, punishment can be limited because verification is imperfect or information arrives only after an agent has been hired for a while.

I think it is important to consider this general setting with both costly verification and limited punishment for two reasons. First, as it will become clear soon, this general setting helps us to identify the role the number of agents plays in shaping optimal mechanisms. In the concluding section, I give a more detailed comparison of the results in this paper with those in previous papers regarding the role played by the number of agents. Second, in practice, it is possible that the principal can obtain more precise information by incurring a higher information acquisition cost, which, in turn, leads to a higher expected punishment. Although, throughout this paper, I take verification cost and punishment level as exogenous, the results in the paper readily extend if the principal can jointly optimize over verification cost and punishment level. The results in this paper help us to understand the interactions between verification cost and punishment level.

Specifically, in the model, there is one principal who has to allocate one indivisible object among a finite number of agents. She would like to give the object to the agent who has the highest value to her. But doing this encourages all agents to exaggerate their values. The principal has at her disposal two devices to discourage agents from exaggeration: first, the principal can ration at the bottom or top of the distribution of values, but this reduces allocative efficiency; second, the principal can verify an agent's claim and punish him if his claim is found to be false, but verification is a costly procedure. The goal of this paper is to find the optimal way to provide incentives via these two devices.

For most parts of this paper, I consider a symmetric environment and characterize an optimal symmetric mechanism in this setting. If the number of agents is sufficiently small, then an *one-threshold mechanism* as in Ben-Porath et al. (2014) is optimal. The allocation rule in this mechanism is efficient at the top of the value distribution, and involves pooling only at the bottom. For intermediate and large numbers of agents, the allocation rule involves pooling at both the top and the bottom as in Mylovanov and Zapechelnyuk (2014). Specifically, the following *two-threshold mechanism* is optimal. If there are agents whose values are above the upper threshold, then one of them is chosen at random. If all agents' values are below the upper threshold, but some are above the lower threshold, then the one with the highest value is chosen. If all agents' values are below the lower threshold, then one of them is chosen at random. Note that an one-threshold mechanism can be viewed as a two-threshold mechanism whose upper threshold is equal to the upper-bound of the value support. For a sufficiently large number of agents, the two thresholds coincide, and the two-threshold mechanism can be implemented using a *shortlisting procedure*. In this shortlisting procedure, agents whose values are above a threshold are shortlisted for sure, and agents whose values are below the threshold are shortlisted with some probability. The principal then chooses one agent from the shortlist at random. The fact that the optimal mechanism depends on the number of agents implies that small and large organizations should behave differently.

To understand the intuition behind these results, consider an agent with the lowest possible value to the principal. Intuitively, as the number of agents increases, this agent gets worse off and has stronger incentives to exaggerate his value in an one-threshold mechanism because now it is more likely that there exists another agent whose value is above the threshold. When punishments are limited, the principal can make exaggeration less attractive only by introducing distortions to the allocation rule at the top of the value distribution.

This distinction between small and intermediate numbers of agents is important as it allows us to establish a connection between Ben-Porath et al. (2014) and Mylovanov and Zapechelnyuk (2014). Note that this distinction is *absent* if either verification is free or punishment is unlimited. In Ben-Porath et al. (2014), an optimal mechanism never involves pooling at the top of value distribution because punishment is unlimited. If punishment is limited, then pooling at the top is part of the optimal mechanism for a sufficiently large number of agents. In Mylovanov and Zapechelnyuk (2014), an optimal mechanism always involves pooling at the top because verification is free. If verification is costly, then pooling at the top disappears for a sufficiently small number of agents.

As an effort to understand the trade-off between verification (or information) cost and punishment level (or information quality), I provide some comparative statics results with respect to verification cost and punishment level in Section 2.4. An increase in verification cost has two opposite effects on the size of the pooling areas. First, when verification becomes more costly, the optimal threshold mechanism sees more pooling at the bottom to save verification cost. Second, the enlarging pooling area at the bottom benefits agents with very low values and reduces their incentives to exaggerating their values, which leads to less pooling at the top or no pooling at the top at all. In the paper, I show that the second effect dominates, and, as a result, one-threshold mechanisms or two-threshold mechanisms remain optimal for a larger number of agents as verification becomes more costly. The impact of a change in punishment level is ambiguous and more interesting. On the one hand, a reduction in punishment effectively makes verification more costly as the principal must inspect agents more frequently to maintain incentive compatibility. Then the above analysis implies that one-threshold mechanisms or two-threshold mechanisms remain optimal for a larger number of agents as punishment becomes less severe. On the other hand, a reduction in punishment level makes it harder to preventing agents from exaggeration through punishments, which leads to larger pooling areas both at the bottom and at the top to restore incentive compatibility. This in turn implies that one-threshold mechanisms or two-threshold mechanisms remain optimal for a smaller number of agents as punishment becomes less severe.In general, the impact of a change in punishment level is not monotonic.

In Section 2.5.1, I study a general model with asymmetric agents. In this setting, threshold mechanisms are still optimal. The analysis, however, is much more complex. Though there is still a unique lower threshold for all agents, different agents may face different upper thresholds. Using this result, I revisit the symmetric environment and characterize the set of all optimal threshold mechanisms. I find that limiting the principal's ability to punish agents also limits her ability to treat agents differently. In particular, when a one-threshold mechanism is optimal, the set of all optimal threshold mechanisms shrinks as punishment becomes more limited. Eventually, the unique optimal threshold mechanism is symmetric. If punishment is sufficiently limited so that a two-threshold mechanism or a shortlisting procedure is optimal, then the principal can once again treat agents differently but to the extent that they share the same set of thresholds. The comparison is less clear in this case because the sets of optimal mechanisms are disjoint for different levels of punishments.

Technically, I follow Vohra (2012) and use tools from linear programming, which allows me to analyze Ben-Porath et al. (2014) and Mylovanov and Zapechelnyuk (2014) in a unified framework. It also allows me to get some results on optimal mechanisms in the asymmetric environment with limited punishments, which are unavailable in Mylovanov and Zapechelnyuk (2014).

The rest of the paper is organized as follows. Section 2.1.1 discusses other related work. Section 2.2 presents the model. Section 2.3 characterizes an optimal symmetric mechanism when agents are ex ante identical. Section 2.4 discusses the properties of this optimal mechanism. Section 2.5.1 studies a general asymmetric environment. Section 2.5.2 considers other variations of verification and punishment technologies. Section 2.6 concludes.

2.1.1. Other related literature

This paper is related to the costly state verification literature. The first contribution in the series is Townsend (1979) who studies a model of a principal and a single agent. In Townsend (1979) verification is deterministic. Border and Sobel (1987) and Mookherjee and Png (1989) generalize it by allowing random inspection. Gale and Hellwig (1985) consider the effects of costly verification in the context of credit markets. These models differ from what I consider here in that in their models there is only one agent and monetary transfers are allowed. Recently, Patel and Urgan (2017) also study the problem of a principal who must allocate a good among multiple agents when transfers are not allowed. As in Ben-Porath et al. (2014), in Patel and Urgan (2017), verification is costly and punishments are unlimited. But, in addition to costly verification, the principal can deploy another instrument: money burning. They show that both instruments are present in the optimal mechanism. Furthermore, the optimal mechanism admits monotonicity in the allocation probability with regards to an agent's value, and takes a threshold form where all the values below a certain threshold are not subject to verification or money burning.

Technically, this paper is related to the literature on reduced form implementation — see, e.g., Maskin and Riley (1984b), Matthews (1984b), Border (1991) and Mierendorff (2011). The most related paper is Pai and Vohra (2014b), who also use reduced form implementation and linear programming to derive optimal mechanisms for financially constrained agents.

2.2. Model

The set of agents is $\mathcal{I} := \{1, \dots, n\}$. There is a single indivisible object to be allocated among them. The value to the principal of assigning the object to agent i is v_i, where v_i is agent i's private information. I assume $\{v_i\}$ are independently distributed, whose density f_i is strictly positive on $V_i := [\underline{v}_i, \overline{v}_i] \subset \mathbb{R}_+$. The assumption that an agent's value to the principal is always non-negative simplifies some statements, but the results in this paper can easily extend to include negative values. I use F_i to denote the corresponding cumulative distribution function. Let $\mathcal{V} := \prod_i V_i$. Agent i gets a payoff of $b_i(v_i)$ if he receives the object, and 0 otherwise. The principal can verify agent i's report at a cost $k_i \geq 0$ if agent i receives the object, and at a cost $k_i^\beta \geq 0$ if agent i does not receive the object. I allow for verification costs to depend on whether an agent gets the object. This is natural in some environments. For example, if the object is a job slot and the private information is about an agent's ability, then it is easier to inspect an agent who is hired.[1] Verification perfectly reveals an agent's type. The cost to an agent to have his report verified is zero. If agent i is inspected, then the principal can impose a penalty $c_i(v_i) \geq 0$ if agent i receives the object, and a penalty $c_i^\beta(v_i) \geq 0$ if agent i does not receive the object. In Ben-Porath et al. (2014), the principal can inspect an agent at the same cost regardless of whether he receives the object or not, i.e., $k_i = k_i^\beta$. However, the principal can only penalize an agent if he receives the object, i.e., $c_i^\beta = 0$. In Mylovanov and Zapechelnyuk (2014), the principal can only inspect and penalize an agent if he receives the object, i.e., $k_i^\beta = \infty$ and $c_i^\beta = 0$. For the rest of the paper, I follow Ben-Porath et al. (2014) and Mylovanov and Zapechelnyuk (2014) and assume that $c_i^\beta = 0$. The interpretation is that the principal can only penalize an agent by taking the object back possibly after a number of periods (e.g., rejecting a job applicant or firing him after a certain length of employment). In Section 2.5.2, I discuss to what extent this assumption can be relaxed.

We say that punishment is *limited* if $c_i(v_i) < b_i(v_i)$ for all v_i. That is, the principal cannot reduce an agent's payoff to his outside option by punishing him. If we interpret verification as acquiring information, then punishment can be limited because information is imperfect. Throughout the paper, I take verification cost and punishment level as exogenous. In practice, it is possible that the principal can get more precise information by incurring a higher information acquisition cost, which, in turn, leads to a severer expected punishment. In other words, by choosing a higher k_i, the principal can obtain a higher c_i. The results in this paper readily extend if the principal can jointly optimize over verification cost and punishment level.

I invoke the Revelation Principle and focus on direct mechanisms in which truth-telling is a Bayes-Nash equilibrium. Clearly, if an agent is inspected, it is optimal to penalize him if and only if he is

[1] I will use the words "verify" and "inspect" exchangeably in this paper.

found to have lied. Using this result, a direct mechanism can be written as a triplet $(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{q}^\beta)$, where $\boldsymbol{p} := (p_1, \dots, p_n) : \mathcal{V} \to [0,1]^n$, $\boldsymbol{q} := (q_1, \dots, q_n) : \mathcal{V} \to [0,1]^n$ and $\boldsymbol{q}^\beta := (q_1^\beta, \dots, q_n^\beta) : \mathcal{V} \to [0,1]^n$. For each i and each profile of reported values $\boldsymbol{v} \in \mathcal{V}$, $p_i(\boldsymbol{v})$ specifies the probability with which i is assigned the object, $q_i(\boldsymbol{v})$ specifies the probability of inspecting i conditional on the object being assigned to agent i, and $q_i^\beta(\boldsymbol{v})$ specifies the probability of inspecting i conditional on the object not being assigned to agent i. The utility of an agent whose true type is v_i and who reports v_i' is $p_i(v_i, v_{-i})b_i(v_i)$ if $v_i' = v_i$ and

$$p_i(v_i', v_{-i})\left(b_i(v_i) - q_i(v_i', v_{-i})c_i(v_i)\right) - (1 - p_i(v_i', v_{-i}))q_i^\beta(v_i', v_{-i})c_i^\beta(v_i)$$

otherwise. A mechanism is *feasible* if $\sum_i p_i(\boldsymbol{v}) \le 1$ for all $\boldsymbol{v} \in \mathcal{V}$. A mechanism satisfies the *incentive compatibility* (IC) constraints if for each agent i,

$$\begin{aligned}&\mathbb{E}_{v_{-i}}\left[p_i(v_i, v_{-i})b_i(v_i)\right]\\ &\ge \mathbb{E}_{v_{-i}}\left[p_i(v_i', v_{-i})\left(b_i(v_i) - q_i(v_i', v_{-i})c_i(v_i)\right) - (1 - p_i(v_i', v_{-i}))q_i^\beta(v_i', v_{-i})c_i^\beta(v_i)\right], \forall v_i, v_i'.\end{aligned}$$

The principal's objective is to maximize her expected gain from allocating the object minus the expected verification cost,

$$\mathbb{E}_{\boldsymbol{v}}\left[\sum_{i=1}^n p_i(\boldsymbol{v})\left(v_i - q_i(\boldsymbol{v})k_i\right) - (1 - p_i(\boldsymbol{v}))q_i^\beta(\boldsymbol{v})k_i^\beta\right], \tag{2.1}$$

subject to the feasibility and IC constraints.

Because $c_i^\beta = 0$, clearly it is optimal to set $q_i^\beta = 0$. In what follows, I abuse notation a bit and denote a mechanism by a pair $(\boldsymbol{p}, \boldsymbol{q})$. The principal's objective function now becomes

$$\mathbb{E}_{\boldsymbol{v}}\left[\sum_{i=1}^n p_i(\boldsymbol{v})\left(v_i - q_i(\boldsymbol{v})k_i\right)\right]. \tag{2.2}$$

The IC constraints become: for each agent i,

$$\mathbb{E}_{v_{-i}}\left[p_i(v_i, v_{-i})b_i(v_i)\right] \geq \mathbb{E}_{v_{-i}}\left[p_i(v_i', v_{-i})\left(b_i(v_i) - q_i(v_i', v_{-i})c_i(v_i)\right)\right], \forall v_i, v_i'. \tag{2.3}$$

Note that if $k_i = 0$, then the above principal's problem reduces to the one considered in Mylovanov and Zapechelnyuk (2014); and if $c_i(v_i) = b_i(v_i)$ for all v_i, then it reduces to the one considered in Ben-Porath et al. (2014).

For each agent i and each $v_i \in V_i$, let $P_i(v_i) := \mathbb{E}_{v_{-i}}\left[p_i(v_i, v_{-i})\right]$ denote the interim probability with which agent i is assigned the object, and let $Q_i(v_i) := \mathbb{E}_{v_{-i}}\left[p_i(v_i, v_{-i})q_i(v_i, v_{-i})\right]/P_i(v_i)$ if $P_i(v_i) \neq 0$ and $Q_i(v_i) := 0$ otherwise. Note that $P_i(v_i)Q_i(v_i)$ is the interim probability with which agent i is inspected. Let $\boldsymbol{P} := (P_1, \dots, P_n)$ and $\boldsymbol{Q} := (Q_1, \dots, Q_n)$. Then the principal's problem can be written in the reduced form:

$$\max_{\boldsymbol{P},\boldsymbol{Q}} \sum_{i=1}^{n} \mathbb{E}_{v_i}\left[P_i(v_i)\left(v_i - Q_i(v_i)k_i\right)\right],$$

subject to

$$P_i(v_i)b_i(v_i) \geq P_i(v_i')\left(b_i(v_i) - Q_i(v_i')c_i(v_i)\right), \forall v_i, v_i', \tag{IC}$$

$$0 \leq Q_i(v_i) \leq 1, \forall v_i, \tag{F1}$$

$$\sum_i \int_{S_i} P_i(v_i)\mathrm{d}F_i(v_i) \leq 1 - \prod_i \left(1 - \int_{S_i} dF_i(v_i)\right), \forall S_i \subset V_i. \tag{F2}$$

In particular, an allocate rule $\boldsymbol{p}$ is feasible if and only if the corresponding reduced form allocation rule $\boldsymbol{P}$ satisfies (F2) by Theorem 2 in Mierendorff (2011), which generalizes the well-known Maskin-Riley-Matthews-Border conditions to asymmetric environments.

I begin solving the principal's problem by solving for the optimal $\boldsymbol{Q}$ for a given $\boldsymbol{P}$. In both Mylovanov and Zapechelnyuk (2014) and Ben-Porath et al. (2014), this exercise is easy. If $k_i = 0$, then $Q(v_i) = 1$ for all $v_i \in V_i$. If $c_i(v_i) = b_i(v_i)$ for all v_i, then (IC) become $P_i(v_i)b_i(v_i) \geq P_i(v_i')b_i(v_i)\left(1 - Q_i(v_i')\right)$

for all v_i and v_i'. Then (IC) hold if and only if

$$\inf_{v_i} P_i(v_i) \geq P_i(v_i')\left(1 - Q_i(v_i')\right), \forall v_i'.$$

Because the principal's objective function is strictly decreasing in Q_i, it is optimal to set $Q_i(v_i) = 1 - \varphi_i / P_i(v_i)$ for all $v_i \in V_i$, where $\varphi_i := \inf_{v_i} P_i(v_i)$. In general, for $k_i > 0$ and $c_i(v_i) \neq b_i(v_i)$, solving for the optimal Q_i is hard.

For tractability, I assume that $c_i(v_i) = c_i b_i(v_i)$ for some $0 < c_i \leq 1$. This assumption is natural in some applications. In the job slot example, this assumption is satisfied if an agent gets a private benefit for each period he is employed and the penalty is being fired after a pre-specified number of periods. In the example of venture capital firms or funding agencies, this assumption is satisfied if agents receive funds periodically and the penalty is cutting off funding after certain periods. Furthermore, this assumption allows us to obtain a clear analysis on the interaction between the verification cost (k) and the level of punishment (c). Lastly, this assumption can be relaxed, and the results in this paper can easily extend if $c_i(v_i)/b_i(v_i)$ is minimized at $\underline{v}_i$.[2]

Under the assumption that the penalty, $c_i(v_i)$, is proportional to the private benefit, $b_i(v_i)$, (IC) become $P_i(v_i) \geq P_i(v_i')\left(1 - c_i Q_i(v_i')\right)$ for all v_i and v_i'. The (IC) constraint holds if and only if

$$\varphi_i \geq P_i(v_i')\left(1 - c_i Q_i(v_i')\right), \forall v_i'. \tag{2.4}$$

Because $Q_i(v_i') \leq 1$, then (2.4) holds only if

$$(1 - c_i)P_i(v_i') \leq \varphi_i, \forall v_i'. \tag{2.5}$$

Remark 1 *Note that if $c_i = 1$ as in Ben-Porath et al. (2014), then* (2.5) *is satisfied automatically.*

---

[2](IC) can be rewritten as: for each agent i,

$$Q_i(v_i') \geq \frac{b_i(v_i)}{c_i(v_i)}\left(1 - \frac{P_i(v_i)}{P_i(v_i')}\right), \forall v_i, v_i'.$$

Suppose that $c_i(v_i)/b_i(v_i)$ is minimized at $\underline{v}_i$ and $P_i(v_i)$ is non-decreasing. Then for any given v_i', the left-hand side of the above inequality is maximized at $\underline{v}_i$. Redefine $c_i := c_i(\underline{v}_i)/b_i(\underline{v}_i)$. Then (IC) hold if and only if (2.4) holds.

This explains why there is no pooling at the top of value distribution in Ben-Porath et al. (2014). In contrast, if $0 < c_i < 1$, then (2.5) imposes an upperbound on P_i, and, as I will show later, there can be pooling at the top for a sufficiently number of agents.

For the rest of the paper, I assume that $0 < c_i < 1$. Suppose (2.5) holds, then it is optimal to set $Q_i(v_i) = \left(1 - \varphi_i / P_i(v_i)\right) / c_i$ for all $v_i \in V_i$. Substituting this into the principal's objective function gives:

$$\sum_{i=1}^{n} \mathbb{E}_{v_i} \left[P_i(v_i) \left(v_i - \frac{k_i}{c_i} \right) \right] + \frac{\varphi_i k_i}{c_i}. \tag{2.6}$$

For the main part of the paper, I assume v_i's are identically distributed, whose density f is strictly positive on $V = [\underline{v}, \overline{v}] \subset \mathbb{R}_+$. I use F to denote the corresponding cumulative distribution function. In addition, I assume $c_i = c$ and $k_i = k$ for all i. In this symmetric setting, there exists an optimal mechanism that is symmetric. Hence, I focus on symmetric mechanisms in Sections 2.3 and 2.4. In what follows, I suppress the subscript i whenever the meaning is clear. The main results of the paper can be extended to environments in which the valuations (v_i) of different agents can follow different distributions (F_i), and both the punishments (c_i) and the verification costs (k_i) can be different for different agents. I discuss this general asymmetric setting in Section 2.5.1.

2.3. Optimal mechanisms

In this section, I show that a simple threshold mechanism is optimal. As an overview of the proof idea, I solve the principal's problem in two steps. In the first step, I characterize an optimal mechanism for any given lowest probability with which an agent receives the object (φ). In the second step, I solve for the optimal φ.

2.3.1. Optimal mechanisms for fixed φ

Fix $\varphi = \inf_v P(v) \leq 1/n$.[3] I first solve the following problem ($OPT - \varphi$):

$$\max_P \mathbb{E}_v \left[P(v) \left(v - \frac{k}{c} \right) \right] + \frac{\varphi k}{c},$$

[3]Note that the problem ($OPT - \varphi$) is feasible only if $\varphi \leq 1/n$.

subject to

$$\varphi \leq P(v) \leq \frac{\varphi}{1-c}, \forall v, \tag{IC$'$}$$

$$n\int_S P(v)\mathrm{d}F(v) \leq 1 - \left(1 - \int_S \mathrm{d}F(v)\right)^n, \forall S \subset V. \tag{F2}$$

Recall that there exists Q such that (F1) and (IC) hold if and only if P satisfies (IC$'$). To solve $(OPT - \varphi)$, I approximate the continuum type space with a finite partition, characterize an optimal mechanism in the finite model and take limits. Later on, I show that the limiting mechanism is optimal in the original model.

Finite case

Fix an integer $m \geq 2$. For $t = 1, \dots m$, let

$$v^t := \underline{v} + \frac{(2t-1)(\overline{v} - \underline{v})}{2m},$$
$$f^t := F\left(\underline{v} + \frac{t(\overline{v} - \underline{v})}{m}\right) - F\left(\underline{v} + \frac{(t-1)(\overline{v} - \underline{v})}{m}\right).$$

Consider the finite model in which v_i can take only m possible different values, i.e., $v_i \in \{v^1, \dots, v^m\}$ and the probability mass function satisfies $f(v^t) = f^t$ for $t = 1, \dots, m$. I abuse notation a bit and let $P := (P^1, \dots, P^m)$, where P^t is the interim probability with which a type v^t agent is assigned the good. Then the corresponding problem of $(OPT - \varphi)$ in the finite model, denoted by $(OPTm - \varphi)$, is given by:

$$\max_P \sum_{t=1}^m f^t P^t \left(v^t - \frac{k}{c}\right) + \frac{\varphi k}{c},$$

subject to

$$\varphi \leq P^t \leq \frac{\varphi}{1-c}, \forall t, \tag{IC$'m$}$$

$$n\sum_{t\in S} f^t P^t \leq 1 - \left(\sum_{t\notin S} f^t\right)^n, \forall S \subset \{1, \dots, m\}. \tag{F2m}$$

To solve $(OPTm - \varphi)$, I first rewrite it as a polymatroid optimization problem. Define $G(S) := 1 - \left(\sum_{t \notin S} f^t\right)^n$ and $H(S) := G(S) - n\varphi \sum_{t \in S} f^t$ for all $S \subset \{1, \dots, m\}$. Define $z^t := f^t(P^t - \varphi)$ for all $t = 1, \dots, m$ and $z := (z^1, \dots, z^m)$. Clearly, $P^t \geq \varphi$ if and only if $z^t \geq 0$ for all $t = 1, \dots, m$. Using these notations, (F2m) can be rewritten as

$$n \sum_{t \in S} z^t \leq H(S), \forall S \subset \{1, \dots, m\}.$$

It is easy to verify that $H(\emptyset) = 0$ and H is submodular. However, H is not monotonic. Define $\overline{H}(S) := \min_{S' \supset S} H(S)$. Then $\overline{H}(\emptyset) = 0$, and $\overline{H}$ is non-decreasing and submodular. Furthermore, by Lemma 16 in Appendix A.1,

$$\left\{ z \,\middle|\, z \geq 0, n \sum_{t \in S} z^t \leq H(S), \forall S \right\} = \left\{ z \,\middle|\, z \geq 0, n \sum_{t \in S} z^t \leq \overline{H}(S), \forall S \right\},$$

Thus, $(OPTm - \varphi)$ can be rewritten as $(OPTm1 - \varphi)$

$$\max_z \sum_{t=1}^m z^t \left(v^t - \frac{k}{c}\right) + \varphi \sum_{t=1}^m f^t v^t,$$

subject to

$$0 \leq z^t \leq \frac{c\varphi f^t}{1-c}, \forall t, \tag{IC$'m1$}$$

$$n \sum_{t \in S} z^t \leq \overline{H}(S), \forall S \subset \{1, \dots, m\}. \tag{F2$m1$}$$

Without the upper-bound on z^t in (IC$'m1$), this is a standard polymatroid optimization problem, and can be solved using the greedy algorithm. With the upper-bound, this is a weighted polymatroid intersection problem and there exist efficient algorithms solving the optima if the weights $(v^t - k/c)$ are rational. See, for example, Cook et al. (2011) and Frank (2011). In this paper, I solve the problem using a "guess-and-verify" approach. Though we cannot directly apply the greedy algorithm to $(OPTm1 - \varphi)$, it is not hard to conjecture the optimal solution. Intuitively, $z^t = 0$ if $v^t < k/c$. Consider $v^t \geq k/c$. Because $\overline{H}$ is non-decreasing and submodular, and the upper-bound on z^t is linear in f^t, the solution found by the greedy algorithm potentially violates the upper-bound for large t. Hence, it is natural to conjecture that there exists a cutoff $\bar{t}$ such that the upper-bounds in (IC$'m1$) bind if and only if $t > \bar{t}$.

Formally, let $S^t := \{t, \ldots, m\}$ for all $t = 1, \ldots, m$, and $S^{m+1} := \emptyset$. If $\varphi \leq (1-c)/n$, let $\bar{t} := 0$; otherwise, I show in the proof of Lemma 1 that there exists a unique $\bar{t} \in \{1, \ldots, m+1\}$ such that

$$\overline{H}(S^{\bar{t}}) \leq n \sum_{\tau=\bar{t}}^{m} \frac{c\varphi f^\tau}{1-c} \text{ and } \overline{H}(S^{\bar{t}+1}) > n \sum_{\tau=\bar{t}+1}^{m} \frac{c\varphi f^\tau}{1-c}.$$

Note that, by definition, if we assign the highest possible value allowed by (F2$m1$) to $\sum_{\tau=\bar{t}+1}^{m} z^\tau$, then (IC$'m1$) must be violated for some $t \geq \bar{t}+1$; but it is possible to assign the highest possible value allowed by (F2$m1$) to $\sum_{\tau=\bar{t}}^{m} z^\tau$ while respecting (IC$'m1$) for all $t \geq \bar{t}$. Hence, it is natural to conjecture that $\bar{t}$ defined above is the cutoff above which the upper-bounds in (IC$'m1$) bind. I can now construct my candidate optimal solution of $(OPTm1 - \varphi)$ as follows

$$\hat{z}^t := \begin{cases} \overline{z}^t \text{ if } v^t \geq \frac{k}{c} \\ 0 \text{ if } v^t < \frac{k}{c} \end{cases}, \tag{2.7}$$

where

$$\overline{z}^t := \begin{cases} \frac{c\varphi f^t}{1-c} & \text{if } t > \bar{t} \\ \frac{1}{n}\overline{H}(S^{\bar{t}}) - \sum_{\tau=\bar{t}+1}^{m} \frac{c\varphi f^t}{1-c} & \text{if } t = \bar{t} \\ \frac{1}{n}\left[\overline{H}(S^t) - \overline{H}(S^{t+1})\right] & \text{if } t < \bar{t} \end{cases},$$

As I have discussed earlier, if $t > \bar{t}$ and $v^t - k/c > 0$, then I conjecture that the upper-bound in (IC$'m1$) binds and let $\hat{z}^t = c\varphi f^t/(1-c)$. If $t \leq \bar{t}$ and $v^t - k/c > 0$, then, in the spirit of greedy algorithms, I start by assigning the highest possible value allowed by (F2$m1$) to $\hat{z}^{\bar{t}}$ and continue to assign values to $\hat{z}^{\bar{t}-1}, \hat{z}^{\bar{t}-2}, \ldots$ in the same fashion. Finally, it is clear that if $v^t - k/c < 0$, then it is optimal to set $\hat{z}^t = 0$. $\hat{z}$ is feasible following from the fact that $\overline{H}(\emptyset) = 0$, and $\overline{H}$ is non-decreasing and submodular. Furthermore, we can verify the optimality of $\hat{z}$ by the duality theorem:

Lemma 1 *$\hat{z}$ defined in* (2.7) *is an optimal solution to* $(OPTm1 - \varphi)$.

For each $t = 1, \ldots, m$, let

$$P_m^t := \frac{\hat{z}^t}{f^t} + \varphi \tag{2.8}$$

The following corollary directly follows from Lemma 1:

Corollary 1 *P_m defined in* (2.8) *is an optimal solution to* $(OPTm - \varphi)$.

Continuum case

I characterize an optimal solution of $(OPT - \varphi)$ by taking m to infinity. Let v^l be such that $F(v^l)^{n-1} = n\varphi$ and

$$v^u := \inf \left\{ v \middle| 1 - F(v)^n - \frac{n\varphi}{1-c} [1 - F(v)] \geq 0 \right\}. \tag{2.9}$$

v^l is chosen so that if all agents whose values are below v^l are pooled together and ranked below any other agents with higher values, then their interim probability of receiving the object $F(v^l)^{n-1}/n$ is equal to the lower-bound in (IC′), φ. The definition of v^u mirrors that of $\bar{t}$. Informally, v^u is chosen so that if all agents whose values are above v^u are pooled together and ranked above any other agents with lower values, then their interim probability of receiving the object $[1 - F(v)^n]/n[1 - F(v)]$ is equal to the upper-bound in (IC′), $\varphi/(1-c)$. Note that if $\varphi \leq (1-c)/n$, then $v^u = \underline{v}$. Let $\overline{P}_\varphi$ be defined as follows: If $v^l < v^u$, let

$$\overline{P}_\varphi(v) := \begin{cases} \frac{\varphi}{1-c} & \text{if } v \geq v^u \\ F(v)^{n-1} & \text{if } v^l < v < v^u \\ \varphi & \text{if } v \leq v^l \end{cases}.$$

If $v^l \geq v^u$, let

$$\hat{v} := \inf \left\{ v \middle| 1 - n\varphi F(v) - \frac{n\varphi}{1-c} [1 - F(v)] \geq 0 \right\} \in [v^u, v^l], \tag{2.10}$$

and

$$\overline{P}_{\varphi}(v) := \begin{cases} \frac{\varphi}{1-c} & \text{if } v \geq \hat{v} \\ \varphi & \text{if } v < \hat{v} \end{cases} .$$

Finally, let

$$P^*_{\varphi}(v) := \begin{cases} \overline{P}(v) & \text{if } v \geq \frac{k}{c} \\ \varphi & \text{if } v < \frac{k}{c} \end{cases} . \tag{2.11}$$

I show in Appendix A.1 that P^*_{φ} is the "pointwise limit" of P_m as $m \to \infty$. Moreover, P^*_{φ} is an optimal solution to $(OPT - \varphi)$.

Theorem 1 *P^*_{φ} defined in* (2.11) *is an optimal solution to* $(OPT - \varphi)$.

2.3.2. Optimal φ

I complete the characterization of an optimal mechanism by solving for the optimal φ. First, if verification is sufficiently costly or the principal's ability to punish an agent is sufficiently limited, then pure randomization is optimal.

Theorem 2 *If $\overline{v} - k/c \leq \mathbb{E}_v[v]$, then pure randomization, i.e., $P^* = 1/n$ is optimal.*

To make the problem more interesting, in what follows, I assume:

Assumption 1 $\overline{v} - k/c > \mathbb{E}_v[v]$.

Recall that given φ, v^l is uniquely pinned down by $F(v^l)^{n-1} = n\varphi$ and v^u is uniquely pinned down by (2.9). Define v^* and v^{**} by the equations (2.12) and (2.13), respectively:

$$\mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^*\}] + \frac{k}{c} = 0, \tag{2.12}$$

$$\mathbb{E}_v[v] - \mathbb{E}_v[\min\{v, v^{**}\}] + (1-c)\left[\mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^{**}\}] + \frac{k}{c}\right] = 0. \tag{2.13}$$

They are well defined under Assumption 1. Furthermore, $v^{**} > v^* \geq k/c$. Finally, let

$$v^\natural := \sup \left\{ v \,\middle|\, \frac{F(v)^{n-1}(1-F(v))}{1-c} - 1 + F(v)^n \leq 0 \right\}. \tag{2.14}$$

An optimal mechanism is characterized by the following theorem:

Theorem 3 *Suppose that Assumption 1 holds. There are three cases.*

1. *If $F(v^*)^{n-1} \geq n(1-c)$, then the optimal $\varphi^* = F(v^*)^{n-1}/n$ and the following allocation rule is optimal:*

$$P^*(v) := \begin{cases} F(v)^{n-1} & \text{if } v \geq v^* \\ \varphi^* & \text{if } v < v^* \end{cases}.$$

2. *If $F(v^*)^{n-1} < n(1-c)$ and $v^{**} \leq v^\natural$, then the optimal $\varphi^* = (1-c)/n(1-cF(v^{**}))$ and the following allocation rule is optimal*

$$P^*(v) := \begin{cases} \frac{\varphi^*}{1-c} & \text{if } v \geq v^{**} \\ \varphi^* & \text{if } v < v^{**} \end{cases}.$$

3. *If $F(v^*)^{n-1} < n(1-c)$ and $v^{**} > v^\natural$, then the optimal φ^* is defined by*

$$\mathbb{E}_v[v] - \mathbb{E}_v[\min\{v, v^u(\varphi^*)\}] + (1-c)\left[\mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^l(\varphi^*)\}] + \frac{k}{c}\right] = 0, \tag{2.15}$$

and the following allocation rule is optimal:

$$P^*(v) := \begin{cases} \frac{\varphi^*}{1-c} & \text{if } v \geq v^u(\varphi^*) \\ F(v)^{n-1} & \text{if } v^l(\varphi^*) < v < v^u(\varphi^*) \\ \varphi^* & \text{if } v \leq v^l(\varphi^*) \end{cases}.$$

To understand the result, consider the following implementation of the optimal mechanism in Theorem 3. There are two thresholds. I abuse notation here and denote them by v^l and v^u with $\underline{v} \leq v^l \leq v^u \leq \overline{v}$. If every agent reports a value below v^l, then an agent is selected uniformly at random and receives the good, and no one is inspected. If any agent reports a value above v^l but all reports are below v^u, then the agent with the highest reported value receives the good, is inspected with some probability (proportional to $1/c$) and is penalized if he is found to have lied. If any agent reports a value above v^u, then an agent is selected uniformly at random among all the agents whose reported values are above v^u, receives the good, is inspected with a probability of 1 and is penalized if he is found to have lied. I call a mechanism a *one-threshold mechanism* if $v^u = \overline{v}$, a *two-thresholds mechanism* if $v^l < v^u < \overline{v}$, and a *shortlisting mechanism* if $v^l = v^u < \overline{v}$.

To understand conditions (2.12), (2.13) and (2.15), consider the impact of a reduction in v^l. Intuitively, this improves allocation efficiency at the bottom of the value distribution. After some algebra, one can verify that the increase in allocation efficiency is proportional to $\mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^l\}]$. However, as v^l decreases, agents with low v's get worse off and have stronger incentives to exaggerate their types. To restore IC, the principal must now inspect agents more frequently, which raises the total verification cost by an amount proportional to k/c. Furthermore, because the principal's ability to penalize an agent is limited, more pooling at the top, i.e., a lower v^u may also be required to restore IC. This reduces the allocation efficiency at the top by an amount proportional to $\left[\mathbb{E}_v[\min\{v, v^u\}] - \mathbb{E}_v[v]\right]/(1-c)$. In an optimal mechanism, the marginal gain from a reduction in v^l (proportional to the left-hand side of (2.16)) must equal the marginal cost (proportional to the right-hand side of (2.16)):

$$\mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^l\}] = \frac{\mathbb{E}_v[\min\{v, v^u\}] - \mathbb{E}_v[v]}{1-c} + \frac{k}{c}. \tag{2.16}$$

This is precisely the case captured by the third part of Theorem 3 (compare (2.16) with (2.15)). If the limited punishment constraint does not bind, i.e., $v^u = \overline{v}$, there is no efficiency loss at the top and $\left[\mathbb{E}_v[\min\{v, v^u\}] - \mathbb{E}_v[v]\right]/(1-c) = 0$. In this case, (2.16) becomes (2.12) ($v^l = v^*$) and an optimal mechanism is characterized by the first part of Theorem 3. If the principal's ability to punish an agent is sufficiently limited so that $v^u = v^l (= v^{**})$, then (2.16) becomes (2.13) and an optimal mechanism is characterized by the first part of Theorem 3.

Remark 2 *If $k = 0$, then $v^* = \underline{v}$ and $F(v^*)^{n-1} = 0 < n(1-c)$ for any $0 < c < 1$. That is, when verification is free, there is always pooling at the top (Mylovanov and Zapechelnyuk (2014)).*

2.4. Properties of optimal mechanisms

Theorem 3 in the previous section shows that either one-threshold mechanisms, two-thresholds mechanisms or shortlisting mechanisms are optimal. In this section, I show that which of the above three kinds of mechanisms are optimal crucially depends on the number of agents (n). Specifically, I show that there exist $n^*(\rho, c)$ and $n^{**}(\rho, c)$ with $n^*(\rho, c) < n^{**}(\rho, c)$ such that if $n \leq n^*(\rho, c)$, then one-threshold mechanisms are optimal; if $n^*(\rho, c) < n < n^{**}(\rho, c)$, then two-thresholds mechanisms are optimal; if $n \geq n^{**}(\rho, c)$, then shortlisting mechanisms are optimal. Here $\rho := k/c \geq 0$ is referred as the *effective verification cost*. The effective verification cost, ρ, is strictly decreasing in c. This is because a smaller c implies a lower level of punishment, which makes verification essentially more costly as the principal must inspect agents more frequently to maintain IC.

Formally, let $n^*(\rho, c) < 1/(1-c)$ be defined by

$$F(v^*)^{n^*(\rho,c)-1} = n^*(\rho, c)(1-c), \tag{2.17}$$

where v^* is defined by

$$\mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^*\}] + \rho = 0. \tag{2.12}$$

Because v^* is independent of n, by Theorem 3, one-threshold mechanisms are optimal if and only if $n \leq n^*(\rho, c)$. Intuitively, for fixed v^*, an agent whose type below v^* gets the object with probability

$$\varphi^* = \frac{1}{n} F(v^*)^{n-1},$$

which is strictly decreasing in n. In particular, an agent with the lowest type gets worse off and has stronger incentives to exaggerate his type when the number of agents, n, increases. For n sufficiently large, IC cannot be sustained without pooling at the top of the value distribution.

Because v^* is strictly increasing in ρ, the left-hand side of (2.17) is strictly decreasing in n, and the right-hand side of (2.18) is strictly increasing in n, we have n^* is strictly increasing in ρ. Intuitively, as the effective verification cost (ρ) increases, the principal optimally reduces the use of verification and instead enlarges the pooling area at the bottom of the value distribution (v^* increases) to maintain IC. As a result, an agent with the lowest type gets better off (φ increases), and therefore IC can be sustained without pooling at the top for a larger number of agents. For a fixed ρ, v^* is independent of c, but the right-hand side of (2.17) is strictly decreasing in c. Hence, n^* is strictly increasing in c. Intuitively, the upper-bound on $\boldsymbol{P}$ in (IC$'$) becomes larger as c increases, and therefore IC can be sustained without pooling at the top for a larger number of agents.

Next, let $n^{**}(\rho, c) < 1/(1-c)$ be defined by

$$\frac{1 - F(v^{**})^{n^{**}(\rho,c)}}{1 - F(v^{**})} = \frac{F(v^{**})^{n^{**}(\rho,c)-1}}{1-c}, \tag{2.18}$$

where v^{**} is given by

$$\mathbb{E}_v[v] - \mathbb{E}_v[\min\{v, v^{**}\}] + (1-c)\left[\mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^{**}\}] + \frac{k}{c}\right] = 0. \tag{2.13}$$

Compare (2.18) with (2.14), and it is easy to see that $v^{**} \leq v^\natural$ if and only if $n \geq n^{**}(\rho, c)$. By Theorem 3, shortlisting mechanisms are optimal if and only if $n \geq n^{**}(\rho, c)$. As I have discussed earlier, an agent with the lowest type gets worse off and has stronger incentives to exaggerate his type when the number of agents, n, increases. As a result, pooling areas at both the bottom and the top of the value distribution must be enlarged to ensure that the mechanism is incentive compatible and to save verification cost. Formally, I show in Appendix 2.4 that $v^l(n, \rho, c)$ is strictly increasing in n and $v^u(n, \rho, c)$ is strictly decreasing in n. Eventually, for a sufficiently large number of agents, the two pooling areas meet and there is a unique threshold such that all agents whose values are above the threshold and all agents whose values are below the threshold are pooled, respectively.

Recall that $v^{**} > v^*$. Hence,

$$\frac{F(v^{**})^{n^*(\rho,c)-1}}{1-c} > \frac{F(v^*)^{n^*(\rho,c)-1}}{1-c} = n^*(\rho,c) \geq \frac{1-F(v^{**})^{n^*(\rho,c)}}{1-F(v^{**})}.$$

Because the left-hand side of (2.18) is strictly increasing in n, and the right-hand side of (2.18) is strictly decreasing in n, we have $n^{**}(\rho,c) > n^*(\rho,c)$. It is easy to see that $v^{**}(\rho,c)$ is strictly increasing in both ρ and c, and independent of n. Recall that $v^\natural$ is independent of ρ. I show in Lemma 18 in Appendix A.1 that if $n(1-c) < 1$, then $v^\natural$ is strictly increasing in n and strictly decreasing in c. Hence, $n^{**}(\rho,c)$ is strictly increasing in both ρ and c.

An increase in c has two opposite impacts on the size of the pooling areas. On the one hand, the upper-bound on P in (IC′) becomes larger as c increases, which reduces the pooling area at the top (v^u increases) needed to sustain IC. On the other hand, it follows from the analysis in Section 2.3 that if v^u increases, then the marginal benefit from an increase in v^l also increases (the right-hand side of (2.16)).[4] Hence, it is optimal for the principal to enlarge the pooling area at the bottom (v^l increases). Formally, I show in Appendix A.2 that both $v^l(n,\rho,c)$ and $v^u(n,\rho,c)$ are strictly increasing in c. The analysis above on n^{**} shows that the first effect dominates, and two-thresholds mechanisms are optimal for a larger number of agents as c increases.

An increase in ρ also has two opposite impacts on the size of the pooling areas. On the one hand, as I have discussed earlier, as the effective verification cost (ρ) increases, the principal optimally reduces the use of verification and instead enlarges the pooling area at the bottom of the value distribution to maintain IC. On the other hand, as the pooling area at the bottom increases, an agent with the lowest type gets better off, and IC can be sustained with less pooling at the top (v^u increases). Formally, I show in Appendix A.2 that both $v^l(n,\rho,c)$ and $v^u(n,\rho,c)$ are strictly increasing in ρ. The analysis above on n^{**} shows that the second effect dominates, and two-thresholds mechanisms are optimal for a larger number of agents as ρ increases.

These results are summarized by the following corollary:

[4] Note that the marginal cost from a reduction in v^l is the marginal benefit from an increase in v^l.

Corollary 2 *Suppose that Assumption 1 holds. Given $k > 0$, $c \in (0,1)$ and $\rho = k/c$, there exists $0 < n^*(\rho, c) < n^{**}(\rho, c) < 1/(1-c)$ such that the following statements are true:*

1. *If $n \leq n^*(\rho, c)$, then one-threshold mechanisms are optimal; if $n^*(\rho, c) < n < n^{**}(\rho, c)$, then two-thresholds mechanisms are optimal; if $n \geq n^{**}(\rho, c)$, then shortlisting mechanisms are optimal.*

2. *$n^*(\rho, c)$ and $n^{**}(\rho, c)$ are strictly increasing in ρ and c.*

3. *$v^*(n, \rho, c)$ is strictly increasing in ρ, and independent of n and c. v^{**} is strictly increasing in ρ and c, and independent of n. If $n^*(\rho, c) < n < n^{**}(\rho, c)$, then $v^l(n, \rho, c)$ is strictly increasing in n, ρ and c, and $v^u(n, \rho, c)$ is strictly decreasing in n, and strictly increasing in ρ and c.*

Corollary 2 gives comparative statics results in terms of (ρ, c). It is also interesting to see the comparative statics results with respect to the model primitives (k, c). The impact of k is straightforward. As k increases, verification becomes more costly. The optimal mechanism given in Theorem 3 sees more pooling at the bottom (measured by φ) to save verification cost. An increase in φ relaxes the upper-bound on P, which leads to less pooling at the top or no pooling at the top at all. The impact of c is ambiguous. On the one hand, given the amount of pooling at the bottom (measured by φ), a reduction in c lowers the upper-bound on P in (IC$'$), which implies more pooling at the top. On the other hand, a reduction in c makes verification more costly. Similar to the case of an increase in k, this change increases the amount of pooling at the bottom (φ increases), and relaxes the upper-bound on P. As a result, there may be less pooling at the top or no pooling at the top at all. The second channel is absent if verification is free ($k = 0$). The non-monotonicity of the pooling area at the top is further illustrated by the following numerical example.

Example 1 *Consider a numerical example in which $\{v_i\}$ are uniformly distributed on $[0, 1]$. There are $n = 8$ agents. The verification cost is $k = 0.4$. I abuse notation a bit and redefine $v^l = v^u = v^{**}$ if $v^l > v^u$. Figure 1 plots v^l and v^u as functions of c. Observe that the change of v^u is not monotonic. As c increases, the pooling area at the top first expands and then shrinks.*

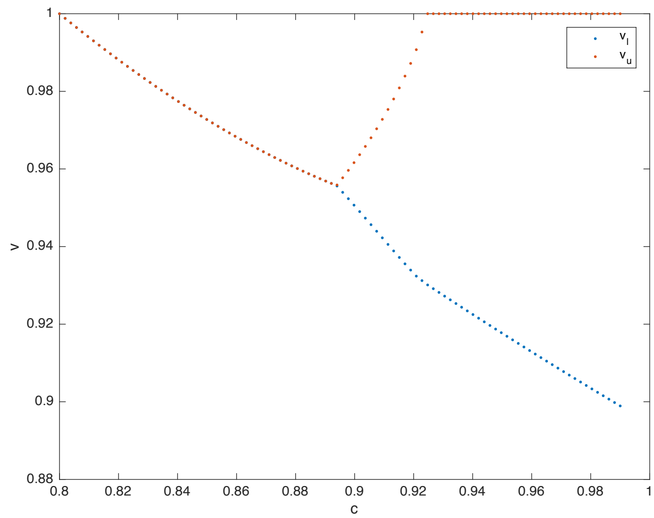


Figure 1: The impact of level of punishment (c)

Finally, a careful examination of (2.17) and (2.12) proves the following corollary:

Corollary 3 $\lim_{c\to 1} n^*(k/c, c) = \infty$ *and* $\lim_{k\to 0} n^*(k/c, c) = 0$.

Corollary 3 shows that as the principal's ability to punish an agent becomes unlimited, the model collapses to Ben-Porath et al. (2014) and as verification cost diminishes, the model collapses to Mylovanov and Zapechelnyuk (2014).

2.5. Extensions

In this section, I consider two extensions. In Section 2.5.1, I consider the general asymmetric environment, and find that a generalized threshold mechanism is optimal in this case. Using this result, I characterize the set of (possibly asymmetric) optimal mechanisms in the symmetric environment and show how limiting the principal's ability to punish agents also limits her ability to treat agents differently. The results in Section 2.5.1 also extend the analysis in Mylovanov and Zapechelnyuk (2014) to the asymmetric environments. In Section 2.5.2, I consider the case in which the principal can get information about and penalize an agent who does not receive the object, and show that threshold mechanism are still optimal in this environment.

2.5.1. Asymmetric environment

In this subsection, I consider the general model with asymmetric agents. Similar to that in Section 2.3, I first characterize an optimal mechanism given the lowest probabilities with which each agent receives the object ($\boldsymbol{\varphi} := (\varphi_1, \dots, \varphi_n)$). Formally, fix $\varphi_i = \inf_{v_i} P_i(v_i)$ for all i and consider the following problem ($OPTA - \varphi$):

$$\max_{\boldsymbol{P},\boldsymbol{Q}} \sum_{i=1}^{n} \mathbb{E}_{v_i}\left[P_i(v_i)\left(v_i - \frac{k_i}{c_i}\right)\right] + \frac{\varphi_i k_i}{c_i},$$

subject to

$$\varphi_i \leq P_i(v_i) \leq \frac{\varphi_i}{1 - c_i}, \forall v_i, \tag{AIC$'$}$$

$$0 \leq Q_i(v_i) \leq 1, \forall v_i, \tag{AF1}$$

$$\sum_i \int_{S_i} P_i(v_i)\mathrm{d}F_i(v_i) \leq 1 - \prod_i \left(1 - \int_{S_i} dF_i(v_i)\right), \forall S_i \subset V_i. \tag{AF2}$$

Clearly, ($OPTA - \varphi$) is feasible only if $\sum_i \varphi_i \leq 1$. As in the symmetric case, I approximate the continuum type space with a finite partition, solve an optimal mechanism in the finite model and take limits. The following theorem gives an optimal solution to ($OPTA - \varphi$):

Theorem 4 *There exist d^l and d_i^u for $i = 1, \dots, n$ such that $\boldsymbol{P}^*$ defined by*

$$P_i^*(v_i) := \begin{cases} \overline{P}_i(v_i) & \text{if } v_i > \frac{k_i}{c_i} \\ \varphi_i & \text{if } v_i < \frac{k_i}{c_i}. \end{cases}, \tag{2.19}$$

where

$$\overline{P}_i(v_i) := \begin{cases} \frac{\varphi_i}{1-c_i} & \text{if } v_i > d_i^u + \frac{k_i}{c_i} \\ \prod_{j \neq i, d_j^u \geq v_i - \frac{k_i}{c_i}} F_j\left(v_i - \frac{k_i}{c_i} + \frac{k_j}{c_j}\right) & \text{if } d^l + \frac{k_i}{c_i} < v_i < d_i^u + \frac{k_i}{c_i} \\ \varphi_i & \text{if } v_i < d^l + \frac{k_i}{c_i} \end{cases}. \tag{2.20}$$

is an optimal solution to $(OPTA - \varphi)$.

Not surprisingly, agents are now ordered by their "net" values $v_i - k_i/c_i$, which is equal to their values to the principal minus the effective verification cost borne by the principal.[5]. As before, there is a unique lower threshold d^l such that all agents whose net values $v_i - k_i/c_i$ below the threshold are pooled. However, there can be up to n distinct upper thresholds d_i^u $(i = 1, \dots, n)$.

To illustrate how an optimal mechanism in Theorem 4 can be implemented, assume that there are two distinct upper thresholds: $d_1^u = \cdots = d_j^u > d_{j+1}^u = \dots d_n^u$. Then the first j agents whose net values are above d_1^u are pooled together, and the rest $n-j$ agents whose net values are above d_{j+1}^u are pooled together and ranked below any of the first j agents whose net value is above d_{j+1}^u. Specifically, the following procedure implements the truth-telling equilibrium in a threshold mechanism: If there exists some agent i $(1 \leq i \leq j)$ whose net value $v_i - k_i/c_i$ is above d_1^u, then one of such agents is selected at random, receives the good and is inspected with probability one. If $v_i - k_i/c_i < d_1^u$ for all $1 \leq i \leq j$ but $v_i - k_i/c_i \geq d_{j+1}^u$ for some $1 \leq i \leq j$, then the agent with the highest reported net value among the first j agents receives the good and is inspected with some probability. If $v_i - k_i/c_i < d_{j+1}^u$ for all $1 \leq i \leq j$ and $v_i - k_i/c_i \geq d_{j+1}^u$ for some $j+1 \leq i \leq n$, then one agent is selected at random among all the agents whose reported net values are above d_{j+1}^u, receives the good and is inspected with some probability. If $v_i - k_i/c_i < d_{j+1}^u$ for all i but $v_i - k_i/c_i \geq d^l$ for some i, then the agent with the highest reported net value receives the good and is inspected with some probability. If $v_i - k_i/c_i < d^l$ for all i, then one agent is selected at random and receives the good, and no one is inspected. Finally, an agent is punished if and only if he is found to have lied.

Because of the complication of pooling areas at the top, it is much harder to find an optimal solution to $(OPTA - \varphi)$. Specifically, d_i^u's are solved recursively from the largest to the smallest. Furthermore, to characterize the set of optimal $\boldsymbol{\varphi}$'s, without priori knowledge of which set of agents share the same upper threshold, one must consider 2^n different cases.[6] Thus, I leave the full characterization

[5]This is consistent with the result in Section 2.3 because when $k_i = k$ and $c_i = c$ for all i, ordering agents by net values is as same as ordering them by values.

[6]Assume, without loss of generality, that $d_1^u \geq \cdots \geq d_n^u$. If there are v distinct upper thresholds, then there are C_n^v possibilities to consider. In total, there are $\sum_{v=1}^n C_n^v = 2^n$ possibilities to consider.

of optimal mechanisms for future research.

Though in general it is extremely hard to characterize the set of optimal $\boldsymbol{\varphi}$, In Appendix A.3.3, I characterize the set of $\boldsymbol{\varphi}$ when the upperbounds on P_i in (AIC$'$) do not bind, i.e., $d_i^u = \overline{v}_i - \frac{k_i}{c_i}$ for all i. If $c_i = 1$ for all i, then these are the unique set of optimal mechanisms found in Ben-Porath et al. (2014).

Symmetric environment revisited

In this section, I revisit the symmetric environment. First, I argue that in the symmetric environment, an optimal mechanism must satisfy: $d_1^u = \cdots = d_n^u$. To understand the intuition behind this result, note first that in the symmetric environment $d_i^u \geq d_j^u$ only if $\varphi_i \geq \varphi_j$. Assume, without loss of generality, that $d_1^u \geq \cdots \geq d_n^u$. Consider, for simplicity, a mechanism in which $\max_j\{\overline{v}_j - k_j/c_j\} > d_1^u > d_2^u > d_3^u$, which implies that $\varphi_1 > \varphi_2$. Then we can construct a new mechanism in which $\varphi_1^* = \varphi_2^* = \sum_{i=1}^2 \varphi_i/2$ and $\varphi_i = \varphi_i^*$ for all $i \geq 3$. In this new mechanism, agents 1 and 2 share the same upper threshold $d^{u*} \in (d_1^u, d_2^u)$ and the upper thresholds of the other agents remain the same. If agents 1 and 2 are ex ante identical, then this new mechanism improves the principal's value by allocating the good between agents 1 and 2 more efficiently when their "net" values, $v_i - k_i/c_i$, lie between (d_1^u, d_2^u).

This property of optimal mechanisms facilitates our analysis of optimal $\boldsymbol{\varphi}$. Theorem 5 below characterizes the set of all optimal $\boldsymbol{\varphi}$. Let v^*, v^{**} and $v^\natural$ be defined by (2.12), (2.13) and (2.14), respectively.

Theorem 5 *Suppose that Assumption 1 holds. There are three cases.*

1. *If $F(v^*)^{n-1} \geq n(1-c)$, then the set of optimal $\boldsymbol{\varphi}$ is the convex hull of*

$$\left\{ \boldsymbol{\varphi} \,\middle|\, \varphi_{i^*} = F(v^*)^{n-1} - (n-1)(1-c), \varphi_j = 1 - c \;\forall j \neq i^*, i^* \in \mathcal{I} \right\}.$$

For each optimal $\boldsymbol{\varphi}^*$*, the following allocation rule is optimal:*

$$P_i^*(v_i) := \begin{cases} F\left(v_i\right)^{n-1} & \text{if } v_i \geq v^* \\ \varphi_i^* & \text{if } v_i < v^* \end{cases}.$$

2. *If* $F(v^*)^{n-1} < n(1-c)$ *and* $v^{**} \leq v^\natural$*, then the set of optimal* $\boldsymbol{\varphi}$ *is the convex hull of*

$$\left\{ \boldsymbol{\varphi} \;\middle|\; \begin{array}{l} \varphi_{i_j} = (1-c)F(v^{**})^{j-1} \text{ if } j \leq h-1, \varphi_{i_h} = \frac{1-c}{1-cF(v^{**})} - \sum_{j=1}^{h-1}(1-c)F(v^{**})^{j-1}, \\ \varphi_{i_j} = 0 \text{ if } j \geq h+1 \text{ and } (i_1, \ldots, i_n) \text{ is a permutation of } (1, \ldots, n) \end{array} \right\},$$

where $1 \leq h \leq n$ *is such that*

$$\frac{1 - F(v^{**})^{h-1}}{1 - F(v^{**})} \leq \frac{1}{1 - cF(v^{**})} < \frac{1 - F(v^{**})^h}{1 - F(v^{**})}.$$

For each optimal $\boldsymbol{\varphi}^*$*, the following allocation rule is optimal:*

$$P_i^*(v_i) := \begin{cases} \frac{\varphi_i^*}{1-c} & \text{if } v_i \geq v^{**} \\ \varphi_i^* & \text{if } v_i < v^{**} \end{cases}.$$

3. *If* $F(v^*)^{n-1} < n(1-c)$ *and* $v^{**} > v^\natural$*, then the the set of optimal* $\boldsymbol{\varphi}$ *is the convex hull of*

$$\left\{ \boldsymbol{\varphi} \;\middle|\; \varphi_{i_j} = (1-c)F(v^u(\varphi^*))^{j-1} \ \forall j \text{ and } (i_1, \ldots, i_n) \text{ is a permutation of } (1, \ldots, n) \right\},$$

where φ^* *is defined by* (2.15) *and, for each* φ*,* v^l *is such that* $F(v^l)^{n-1} = \varphi$ *and* v^u *is defined by* (2.9)*. For each optimal* $\boldsymbol{\varphi}^*$*, the following allocation rule is optimal:*

$$P_i^*(v_i) := \begin{cases} \frac{\varphi_i^*}{1-c} & \text{if } v_i \geq v^u(\varphi^*) \\ F\left(v_i\right)^{n-1} & \text{if } v^l(\varphi^*) < v_i < v^u(\varphi^*) \\ \varphi_i^* & \text{if } v_i \leq v^l(\varphi^*) \end{cases}.$$

Theorem 5 illustrates how limiting the principal's ability to punish agents restricts the principal's ability to treat agents differently. Suppose $F(v^*)^{n-1} \geq n(1-c)$, then the upperbounds on P_i do not bind in an optimal mechanism. This inequality is trivially satisfied if $c = 1$ as in Ben-Porath et al. (2014). In Ben-Porath et al. (2014), there is a class of optimal mechanisms called *favored-agent mechanisms*. In a favored-agent mechanism, there exists a favored-agent i^* whose $\varphi_{i^*} = F(v^*)^{n-1}$ and $\varphi_i = 0$ for any other agent $i \neq i^*$. However, if $c < 1$, then in an optimal mechanism it must be that $\varphi_i \geq 1 - c$ for all i because otherwise some upperbounds on P_i would be violated. Intuitively, the worse an agent is treated when he reports a low type, the stronger incentive he has to exaggerate his type. As a result, as the level of punishment declines, the extent to which the principal can favor one agent at the cost of others without violating the (IC) constraints also declines. Fix the ratio of $\rho = k/c$ so that v^* remains the same. The optimal set of $\boldsymbol{\varphi}$ shrinks as c becomes smaller. When c is such that $F(v^*)^{n-1} = n(1-c)$, the unique optimal $\boldsymbol{\varphi}^*$ is such that $\varphi_1^* = \cdots = \varphi_n^*$. These results are summarized in Corollary 4.

Corollary 4 *Suppose that Assumption 1 holds. Suppose Let $\Phi(\rho, c)$ denote the set of optimal $\boldsymbol{\varphi}^*$. If $c \geq 1 - F(v^*)^{n-1}/n$, then $c < c'$ implies that $\Phi(\rho, c) \subsetneq \Phi(\rho, c')$ and*

$$\lim_{c \searrow 1 - F(v^*)^{n-1}/n} \Phi(\rho, c) = \left\{ \left(\frac{F(v^*)^{n-1}}{n}, \ldots, \frac{F(v^*)^{n-1}}{n} \right) \right\},$$

where v^* is given by (2.12).

If c is small enough so that $F(v^*)^{n-1} < n(1-c)$, then the comparison is less clear because the sets of optimal mechanisms are disjoint for different punishments. In this case, the principal can again treat agents differently but to the extent that they share the same upper threshold. Assume, without loss of generality, that an agent with a smaller index is more favored by the principal in terms of a larger φ_i. Then, in an optimal mechanism, the first h agents cannot be favored too much in the sense that $\sum_{i=1}^{h} \varphi_i \leq (1-c) \sum_{i=1}^{h} F(v^u)^{i-1}$ for all $h = 1, \ldots, n$.

2.5.2. Other verification and punishment technologies

In this subsection, I consider a variation of the model in which I allow for $k_i^\beta < \infty$ and $c_i^\beta > 0$. This means that the principal can get information about and penalize an agent who does not receive the object. I show that threshold mechanisms are still optimal in this environment.

In general, given $p_i(\boldsymbol{v})$ and the expected punishment $\bar{c}_i(\boldsymbol{v}) \leq p_i(\boldsymbol{v})c_i(v_i) + (1 - p_i(\boldsymbol{v}))c_i^\beta(v_i)$, it is optimal for the principal to minimize the expected verification cost:

$$\min_{q_i(\boldsymbol{v}), q_i^\beta(\boldsymbol{v})} p_i(\boldsymbol{v})q_i(\boldsymbol{v})k_i + (1 - p_i(\boldsymbol{v}))q_i^\beta(\boldsymbol{v})k_i^\beta$$

subject to

$$p_i(\boldsymbol{v})q_i(\boldsymbol{v})c_i(v_i) + (1 - p_i(\boldsymbol{v}))q_i^\beta(\boldsymbol{v})c_i^\beta(v_i) = \bar{c}_i(\boldsymbol{v}). \tag{2.21}$$

Clearly, depending on the relative magnitudes of the effective verification costs when an agent receives the object ($k_i/c_i(v_i)$) and that when an agent does not receive the object ($k_i^\beta/c_i^\beta(v_i)$), there are three cases: (i) If $k_i/c_i(v_i) < k_i^\beta/c_i^\beta(v_i)$, then it is optimal to set $q_i(\boldsymbol{v}) = \min\{\bar{c}_i(\boldsymbol{v})/p_i(\boldsymbol{v})c_i(v_i), 1\}$ and $q_i^\beta(\boldsymbol{v}) = \max\{0, (\bar{c}_i(\boldsymbol{v}) - p_i(\boldsymbol{v})c_i(v_i))/(1 - p_i(\boldsymbol{v}))c_i^\beta(v_i)\}$. (ii) If $k_i/c_i(v_i) > k_i^\beta/c_i^\beta(v_i)$, then it is optimal to set $q_i^\beta(\boldsymbol{v}) = \min\{1, \bar{c}_i(\boldsymbol{v})/(1 - p_i(\boldsymbol{v}))c_i^\beta(v_i)\}$ and $q_i(\boldsymbol{v}) = \max\{0, (\bar{c}_i(\boldsymbol{v}) - (1 - p_i(\boldsymbol{v}))c_i^\beta(v_i))/p_i(\boldsymbol{v})c_i(v_i)\}$. (iii) If $k_i/c_i(v_i) = k_i^\beta/c_i^\beta(v_i)$, then any $q_i(\boldsymbol{v})$ and $q_i^\beta(\boldsymbol{v})$ satisfying (2.21) are optimal.

For tractability, I assume in what follows that $c_i(v_i) = c_i b_i(v_i)$ and $c_i^\beta(v_i) = c_i^\beta b_i(v_i)$ for all v_i. For simplicity, I also assume that $k_i^\beta = k_i$ and $c_i^\beta = c_i$ for all i. The results in this subsection can readily extend to more general cases, when, for example, it is more costly for the principal to get information about an agent who does not receive the object ($k_i^\beta \geq k_i$), and the punishment is also less severe for an agent who does not receive the object ($c_i^\beta \leq c_i$). Given $(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{q}^\beta)$, let $P_i(v_i) := \mathbb{E}_{v_{-i}}\left[p_i(v_i, v_{-i})\right]$ be the interim probability with which an agent receives the object and $\hat{P}(v_i) := \mathbb{E}_{v_{-i}}\left[p_i(v_i, v_{-i})q_i(v_i, v_{-i}) + (1 - p_i(v_i, v_{-i}))q_i^\beta(v_i, v_{-i})\right]$ be the interim probability with which an agent is inspected. The principal's problem can be written in the reduced form:

$$\max_{\boldsymbol{P},\hat{\boldsymbol{P}}} \sum_{i=1}^{n} \mathbb{E}_{v_i} \left[P_i(v_i)v_i - \hat{P}(v_i)k_i \right],$$

subject to

$$P_i(v_i) \geq P_i(v_i') - \hat{P}_i(v_i')c_i, \forall v_i, v_i', \tag{IC-OT}$$

$$0 \leq \hat{P}_i(v_i) \leq 1, \forall v_i, \tag{F1-OT}$$

$$\sum_i \int_{S_i} P_i(v_i)\mathrm{d}F_i(v_i) \leq 1 - \prod_i \left(1 - \int_{S_i} dF_i(v_i)\right), \forall S_i \subset V_i. \tag{AF2}$$

Note that the (IC-OT) constraints hold if and only if

$$\varphi_i \geq P_i(v_i') - \hat{P}(v_i')c_i, \forall v_i'. \tag{2.22}$$

Because $\hat{P}_i(v_i') \leq 1$, (2.22) holds only if

$$P_i(v_i') \leq \varphi_i + c_i, \forall v_i'. \tag{2.23}$$

Suppose that (2.23) holds, then it is optimal to set $\hat{P}_i(v_i) = (P_i(v_i) - \varphi_i)/c_i$ for all $v_i \in V_i$. Substituting this into the principal's objective function yields:

$$\sum_{i=1}^{n} \mathbb{E}_{v_i} \left[P_i(v_i) \left(v_i - \frac{k_i}{c_i} \right) \right] + \frac{\varphi_i k_i}{c_i}. \tag{2.6}$$

Note that, given $\{\varphi_i\}$, the principal's objective function is as same as that in the case of $c_i^{\beta} = 0$. The only difference between the principal's two problems is the upperbound on P_i. (Compare (2.23) with (2.5).)

There are two interesting observations. First, the upperbound on P_i does not bind in the original problem ($\varphi_i/(1 - c_i) \geq 1$) if and only if it does not bind in the new problem ($\varphi_i + c_i \geq 1$). This implies that part 1 of Theorem 3 still applies here. Second, If the upperbound binds in the original problem, i.e., $\varphi_i/(1-c_i) < 1$, then the new upperbound is larger:

$$\varphi_i + c_i - \frac{\varphi_i}{1-c_i} = c_i\left(1 - \frac{\varphi_i}{1-c_i}\right) > 0.$$

This is intuitive because allowing for the principal to penalize an agent who does not receive the object clearly relaxes the principal's problem. Hence, any feasible solution to the new problem is also feasible in the original problem.

In the interest of the length of the paper, I only characterize an optimal mechanism in the symmetric environment. In what follows, I assume that $\{v_i\}$ are identically distributed and $c_i = c$ and $k_i = k$ for all i. Without loss of generality, I can focus on symmetric mechanisms. In what follows, I suppress the subscript i whenever the meaning is clear.

First, as in the case of $k_i^\beta = 0$, if verification is sufficiently costly or the principal's ability to punish an agent is sufficiently limited, then pure randomization is optimal. In particular, Theorem 2 still applies here. To make the problem more interesting, in what follows, I assume that Assumption 1 holds. Let v^* be defined by (2.12) and redefine v^{**} by

$$\mathbb{E}_v[v] - \mathbb{E}_v[\min\{v, v^{**}\}] + \mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^{**}\}] + \frac{k}{c} = 0. \tag{2.24}$$

v^* and v^{**} are well defined under Assumption 1. Furthermore, $v^{**} > v^* \geq k/c$. Finally, redefine

$$v^\natural := \sup\left\{ v \,\middle|\, \left(F(v)^{n-1} + nc\right)(1 - F(v)) - 1 + F(v)^n \leq 0 \right\}. \tag{2.25}$$

Theorem 6 below characterizes an optimal symmetric mechanism. The proof is similar to that in Section 2.3 and neglected here.

Theorem 6 *Suppose Assumption 1 holds. There are three cases.*

1. *If $F(v^*)^{n-1} \geq n(1-c)$, then the optimal $\varphi^* = F(v^*)^{n-1}/n$ and the following allocation rule*

is optimal:

$$P^*(v) := \begin{cases} F(v)^{n-1} & \text{if } v \geq v^* \\ \varphi^* & \text{if } v < v^* \end{cases}.$$

2. *If* $F(v^*)^{n-1} < n(1-c)$ *and* $v^{**} \leq v^\natural$, *then the optimal* $\varphi^* = (1-nc)/n(1-nc+ncF(v^{**}))$ *and the following allocation rule is optimal:*

$$P^*(v) := \begin{cases} \varphi^* + c & \text{if } v \geq v^{**} \\ \varphi^* & \text{if } v < v^{**} \end{cases}.$$

3. *If* $F(v^*)^{n-1} < n(1-c)$ *and* $v^{**} > v^\natural$, *then the optimal* φ^* *is defined by*

$$\mathbb{E}_v[v] - \mathbb{E}_v[\min\{v, v^u(\varphi^*)\}] + \mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^l(\varphi^*)\}] + \frac{k}{c} = 0, \tag{2.26}$$

and the following allocation rule is optimal:

$$P^*(v) := \begin{cases} \varphi^* + c & \text{if } v \geq v^u(\varphi^*) \\ F(v)^{n-1} & \text{if } v^l(\varphi^*) < v < v^u(\varphi^*) \\ \varphi^* & \text{if } v \leq v^l(\varphi^*) \end{cases}.$$

Note that the optimal mechanism obtained here is very similar to what obtained if the principal can only penalize an agent who receives the object, but has different thresholds when the limited punishment constraint is binding.

2.6. Concluding remarks

In this paper, I study the problem of a principal who has a single indivisible object to allocate among a number of agents. Each agent has private information about the principal's payoff of allocating the object to him. There are no monetary transfers. The principal can inspect agents' reports at a cost and punish them, but the punishments are limited. I show that some simple threshold mechanisms are optimal in this setting. This paper includes Ben-Porath et al. (2014) and Mylovanov and Zapechelnyuk (2014) as special cases and bridges their gaps. Specifically, if the number of agents is small, then the optimal mechanism only involves pooling area at the bottom of value distribution as in Ben-Porath et al. (2014). As the number of agents increases, pooling at the top is required to guarantee incentive compatibility as in Mylovanov and Zapechelnyuk (2014). These results highlight the role played by the number of agents in shaping optimal mechanisms, which is absent or overlooked in previous work on mechanism design.

First, earlier mechanism design papers studying an allocation problem often focus on mechanisms with monetary transfers and ignore the possibility of the principal verifying agents' information. In these papers, a robust feature of optimal mechanisms is that they are independent of the number of agents. For example, in the seminal work of Myerson (1981), under some regularity conditions, the revenue-maximizing mechanism can be implemented by a first-price or second-price auction with a reserve price, and, in particular, this optimal reserve price is independent of the number of agents. This difference is mainly because the kinds of binding IC constraints are different in the two settings. In Myerson (1981), the binding IC constraints are between adjacent types, and the difference between two adjacent types' allocation rules is insensitive to a change in the number of agents. But, in this paper, the binding IC constraints correspond to those of the lowest possible type misreports as higher types. Note that, as the number of agents increases, the lowest possible type's probability of receiving the object declines much faster compared with a much higher type.

Second, in Ben-Porath et al. (2014), the optimal mechanisms are also independent of the number of agents. Recall that when punishment is unlimited, a one-threshold mechanism is optimal, and the threshold is independent of the number of agents (see the third part in Corollary 2). This difference is because when the level of punishment is sufficiently high, although the difference in the probabilities of receiving the object between the highest possible type and the lowest possible type increases as the number of agents increases, the principal can always guarantee IC by verifying an agent's information and punishing him. But, in this paper, the level of punishment is limited. In this case, as the number of agents increase, rationing becomes indispensable to guarantee IC, and the required rationing areas also increase.

CHAPTER 3 : MECHANISM DESIGN WITH FINANCIALLY CONSTRAINED AGENTS AND COSTLY VERIFICATION

3.1. Introduction

Governments around the world allocate a variety of valuable resources to agents who are financially constrained. In Singapore, for example, 80% of the population's housing needs are met by the Housing and Development Board (HDB), a government agency founded in 1960 to provide affordable housing.[1] In the United States, Medicaid has provided health care to individuals and families with low income and limited resources since 1965. Medicaid currently accounts for 16.1% of the state general funds[2] and provides health coverage to 80 million low-income people.[3] Similar public housing and social health care programs prevail in many other countries.[4,5] In China, several cities limit the supply of vehicle licenses to curb the growth in private vehicles, and different cities have implemented different mechanisms. For example, Shanghai allocates vehicle licenses through an auction-like mechanism, while Beijing uses a vehicle license lottery (see Rong et al. 2015). The evaluation of existing mechanisms has attracted attention from researchers and policymakers. In comparison to lotteries, an auction-like mechanism is considered more efficient but favors high-income families more.

One justification for this role of a government is that a competitive market outcome will not maximize social surplus if agents are financially constrained. Financial constraints mean that in a competitive market some agents with high valuations will not obtain goods, while agents with low valuations but access to cash will. The natural question arises as to what the surplus-maximizing (or optimal) mechanism is in these circumstances when both valuations and financial constraints are the agents' private information.

The mechanism design literature concerning this question has focused on mechanisms with only

[1] http://www.hdb.gov.sg/fi10/fi10320p.nsf/w/AboutUsPublicHousing?OpenDocument

[2] http://ccf.georgetown.edu/wp-content/uploads/2012/03/Medicaid-state-budgets-2005.pdf

[3] http://www.cbpp.org/research/health/policy-basics-introduction-to-medicaid?fa=view&id=2223

[4] https://en.wikipedia.org/wiki/Public_housing

[5] https://en.wikipedia.org/wiki/Universal_health_coverage_by_country

monetary transfers and has ignored the possibility of the principal verifying the agents' reported information about their abilities to pay. Indeed, in many instances, the principal relies on agents' reports of their ability to pay, and the principal can verify this information and punish an agent who makes a false statement. For example, applicants for HDB flats in Singapore and Medicaid in the United States are subject to a set of eligibility conditions on age, family nucleus, monthly income, and so on. The verification process can be costly, though. First, in some developing countries, verifiable records on household income or wealth are rarely available, and governments lack the administrative capacity to process this information. In such cases, alternative verification methods such as a visit to the household to inspect the visible living conditions are not uncommon but are often costly (see Coady et al. 2004). Second, certain types of income such as tips, side-jobs and cash receipts are costly to verify. Similarly, governments have few ways to verify the income reports by individuals who are self-employed or run small business without performing a costly investigation. Third, agents may be financially constrained due to limited access to the financial market or high expenditures, such as medical expenses or education costs. This information is often costly for governments to verify. Last but not least, even when the verification cost for one individual is low, the total cost can be substantial for a large population.

Hence, it is important to explore how the option of costly verification affects the optimal mechanism. Verification allows the principal to better target low-budget agents and potentially improve their welfare. However, verification is costly and reduces the amount of money available for subsidies. The principal must now trade allocative efficiency for verification cost. The cost of verification also influences whether the principal chooses to use *cash subsidies* or *in-kind subsidies* (the provision of goods at discounted prices). The latter is less efficient because it often involves rationing, but saves verification cost because it only benefits low-budget agents with high valuations. Finally, introducing costly verification also complicates the analysis because it is no longer sufficient to consider "local" incentive compatibility (IC) constraints. Because the IC constraints between distant types can also bind, one cannot anticipate a priori the set of binding IC constraints.

To study these questions, I consider a mechanism design problem in which there is a unit mass of a continuum of agents and a limited supply of indivisible goods. Each agent has two-dimensional private information — his valuation of the good $v \in [\underline{v}, \overline{v}]$ and his exogenous budget constraint b. The budget constraint is a hard one in the sense that agents cannot be compelled to pay more than their budgets. For simplicity, I assume there are only two possible types of budgets, $b_2 > b_1$. The principal can inspect an agent at a cost, perfectly revealing his budget, and impose a penalty on detected misreporting. The principal is also subject to a budget balance constraint which requires that the revenue from selling the good must exceed the inspection cost. This constraint rules out the possibility that the principal can inject money and relieve all budget constraints. I focus on direct mechanisms in which each agent reports private information directly and is punished if and only if found to have lied about the budget. Given the report, the mechanism specifies for each agent his probability of getting the good, his payment and his probability of being inspected.

I characterize the optimal direct mechanism which maximizes utilitarian efficiency among all mechanisms that are incentive compatible and individually rational, and that satisfy the resource constraint, agents' budget constraints and the principal's budget balance constraint.

Let $u(\underline{v}, b)$ denote the utility of an agent with the lowest valuation $\underline{v}$ and budget b, which is also the amount of cash subsidies received by agents with budget b. There exist three cutoffs $v_1^* \leq v_2^* \leq v_2^{**}$. Firstly, low-budget agents whose valuations are below v_1^* and high-budget agents whose valuations are below v_2^* only receive only cash subsidies. Not surprisingly, these low-budget agents receive higher cash subsidies and are inspected with probability proportional to the difference in cash subsidies $u(\underline{v}, b_1) - u(\underline{v}, b_2)$. Secondly, low-budget agents whose valuations exceed v_1^* receive the good with probability $a^* \leq 1$ and make a payment of $a^* v_1^* - u(\underline{v}, b_1)$. They receive both cash and in-kind subsidies. High-budget agents whose valuations lie in $[v_2^*, v_2^{**}]$ are pooled with low-budget agents whose valuations are above v_1^*. They also receive the good with probability a^*, but they make a payment of $a^* v_2^* - u(\underline{v}, b_2)$. The difference in in-kind subsidies is given by $a^*(v_2^* - v_1^*)$, and these low-budget agents are inspected with probability proportional to the sum of differences in cash and in-kind subsidies $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^*(v_2^* - v_1^*)$. Finally, high-budget agents receive the good for sure and make a payment of $v_2^{**} - u(\underline{v}, b_2)$ if their valuations exceed v_2^{**}.

If budgets are common knowledge, then the principal can without cost target low-budget agents and provides cash subsidies and in-kind subsidies only to low-budget agents. If budgets are agents' private information and cannot be verified, then high-budget agents whose valuations are below v_2^* have incentives to misreport as low-budget types to receive cash subsidies; and high-budget agents whose valuations are slightly above v_2^* have incentives to misreport as low-budget types to receive the good at a lower payment. As a result, in this case, agents with both budgets receive the same amount of cash subsidies ($u(\underline{v}, b_1) = u(\underline{v}, b_2)$) and in-kind subsidies ($v_1^* = v_2^*$).

The optimal direct mechanism can be implemented by a simple two-stage mechanism. Specifically, all agents are asked to report their budgets in the first stage. The principal then provides budget-dependent cash subsidies to agents and assigns the goods randomly (with uniform probability) at budget-dependent prices. Agents who report low budgets receive higher cash subsidies and lower prices. In the second stage, a resale market opens, but is regulated with budget-dependent sales taxes. Agents who report low budgets are subject to higher sales taxes. Only agents who report low budgets are inspected randomly. Unlike the case without inspection, in which all agents are subsidized and regulated equally regardless of their budgets, the two-stage mechanism provides more subsidies to low-budget agents in their initial purchases (the first stage) and imposes more restrictions on them in the resale market (the second stage). Although in my analysis the principal's objective is to maximize social surplus, I conjecture that these features would continue to apply when the principal wants to benefit only low-budget agents.

This implementation exhibits some features of the public housing program in Singapore, as shown in Table 1. In Singapore, buyers of resale HDB flats can apply for additional housing grants. If these flats are purchased with housing grants, these buyers are required to reside in their flats for at least 5 years before they could resell or sublet. In contrast, flats purchased without housing grants are subject to no requirement or a shorter one.

It is interesting to see how verification cost, the supply of goods and other parameters affect the optimal mechanism and welfare. I provide analytic results of comparative statics for extreme cases, such as when verification cost is sufficiently large and the supply of goods is sufficiently large or

Table 1: Minimum occupation periods (MOP) of housing and development board (HDB) flats

Types of HDB flats	MOP	
	Sell	Sublet
Resale flats w/ Grants	5–7 years	5–7 years
Resale flats w/o Grants	0–5 years	3 years

Sources. — Sell: http://www.hdb.gov.sg/cs/infoweb/residential/selling-a-flat/eligibility; and Sublet: http://www.hdb.gov.sg/cs/infoweb/residential/renting-out-a-flat-bedroom/renting-out-your-flat/eligibility.

small, and I explore the intermediate case numerically.

Verification allows the principal to better target low-budget agents and improves their welfare. Intuitively, as verification becomes costly, the principal tends to provide relatively smaller subsidies to low-budget agents and inspect them less frequently. More interestingly, the optimal mechanism makes use of both cash and in-kind subsidies, and the change in verification cost affects that mechanism's reliance on each of them. If verification is cheap, then the principal achieves efficiency mainly by offering more cash subsidies to low-budget agents. As verification becomes costly, the difference in cash subsidies declines but the difference in in-kind subsidies increases. This is because in-kind subsidies are attractive only to high-valuation agents, which is cheaper in terms of verification cost. Eventually, the difference in in-kind subsidies also declines as verification becomes sufficiently costly. Though reducing verification cost improves the welfare of low-budget agents, it may hurt high-budget agents as more subsidies are diverted to low-budget agents.

Another interesting observation is that although an increase in the supply of goods improves the total welfare, its impact on the welfare of each budget type is not monotonic. This is because an increase in the supply has two opposite effects. On the one hand, the principal becomes less budget constrained, and can direct more subsidies to low-budget agents and inspect them more frequently. On the other hand, low-budget agents also become less budget constrained, which reduces the needs to subsidize and inspect them. As a result, the differences in cash and in-kind subsidies and the inspection probability are hump-shaped. Initially, the welfare of both budget types increases as the supply increases. When the supply is large enough that the principal can afford to provide more subsidies to low-budget agents, the welfare of high-budget agents begins to decrease. Eventually, the need to subsidize low-budget agents decreases as the supply increases while the welfare of low-budget agents begins to decrease and that of high-budget agents begins to increase, until they coincide.

Technically, this paper develops a novel method that can potentially be used in solving other mechanism design problems with multidimensional types. If each agent has only one-dimensional private information, i.e., valuation, then it is sufficient to consider adjacent IC constraints; if each agent has two-dimensional private information but the principal cannot inspect budgets, then it is sufficient to consider two one-dimensional deviations. These, however, no longer apply in the case that each agent has two dimensional private information and the principal can inspect budget at a cost. In this case, in addition to downward adjacent IC constraints of misreporting values, one must consider deviations in which an agent can misreport both dimensions of his private information. As a result, the local approach commonly used does not work here.

To overcome this difficulty, I first restrict attention to a class of allocation rules that have enough structures to help me keep track of binding IC constraints, and that are also rich enough to approximate any general allocation rule well. Specifically, I approximate the allocation rule of each budget type using step functions. When restricting attention to step functions, binding IC constraints corresponding to the under-reporting of budgets are between different budget types whose values are the jump discontinuity points of their allocation rules. This structure allows me to write the optimal inspection rule as a function of the possible values and jump discontinuity points of the allocation rule. I then solve a modification of the principal's problem in which the allocation rule of low-budget types are restricted to take at most M distinct values. Because for M sufficiently large step-functions can approximate the optimal allocation rule arbitrarily well, I can obtain a characterization of the optimal mechanism in the limit.

The rest of the paper is organized as follows. Section 3.1.1 discusses related work. Section 3.2 presents the model. Section 3.3 characterizes the direct optimal mechanism when all agents' budget constraints are common knowledge. Section 3.4 characterizes the direct optimal mechanism when an agent's budget is his private information. Section 3.5 provides a simple implementation. Section

3.6 studies the properties of the optimal mechanism. Section 3.7 considers various extensions of the model. Section 3.8 concludes. All the proofs are relegated to the appendix.

3.1.1. Related literature

This paper is related to two branches of literature. First, it contributes to the literature studying mechanism design problems when agents are financially constrained by incorporating costly verification. Prior work analyzes the revenue or efficiency of a given mechanism or the design of an optimal mechanism when either budgets are common knowledge, or budgets are agents' private information but cannot be verified. See Che and Gale (1998, 2006, 2000), Laffont and Robert (1996), Maskin (2000), Benoit and Krishna (2001), Brusco and Lopomo (2008), Malakhov and Vohra (2008) and Pai and Vohra (2014a).

In this first branch of literature, the two closest papers to the current paper are Che et al. (2013a) and Richter (2015). In Che et al. (2013a) and Richter (2015), like in this paper, there is a unit mass of a continuum of agents and a limited supply of goods. In Richter (2015) agents have linear preferences for an unlimited supply of the goods. He finds that both the revenue-maximizing mechanism and surplus-maximizing mechanism feature a linear price for the good. In addition, the surplus-maximizing mechanism has a uniform cash subsidy. In both Che et al. (2013a) and this paper, each agent has a unit demand for an indivisible good, and the surplus-maximizing mechanism can be implemented via a random assignment with a regulated resale and cash subsidy scheme. However, Che et al. (2013a) does not consider the possibility that the principal can verify an agent's budget at a cost. This feature also distinguishes the current paper from all the other papers on mechanism design with financially constrained agents. Che et al. (2013a) first compare three different methods of assigning the goods when agents have a continuum of possible valuations and a continuum of possible budgets, and then characterize the optimal mechanism in a simple 2×2 model, in which each agent has two possible valuations of the good and two possible budgets. In the presence of costly verification, unlike Che et al. (2013a), in which all agents are subsidized and regulated equally regardless of their budgets in an optimal mechanism, I show that an optimal mechanism provides more subsidies to low-budget agents in their initial purchases and imposes more restrictions on them in the resale market.

Second, this paper is related to the costly state verification literature. The first significant contribution to this series is from Townsend (1979), who studies a model of a principal and a single agent. In Townsend (1979) verification is deterministic. Border and Sobel (1987) and Mookherjee and Png (1989) generalize it by allowing random inspection. Gale and Hellwig (1985) consider the effects of costly verification in the context of credit markets. Recently, Ben-Porath et al. (2014) study the allocation problem in the costly state verification framework when there are multiple agents and monetary transfer is not possible. Li (2016) extends Ben-Porath et al. (2014) to environments in which the principal's ability to punish an agent is limited. These models differ from what I consider here in that in their models each agent has only one-dimensional private information.

This paper is also somewhat related to the literature on costless or ex-post verification. Glazer and Rubinstein (2004) can be interpreted as a model of a principal and one agent with limited but costless verification and no monetary transfers. Mylovanov and Zapechelnyuk (2014) study a model of multiple agents with costless verification but limited punishments. This paper differs from these earlier studies in that each agent has two-dimensional private information, verification is costly and there are monetary transfers.

In the literature discussed above, one can anticipate a priori the set of binding IC constraints, which is no longer true here. Instead, I use new techniques for keeping track of binding IC constraints.

3.2. Model

There is a unit mass of a continuum of agents. There is a mass $S \in (0,1)$ of indivisible goods.[6] Each agent has a private valuation of the good $v \in V := [\underline{v}, \overline{v}] \subset \mathbb{R}_+$, and a privately known budget $b \in B := \{b_1, b_2\}$. I assume that $b_1 > \underline{v}$ and $b_2 > \overline{v}$.[7] Thus, a high-budget agent is never budget constrained in an individually rational mechanism. The type of an agent is a pair consisting of his valuation and his budget: $t := (v, b)$; and the type space is $T := V \times B$.

[6]The model is also applicable to divisible goods when an agent's per-unit value for the good is constant up to an upper bound.

[7]All the results can be easily extended to any $b_1 \geq 0$. In the paper, I assume $b_1 > \underline{v}$ to make the statement more concise.

I assume v and b are independent. Each agent has a high budget with probability π and a low budget with probability $1 - \pi$. The valuation v is distributed with cumulative distribution function F and strictly positive density f.

The principal can inspect an agent's budget at a cost $k \geq 0$, and can impose a penalty $c > 0$. Inspection perfectly reveals an agent's budget.[8] I assume that the penalty c is large enough that an agent never find it optimal to misreport his budget if he is certain he will be inspected. For the main body of the paper, I assume that the penalty is not transferable. In Section 3.7.2, I study the case in which penalty is transferable and show that all results hold in that case. For later use, let $\rho := k/c$. As it will become clear, ρ measures the "effective" inspection cost to the principal. The cost to an agent to have his report verified is zero. This assumption is reasonable if the goods are valuable to agents and disclosure costs are negligible. In Section 3.7.3, I discuss what happens if it is also costly for an agent to have his report verified.

The usual version of the revelation principle (see, e.g., Myerson 1979 and Harris and Townsend 1981) does not apply to models with verification. However, it is not hard to extend the argument to this type of environment.[9] Specifically, I show in Appendix B.1 that it is without loss of generality to restrict attention to direct mechanisms. Furthermore, I assume that the principal can only punish an agent who is inspected and found to have lied about his budget. This assumption, however, is not without loss of generality. I discuss what happens if the principal is allowed to punish an agent without verifying his budget or to punish an agent who is found to have reported his budget truthfully in Section 3.7.4.

A direct mechanism is a triple (a, p, q), where $a : T \to [0, 1]$ denotes the probability an agent obtains the good, $p : T \to \mathbb{R}$ denotes the payment an agent must make and $q : T \to [0, 1]$ denotes the probability of inspection. In this definition, I implicitly assume that payment rules are deterministic. I discuss random payment rules at the end of this section and show that it is without loss of generality to focus on deterministic payment rules.

[8]The paper's results will not change if the principal cannot detect a lie with some probability.

[9]See Townsend (1988) and Ben-Porath et al. (2014) for more discussion and extension of the revelation principle to various verification models, not including the environment considered in this paper.

The utility of an agent who has type $t := (v, b)$ and reports $\hat{t}$ is

$$u(\hat{t}, t) = \begin{cases} a(\hat{t})v - p(\hat{t}) & \text{if } \hat{b} = b \text{ and } p(\hat{t}) \leq b, \\ a(\hat{t})v - q(\hat{t})c - p(\hat{t}) & \text{if } \hat{b} \neq b \text{ and } p(\hat{t}) \leq b, \\ -\infty & \text{if } p(\hat{t}) > b. \end{cases}$$

An agent has a standard quasi-linear utility up to his budget constraint, and cannot pay more than his budget.

The welfare criterion I use is utilitarian efficiency. For why utilitarian efficiency is a reasonable welfare criterion, see Vickrey (1945) and Harsanyi (1955). Given quasi-linear preferences, the total value realized minus total inspection cost is an equivalent criterion.[10] The principal's problem is[11]

$$\max_{a,p,q} \mathbb{E}_t \left[a(t)v - q(t)k \right], \tag{$\mathcal{P}$}$$

subject to

$$u(t) \equiv u(t, t) \geq 0, \ \forall t \in T, \tag{IR}$$

$$u(t) \geq u(\hat{t}, t), \ \forall t \in T, \hat{t} \in \left\{ \hat{t} \in T \,\middle|\, p(\hat{t}) \leq b \right\}, \tag{IC}$$

$$p(t) \leq b, \ \forall t \in T, \tag{BC}$$

$$\mathbb{E}_t \left[p(t) - q(t)k \right] \geq 0, \tag{BB}$$

$$\mathbb{E}_t \left[a(t) \right] \leq S. \tag{S}$$

The individual rationality (IR) constraint requires that each agent gets a non-negative expected payoff from participating in the mechanism. The incentive compatibility (IC) constraint requires that it is

[10]To see this, consider a feasible mechanism (a, p, q). Note that if (a, p, q) maximizes welfare, then (BB) must hold with equality. Otherwise the principal can improve welfare through lump-sum transfers. Then the principal's objective function becomes

$$\mathbb{E}[u(t)] = \mathbb{E}[va(t) - p(t)] = \mathbb{E}[va(t) - q(t)k],$$

where the last equality holds since (BB) holds with equality.

[11]There are some subtle issues with a continuum of random variables. See Judd (1985). However, if we interpret the continuum model as an approximation of a large economy, then Al-Najjar (2004) makes the limiting argument rigorous.

weakly better for an agent to report his true type than any other type whose transfers he can afford. The budget constraint (BC) states that an agent cannot be ask to make a payment larger than his budget b. To be clear, note that (BC) follows from (IR). This budget constraint is the same as that found in Che and Gale (2000) and Pai and Vohra (2014a), but different from Che et al. (2013a), who use a per unit price constraint.[12] I discuss the differences of the two frameworks in Section 3.7.1. The principal's budget balance (BB) constraint requires that the revenue raised from selling the goods must exceed the inspection cost. (BB) rules out the possibility that the principal can inject money and relieve all budget constraints. Finally, the limited supply (S) constraint, which requires that the amount of good assigned cannot exceed the supply. We say a mechanism (a, p, q) is *feasible* if it satisfies constraints (IR), (IC), (BC), (BB) and (S).

Throughout the paper, I assume that $S < 1 - F(b_1)$ since otherwise the first-best can be achieved via a competitive market. I also impose the following two assumptions throughout the paper.

Assumption 2 $\frac{1-F}{f}$ *is non-increasing.*

Assumption 3 f *is non-increasing.*

Assumption 2 is the standard monotone hazard rate condition, which is often adopted in the mechanism design literature. This assumption ensures that allocating more good to agents with higher valuations from those with lower valuations generates higher revenues for the principal. Assumption 3 says that agents are less likely to have higher valuations than to have lower valuations. These two assumptions are also imposed in Richter (2015) and Pai and Vohra (2014a). These two assumptions are satisfied by some commonly used distributions such as uniform distributions, exponential distributions and left truncation of a normal distribution.

I conclude this section with a discussion of random payment rules.

[12]This constraint is called ex-post budget constraint in Che et al. (2013a).

3.2.1. *Random payment rules*

When defining a direct mechanism, I implicitly assume that the payment rule is deterministic. I argue that this is without loss of generality. Consider a random payment rule $\tilde{p} : T \to \Delta(\mathbb{R})$. Let $\text{supp}(\tilde{p}(t))$ denote the supremum of payments in the support of $\tilde{p}(t)$. The utility of an agent who has type t and report $\hat{t}$ is

$$u(\hat{t}, t) = \begin{cases} a(\hat{t})v - \mathbb{E}[\tilde{p}(\hat{t})] & \text{if } \hat{b} = b \text{ and } \text{supp}(\tilde{p}(\hat{t})) \leq b, \\ a(\hat{t})v - q(\hat{t})c - \mathbb{E}[\tilde{p}(\hat{t})] & \text{if } \hat{b} \neq b \text{ and } \text{supp}(\tilde{p}(\hat{t})) \leq b, \\ -\infty & \text{if } \text{supp}(\tilde{p}(\hat{t})) > b. \end{cases}$$

In other words, an agent suffers an unbounded dis-utility if his budget constraint is violated with a positive probability. The IC constraints become

$$u(t) \geq u(\hat{t}, t), \ \forall t \in T, \hat{t} \in \left\{ \hat{t} \in T \,\middle|\, \text{supp}(\tilde{p}(\hat{t})) \leq b \right\}, \tag{IC}$$

The principal's objective function and all the other constraints remain intact.

By a similar argument to that used in Pai and Vohra (2014a), for any feasible mechanism $(a, \tilde{p}, q)$, one can construct another feasible mechanism $(a, \tilde{p}, q)$ by setting

$$\hat{p}(t) = \begin{cases} \mathbb{E}[\tilde{p}(t)] - \epsilon & \text{with propability } \frac{b - \mathbb{E}[\tilde{p}(t)]}{b - \mathbb{E}[\tilde{p}(t)] + \epsilon}, \\ b & \text{with propability } \frac{\epsilon}{b - \mathbb{E}[\tilde{p}(t)] + \epsilon}, \end{cases}$$

for some $\epsilon > 0$ sufficiently small. Furthermore, both mechanisms have the same welfare. Observe that, under this construction, IC constraints corresponding to over reporting of budget are satisfied "for free". Given these observations, it is not hard to see that one can solve the principal's problem (allowing for random payment rules) by restricting attention to deterministic payment rules but relaxing IC constraints corresponding to the over reporting of budget. As I will show later, in the optimal mechanism of $\mathcal{P}$ no low-budget agent has any incentive to over report his budget. Hence, it is without loss of generality to focus on deterministic payment rules.

3.3. Common knowledge budgets

As a benchmark, I first analyze the case in which all agents' budget constraints are common knowledge. This case can be viewed as the situation in which the principal can inspect an agent's budget for free (i.e., $\rho = k/c = 0$).

Since budgets are common knowledge, the IC constraints hold as long as for each $b \in \boldsymbol{B}$, no agent has incentive to misreport his value:

$$a(v,b)v - p(v,b) \leq a(\hat{v},b)v - p(\hat{v},b), \ \forall v, \hat{v}. \tag{IC-v}$$

The principal's problem becomes

$$\max_{a,p,q} \mathbb{E}_t \left[a(t)v\right], \tag{$\mathcal{P}_{CB}$}$$

subject to (IR), (IC-v), (BC), (S) and

$$\mathbb{E}_t[p(t)] \geq 0, \forall t \in \boldsymbol{T}. \tag{BB}$$

By the standard argument, (IC-v) holds if and only if for all $b \in \boldsymbol{B}$, $a(v,b)$ is non-decreasing in v and $p(v,b) = a(v,b)v - \int_{\underline{v}}^{v} a(\nu,b)\mathrm{d}\nu - u(\underline{v},b)$. Since $a(v,b)$ is non-decreasing in v, the payment $p(v,b)$ is also non-decreasing in v. Hence, (BC) holds if and only if $p(\overline{v},b) \leq b$ for all b.

Let χ denote the characteristic function. The following theorem characterizes the optimal mechanism.

Theorem 7 *Suppose Assumption 3 holds, and budgets are common knowledge. There exist* $v_1^*(0)$, $v_2^*(0)$, $u_1^*(0)$ *and* $u_2^*(0)$ *such that an optimal mechanism of* $\mathcal{P}_{CB}$ *is given by*

$$\begin{aligned} a(v,b_1) &= \chi_{\{v \geq v_1^*(0)\}} a^*(0), \quad & p(v,b_1) &= \chi_{\{v \geq v_1^*(0)\}}(u_1^*(0)+b_1) - u_1^*(0), \\ a(v,b_2) &= \chi_{\{v \geq v_2^*(0)\}} 1, & p(v,b_2) &= \chi_{\{v \geq v_2^*\}} v_2^*(0), \end{aligned}$$

where $a^*(0) = \left[u_1^*(0)+b_1\right]/v_1^*(0)$, $b_1 < v_1^*(0) \leq v_2^*(0) < \overline{v}$ *and* $0 = u_2^*(0) < u_1^*(0) \leq v_1^*(0) - b_1$.

In notations $a^*(0)$, $v_i^*(0)$ and $u_i^*(0)$ $(i = 1, 2)$, subscript i indicates the corresponding budget b_i and argument 0 indicates that this can be viewed as an optimal mechanism when $\rho = 0$.

As expected, when budgets are common knowledge, the two budget group can be treated separately. Only low-budget agents receive positive cash subsidies aiming to relax their budget constraints: $u(\underline{v}, b_1) = u_1^*(0) > 0 = u_2^*(0) = u(\underline{v}, b_2)$. There are two cutoffs: $v_1^*(0) \le v_2^*(0)$. All high-budget agents whose valuations are above $v_2^*(0)$ receive the good with probability one. This allocation can be implemented by posting a price $v_2^*(0)$ for high-budget agents, which is the efficient mechanism when agents are not financially constrained. All low-budget agents whose valuations are above $v_1^*(0)$ receive the good with positive probability but are possibly rationed. The intuition for rationing is familiar from the literature. Increasing allocations to low value agents reduces the payment of high value agents and therefore "relaxes" their budget constraints.

Clearly, a high-budget agent whose value is below $v_1^*(0)$ has a strict incentive to misreport as a low-budget agent since $u(\underline{v}, b_1) > 0 = u(\underline{v}, b_2)$. A high-budget agent whose value is slightly above $v_1^*(0)$ also has strict incentives to misreport as a low-budget agent:

$$\frac{v\left(u(\underline{v}, b_1) + b_1\right)}{v_1^*(0)} - b_1 > \frac{\left(v - v_1^*(0)\right) b_1}{v_1^*(0)} \ge \max\left\{v - v_2^*(0), 0\right\}.$$

The last inequality holds for $v > v_1^*(0)$ sufficiently close to $v_1^*(0)$. As it will become clear in Section 3.4.1, when budgets are agents' private information and the principal does not inspect, to discourage agents from under reporting their budgets, it must be that $u(\underline{v}, b_1) = u(\underline{v}, b_2)$ and a high-budget agent must receive the good with a probability no less than that of a low-budget agent who has the same valuation.

3.4. Privately known budgets

In this section, I analyze the case in which an agent's budget is his private information. In this case, IC constraints can be separated into two categories:

$$\text{Misreport value: } a(v,b)v - p(v,b) \geq a(\hat{v},b)v - p(\hat{v},b),\ \forall v,\hat{v},b, \tag{IC-v}$$

$$\text{Misreport both: } a(v,b)v - p(v,b) \geq \chi_{\{p(\hat{v},\hat{b})\leq b\}} \left(a(\hat{v},\hat{b})v - q(\hat{v},\hat{b})c - p(\hat{v},\hat{b})\right),\ \forall v,\hat{v},b,\hat{b}. \tag{3.1}$$

As I stated in the previous section, (IC-v) holds if and only if for all $b \in \boldsymbol{B}$, $a(v,b)$ is non-decreasing in v and $p(v,b) = va(v,b) - \int_{\underline{v}}^{v} a(\nu,b)\mathrm{d}\nu - u(\underline{v},b)$. The difficulty arises from (3.1). In what follows, I first consider a relaxed problem by replacing (3.1) with the following constraint:

$$a(v,b_2)v - p(v,b_2) \geq a(\hat{v},b_1)v - q(\hat{v},b_1)c - p(\hat{v},b_1),\ \forall v,\hat{v}. \tag{IC-b}$$

This relaxation formalizes the intuition that the principal's main concern is to prevent high-budget agents from falsely claiming to be low-budget agents. Later, I verify that an optimal mechanism of the relaxed problem automatically satisfies IC constraints corresponding to over-reporting of budgets. In other words, it also solves the original problem.

To summarize, the principal's relaxed problem is

$$\max_{a,p,a} \mathbb{E}_t[a(t)v - q(t)k], \tag{$\mathcal{P}'$}$$

subject to (IR), (IC-v), (IC-b), (BC), (BB) and (S).

3.4.1. No verification

In this section, I consider the case in which the principal does not inspect agents, i.e., $q \equiv \mathbf{0}$. In this case, as will become clear in the discussion below, it is sufficient to consider two *one-dimensional deviations*, which greatly simplifies the analysis. Although some of the results may be familiar, it highlights the differences in my approach. Denote the principal's problem in this case by $\mathcal{P}_{NI}$ and the corresponding relaxed problem by $\mathcal{P}'_{NI}$. As will become clear in Section 3.6, if the inspection cost, k, is sufficiently high relative to the punishment, c, then it is optimal for the principal not to use inspection. In particular, this is the case when the principal's inspection cost is infinity (i.e., $\rho = k/c = \infty$).

Observe first that in this case (IC-b) holds if and only if (IC-v) holds and

$$a(v, b_2)v - p(v, b_2) \geq a(v, b_1)v - p(v, b_1), \ \forall v. \tag{3.2}$$

To see this, note that if (3.2) holds, then

$$\begin{aligned} a(v, b_2)v - p(v, b_2) &\geq a(v, b_1)v - p(v, b_1) \\ &\geq a(\hat{v}, b_1)v - p(\hat{v}, b_1), \end{aligned}$$

where the second inequality follows from (IC-v). Thus, it is sufficient to consider the two one-dimensional deviations: only misreport value and only misreport budget. The above inequality says that if a type (v, b_2) agent has no incentive to misreport (v, b_1), then he has no incentive to misreport $(\hat{v}, b_1)$. This argument is not true when there is verification because it is possible that types (v, b_1) and $(\hat{v}, b_1)$ are inspected with different probabilities. Instead, one must identify for each type $(\hat{v}, b_1)$ the high-budget type who benefits most from misreporting $(\hat{v}, b_1)$ in the absence of inspection, which determines the set of binding (IC-b) constraints.

Using the envelope condition, (3.2) can be rewritten as

$$u(\underline{v}, b_2) + \int_{\underline{v}}^{v} a(\nu, b_2)\mathrm{d}\nu \geq u(\underline{v}, b_1) + \int_{\underline{v}}^{v} a(\nu, b_1)\mathrm{d}\nu, \ \forall v. \tag{3.3}$$

If $v = \underline{v}$, then (3.3) implies that $u(\underline{v}, b_2) \geq u(\underline{v}, b_1)$. If $u(\underline{v}, b_2) > u(\underline{v}, b_1)$, then one can construct another feasible mechanism by reducing cash subsidies to high-budget agents while increasing their probabilities of receiving the goods, which generates the same welfare. Hence, it is without loss of generality to assume that $u(\underline{v}, b_1) = u(\underline{v}, b_2)$. This result is summarized in Lemma 2, and a complete proof can be found in the appendix.[13]

Lemma 2 *Suppose Assumption 3 holds, and the principal does not inspect agents. In an optimal mechanism of $\mathcal{P}'_{NI}$, it is without loss of generality to assume that $u(\underline{v}, b_1) = u(\underline{v}, b_2)$.*

One implication of Lemma 2 is that in an optimal mechanism agents receive positive cash subsidies regardless of their budgets. This result contrasts the case of common knowledge budgets in which only low-budget agents receive positive cash subsidies.

Next, I show that, for any given v, an optimal mechanism on average allocates weakly more resources to high-budget agents whose valuations are below v than to low-budget agents whose valuations are below v.

Lemma 3 *Suppose Assumptions 2 and 3 hold, and the principal does not inspect agents. In an optimal mechanism of $\mathcal{P}'_{NI}$, the allocation rule satisfies*

$$\int_{\underline{v}}^{v} a(\nu, b_2) f(\nu) \mathrm{d}\nu \geq \int_{\underline{v}}^{v} a(\nu, b_1) f(\nu) \mathrm{d}\nu, \ \forall v. \tag{3.4}$$

Given Lemma 2, (3.4) follows immediately from (3.3) if v is uniformly distributed. Lemma 3 shows that the result holds more generally for any distribution with non-increasing density. Using Lemmas 2 and 3, one can prove the following theorem, which characterizes the optimal direct mechanism.

Theorem 8 *Suppose Assumptions 2 and 3 hold, and the principal does not inspect agents. There exist $v_1^*(\infty)$, $v_2^*(\infty)$, $v_2^{**}(\infty)$, $u_1^*(\infty)$ and $u_2^*(\infty)$ such that an optimal mechanism of $\mathcal{P}_{NI}$ with no inspection satisfies*

$$\begin{aligned}
&a(v, b_1) = \chi_{\{v \geq v^*(\infty)\}} a^*(\infty), \ p(v, b_1) = \chi_{\{v \geq v_1^*(\infty)\}} (u_1^*(\infty) + b_1) - u_1^*(\infty), \\
&a(v, b_2) = \chi_{\{v \geq v_2^*(\infty)\}} a^*(\infty) + \chi_{\{v \geq v_2^{**}(\infty)\}} \left(1 - a^*(\infty)\right), \\
&p(v, b_2) = \chi_{\{v \geq v_2^*(\infty)\}} (u_2^*(\infty) + b_1) + \chi_{\{v \geq v_2^{**}(\infty)\}} \left(1 - a^*(\infty)\right) v_2^{**}(\infty) - u_2^*(\infty),
\end{aligned}$$

[13] It is immediate that $u(\underline{v}, b_1) = u(\underline{v}, b_2)$ if one also requires that a low-budget agent has no incentive to misreport as a high-budget agent.

where

$$a^*(\infty) = \frac{u_1^*(\infty) + b_1}{v_1^*(\infty)},$$

$b_1 < v_1^*(\infty) = v_2^*(\infty) \leq v_2^{**}(\infty) \leq \overline{v}$ *and* $0 < u_1^*(\infty) = u_2^*(\infty) \leq v_1^*(\infty) - b_1$.

In notations $a^*(\infty)$, $v_i^*(\infty)$, $v_2^{**}(\infty)$ and $u_i^*(\infty)$ $(i = 1, 2)$, subscript i indicates the corresponding budget b_i and argument ∞ indicates that this can be viewed as an optimal mechanism when $\rho = \infty$.

Not surprisingly the optimal allocation rule obtained here shares similar features with the one found in Pai and Vohra (2014a).There are three cutoffs: $v_1^*(\infty) = v_2^*(\infty) < v_2^{**}(\infty)$. All high-budget agents whose valuations are above $v_2^{**}(\infty)$ receive the good with probability one. All low-budget agents whose valuations are above $v_1^*(\infty)$ receive the good with positive probability but may be rationed. In addition, high-budget agents whose valuations are in $[v_2^*(\infty), v_2^{**}(\infty)]$ are pooled with low-budget agents whose valuations are at least $v_1^*(\infty)(= v_2^*(\infty))$. To understand this pooling, consider two agents with the same valuation v, but different budgets $b_2 > b_1$. Then (IC-b) implies that as long as agent (v, b_2)'s payment is less than b_1, he must receive the good with the same probability as (v, b_1) does.

The proof of Theorem 8 follows a weight-shifting argument similar to that of Lemma 1 in Richter (2015). Consider a feasible mechanism $(a, p, \mathbf{0})$ whose allocation rule is indicated by the two thick dotted curves in Figure 2. One can construct another feasible mechanism $(a^*, p^*, \mathbf{0})$, whose allocation rule is indicated by the thick solid lines, in the following way. Find a v_1^* and shift the allocation mass of low-budget agents from the region to the left of v_1^* to the region to the right of v_1^*. The choice of v_1^* is uniquely determined so that the supply to low-budget agents remains unchanged. Let $\hat{v}$ denote the minimum valuation of high-budget agents who receive the good with a probability of at least $a(\overline{v}, b_1) = a^*(\overline{v}, b_1)$. Find v_2^* and v_2^{**} such that $v_2^* \leq \hat{v} \leq v_2^{**}$. Shift the allocation mass of high-budget agents from the region to the left of v_2^* to $[v_2^*, \hat{v}]$ and from $[\hat{v}, v_2^{**}]$ to the region to the right of v_2^{**}. The choice of v_2^* (and v_2^{**}, respectively) is uniquely determined so that the supply to high-budget agents whose valuations are in $[\underline{v}, \hat{v}]$ (and $[\hat{v}, \overline{v}]$, respectively) remains unchanged. Finally, define the new payment rule using the envelope condition. If f is "regular", i.e., satisfies Assumptions 2

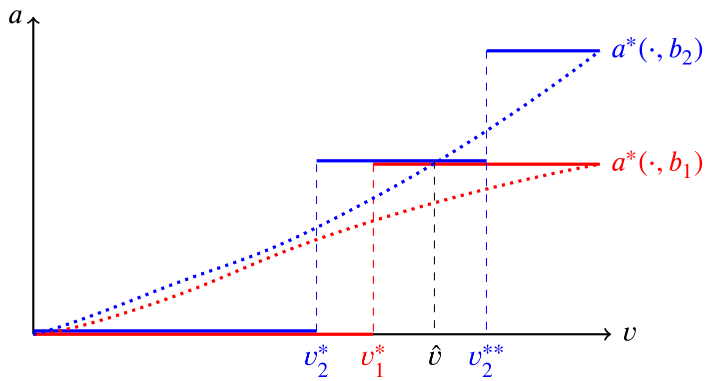


Figure 2: Proof sketch of Theorem 8

and 3, then the new mechanism improves welfare and revenue while remaining affordable. Lemma 3 guarantees that $v_2^* \leq v_1^*$. Thus, no high-budget agent has incentive to misreport his budget. It is easy to see that one can further improve welfare by increasing v_2^* and reducing v_1^*. Hence, in an optimal mechanism $v_1^*(\infty) = v_2^*(\infty)$.

3.4.2. The general case

I now turn to the general problem of the principal. Using the envelope condition, (IC-b) becomes the following: For all v and $\hat{v}$,

$$u(\underline{v}, b_2) + \int_{\underline{v}}^{v} a(\nu, b_2)\mathrm{d}\nu \geq u(\underline{v}, b_1) + a(\hat{v}, b_1)(v - \hat{v}) - q(\hat{v}, b_1)c + \int_{\underline{v}}^{\hat{v}} a(\nu, b_1)\mathrm{d}\nu. \quad \text{(IC-b)}$$

First, for each $\hat{v}$, I identify the type of high-budget agents whose gains from falsely claiming to be a type $(\hat{v}, b_1)$ agent are the largest. (IC-b) holds if and only if for each $\hat{v} \in V$, $q(\hat{v}, b_1)c \geq \sup_v \Delta(v, \hat{v})$, where

$$\Delta(v, \hat{v}) \equiv u(\underline{v}, b_1) - u(\underline{v}, b_2) - \int_{\underline{v}}^{v} a(\nu, b_2)\mathrm{d}\nu + a(\hat{v}, b_1)(v - \hat{v}) + \int_{\underline{v}}^{\hat{v}} a(\nu, b_1)\mathrm{d}\nu.$$

Since $\partial \Delta(v, \hat{v})/\partial v = -a(v, b_2) + a(\hat{v}, b_1)$ is non-increasing in v, $\Delta(v, \hat{v})$ is concave in v and achieves its maximum at $v = v^d(\hat{v})$, where

$$v^d(\hat{v}) \equiv \inf \left\{ v | a(v, b_2) \geq a(\hat{v}, b_1) \right\}. \quad (3.5)$$

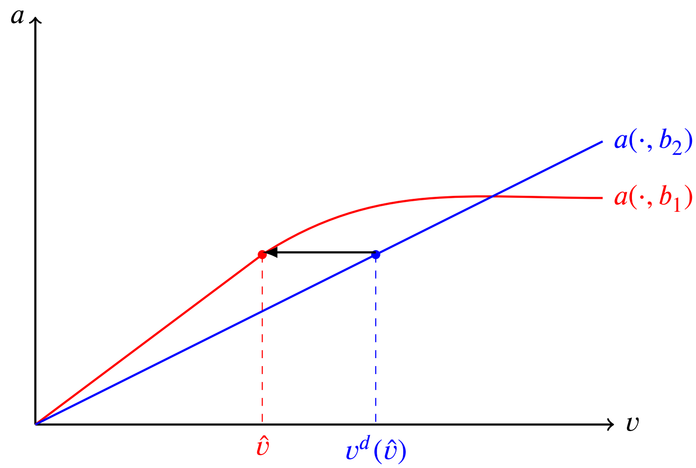


Figure 3: The set of binding (IC-b) constraints

Suppose the allocation rules for both budget types are continuous in value v. Then the high-budget agents who benefit most from falsely claiming to be $(\hat{v}, b_1)$ are those who get the goods with the same probability as type $(\hat{v}, b_1)$ agents do. This point is illustrated by Figure 3, which plots an allocation rule for high-budget agents, $a(\cdot, b_2)$, and an allocation rule for low-budget agents, $a(\cdot, b_1)$, as a function of their valuations v.

Since the principal's objective function is strictly decreasing in q, the optimal inspection rule satisfies

$$q(\hat{v}, b_1) = \frac{1}{c} \max \left\{0, \Delta(v^d(\hat{v}))\right\}. \tag{3.6}$$

Note that $v^d(\cdot)$ is defined using the allocation rule. As a result, one cannot anticipate, a priori, which (IC-b) constraint binds. Furthermore, (IC-b) constraints are frequently binding not only among local types. These difficulties are inherent in all multidimensional problems, and as a result the existing approaches in the mechanism literature do not apply to this problem.[14]

In order to keep track of the binding (IC-b) constraints, we solve the principal's problem by approximating the allocation rule using step functions. Fix $M \geq 2$. Let $\underline{v} = v_1^0 < v_1^1 < \cdots < v_1^M = \overline{v}$ and $0 = a^0 \leq a^1 < a^2 < \cdots < a^M \leq a^{M+1} = 1$. Suppose the allocation rule for type b_1 agents takes M distinct values: $a(v, b_1) = a^m$ if $v \in (v_1^{m-1}, v_1^m)$ for $m = 1, \ldots, M$. The next lemma shows that the

[14]See Rochet and Stole (2003) for a survey on multidimensional mechanism design problem.

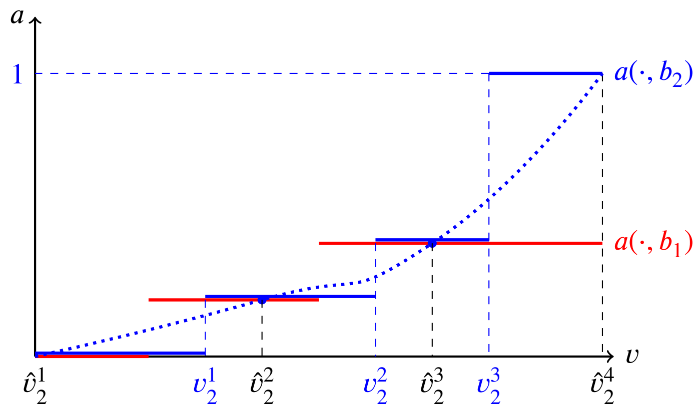


Figure 4: Proof Sketch of Lemma 4

optimal allocation rule for type b_2 agents can take at most $M + 2$ distinct values: $a^0, a^1, \ldots, a^{M+1}$.

Lemma 4 *Suppose Assumptions 2 and 3 hold. Suppose* $a(v, b_1) = a^m$ *if* $v \in (v_1^{m-1}, v_1^m)$ *for* $m = 1, \ldots, M$. *Then there exists* $\underline{v} \leq v_2^0 \leq v_2^1 \leq \cdots \leq v_2^M \leq \overline{v}$ *such that an optimal allocation rule for* b_2 *satisfies* $a(v, b_2) = a^m$ *if* $v \in (v_2^{m-1}, v_2^m)$ *for* $m = 1, \ldots, M$, $a(v, b_2) = 0$ *if* $v < v_2^0$ *and* $a(v, b_2) = 1$ *if* $v > v_2^M$.

The proof of Lemma 4 is similar to that of Theorem 8 and illustrated by Figure 4, where the allocation rule for low-budget agents (the solid red line) takes three distinctive values: $a^1 < a^2 < a^3$. Consider a feasible allocation rule for high-budget agents indicated by the dotted blue curve. Suppose there exist a payment rule and an inspection to be used in conjunction with the allocation rule so that the resulting mechanism is feasible. For ease of exposition, suppose $a(\cdot, b_2)$ is continuous and let $\hat{v}_2^m$ be such that $a(\hat{v}_2^m, b_2) = a^m$ for $m = 1, 2, 3$. For each $m = 1, 2, 3$, find v_2^m and move the allocation mass of high-budget agents from $[\hat{v}_2^m, v_2^m]$ to $[v_2^m, \hat{v}_2^{m+1}]$, where $\hat{v}_2^4 = \overline{v}$. The choice of v_2^m is uniquely determined so that the supply to high-budget agents whose value is in $[\hat{v}_2^m, \hat{v}_2^{m+1}]$ remains unchanged. Redefine the payment rule using the envelope condition and let the inspection rule remain the same. One can verify that the new mechanism is feasible and clearly improves welfare.

We say an allocation rule a is an *M-step allocation rule* if there exist $\underline{v} = v_1^0 < v_1^1 < \cdots < v_1^M = \overline{v}$, $\underline{v} \leq v_2^0 \leq v_2^1 \leq \cdots \leq v_2^M \leq \overline{v}$ and $0 = a^0 \leq a^1 < a^2 < \cdots < a^M \leq a^{M+1} = 1$ for some $M \geq 2$ such that $a(v, b_1) = a^m$ if $v \in (v_1^{m-1}, v_1^m)$ for $m = 1, \ldots, M$ and $a(v, b_2) = a^m$ if $v \in (v_2^{m-1}, v_2^m)$

for $m = 0, 1, \dots, M+1$. Lemma 4 shows that it is without loss of generality to focus on M-step-allocation rules among all step allocation rules.

Consider a mechanism using a M-step allocation rule. It is easy to see that for $v \in (v_1^{m-1}, v_1^m)$, the type b_2 agents who benefit most from falsely claiming to be type (v, b_1) have valuations $v^d(v) = v_2^{m-1}$. Hence, we can keep track of the binding (IC-b) constraints by keeping track of the jump points of the allocation rule. In this case, the optimal inspection rule satisfies $q(v, b_1) = q^m$ for all $v \in (v_1^{m-1}, v_1^m)$ and

$$q^m = \frac{1}{c} \max \left\{ 0, u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m} (a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) \right\} \tag{3.7}$$

for $m = 1, \dots, M$.

Consider the principal's problem ($\mathcal{P}'$) with two modifications:

$$\max_{a,p,q} \mathbb{E}_t[a(t)v - q(t)k], \tag{$\mathcal{P}'(M,d)$}$$

subject to (IR), (IC-v), (IC-b), (BC), (S),

$$a \text{ is a } M'\text{-step allocation rule for some } M' \leq M,$$

$$\mathbb{E}[p(t) - q(t)k] \geq -d. \tag{BB-d}$$

The second modification is to relax the government's budget balance constraint by $d \geq 0$. As it will become clear later, any feasible mechanism of $\mathcal{P}'$ can be approximated arbitrarily well by a feasible mechanism of $\mathcal{P}'(M, d)$ for M sufficiently large and d sufficiently small.

Next, I show that in an optimal mechanism of $\mathcal{P}'(M, d)$, in the absence of verification, either no high-budget agent has incentives to misreport as low budget, or all high-budget agents weakly prefer to misreport as low budget.

Lemma 5 *Suppose Assumptions 2 and 3 hold. An optimal mechanism of $\mathcal{P}'(M, d)$ satisfies one of the following two conditions:*

(C1) For all $m = 1, \dots, M$,

$$u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) \geq 0. \tag{3.8}$$

(C2) For all $m = 1, \dots, M$,

$$u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) \leq 0. \tag{3.9}$$

The basic intuition underlying Lemma 5 is as follows: As long as a mechanism satisfies neither (C1) nor (C2), one can strictly improve welfare by adjusting the allocation rule in regions in which high-budget agents find it strictly optimal to report their budgets truthfully. I provide only a proof sketch of Lemma 5 here. The full proof can be found in the appendix.

Proof Sketch. The proof is by contradiction. Let (a, p, q) be a feasible mechanism, where a is a M-step allocation rule. Suppose (a, p, q) satisfies neither (C1) nor (C2). I show that one can construct another feasible mechanism (a^*, p^*, q^*), which strictly improves welfare and satisfies one of the two conditions. Furthermore, a^* is a M'-step function for some $M' \leq M$. I break the proof into two steps.

Step 1. I show that it is without loss of generality to assume that (3.8) holds for $m = 1$. Suppose, on the contrary, that $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1 v_2^0 < 0$. Then there exists $m > 1$ such that $v_2^{m'-1} - v_1^{m'-1} \leq 0$ for all $m' < m$ and $v_2^{m-1} - v_1^{m-1} > 0$. One can construct another feasible mechanism by redirecting cash subsidies from high-budget agents to low-budget agents, and shifting the allocation mass from low-budget agents in $[v_1^{m-1}, \tilde{v}_1^{m-1}]$ to high-budget agents in $[\tilde{v}_2^{m-1}, v_2^{m-1}]$ for some $v_1^{m-1} \leq \tilde{v}_1^{m-1} \leq \tilde{v}_2^{m-1} \leq v_2^{m-1}$.

Step 2. Suppose $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1 v_2^0 \geq 0$. There exists $m > 1$ such that (3.8) holds for all $m' < m$ and

$$u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) < 0.$$

It must be the case that $v_2^{m-1} < v_1^{m-1}$. For ease of exposition, assume that $v_2^m > v_2^{m-1}$.[15] One can construct another feasible mechanism by either shifting the allocation mass from high-budget agents in $[v_2^{m-1}, \hat{v}]$ to high-budget agents in $[\hat{v}, v_2^m]$ for some $v_2^{m-1} < \hat{v} < v_2^m$, or shifting the allocation mass from high-budget agents in $[v_2^{m-1}, \tilde{v}_2^{m-1}]$ to low-budget agents in $[\tilde{v}_1^{m-1}, v_1^{m-1}]$ for some $v_2^{m-1} \leq \tilde{v}_2^{m-1} \leq \tilde{v}_1^{m-1} \leq v_1^{m-1}$. ■

If (C2) holds, then the optimal inspection rule is $q \equiv \mathbf{0}$. The optimal mechanism of $\mathcal{P}'$ in this case, which is characterized in Section 3.4.1, is a feasible mechanism of $\mathcal{P}'(M, d)$ and satisfies (C1) with equality. Thus, I can conclude that an optimal mechanism of $\mathcal{P}'(M, d)$ satisfies (C1).

Corollary 5 *Suppose Assumptions 2 and 3 hold. An optimal mechanism of $\mathcal{P}'(M, d)$ satisfies (C1).*

Hence, an optimal inspection rule satisfies $q(v, b_1) = q^m$ for all $v \in (v_1^{m-1}, v_1^m)$, where

$$q^m = \frac{1}{c}\left[u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1})\right] \tag{3.10}$$

for $m = 1, \ldots, M$. Now the principal's problem $\mathcal{P}'(M, d)$ can be written as follows, where the Greek letters in parentheses denote the corresponding Lagrangian multipliers.

$$\max_{\substack{u(\underline{v},b_1),u(\underline{v},b_2),\\ \{a^m\}_{m=1}^M,\{v_1^m\}_{m=1}^{M-1},\{v_2^m\}_{m=0}^M}} \quad \pi \sum_{m=1}^{M+1} \int_{v_2^{m-1}}^{v_2^m} a^m v f(v)\mathrm{d}v + (1-\pi)\sum_{m=1}^{M} \int_{v_1^{m-1}}^{v_1^m} a^m v f(v)\mathrm{d}v$$
$$-(1-\pi)\frac{k}{c}\sum_{m=1}^{M}\int_{v_1^{m-1}}^{v_1^m}\left[u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1})\right] f(v)\mathrm{d}v,$$

[15] In the appendix, I break the proof in three steps. I consider the case in which $v_2^{m-1} < v_2^M$ in Step 2 and the case $v_2^{m-1} = v_2^M$ in Step 3.

subject to

$$\pi \sum_{m=1}^{M+1} a^m[F(v_2^m) - F(v_2^{m-1})] + (1-\pi) \sum_{m=1}^{M} a^m[F(v_1^m) - F(v_1^{m-1})] \leq S, \tag{β}$$

$$a^M v_1^{M-1} - \sum_{j=1}^{M-1} a^j(v_1^j - v_1^{j-1}) - u(\underline{v}, b_1) \leq b_1, \tag{η}$$

$$\begin{aligned} &-(1-\pi)u(\underline{v}, b_1) + (1-\pi)\sum_{m=1}^{M}\int_{v_1^{m-1}}^{v_1^m} a^m\left[v - \frac{1-F(v)}{f(v)}\right] f(v)\mathrm{d}v \\ &-(1-\pi)\frac{k}{c}\sum_{m=1}^{M}\int_{v_1^{m-1}}^{v_1^m}\left[u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1})\right] f(v)\mathrm{d}v \\ &-\pi u(\underline{v}, b_2) + \pi\sum_{m=1}^{M+1}\int_{v_2^{m-1}}^{v_2^m} a^m\left[v - \frac{1-F(v)}{f(v)}\right] f(v)\mathrm{d}v \geq -d, \end{aligned} \tag{λ}$$

$$u(\underline{v}, b_1) \geq 0, u(\underline{v}, b_2) \geq 0, \tag{ξ_1, ξ_2}$$

$$u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) \geq 0,\ m = 1, \ldots, M, \tag{μ^m}$$

$$0 = a^0 \leq a^1 \leq a^2 \leq \cdots \leq a^M \leq a^{M+1} = 1, \tag{$\alpha^1, \ldots, \alpha^{M+1}$}$$

$$\underline{v} = v_1^0 \leq v_1^1 \leq \cdots \leq v_1^M = \overline{v}, \tag{$\gamma_1^1, \ldots, \gamma_1^M$}$$

$$\underline{v} \leq v_2^0 \leq v_2^1 \leq \cdots \leq v_2^M \leq \overline{v}. \tag{$\gamma_2^0, \ldots, \gamma_2^{M+1}$}$$

To solve this problem, I first show that in an optimal mechanism of $\mathcal{P}'(M, d)$, the inspection probability is non-decreasing in a low-budget agent's reported value:

Lemma 6 *Suppose Assumptions 2 and 3 hold. In an optimal mechanism of $\mathcal{P}'(M, d)$, $v_2^1 - v_1^1 \geq 0$. Suppose in addition that $V(M, d) > V(M-1, d)$ for $M \geq 3$, then*

$$v_2^{M-1} - v_1^{M-1} > \cdots > v_2^1 - v_1^1 \geq 0.$$

As a result, the inspection probability in an optimal mechanism of $\mathcal{P}'(M, d)$ is non-decreasing in reported value, i.e., $q^M \geq \cdots \geq q^1 \geq 0$.

To understand the intuition behind the monotonicity of inspection probability, consider a low-budget agent and a high-budget agent both receiving the good with probability a^m. Let p_1^m and p_2^m denote their payments respectively. The difference in their payments, to which the inspection probability is proportional, is

$$p_2^m - p_1^m = u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}).$$

Clearly, this difference is non-decreasing in m since $v_2^{m-1} - v_1^{m-1} \geq 0$. Suppose, on the contrary, that $q^{m-1} > q^m$. Then the principal can shift allocation from low-budget agents in $[v_1^{m-2}, v_1^{m-1}]$ to low-budget agents in $[v_1^{m-1}, v_1^m]$, which clearly improves allocation efficiency and revenue. This shift also strictly reduces inspection cost because more low-budget agents are inspected with probability q^m rather than q^{m-1} and $q^{m-1} > q^m$.

The inequality constraints corresponding to μ^m's in $\mathcal{P}'(M, d)$ are non-negativity constraints on inspection probabilities. As shown in Lemma 6, in an optimal mechanism of $\mathcal{P}'(M, d)$, the inspection probability is non-decreasing in a low-budget agent's reported value. As a result, it is sufficient to consider the inequality constraint corresponding to μ^1:

$$u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) \geq 0.$$

Note that for fixed jump discontinuity points v_i^m's, the principal's problem $\mathcal{P}'(M, d)$ is linear in $u(\underline{v}, b_1)$, $u(\underline{v}, b_2)$ and a^m's. Hence, an optimal solution can be obtained at an extreme point of the feasible region. The monotonicity of inspection probability implies that in addition to the monotonicity constraints on a^m's there are only finitely many other constraints binding. As a result, for an M sufficiently large, there are finitely many distinct a^m's in an optimal mechanism. More formally, let $V(M, d)$ denote the value of $\mathcal{P}'(M, d)$. Then $V(M, d) = V(M-1, d)$ for M sufficiently large. This result still holds if I replace (BC) with a per-unit price constraint, as shown in Section 3.7.1. If I impose only (BC), then I can further prove that in an optimal mechanism of $\mathcal{P}'(M, d)$ the allocation rule is a 2-step allocation rule, i.e., $V(M, d) = V(M-1, d)$ for $M \geq 3$.

Lemma 7 *Suppose Assumptions 2 and 3 hold. Then $V(M, d) = V(2, d)$ for all $M \geq 2$ and $d \geq 0$.*

Furthermore, for all $M \geq 2$, in an optimal mechanism of $\mathcal{P}'(M,d)$ the allocation rule is a 2-step allocation rule.

Proof Sketch. I provide a proof sketch of Lemma 7. Assume, for ease of exposition, that

$$u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) > 0, \qquad (\mu^1)$$

$$\underline{v} = v_1^0 < v_1^1 < \cdots < v_1^M = \overline{v}, \qquad (\gamma_1^1, \ldots, \gamma_1^M)$$

$$0 \leq v_2^0 < v_2^1 < \cdots < v_2^M < \overline{v}. \qquad (\gamma_2^0, \ldots, \gamma_2^{M+1})$$

Then $\mu^1 = \cdots = \mu^M = 0$, $\gamma_1^1 = \cdots = \gamma_1^M = 0$ and $\gamma_2^1 = \cdots = \gamma_2^{M+1} = 0$. The first-order conditions for v_1^m and v_2^m ($m = 1, \ldots, M-1$) are

$$(1-\pi)\left[(\beta - v_1^m - \lambda\varphi(v_1^m))f(v_1^m) + (1+\lambda)\rho[1-F(v_1^m)] + (1+\lambda)\rho(v_2^m - v_1^m)f(v_1^m)\right] - \eta = 0,$$

$$\pi(\beta - v_2^m - \lambda\varphi(v_2^m))f(v_2^m) - (1-\pi)(1+\lambda)\rho[1-F(v_1^m)] = 0,$$

where $\varphi(v) := v - [1-F(v)]/f(v)$ denotes the virtual value function. I show in the appendix that If f is "regular", which is to say that it satisfies Assumptions 2 and 3, then the above system of equations has at most one solution. This result is illustrated by Example 2. Hence, I can conclude that $V(M,d) = V(2,d)$. ■

Example 2 *Let v be uniformly distributed on $[0,1]$ and $\rho < \frac{\pi+\sqrt{\pi}}{1-\pi}$. Then the first-order conditions for v_1^m and v_2^m $(m = 1, \ldots, M-1)$*

$$(1-\pi)\left[\beta + \lambda + (1+\lambda)\rho - (1+2\lambda+2(1+\lambda)\rho)v_1^m + (1+\lambda)\rho v_2^m\right] - \eta = 0, \qquad (3.11)$$

$$\pi(\beta+\lambda) - (1-\pi)(1+\lambda)\rho + (1-\pi)(1+\lambda)\rho v_1^m - \pi(1+2\lambda)v_2^m = 0. \qquad (3.12)$$

Given β, η and λ, (3.11) *and* (3.12) *define v_2^m as functions of v_1^m, denoted by g_1 and g_2, respectively. Then*

$$g_1'(v_1^m) = 2 + \frac{1+2\lambda}{\rho(1+\lambda)} > \frac{(1-\pi)(1+\lambda)\rho}{\pi(1+2\lambda)} = g_2'(v_1^m).$$

This inequality implies that g_1 *can cross* g_2 *at most once from below. Hence,* (3.11) *and* (3.12) *have at most one solution.*

The main result of this section is Theorem 9, which characterizes an optimal mechanism of the original problem $\mathcal{P}$. In particular, I show that an optimal mechanism of $\mathcal{P}'(2,0)$ is also an optimal mechanism of $\mathcal{P}$. In other words, in an optimal mechanism of $\mathcal{P}$, the allocation rule is a 2-step allocation rule.

Let V denote the value of $\mathcal{P}'$. We prove Theorem 9 by first showing that for any $d > 0$ there exists $\overline{M}(d) > 0$ such that for all $M > \overline{M}(d)$

$$V - V(M,d) \leq (1-\pi)(1+\rho)\frac{\mathbb{E}[v]}{M}.$$

The proof is by construction. Fix $d > 0$ and an integer $M > \overline{M}(d)$. I can construct a feasible mechanism of $\mathcal{P}'(M,d)$ that possibly violates (BB) by at most d and generates welfare which is at least $V - (1-\pi)(1+\rho)\mathbb{E}[v]/M$. By Lemma 6, $V - V(M,d) = V - V(2,d) \leq (1-\pi)(1+k/c)\mathbb{E}[v]/M$ for all $d > 0$ and $M > \overline{M}(d)$. Fixing $d > 0$ and taking M to infinity yields $V \leq V(2,d)$ for all $d > 0$. By definition, $V \geq V(2,0)$. Hence, $V = V(2,0)$ by the continuity of $V(2,\cdot)$. Thus, an optimal mechanism of $\mathcal{P}'$ also solves $\mathcal{P}'$. It is easy to verify that an optimal solution to $\mathcal{P}'(2,0)$ satisfies (IC) constraints corresponding to agents over reporting their budgets and therefore solves $\mathcal{P}$. Finally, I show that $v_2^0 = \underline{v}$ and $a^2 = 0$ in an optimal mechanism. Let $a^*(\rho) = a^2$, $v_1^*(\rho) = v_1^1$, $v_2^*(\rho) = v_2^1$, $v_2^{**}(\rho) = v_2^2$, $u_1^*(\rho) = u(\underline{v}, b_1)$ and $u_2^*(\rho) = u(\underline{v}, b_2)$, then an optimal mechanism is characterized by the following Theorem 9.

Theorem 9 *Suppose Assumptions 2 and 3 hold. There exist* $a^*(\rho)$, $v_1^*(\rho)$, $v_2^*(\rho)$, $v_2^{**}(\rho)$, $u_1^*(\rho)$ *and*

$u_2^*(\rho)$ *such that an optimal mechanism of* $\mathcal{P}$ *is given by*

$$\begin{aligned}
a(v, b_1) &= \chi_{\{v \geq v_1^*(\rho)\}} a^*(\rho),\ \ p(v, b_1) = \chi_{\{v \geq v_1^*(\rho)\}} a^*(\rho) v_1^*(\rho) - u_1^*(\rho), \\
q(v, b_1) &= \frac{1}{c} \left[\chi_{\{v \geq v_1^*(\rho)\}} a^*(\rho) \left(v_2^*(\rho) - v_1^*(\rho) \right) + u_1^*(\rho) - u_2^*(\rho) \right], \\
a(v, b_2) &= \chi_{\{v \geq v_2^*(\rho)\}} a^*(\rho) + \chi_{\{v \geq v_2^{**}(\rho)\}} \left(1 - a^*(\rho) \right), \\
p(v, b_2) &= \chi_{\{v \geq v_2^*(\rho)\}} a^*(\rho) v_2^*(\rho) + \chi_{\{v \geq v_2^{**}(\rho)\}} \left(1 - a^*(\rho) \right) v_2^{**}(\rho) - u_2^*(\rho), \\
q(v, b_2) &= 0,
\end{aligned}$$

where $a^*(\rho) = \left[u_1^*(\rho) + b_1 \right] / v_1^*(\rho)$, $\underline{v} < v_1^*(\rho) \leq v_2^*(\rho) \leq v_2^{**}(\rho) \leq \overline{v}$, $0 < a^*(\rho) \leq 1$ *and* $u_1^*(\rho) \geq u_2^*(\rho)$.

In notations $a^*(\rho)$, $v_i^*(\rho)$, $v_2^{**}(\rho)$ and $u_i^*(\rho)$ ($i = 1, 2$), subscript i indicates the corresponding budget b_i and argument ρ indicates their dependence on ρ. In an optimal mechanism, low-budget agents receive more cash subsidies (as in the case of common knowledge budgets), but high-budget agents may also receive strictly positive cash subsidies (as in the case of private budgets without inspection). There are three cutoffs: $v_1^*(\rho) \leq v_2^*(\rho) \leq v_2^{**}(\rho)$. All high-budget agents whose valuations are above $v_2^{**}(\rho)$ receive the good with probability 1. All low-budget agents whose valuations are above $v_1^*(\rho)$ receive the good with positive probability but may be rationed. Similar to the case of private budgets without inspection, some high-budget agents (whose valuations are in $[v_2^*(\rho), v_2^{**}(\rho)]$) are pooled with low-budget agents. However, $v_1^*(\rho) \leq v_2^*(\rho)$. This difference between $v_1^*(\rho)$ and $v_2^*(\rho)$, together with budget dependent cash subsidies, creates an incentive for high-budget agents to under report their budgets. All agents who report low budgets are inspected with non-negative probability and those who receive the goods are more likely to be inspected.

I note here that if $\rho = 0$, then $u_2^*(0) = 0$ and $v_2^*(0) = v_2^{**}(0)$, which is the case in Theorem 7. If $\rho = \infty$, then $u_1^*(\infty) = u_2^*(\infty)$ and $v_1^*(\infty) = v_2^*(\infty)$, which is the case in Theorem 8. To simplify notation, in what follows, I suppress the dependence of u_1^*, u_2^*, v_1^*, v_2^*, v_2^{**} and a^* on ρ whenever it is clear.

Theorem 9 also greatly simplifies the analysis. Now the principal's problem can be reduced to:

$$\max_{\substack{u(\underline{v},b_1),u(\underline{v},b_2),\\ a^2,v_1^1,v_2^1,v_2^2}} \pi \left[\int_{v_2^1}^{v_2^2} a^2 v f(v)\mathrm{d}v + \int_{v_2^2}^{\overline{v}} v f(v)\mathrm{d}v\right] + (1-\pi)\int_{v_1^1}^{\overline{v}} a^2 v f(v)\mathrm{d}v$$

$$-(1-\pi)\frac{k}{c}\left[u(\underline{v},b_1)-u(\underline{v},b_2)\right]F(v_1^1) - (1-\pi)\frac{k}{c}\int_{v_1^1}^{\overline{v}}\left[u(\underline{v},b_1)-u(\underline{v},b_2)+a^2(v_2^1-v_1^1)\right]f(v)\mathrm{d}v,$$

subject to

$$\pi a^2\left[F(v_2^2)-F(v_2^1)\right] + \pi\left[1-F(v_2^2)\right] + (1-\pi)a^2[1-F(v_1^1)] \le S, \tag{β}$$

$$a^2 v_1^1 - u(\underline{v},b_1) \le b_1, \tag{η}$$

$$\begin{aligned}
&-(1-\pi)u(\underline{v},b_1) + (1-\pi)\int_{v_1^1}^{v_1^2} a^2\left[v - \frac{1-F(v)}{f(v)}\right]f(v)\mathrm{d}v \\
&-(1-\pi)\frac{k}{c}\left[u(\underline{v},b_1)-u(\underline{v},b_2)\right]F(v_1^1) - (1-\pi)\frac{k}{c}\int_{v_1^1}^{\overline{v}}\left[u(\underline{v},b_1)-u(\underline{v},b_2)+a^2(v_2^1-v_1^1)\right]f(v)\mathrm{d}v \\
&-\pi u(\underline{v},b_2) + \pi\int_{v_2^1}^{v_2^2} a^2\left[v-\frac{1-F(v)}{f(v)}\right]f(v)\mathrm{d}v + \pi\int_{v_2^2}^{\overline{v}}\left[v-\frac{1-F(v)}{f(v)}\right]f(v)\mathrm{d}v \ge 0,
\end{aligned} \tag{λ}$$

$$u(\underline{v},b_1)\ge 0, u(\underline{v},b_2)\ge 0, \tag{ξ_1,ξ_2}$$

$$u(\underline{v},b_1)-u(\underline{v},b_2)\ge 0, \tag{μ^1}$$

$$u(\underline{v},b_1)-u(\underline{v},b_2)+a^2(v_2^1-v_1^1)\ge 0, \tag{μ^2}$$

$$0\le a^2\le a^3=1, \tag{α^2,α^3}$$

$$\underline{v}\le v_1^1\le\overline{v}, \tag{γ_1^1,γ_1^2}$$

$$\underline{v}\le v_2^1\le v_2^2\le\overline{v}. \tag{$\gamma_2^1,\gamma_2^2,\gamma_2^3$}$$

Furthermore, the optimal mechanism is *unique*. Suppose, on the contrary, that there are two optimal mechanism. Since $\mathcal{P}'$ is linear in (a,p,q),[16] the convex combination of these two optimal mechanisms is also optimal. However, the convex combination of two 2-step allocation rules is not a 2-step allocation rule in general, which cannot be optimal by Lemma 4. Hence, there exists a unique optimal mechanism.

[16] $\mathcal{P}'(2,0)$ is not linear in $u(\underline{v},b_1)$, $u(\underline{v},b_2)$, a^2, v_1^1, v_2^1 and v_2^2.

Corollary 6 *Suppose Assumptions 2 and 3 hold. There exists a unique optimal mechanism of $\mathcal{P}$. Furthermore, u_1^*, u_2^*, v_1^*, v_2^*, v_2^{**} and a^* are continuous in k, c, π, b_1 and S.*

3.4.3. Subsidies in cash and in kind

I complete this section by a discussing subsidies in cash and in kind. In the optimal mechanism, compared with the high-budget agents who do not receive the goods, high-budget agents whose valuations exceed v_2^{**} receive the good with probability 1 by making an additional payment $a^* v_2^* + (1-a^*)v_2^{**}$. All high-budget agents whose valuations lie in $[v_2^*, v_2^{**}]$ receive the good with probability a^* by making an additional payment $a^* v_2^*$, which is an *in-kind subsidy*. In the literature, the value of an in-kind subsidy is often measured by its market value. In this paper, I do not model the private market explicitly, so I use the additional payment, $a^* v_2^* + (1 - a^*)v_2^{**}$, made by high-budget high-value agents as a measure of "price". Then the amount of in-kind subsidies offered to a high-budget agent is $a^* \left[a^* v_2^* + (1 - a^*)v_2^{**}\right] - a^* v_2^*$. Note that high-budget agents do not receive any in-kind subsidies if $v_2^* = v_2^{**}$, as in the case when budgets are common knowledge. Similarly, the amount of in-kind subsidies offered to a low-budget agent is $a^* \left[a^* v_2^* + (1 - a^*)v_2^{**}\right] - a^* v_1^*$. The difference in in-kind subsidies offered to the two budget types is $a^*(v_2^* - v_1^*)$.

In-kind subsidies are widespread around the world. The conventional wisdom rationalizing the prevalence of in-kind subsidies is paternalism. A more recent justification is based on the idea that agents have private information about their financial constraints, and governments cannot accurately identify low-budget agents in need of help. As a result, in-kind subsidies will be part of a surplus-maximizing mechanism as it is less susceptible to mimicking by high-budget agents. One difficulty with this justification is that, in many transfer programs, governments first "verify income, and then give benefits in kind, which would seem to rule out self-targeting as the primary reason for supplying benefits in-kind".[17] Moreover, governments "generally expend considerable resources determining eligibility".[18] In this paper, I formalize the idea that governments can verify agents' private information about their financial constraints via a costly procedure, and show that in such an environment

[17] Currie and Gahvari (2008)
[18] Currie and Gahvari (2008)

the optimal mechanism still makes use of both cash and in-kind subsidies.

3.5. Implementation

In this section, I provide one simple implementation of the direct optimal mechanism characterized in Section 3.4. This implementation exhibits some of the features of Singapore's housing and development board (HDB).

Consider the following *random assignment with regulated resale and cash subsidy* (RwRRC) scheme, which consists of two stages.

1. In the first stage, each agent reports his budget. Agents who report low budget are inspected with a probability of $(u_1^* - u_2^*)/c$. The principal offers cash subsidies u_1^* to low-budget agents and u_2^* to high-budget agents. The principal also offers low-budget agents the choice of participating in a lottery at price $p_1^* := a^* v_1^*$ and high budget agents the choice of participating in the same lottery at price $p_2^* := a^* v_2^*$. The principal distributes the goods randomly with uniform probability among all participants of the lottery. Each participant receives one unit of good with a probability no more than a^*.

2. In the second stage, the resale market opens, in which agents can purchase goods from each other and the principal if not all the goods are distributed in the first stage. The per-unit sales taxes are $\tau_1^* := v_2^{**} - v_1^*$ for low-budget sellers and $\tau_2^* := v_2^{**} - v_2^*$ for high-budget sellers. Agents who report low budget in the first stage and choose not to sell the good in the second stage are inspected with probability $(v_2^* - v_1^*)/c$.

Let a denote a lottery participant's expected probability of receiving the good in the first stage, and p^s denote the expected price a buyer pays in the second stage. Assume without loss of generality that $p^s > b_1$ so that a low-budget agent cannot afford it. Consider a low-budget agent whose type is (v, b_1) and who reports his budget truthfully. Then his payoff is u_1^* if he does not enter the lottery. If he buys the lottery, there are two possibilities. If he keeps the good when he receives it in the first stage, then his payoff is $u_1^* + av - a^* v_1^*$; otherwise his payoff is

$$u_1^* - a^* v_1^* + a(p^s - v_2^{**} + v_1^*).$$

Clearly, in the second stage, it is optimal for him to keep the good if and only if $v \geq p^s - v_2^{**} + v_1^*$. In the first stage, it is optimal for him to purchase the lottery if and only if

$$a \max \left\{ v, p^s - v_2^{**} + v_1^* \right\} \geq a^* v_1^*.$$

Similarly, consider a high-budget agent whose type is (v, b_2) and who reports his budget truthfully. It is easy to see that if it is optimal for an agent not to buy the lottery in the first stage, then it is also optimal for him not to buy the good in the second stage. If it is optimal for an agent to sell the good he receives in the first stage, then it is optimal for him not to buy the good in the second stage when he does not receive it in the first stage. Then his payoff is u_2^* if he does not buy the lottery. If he buys the lottery, there are three possibilities. If he buys the lottery, keeps the good when he receives it and buys it when he does not receive it, his payoff is

$$u_2^* - a^* v_2^* + av + (1 - a)(v - p^s);$$

if he buys the lottery, keeps the good when he receives it and does not buy when he does not receive it, his payoff is $u_1^* + a^*(v - v_1^*)$; if he buys the lottery and sells the good when he receives it, then his payoff is

$$u_2^* - a^* v_2^* + a(p^s - v_2^{**} + v_2^*).$$

Clearly, in the second stage, it is optimal for him to keep the good if and only if $v \geq p^s - v_2^{**} + v_2^*$ and buy the good if and only if $v \geq p^s$. In the first stage, it is optimal for him to purchase the lottery if and only if

$$a \max \left\{ v, p^s - v_2^{**} + v_2^* \right\} \geq a^* v_2^*.$$

Hence, in the second stage, the demand of the goods is $\pi(1 - a) [1 - F(p^s)]$ and the supply of the goods is

$$S - a(1-\pi)\left[1 - F\left(\max\left\{p^s - v_2^{**} + v_1^*, \frac{a^* v_1^*}{a}\right\}\right)\right] - a\pi\left[1 - F\left(\max\left\{p^s - v_2^{**} + v_2^*, \frac{a^* v_2^*}{a}\right\}\right)\right].$$

It is not hard to verify that $a = a^*$ and $p^s = v_2^{**}$ is the unique equilibrium.[19] Note that in this equilibrium, an agent is indifferent between not buying the lottery, and buying the lottery but selling the good when he receives it. All low-budget agents whose valuations are above v_1^* strictly prefer to participate in the lottery and keep the good they receive. All high-budget agents whose valuations are above v_2^* strictly prefer to participate in the lottery and keep the goods they receive. In addition, all high-budget agents whose valuations are above v_2^{**} will buy the goods in the second stage if they do not receive any in the first stage. These arguments prove the following result.

Proposition 1 *Suppose Assumptions 2 and 3 hold. The optimal mechanism is implemented by RwRRC with* $\underline{v} \leq v_1^* \leq v_2^* \leq v_2^{**} \leq \overline{v}$, $u_1^* \geq u_2^*$ *and* $0 \leq a^* \leq 1$ *given by Theorem 9.*

If inspection is sufficiently costly or the principal cannot inspect agents, then in the RwRRC scheme agents receive the same amount of cash subsidies $u_1^* = u_2^*$ and the same price $p_1^* = p_2^*$ in the first stage and face the same sales taxes $\tau_1^* = \tau_2^*$ in the second stage regardless of their budgets. This is consistent with the findings in Che et al. (2013a). If inspection is not too costly, then the principal provides financial aids to low-budget agents ($u_1^* \leq u_2^*$, $p_1^* \geq p_2^*$) in the first stage and discourages them from reselling by imposing a higher sales tax in the second stage.

This implementation exhibits some of the features of Singapore's HDB. HDB develops new flats and sells them to eligible buyers.[20] Buyers can purchase new flats directly from HDB or resale flats from existing owners in the open market. Buyers must have resided in their flats for a period of time, referred to as the minimum occupation period (MOP), before they are eligible to resell or sublet their flats. Buyers of resale HDB flats can apply for a CPF housing grant, which is a housing subsidy to

[19] Clearly, for each a, there is a unique p^s such that demand is equal to supply in the second stage. By construction, $a \leq a^*$. Suppose $a < a^*$, then the market clearing condition in the second stage implies that $p^s < v_2^{**}$. This implies that a low-budget agent buys the lottery only if $v > v_1^*$ and a high-budget agent buys the lottery only if $v > v_2^*$, which in turn implies that $a = a^*$, a contradiction. Hence, $a = a^*$ and $p^s = v_2^{**}$.

[20] 90% of HDB flats are owned by their residents. The remainder are rental flats for people who cannot afford to purchase the cheapest form of HDB flats despite financial aid.

Table 1: Minimum occupation periods (MOP) of housing and development board (HDB) flats

Types of HDB flats	MOP	
	Sell	Sublet
Resale flats w/ Grants	5–7 years	5–7 years
Resale flats w/o Grants	0–5 years	3 years

Sources. — Sell: http://www.hdb.gov.sg/cs/infoweb/residential/selling-a-flat/eligibility; and Sublet: http://www.hdb.gov.sg/cs/infoweb/residential/renting-out-a-flat-bedroom/renting-out-your-flat/eligibility.

help eligible households. HDB flats purchased with CPF housing grants are subject to longer MOPs as illustrated by Table 1.

3.6. Properties of the optimal mechanism

Having derived the optimal mechanism, I would like to investigate the following questions. Is it optimal for the principal to limit the supply of goods? When can the first-best outcome be achieved? What is the effect of a decrease in verification cost as, for example, a government's bureaucratic efficiency improves? What is the effect of an increase in the supply as, for example, a government builds more houses? What if agents become less budget-constrained? This loosening of constraints could happen as more agents are admitted into the formal financial system (π increases) or if their wealth increases as the economy grows (b_1 increases). What if the principal becomes less budget-constrained as a government increases expenditures on transfer programs? In what follows, I investigate each of these questions in turn.

Firstly, I show that it is not optimal for the principal to limit the supply of goods.

Proposition 2 *Suppose Assumptions 2 and 3 hold. In an optimal mechanism, (S) holds with equality.*

This result is straightforward if agents are unconstrained. However, it is not immediate from the principal's concern for efficiency if agents are budget-constrained. Recall that the principal also has a budget constraint, and this constraint may cause her to restrict supply. To see why, consider the extreme case in which low-budget agents have no money, i.e., $b_1 = 0$. In this case, the principal needs to raise all money from selling to high-budget agents. On the one hand, as she increases the amount of goods sold to high-budget agents, the revenue will start declining at some point. On the other hand, increasing the amount of goods allocated to low-budget agents raises the inspection cost. Thus, it is not obvious that in an optimal mechanism all the goods are distributed to agents. In the proof of Proposition 2, I show that if not all the goods are distributed to agents yet, then the principal can increase the amount of goods allocated to high-budget and low-budget agents simultaneously. For an appropriately chosen allocation rule, the resulting mechanism is feasible and strictly improves welfare.[21]

Secondly, I give a necessary and sufficient condition under which the first-best is achieved.

Proposition 3 *Suppose Assumptions 2 and 3 hold. The first-best is achieved if and only if* $S \geq \hat{S}(b_1)$*, where* $\hat{S}(b_1)$ *is the solution to*

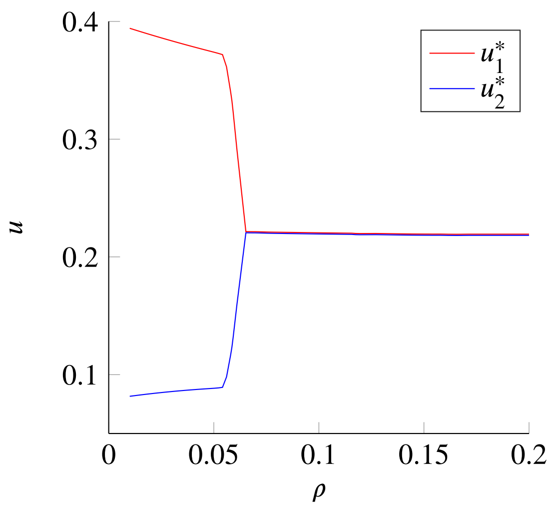
$$b_1 - v^* F(v^*) = 0$$

with $v^* = F^{-1}(1 - S)$*. Furthermore,* $\hat{S}(b_1)$ *is strictly decreasing in* b_1*.*

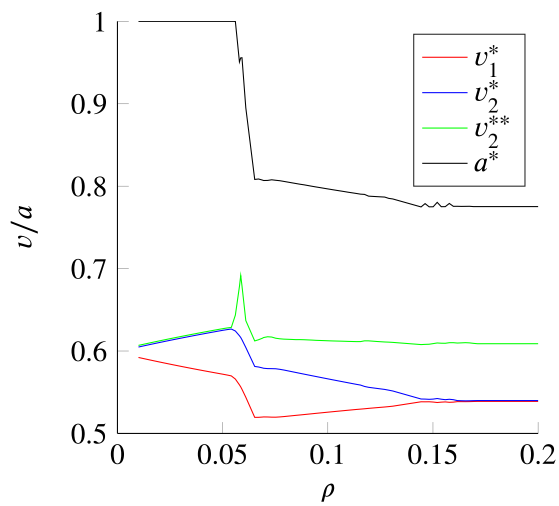
Intuitively, the first-best is achieved if the supply of the good is abundant or agents have ample budgets. Note that the condition given in Proposition 3 is independent of inspection cost k, punishment c and the percentage of high-budget agents π. This is because when the first-best is achieved, agents of both budget types receive the same amount of cash subsidies and the same allocation rule, and the inspection probability is zero. For the rest of this section, I assume that the first-best cannot be achieved, i.e., $S < \hat{S}(b_1)$.

Thirdly, I study the impact of changes in effective inspection cost ($\rho = k/c$), supply (S), budget (b_1) and the percentage of high-budget agents (π) on the optimal mechanism as well as welfare. The optimal mechanism is characterized by u_1^*, u_2^*, v_1^*, v_2^*, v_2^{**} and a^*, which (together with the corresponding Lagrangian multipliers) are solutions to a system of non-linear equations. As a result, it is hard to perform all comparative statics analysis analytically. In what follows, I give some analytic results for extreme cases such as when effective inspection cost is sufficiently large and explore the intermediate case numerically.

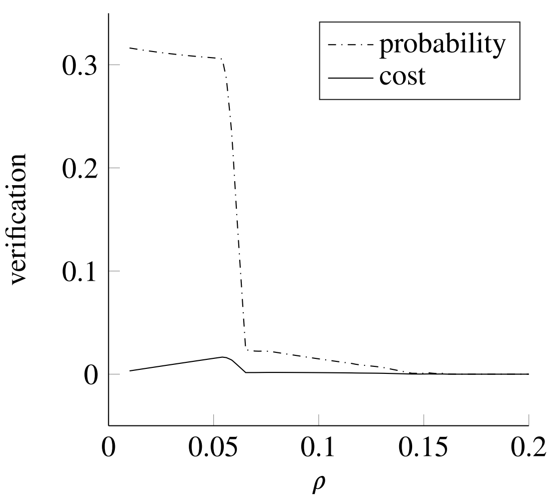
[21] I thank Michael Richter for suggesting this proof.



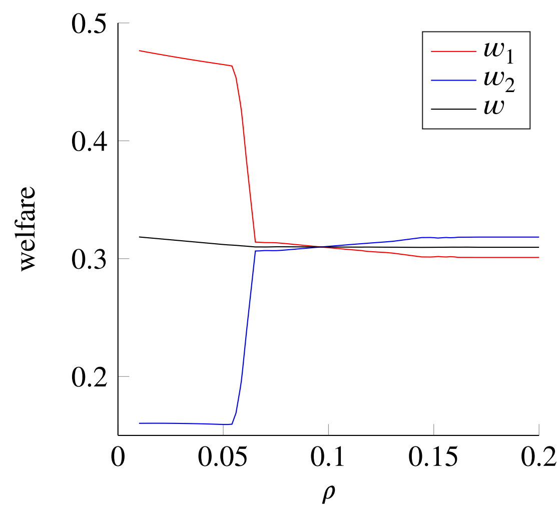
(a) Cash Subsidies



(b) Allocation



(c) Verification



(d) Welfare

Figure 5: The impact of an increase in effective inspection cost (ρ) on cash subsidies, allocation, inspection and welfare. In this numerical example, v is uniformly distributed on $[0, 1]$, $S = 0.4$, $b_1 = 0.2$, $\pi = 0.5$ and $\rho \in [0, 0.2]$.

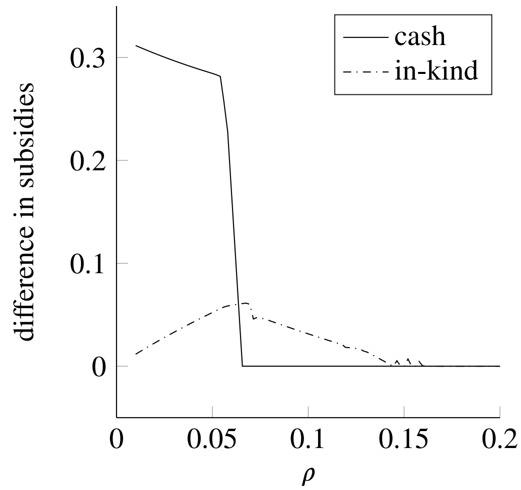


Figure 6: The impact of an increase in effective inspection cost (ρ) on the differences in cash and in-kind subsidies. In this numerical example, v is uniformly distributed on [0, 1], $S = 0.4$, $b_1 = 0.2$, $\pi = 0.5$ and $\rho \in [0, 0.2]$.

Effective Verification Cost (ρ). Intuitively, as verification becomes more costly (ρ increases), the principal tends to inspect agents less frequently in the optimal mechanism. To maintain incentive compatibility, the principal needs to reduce the differences in cash and in-kind subsidies offered to agents with different budgets. Proposition 4 shows that, for a large ρ, agents of both budget types receive the same amount of cash subsidies. Eventually, for ρ sufficiently large, verification is never used. The two lower bounds given in Proposition 4 are not tight, as illustrated in the numerical example in Figure 5. If v is uniformly distributed, then one can further prove that, for fixed punishment c, the verification probability is non-increasing in verification cost k. However, the change in total verification cost may not be monotonic as illustrated by Figure 5c.

Proposition 4 *Suppose Assumptions 2 and 3 hold.*

1. *If $\rho \geq \frac{\pi}{1-\pi}$, then agents of both budget types receive the same amount of cash subsidies, i.e., $u_1^* = u_2^*$.*

2. *There exists $\overline{\rho} \leq \frac{\pi}{S(1-\pi)}$ such that the verification probability in an optimal mechanism is zero, i.e., $q(v, b) = 0$ for all v and b, if and only if $\rho \geq \overline{\rho}$. Furthermore, the total welfare is strictly decreasing in ρ over $[0, \overline{\rho}]$ and constant in ρ over $[\overline{\rho}, \infty)$.*

3. *If v is uniformly distributed, then the verification probability is non-increasing in k.*

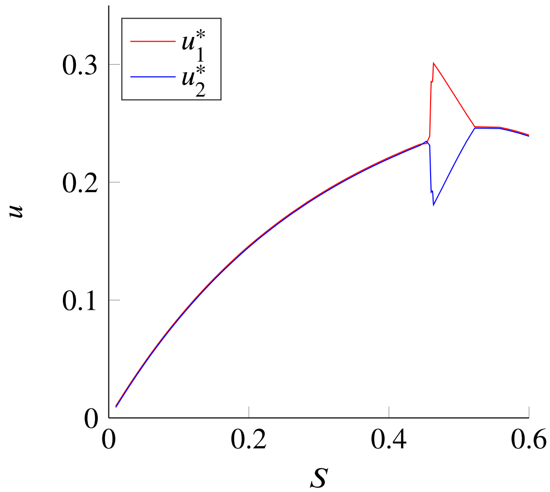
Figure 5 plots the impact of an increase in effective verification cost (ρ) on cash subsidies, allocation, verification and welfare in a numerical example. It is straightforward that an increase in ρ reduces the total welfare but its impacts on different budget types are different. Verification allows the principal to more accurately target low-budget agents and improves their welfare. As a result, as ρ increases, the welfare of low-budget agents declines while that of high-budget agents rises, as seen in Figure 5d.

More interestingly, the optimal mechanism makes use of both cash and in-kind subsidies, and the change in verification cost affects that mechanism's reliance on each of them as shown in Figure 6. If ρ is sufficiently small, then the principal helps low-budget agents mainly by offering them more cash subsidies. As ρ increases, the difference in cash subsidies declines but the difference in in-kind subsidies increases. This is because even though cash subsidy is more efficient in the sense that it does not introduce any distortion in allocation, it is more expensive in terms of verification cost. Cash subsidy is attractive to everyone regardless of their valuations. In contrast, in-kind subsidy is attractive only to agents whose valuations are high enough. Eventually, the difference in in-kind subsidies also declines as verification becomes sufficiently costly.

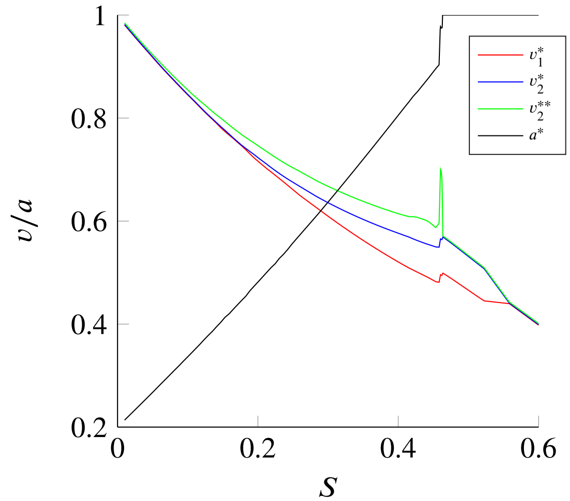
Supply (S). The impact of an increase in the supply (S) on the optimal mechanism is complicated. On the one hand, the principal becomes less budget constrained, and can direct more subsidies to low-budget agents and inspect them more frequently. On the other hand, low-budget agents also become less budget constrained as S increases,[22] which reduces the needs to subsidize and inspect them. As shown in Propositions 4 and 5, for sufficiently large and small S, agents of both budget types receive the same amount of cash subsidies.

Proposition 5 *Suppose Assumptions 2 and 3 hold. If S is sufficiently small, then agents of both budget types receive the same amount of cash subsidies, i.e., $u_1^* = u_2^*$.*

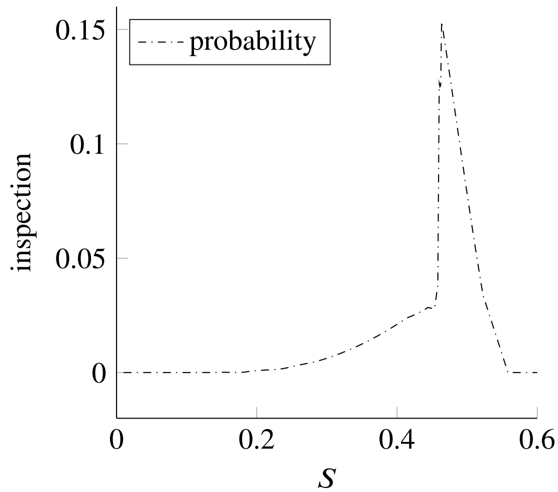
[22] As in Section 3.4.3, I use the additional payment $a^* v_2^* + (1-a^*) v_2^{**}$ made by a high-budget high-valuation agent a measure of "price". Then this price generally declines as S increases and low-budget agents become less budget constrained in the sense that the gap between this price and their budgets shrinks.



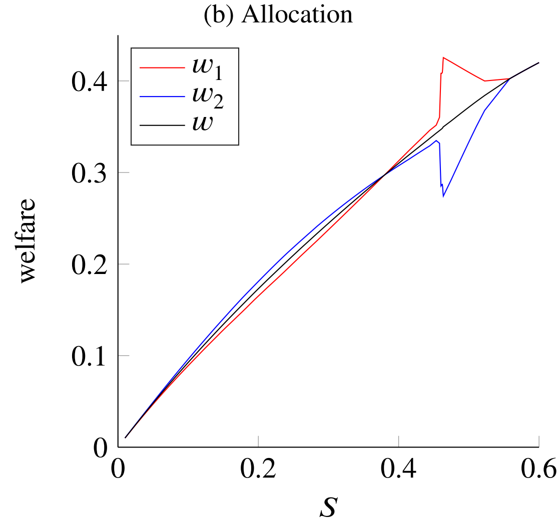
(a) Cash Subsidies



(b) Allocation

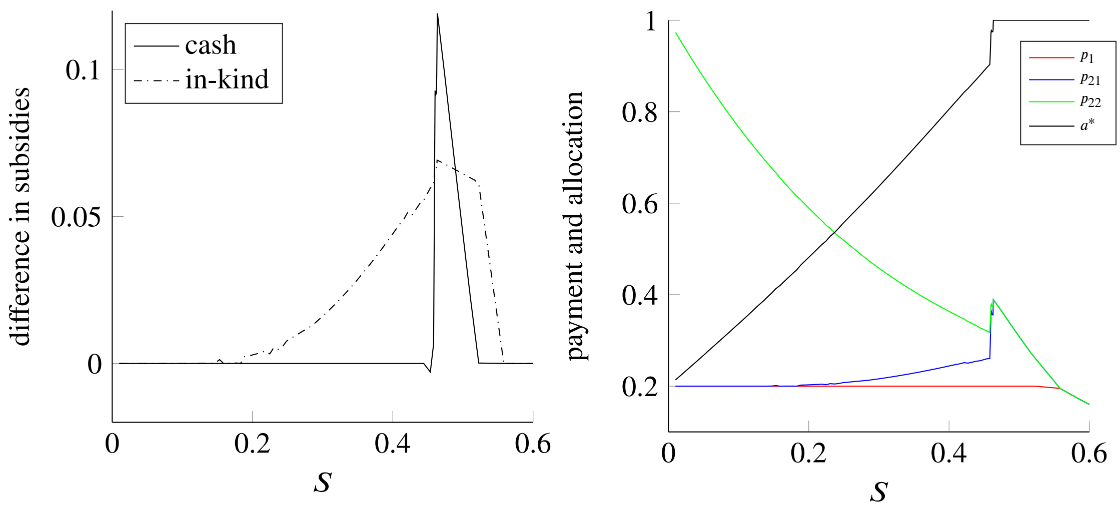


(c) Verification Probability



(d) Welfare

Figure 7: The impact of an increase in the supply (S) on cash subsidies, allocation, verification and welfare. In this numerical example, v is uniformly distributed on $[0, 1]$, $\rho = 0.08$, $b_1 = 0.2$, $\pi = 0.5$ and $S \in [0, 0.6]$.



(a) Differences in Cash and In-kind Subsidies

(b) Payment and Allocation

Figure 8: The impact of an increase in the supply (S) on the differences in cash and in-kind subsidies, allocation and payment. In the right panel, the red line (p_1) denotes the payment by a low-budget agent who receives the good with probability a^*, the blue line (p_{21}) denotes the payment by a high-budget agent who receives the good with probability a^*, and the red line (p_{22}) denotes the payment by a high-budget agent who receives the good with probability one. In this numerical example, υ is uniformly distributed on $[0, 1]$, $\rho = 0.08$, $b_1 = 0.2$, $\pi = 0.5$ and $S \in [0, 0.6]$.

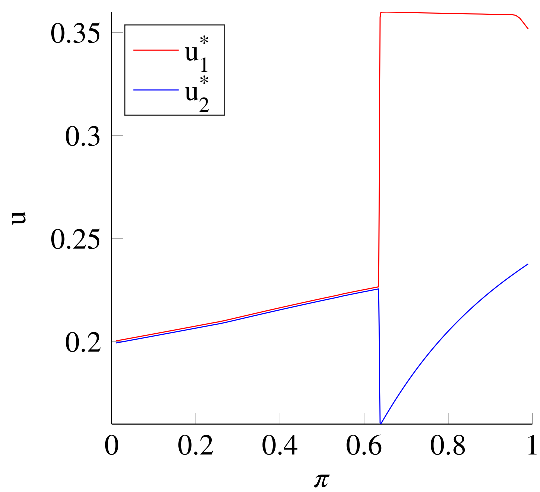
These effects can also be seen in Figures 7 and 8, which plot the impact of an increase in the supply (S) on cash subsidies, allocation, verification and welfare in a numerical example. Specifically, Figure 8a plots the differences in cash and in-kind subsidies between high-budget and low-budget agents. If S is sufficiently small, then agents receive the same amount of subsidies regardless of their budgets. As S increases, the principal raises first the difference in in-kind subsidies and then that in cash subsidies. This order occurs because it is less expensive to target only low-budget high-valuation agents than all low-budget agents. Eventually, the differences in both cash and in-kind subsidies decline as the need to subsidize low-budget agents declines. As a result, the verification probability is hump-shaped as shown in Figure 7c.

Intuitively, the total welfare is strictly increasing in S. More interestingly, the welfare of each type is not monotonic in S. Figure 7d plots the total welfare and the average utility of each budget type as a function of S. Initially, the average utilities of both budget types increase as S increases. When S is large enough that the principal begins to divert more cash subsidies and goods to low-budget agents, the average utility of high-budget agents begins to decrease as S increases. Eventually, the need to subsidize low-budget agents decreases as S increases, and the average utility of low-budget agents begins to decrease while that of high-budget agents begins to increase, until they coincide. Specifically, low-valuation agents of both budget types can get worse off as they receive less cash subsidies. Interestingly, high-budget high-valuation agents can also get worse off because their payments can increase as S increases (see Figure 8b). These increases in payments occur because disproportionately more goods will be allocated to low-budget agents and there will be less pooling when S increases.

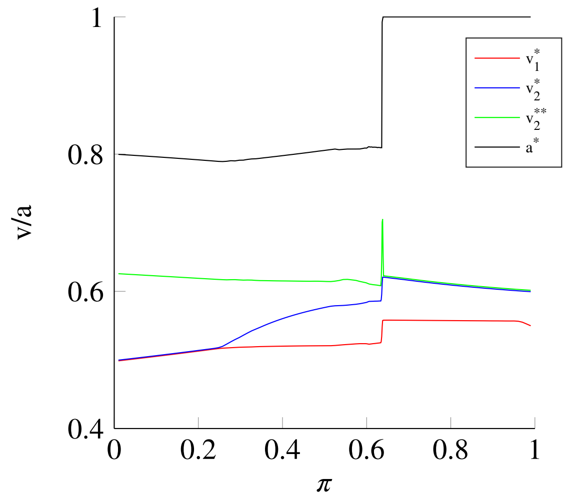
Percentage of high-budget agents (π). Proposition 4 also proves that for small π, agents of both budget types receive the same amount of cash subsidies. Eventually, for π sufficiently small, verification is never used. This result is intuitive because a smaller π means a larger population of low-budget agents and therefore higher total verification cost given the same mechanism. Hence, the principal tends to inspect agents less frequently as π decreases. However, this change in verification probability is not monotonic in π, because an increase in π not only makes verification less costly but also makes the economy wealthier. If π is sufficiently large, then the principal becomes less budget-constrained and can afford to maintain incentive compatibility by subsidizing high-budget agents directly rather than inspect low-budget agents. This is illustrated by the numerical example in Figure 9.

The total welfare as well as the welfare of low-budget agents are strictly increasing in π, but the welfare of high-budget agents is not monotonic in π. Initially, as π increases, the welfare of high-budget agents declines as the principal provides more subsidies to low-budget agents. Eventually, the welfare of high-budget agents rises as the principal subsidizes high-budget agents rather than inspecting low-budget agents.

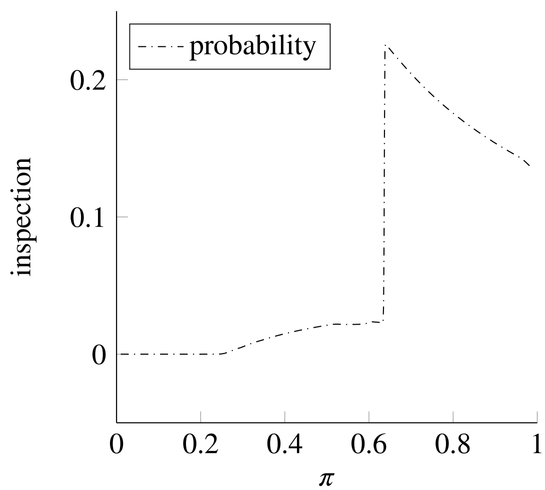
Budget (b_1) Low-budget agents become less budget constrained as b_1 increases. This change reduces the need for subsidies, which leads to a decline in verification probability. Proposition 6 proves that for large b_1, agents of both budget types receive the same amount of cash subsidies. Figure 10



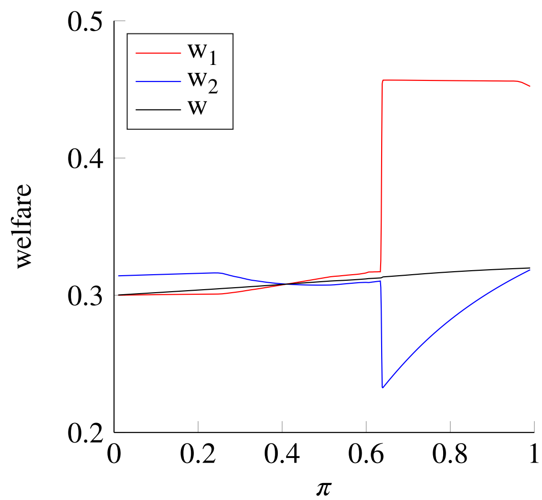
(a) Cash Subsidies



(b) Allocation

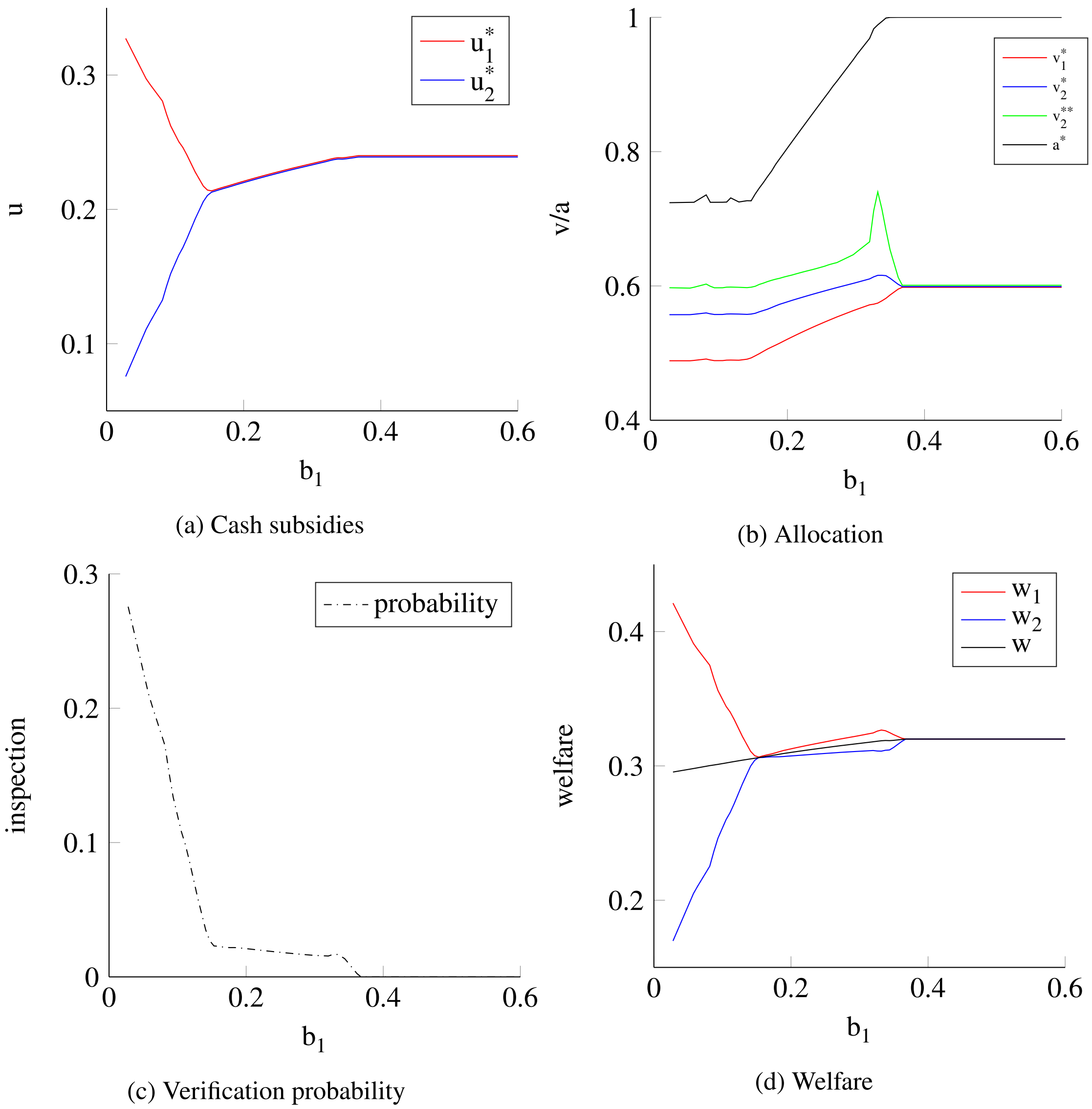


(c) Verification Probability



(d) Welfare

Figure 9: The impact of an increase in the percentage of high-budget agents (π) on cash subsidies, allocation, verification and welfare. In this numerical example, v is uniformly distributed on $[0, 1]$, $\rho = 0.08$, $b_1 = 0.2$, $S = 0.4$ and $\pi \in [0, 1]$.



(a) Cash subsidies

(b) Allocation

(c) Verification probability

(d) Welfare

Figure 10: The impact of relaxing low-budget agents' budget constraint (b_1) on cash subsidies, allocation, verification and welfare. In this numerical example, v is uniformly distributed on $[0, 1]$, $\rho = 0.08$, $S = 0.4$, $\pi = 0.5$ and $b_1 \in [0, 0.6]$.

plots the impact of an increase in b_1 on cash subsidies, allocation, verification and welfare in a numerical example. In this numerical example, the total verification probability is non-increasing in b_1 and zero for b_1 sufficiently large.

Proposition 6 *Suppose Assumptions 2 and 3 hold. If b_1 is sufficiently large, then agents of both budget types receive the same amount of cash subsidies, i.e., $u_1^* = u_2^*$.*

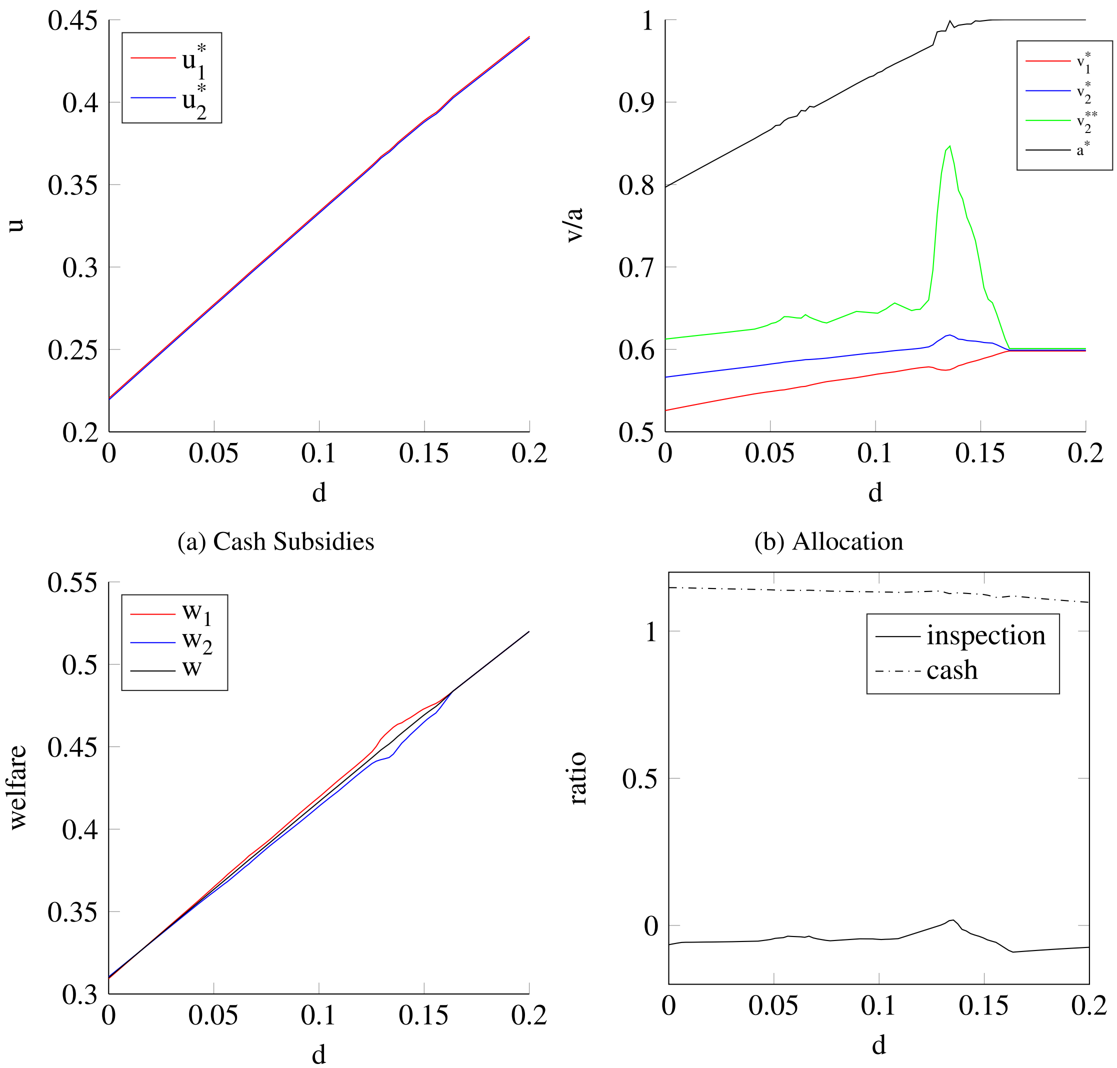
The total welfare, as well as, the welfare of high-budget agents is strictly increasing in b_1, but the welfare of low-budget agents is not monotonic in b_1. On the one hand, low-budget agents become less budget-constrained as b_1 increases. On the other hand, the principal provides lower cash and in-kind subsidies to low-budget agents as b_1 increases. As shown in Figure 10d, either effect can dominate the other. Hence, the welfare of low-budget agents may either increase or decrease as b_1 increases.

Lastly, I study the impact of relaxing the principal's budget-balanced constraint on the optimal mechanism and welfare. Specifically, I reformulate the principal's budget constraint as follows:

$$\mathbb{E}_t \left[p(t) - q(t)k \right] \geq -d. \tag{BB}$$

In the main part of the paper, I assume $d = 0$. But it is easy to see that all the results in Sections 3.3 and 3.4 extend to the case of $d \geq 0$.

Figure 11 plots the impact of an increase in the principal's budget (d) on cash subsidies, allocation, verification and welfare in a numerical example. Note that an increase in d leads to an increase in the total cash subsidies by more than one-hundred percent. This is easy to see when there is no verification. An increase in d raises cash subsidies to low-budget agents, which relax their budget constraints and improve allocation efficiency. Under Assumption 2, this in turn improves the principal's revenue and allows her to further raise cash subsidies. The numerical example suggests this is still true when verification is possible.



(a) Cash Subsidies

(b) Allocation

(c) Welfare

(d) Change in Verification Cost/Cash Subsidies

Figure 11: The impact of an increase in the principal's budget (d) on cash subsidies, allocation, verification and welfare. In this numerical example, v is uniformly distributed on $[0, 1]$, $\rho = 0.04$, $b_1 = 0.2$, $S = 0.4$, $\pi = 0.5$ and $d \in [0.0.2]$.

3.7. Extensions and discussions

In this section, I discuss several issues. Section 3.7.1 shows that some of the analysis extends if I replace the budget constraint by a more stringent per-unit price constraint. Section 3.7.2 shows that the analysis extends to the case where punishment is transferable. Sections 3.7.3 and 3.7.4 discuss the robustness of my analysis to weakening the assumptions on verification and punishment, respectively. Section 3.7.5 discusses why it is necessary to explicitly model budget constraints.

3.7.1. Per-unit price constraint

In the optimal direct mechanism, agents make payments to the principal regardless of whether they receive the goods,[23] which some may consider unrealistic. The question, then, is whether this direct mechanism can be implemented by a mechanism in which agents pay if and only if they receive the goods and their payments are within their budgets. Such an implementation is impossible if $a^* < 1$. I can guarantee that such an implementation always exists if I replace (BC) by the following *per-unit price constraint*:

$$p(t) \leq a(t)b, \ \forall t = (v, b). \tag{PC}$$

(BC) is the same as that found in Che and Gale (2000) and Pai and Vohra (2014a), but different from Che et al. (2013a), which uses (PC).

Nevertheless, I assume (BC) in the main body of the paper for the following reasons. First, as will become clear, the optimal mechanisms in these two settings share qualitatively similar features. Second, for some parameter values (e.g., verification cost is low, resources are relatively abundant or the percentage of budget constrained agents is small), there is no rationing in the optimal mechanism ($a^*(\rho) = 1$). Third, rationing is realistic if b_1 is close to zero. For example, families with very low income may receive free coverage from Medicaid.

In the rest of this subsection, I consider the principal's problem in which (BC) is replaced by (PC), denoted by $\mathcal{P}_{PC}$. I first make the observation that if (PC) holds for v' then it holds for all $v < v'$.

[23] For a finite number of agents, this is similar to an all-pay auction.

This is trivial if $a(v,b)=0$. If $a(v,b)>0$, then by the envelope condition we have

$$\frac{p(v',b)}{a(v',b)}-\frac{p(v,b)}{a(v,b)}=\int_{v}^{v'}\left(1-\frac{a(\nu,b)}{a(v',b)}\right)\mathrm{d}\nu-\int_{\underline{v}}^{v}\left(\frac{a(\nu,b)}{a(v',b)}-\frac{a(\nu,b)}{a(v,b)}\right)\mathrm{d}\nu-\frac{u(\underline{v},b)}{a(v',b)}+\frac{u(\underline{v},b)}{a(v,b)}$$
$$\geq 0,$$

where the last inequality holds since $a(v,b)$ is non-decreasing in v. Hence, (PC) holds if and only if $p(\overline{v},b)\leq a(\overline{v},b)b$ for all b.

Given this observation, it is straightforward to extend the results of Theorems 7 and 8 to the current setting. Theorem 10 characterizes an optimal mechanism of $\mathcal{P}_{PC}$ when budgets are common knowledge ($\rho=0$). Theorem 11 characterizes an optimal mechanism of $\mathcal{P}_{PC}$ when budget is an agent's private information and the principal cannot verify this information ($\rho=\infty$). The latter theorem extends the results in Section 3 of Che et al. (2013a) to the setting of a continuum of values under the regularity conditions.

Theorem 10 *Suppose Assumption 3 holds, and budgets are common knowledge. There exists* $v_1^*(0)$*,* $v_2^*(0)$*,* $u_1^*(0)$ *and* $u_2^*(0)$ *such that an optimal mechanism of* $\mathcal{P}_{PC,CB}$ *is given by*

$$\begin{aligned}
a(v,b_1)&=\chi_{\{v\geq v_1^*(0)\}}a^*(0), \quad & p(v,b_1)&=\chi_{\{v\geq v_1^*(0)\}}\left(u_1^*(0)+a^*(0)b_1\right)-u^*(0),\\
a(v,b_2)&=\chi_{\{v\geq v_2^*(0)\}}, & p(v,b_2)&=\chi_{\{v\geq v_2^*(0)\}}v_2^*(0),
\end{aligned}$$

where $a^*(0)=u_1^*(0)/\left[v_1^*(0)-b_1\right]$, $b_1<v_1^*(0)\leq v_2^*(0)<\overline{v}$ *and* $0=u_2^*(0)<u_1^*(0)\leq v_1^*(0)-b_1$.

Theorem 11 *Suppose Assumptions 2 and 3 hold, and the principal does not inspect agents. There exists* $v_1^*(\infty)$*,* $v_2^*(\infty)$*,* $v_2^{**}(\infty)$*,* $u_1^*(\infty)$ *and* $u_2^*(\infty)$*such that an optimal mechanism of* $\mathcal{P}_{PC,NI}$ *with no verification satisfies*

$$\begin{aligned}
a(v,b_1)&=\chi_{\{v\geq v_1^*(\infty)\}}a^*(\infty),\ p(v,b_1)=\chi_{\{v\geq v_1^*(\infty)\}}a^*(\infty)v_1^*(\infty)-u_1^*(\infty),\\
a(v,b_2)&=\chi_{\{v\geq v_1^*(\infty)\}}a^*(\infty)+\chi_{\{v\geq v_2^{**}(\infty)\}}\left(1-a^*(\infty)\right),\\
p(v,b_2)&=\chi_{\{v\geq v_2^*(\infty)\}}a^*(\infty)v_2^*(\infty)+\chi_{\{v\geq v_2^{**}\}}\left(1-a^*(\infty)\right)v_2^{**}(\infty)-u_2^*(\infty),
\end{aligned}$$

where $a^*(\infty) = u_1^*(\infty)/\left[v_1^*(\infty) - b_1\right]$, $b_1 < v_1^*(\infty) = v_2^*(\infty) \leq v_2^{**}(\infty) \leq \overline{v}$ *and* $0 < u_1^*(\infty) = u_2^*(\infty) \leq v_1^*(\infty) - b_1$.

The analysis is more complex if budget is an agent's private information and the principal can verify this information at a cost. As before, I first consider the principal's relaxed problem $\mathcal{P}'_{PC}$ in which I relax (IC) corresponding to over-reporting of budgets. One can show that Lemmas 4 and 5 and Corollary 5 still hold. Next, I consider the principal's relaxed problem with two modifications: (i) The allocation rule is an M'-step allocation rule for some integer $M' \leq M$ and $M \geq 2$ is a fixed integer; and (ii) the principal's budget balance constraint is relaxed by a constant $d \geq 0$. Denote this problem by $\mathcal{P}'_{PC}(M, d)$ and its value by $V_{PC}(M, d)$. Then $\mathcal{P}'_{PC}(M, d)$ is identical to $\mathcal{P}'(M, d)$ if I replace (BC) by the following (PC) constraint:

$$a^M v_1^{M-1} - \sum_{j=1}^{M-1} a^j (v_1^j - v_1^{j-1}) - u(\underline{v}, b_1) \leq b_1 a^M. \qquad (\eta)$$

One can readily extend the results of Lemma 6 to the current setting, which says that, in an optimal mechanism of $\mathcal{P}'_{PC}(M, d)$, the verification probability is non-decreasing in a low-budget agent's reported value. Using the monotonicity of verification probability and the linearity of $\mathcal{P}'_{PC}(M, d)$ in $u(\underline{v}, b_1)$, $u(\underline{v}, b_2)$ and a^m's, we have $V_{PC}(M, d) = V_{PC}(M-1, d)$ for M sufficiently large. By a similar approximation argument to that in the proof of Theorem 9, one can prove the following theorem, which characterizes an optimal mechanism of $\mathcal{P}_{PC}$.

Theorem 12 *Suppose Assumptions 2 and 3 hold. There exists an integer* $2 \leq M \leq 5$, $\underline{v} < v_1^1 < \cdots < v_1^{M-1} < \overline{v}$, $\underline{v} \leq v_2^0 \leq v_2^1 < \cdots < v_2^{M-1} \leq v_2^M < \overline{v}$ *and* $0 \leq a^1 < a^2 < \ldots a^M \leq 1$ *such that an* *optimal mechanism of $\mathcal{P}_{PC}$ is given by*

$$a(v, b_1) = \sum_{m=1}^{M} \chi_{\{v_1^{m-1} < v \le v_1^m\}} a^m,$$

$$p(v, b_1) = \sum_{m=1}^{M-1} \chi_{\{v \ge v_1^m\}} \left(a^{m+1} - a^m\right) v_1^m - u(\underline{v}, b_1),$$

$$q(v, b_1) = \frac{1}{c}\left[u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1 \left(v_2^0 - \underline{v}\right) + \sum_{m=1}^{M-1} \chi_{\{v \ge v_1^m\}} \left(a^{m+1} - a^m\right)\left(v_2^m - v_1^m\right)\right]$$

$$a(v, b_2) = \sum_{m=1}^{M} \chi_{\{v_2^{m-1} < v \le v_2^m\}} a^m + \chi_{\{v \ge v_2^M\}},$$

$$p(v, b_2) = \sum_{m=0}^{M-1} \chi_{\{v \ge v_2^m\}} \left(a^{m+1} - a^m\right) v_2^m + \chi_{\{v \ge v_2^M\}} \left(1 - a^M\right) - u(\underline{v}, b_2).$$

However, it is hard to further improve this result, as in Section 3.4.2 when we require only the weaker (BC) constraint. In particular, the proof of Lemma 7 does not apply here. It holds if we also make the following assumption.

Assumption 4 $a(v, b) = 0$ *for all* $v < b_1$.

Assumption 4 requires that an agent whose valuation is too low (lower than b_1) receives the good with probability zero. Note that optimal mechanisms in Theorem 10 and Theorem 11 satisfy this condition. I conjecture this condition also holds in the general case, although I cannot prove it. Under this additional assumption, we have

Lemma 8 *Suppose Assumptions 2, 3 and 4 hold. Then* $V_{PC}(M, d) = V_{PC}(2, d)$ *for all* $M \ge 2$ *and* $d \ge 0$.

Given Lemma 8, it is easy to extend the result of Theorem 9 to this setting.

Theorem 13 *Suppose Assumptions 2, 3 and 4 hold. There exist* $a^*(\rho)$, $v_1^*(\rho)$, $v_2^*(\rho)$, $v_2^{**}(\rho)$, $u_1^*(\rho)$

and $u_2^*(\rho)$ *such that an optimal mechanism of* $\mathcal{P}_{PC}$ *is given by*

$$
\begin{aligned}
&a(v,b_1) = \chi_{\{v \ge v_1^*(\rho)\}} a^*(\rho),\ \ p(v,b_1) = \chi_{\{v \ge v_1^*(\rho)\}} a^*(\rho) v_1^*(\rho) - u_1^*(\rho),\\
&q(v,b_1) = \frac{1}{c}\left[\chi_{\{v \ge v_1^*(\rho)\}} a^*(\rho)\left(v_2^*(\rho) - v_1^*(\rho)\right) + u_1^*(\rho) - u_2^*(\rho)\right],\\
&a(v,b_2) = \chi_{\{v \ge v_2^*(\rho)\}} a^*(\rho) + \chi_{\{v \ge v_2^{**}(\rho)\}}\left(1 - a^*(\rho)\right),\\
&p(v,b_2) = \chi_{\{v \ge v_2^*(\rho)\}} a^*(\rho) v_2^*(\rho) + \chi_{\{v \ge v_2^{**}(\rho)\}}\left(1 - a^*(\rho)\right) v_2^{**}(\rho) - u_2^*(\rho),\\
&q(v,b_2) = 0,
\end{aligned}
$$

where $a^*(\rho) = u_1^*(\rho)/\left[v_1^*(\rho) - b_1\right]$, $\underline{v} \le v_1^*(\rho) \le v_2^*(\rho) \le v_2^{**}(\rho) \le \overline{v}$, $0 < a^*(\rho) \le 1$ *and* $u_1^*(\rho) \ge u_2^*(\rho)$*.*

3.7.2. Monetary penalty

In this subsection, I discuss what happens if penalty is transferable. Specifically, the principal can inspect an agent's budget at a cost $k > 0$, and can impose a monetary penalty of up to $c \ge 0$. I also allow the principal to punish an innocent agent and an agent without verification. Nonetheless, as I will show later, it is optimal for the principal to punish an agent if and only if he is found to have lied about his budget. Using this result, the principal's problem can be reduced to the one stated in Section 3.2, when penalty is not transferable. Hence, all results in previous sections also hold in the case of monetary penalty.

I also relax the assumption that an agent is punished if and only if he is found to have lied. In this case, a direct mechanism is a quadruple (a, p, q, θ), where a, p and q are defined as before and $\theta : T \times \{b_1, b_2, n\} \to [0, c]$ denotes the penalty imposed on an agent. In particular, $\theta(\hat{t}, n) \in [0, c]$ denotes the penalty imposed on an agent who reports $\hat{t}$ and is not inspected, and $\theta(\hat{t}, b) \in [0, c]$ denotes the penalty imposed on an agent who reports $\hat{t}$ and is inspected, and whose budget is revealed to be b. The utility of an agent who has type $t := (v, b)$ and reports $\hat{t}$ is

$$
u(\hat{t}, t) = \begin{cases} a(\hat{t})v - p(\hat{t}) - (1 - q(\hat{t}))\theta(\hat{t}, n) - q(\hat{t})\theta(\hat{t}, b) & \text{if } p(\hat{t}) + \theta(\hat{t}, b) \le b \text{ and } p(\hat{t}) + \theta(\hat{t}, n) \le b,\\ -\infty & \text{otherwise.} \end{cases}
$$

The principal's problem is

$$\max_{a,p,q,\theta} \mathbb{E}_t \left[a(t)v - q(t)k\right], \tag{$\mathcal{P}$}$$

subject to

$$u(t) \equiv u(t,t) \geq 0, \ \forall t \in T, \tag{IR}$$

$$u(t) \geq u(\hat{t},t), \ \forall t \in T, \hat{t} \in \left\{\hat{t} \in T \,\middle|\, p(\hat{t}) + \max\{\theta(\hat{t},n), \theta(\hat{t},b)\} \leq b\right\}, \tag{IC}$$

$$p(t) + \max\{\theta(t,n), \theta(t,b)\} \leq b, \ \forall t \in T, \tag{BC}$$

$$\mathbb{E}_t \left[p(t) + (1-q(t))\theta(t,n) + q(t)\theta(t,b) - q(t)k\right] \geq 0, \tag{BB}$$

$$\mathbb{E}_t \left[a(t)\right] \leq S. \tag{S}$$

Note that (BC) requires that an agent must be able to afford the payment and the punishment. I show that it is without loss of generality to focus on mechanisms in which an agent is penalized if and only if he is found to have lied about his budget, and whenever he is found to have lied he has the maximum possible monetary penalty c imposed upon him.

Lemma 9 *It is without loss of generality to focus on mechanisms in which* $\theta(\hat{t},n) = 0$*,* $\theta(\hat{t},b) = 0$ *if* $\hat{b} = b$ *and* $\theta(\hat{t},b) = c$ *if* $\hat{b} \neq b$*.*

Using Lemma 9, the principal's problem can be reduced to the one stated in Section 3.2 when penalty is not transferable. Hence, all results in previous sections also hold in the case of monetary penalty.

3.7.3. Costly disclosure

In this subsection, I study what happens if it is also costly for an agent to have his report verified. For example, agents may bear some costs of providing documentation. Assume that an agent incurs a non-monetary cost only when his report is verified. Let $c^T \geq 0$ denote the cost incurred by an agent from being verified by the principal if he reported his type truthfully, and let $c^F \geq c^T$ be his cost if he lied.[24] As I will show below, disclosure costs have three effects. Firstly, similar to monetary transfers, disclosure costs can also be used to screen agents with different valuations and help relax agents' budget constraints. Secondly, it is more costly for an agent to lie about his budget because $c^F \geq c^T$. Finally, disclosure costs make verification more costly for the principal whose concern is welfare. Even though it is difficult to solve the optimal mechanism, I show that if the difference between c^F and c^T is sufficiently large, then the first two effects dominate and introducing disclosure costs improves welfare.

The utility of an agent who has type $t = (v, b)$ and reports $\hat{t}$ is

$$u(\hat{t}, t) = \begin{cases} a(\hat{t})v - p(\hat{t}) - q(\hat{t})c^T & \text{if } \hat{b} = b \text{ and } p(\hat{t}) \leq b, \\ a(\hat{t})v - p(\hat{t}) - q(\hat{t})\left(c^F + c\right) & \text{if } \hat{b} \neq b \text{ and } p(\hat{t}) \leq b, \\ -\infty & \text{if } p(\hat{t}) > b. \end{cases}$$

The principal's problem is

$$\max_{a,p,q} \mathbb{E}_t \left[a(t)v - q(t)k - q(t)c^T\right], \tag{$\mathcal{P}_{DC}$}$$

subject to (IR), (IC), (BC), (BB) and (S). Note that if $c^T = 0$, then ($\mathcal{P}_{DC}$) is equivalent to the original problem ($\mathcal{P}$) in which the punishment is $c^F + c$.

Consider the more general case in which $c^T \geq 0$. Define $p^e(t) := p(t) + q(t)c^T$, $k^e := k + c^T$ and $c^e := c + c^F - c^T$. As in Section 3.4, I separate (IC) into two categories and ignore those corresponding to over-reporting of budgets. Then the principal's relaxed problem can be written as:

$$\max_{a,p,q} \mathbb{E}_t \left[a(t)v - q(t)k^e\right], \tag{$\mathcal{P}'_{DC}$}$$

[24]The analysis goes through as long as $c^F + c \geq c^T$.

subject to

$$a(t)v - p^e(t) \geq 0, \ \forall t \in T, \tag{IR}$$

$$a(v,b)v - p^e(v,b) \geq a(\hat{v},b)v - p^e(\hat{v},b), \ \forall v, \hat{v}, b, \tag{IC-v}$$

$$a(v,b_2)v - p^e(v,b_2) \geq a(\hat{v},b_1)v - q(\hat{v},b_1)c^e - p^e(\hat{v},b_1), \ \forall v, \hat{v}. \tag{IC-b}$$

$$p^e(t) \leq b + q(t)c^T, \ \forall t \in T, \tag{BC}$$

$$\mathbb{E}_t \left[p^e(t) - q(t)k^e\right] \geq 0,, \tag{BB}$$

$$\mathbb{E}_t \left[a(t)\right] \leq S. \tag{S}$$

Compare $\mathcal{P}'_{DC}$ with $\mathcal{P}'$. It is easy to see that the two problems are identical except for the (BC) constraint. In $\mathcal{P}'_{DC}$, a low-budget agent faces a less stringent budget constraint if he expects to be inspected with a non-zero probability. This is because in the presence of disclosure cost the *effective payment* made by an agent who reports his type truthfully is $p^e(t) = p(t) + q(t)c^T$. In addition to the monetary transfer $p(t)$, disclosure cost $q(t)c^T$ can also be used to screen agents with different valuations. Intuitively, an agent with a higher valuation is also willing to bear a higher disclosure cost. Though disclosure cost can be used to relax an agent's budget constraint, it reduces an agent's utility which makes verification more costly from the principal's perspective, i.e., $k^e = k + c^T \geq k$. As a result, the total welfare effect of introducing c^T is ambiguous.

The *effective punishment* perceived by an agent is now $c + c^F - c^T$, the original punishment plus the additional disclosure cost one must incur when lying about his budget. Hence, an increase in c^F is always welfare-enhancing, as it discourages agents from misreporting their budgets.

Though solving $\mathcal{P}_{DC}$ is beyond the scope of this paper, Proposition 7 provides a sufficient condition under which introducing disclosure costs c^T and c^F improve the total welfare. Let $V(k, c, b_1)$ denote the value of the principal's original problem, in which verification cost is k, punishment is c and low-budget agent's budget is b_1; and let $V_{DC}(k, c, b_1, c^T, c^F)$ denote the value of the principal's problem in which verification cost is k, punishment is c, low-budget agent's budget is b_1 and disclosure costs are c^T and c^F. Then

Proposition 7 *Suppose Assumptions 2 and 3 hold. If* $k/c \geq c^T/(c^F - c^T)$*, then* $V_{DC}(k, c, b_1, c^F, c^T) \geq V(k, c, b_1)$*. Furthermore, if* $q(\overline{v}, b_1) > 0$ *in the optimal mechanism of* $\mathcal{P}(k + c^T, c + c^F - c^T, b_1)$*, then* $V_{DC}(k, c, b_1, c^F, c^T) > V(k, c, b_1)$*.*

3.7.4. Punishing the innocent or without verification

In Appendix B.1, I show that it is without loss of generality to focus on a direct mechanism (a, p, q, θ), where $a : T \to [0, 1]$ denotes the probability an agent obtains the good, $p : T \to \mathbb{R}$ denotes the payment an agent must make, $q : T \to [0, 1]$ denotes the probability of inspecting and $\theta : T \times \{b_1, b_2, n\} \to [0, 1]$ denotes the probability of punishment. In particular, $\theta(\hat{t}, n) \in [0, 1]$ denotes the probability of punishing an agent who reports $\hat{t}$ and is not inspected, and $\theta(\hat{t}, b) \in [0, 1]$ denotes the probability of punishing an agent who reports $\hat{t}$ and is inspected and whose budget is revealed to be b. In the main part of the paper, I assume that $\theta((v, b), n) = \theta((v, b), b) = 0$. In other words, the principal is not allowed to punish an agent without verifying his budget or an agent who is found to have reported his budget truthfully. This assumption is not without loss of generality.

In this case, the utility of an agent who has type $t = (v, b)$ and reports $\hat{t} = (\hat{v}, \hat{b})$ is

$$u(\hat{t}, t) = \begin{cases} a(\hat{t})v - p(\hat{t}) - (1 - q(\hat{t}))\theta(\hat{t}, n)c - q(\hat{t})\theta(\hat{t}, b)c & \text{if } p(\hat{t}) \leq b \\ -\infty & \text{if } p(\hat{t}) > b \end{cases}$$

Then the principal's problem is

$$\max_{a,p,q,\theta} \mathbb{E}_t \left[a(t)v - q(t)k - (1 - q(t))\theta(t, n)c - q(t)\theta(t, b)c \right], \tag{$\mathcal{P}_{PI}$}$$

subject to (IR), (IC), (BC), (BB) and (S). Lemma 10 shows that the principal finds it optimal to *always* punish an agent who is found to have lied about his budget and *never* punish an agent who is found to have reported his budget truthfully.

Lemma 10 *An optimal mechanism of* $\mathcal{P}_{PI}$ *satisfies that (i)* $\theta((\hat{v}, \hat{b}), b) = 1$ *if* $\hat{b} \neq b$ *and (ii)* $\theta((v, b), b) = 0$ *for almost all* (v, b)*.*

Define $p^e(t) := p(t) + (1 - q(t))\theta(t, n)c$, which is the *effective payment* made by an agent who reports his type truthfully. As in Section 3.4, I separate (IC) constraints into two categories and ignore those corresponding to over-reporting of budgets. Using Lemma 10, the principal's problem can be written as:

$$\max_{a,p,q,\theta} \mathbb{E}_t \left[a(t)v - q(t)k - (1 - q(t))\theta(t, n)c \right], \qquad (\mathcal{P}'_{PI})$$

subject to

$$a(t)v - p^e(t) \geq 0, \ \forall t \in T, \qquad \text{(IR)}$$

$$a(v, b)v - p^e(v, b) \geq a(\hat{v}, b)v - p^e(\hat{v}, b), \ \forall v, \hat{v}, b, \qquad \text{(IC-v)}$$

$$a(v, b_2)v - p^e(v, b_2) \geq a(\hat{v}, b_1)v - q(\hat{v}, b_1)c - p^e(\hat{v}, b_1), \ \forall v, \hat{v}. \qquad \text{(IC-b)}$$

$$p^e(t) \leq b + (1 - q(t))\theta(t, n)c, \ \forall t \in T, \qquad \text{(BC)}$$

$$\mathbb{E}_t \left[p^e(t) - q(t)k - (1 - q(t))\theta(t, n)c \right] \geq 0, \qquad \text{(BB)}$$

$$\mathbb{E}_t \left[a(t) \right] \leq S. \qquad \text{(S)}$$

Compare $\mathcal{P}'_{PI}$ with $\mathcal{P}'$. Note that by punishing an agent without verifying his budget, the principal relaxes the agent's budget constraint. However, it is costly, as reflected in the principal's objective function and (BB). Hence, in general, it is unclear whether it is optimal for the principal to do so.

3.7.5. Modified type

In the standard environment, when agents are not budget-constrained, an agent's valuation is defined as the maximum amount of money he is willing to pay for the good. When agents are budget constrained, the natural analogue is the minimum of an agent's valuation v and budget b. I follow Pai and Vohra (2014a) and redefine $t := \min\{v, b\}$ as an agent's *modified type*. In this subsection, I show why it is necessary to explicitly model budget constraint rather than accommodate budgets in the above way.

Let G denote the distribution of the modified type. Then

$$G(t) = \begin{cases} F(t) & \text{if } t < b_1, \\ \pi F(b_1) + 1 - \pi & \text{if } t = b_1, \\ \pi F(t) & \text{if } t > b_1. \end{cases}$$

The principal's ability to inspect an agent's budget implies that she can perfectly learn a low-budget agent's modified type if his valuation exceeds his budget. I first solve the principal's problem by assuming common-knowledge budgets and then verify that no agent has any incentive to misreport his modified type. In other words, there is no inspection in the optimal mechanism. Denote the principal's problem by $\mathcal{P}_{MT}$.

Proposition 8 *Suppose an agent's budget is common knowledge. (i) If* $\pi\left[1 - F(b_1)\right] \leq S < 1 - F(b_1)$*, then the optimal mechanism of* $\mathcal{P}_{MT}$ *is given by*

$$a(t) = \chi_{\{t=b_1\}} \frac{S - \pi\left[1 - F(b_1)\right]}{1-\pi} + \chi_{\{t>b_1\}}, \quad p(t) = \chi_{\{t=b_1\}} \frac{S - \pi\left[1 - F(b_1)\right]}{1-\pi} b_1 + \chi_{\{t\geq b_1\}} b_1.$$

(ii) If $S < \pi\left[1 - F(b_1)\right]$*, then the optimal mechanism is given by*

$$a(t) = \chi_{\{t>t^*\}}, \quad p(t) = \chi_{\{t\geq b_1\}} t^*,$$

where t^* *is such that* $\pi\left[1 - F(t^*)\right] = S$.

The following corollary is a straightforward corollary of Proposition 8.

Corollary 7 *Suppose an agent's budget is his private information. The mechanism given in Proposition 8 is incentive compatible and therefore optimal in* $\mathcal{P}_{MT}$.

Compared with Theorem 9, the mechanism given in Proposition 8 is sub-optimal because (i) it allocates too many resources to high-budget agents; and (ii) it has "too little" rationing for high-budget agents but "too much" rationing for low-budget agents.

What went wrong here? First, consider a low-budget agent with modified type $t = b_1$ and a high-budget agent with modified type $t = b_1 + \epsilon$ for some $\epsilon > 0$. Then the low-budget agent's expected valuation is higher than the high-budget agent's valuation, i.e., $\mathbb{E}[v|t = b_1, b = b_1] > b_1 + \epsilon$, for $\epsilon > 0$ sufficiently small. This implies that the low-budget agent should receive the good with higher probability as in Theorem 9, i.e., $v_1^* \leq v_2^*$. However, in the current mechanism it is the high-budget agent who receives the good with higher probability. Second, consider two low-budget agents with valuations $v = b_1$ and $v' = b_1 + \epsilon$ for $\epsilon > 0$ sufficiently small, respectively. In the current mechanism, they are pooled. However, their payments are $p(b_1) < b_1$, which suggests that they should be separated as in Theorem 9, i.e., $v_1^* > b_1$. The second observation is also made in Pai and Vohra (2014a) in which the principal's objective is maximizing revenue.

3.8. Conclusion

In this paper, I study the problem of a principal who wishes to distribute an indivisible good to a population of budget-constrained agents. Both valuation and budget are an agent's private information. The principal can inspect an agent's budget through a costly verification process and punish an agent who makes a false statement. I characterize the direct surplus-maximizing mechanism. This direct mechanism can be implemented by a two-stage mechanism that exhibits some of the features of Singapore's housing and development board.

Throughout the paper, I impose two regularity assumptions on the distribution of an agent's valuation: monotone hazard rate condition and decreasing density condition. These two assumptions are commonly used in the literature studying mechanism design problem with financially constrained agents. Their primary role is to rule out complicated pooling regions in an optimal mechanism, which greatly simplifies analysis. Several of the paper's results (Lemmas 4, 5 and 6) still hold if I replace these two assumptions by weaker conditions. However, Lemma 7 may not hold anymore as an optimal mechanism is expected to involve more complicated pooling regions.

Another simplifying assumption I make in the paper is that valuation and budget are independent. In some environments, this assumption is reasonable. For example, an individual's valuation of health insurance is largely affected by his or her health risk, which is relatively independent of his or her wealth. In general, an individual's valuation and budget can be either positively or negatively correlated, depending on whether the goods are considered normal goods or inferior goods. The independence assumption is much harder to relax. As Pai and Vohra (2014a) show, if valuation and budget are correlated, an optimal mechanism may involve more complicated pooling regions.

CHAPTER 4 : EFFICIENT MECHANISMS WITH INFORMATION ACQUISITION

4.1. Introduction

Most literature on mechanism design assumes that the amount of information possessed by agents is exogenous. In many important settings, however, this assumption does not apply. For example, in auctions for offshore oil and gas leases in the U.S., companies use seismic surveys to collect information about the tracts offered for sale before participating in the auctions. Another example is the sale of financial or business assets, in which buyers perform due diligence to investigate the quality and compatibility of the assets before submitting offers. In other words, in these settings the information held by agents is endogenous. Moreover, it is costly to acquire information. In the example of U.S. auctions for offshore oil and gas leases (see Haile et al. (2010)), companies can choose to conduct two-dimensional (2-D) or three dimensional (3-D) seismic surveys. 3-D surveys can produce more accurate information, and thus were used in 80% of wells drilled in the Gulf of Mexico by 1996. However, this number was only 5% in 1989 when 3-D surveys were more expensive than 2-D surveys.[1] Similarly, the legal and accounting costs of performing due diligence often amount to millions of dollars in the sale of a business asset (see Quint and Hendricks (2013) and Bergemann et al. (2009)).

Furthermore, the incentives for agents to acquire information depend on the design of a mechanism. This can be seen from earlier studies that compare first price auctions with second price auctions in terms of the incentives they provide for agents to collect information in advance (among them Matthews (1984a), Stegeman (1996) and Persico (2000)). More recently, Bergemann and Välimäki (2002) consider the socially optimal information acquisition in the context of general mechanism design. They focus on mechanisms that implement the ex post efficient allocations given acquired private information, and find that ex ante efficient information acquisition can be achieved if agents have independent private values. However, if agents' values are interdependent, then ex post efficient mechanisms will result in socially sub-optimal information acquisition. In a follow-up paper,

[1] For instance, it costs $1 million to examine a 50 square mile 3-D seismic survey in 1990, while this number was less than $100, 000 in 2000.

Bergemann et al. (2009) study the equilibrium level of information acquisition when agents face binary information decisions, and find that ex post efficient mechanisms result in excessive information acquisition in equilibrium. In summary, there is a conflict between the provision of ex ante efficient incentives to acquire information and the ex post efficient use of information. The question regarding how to design an ex ante efficient mechanism to balance the two trade-offs remains open.

This paper studies the design of ex ante efficient mechanisms in the sale of a single object when agents' values are positively interdependent. The true value of the object to each agent is initially unknown. Before participating in a mechanism agents can simultaneously and independently decide how much information to acquire, and the private information they acquire is independent. Information acquisition is costly and the information choice of each agent is his private information. In this paper, I assume that the information structures are *supermodular ordered*. Ganuza and Penalva (2010) first introduce the notion of "supermodular precision", and Shi (2012) later uses it when studying revenue-maximizing mechanisms in the independent private value setting with endogenous information.

In the main body of the paper, I focus on symmetric mechanisms and symmetric equilibria in which agents acquire the same amount of information before participating in a mechanism. Firstly, I show that the social planner never withholds the item in an ex ante efficient mechanism. Intuitively, whenever the social planner withholds the object, she can also allocate it randomly. By doing so, the allocative efficiency increases while an agent's ex ante incentive to acquire information remains unaffected. Though intuitive, the proof of this result is non-trivial because of the presence of the non-standard information acquisition constraint. In addition, this result facilitates the analysis by allowing us to work with interim allocation rules directly.

Secondly, I show that it is socially optimal for agents to acquire no more information than what they would when the ex post efficient mechanism is used. This is consistent with Bergemann and Välimäki (2002) and Bergemann et al. (2009). For any given information choice satisfying the above condition, I fully characterize the mechanisms that implement this choice and maximize the net social surplus. An ex ante efficient mechanism discourages agents from excessive information acquisition by sometimes randomly allocating the item. Specifically, an ex ante efficient interim allocation rule randomizes in areas in which the accuracy of an agent's posterior estimate can be significantly improved if an additional piece of information is acquired. In the special case in which the improvements are the same for all possible posterior estimates, i.e., the information structures are *uniformly supermodular ordered*, it is optimal to allocate the object uniformly at random with some probability.

Methodologically, when characterizing the optimal mechanisms, I use an approach first proposed by Reid (1968) and later introduced into the mechanism design problem by Mierendorff (2009). My proof, however, is not a straightforward modification of Mierendorff (2009). In Mierendorff (2009), the interim allocation rule is discontinuous at one known point. In my model, the interim allocation rule could be discontinuous at most countably many times, at unknown points.

I also study how the socially optimal information choice is affected by model primitives such as the interdependency of agents' values and the number of agents. In general, it is difficult to solve the optimal information choice analytically. Hence, I restrict attention to the special case in which the information structures are uniformly supermodular ordered. Under this restriction, I show that the optimal level of information gathering decreases as the interdependency of agents' values increases, and gathering no information is optimal in the case of pure common value. Furthermore, when the ex post efficient mechanism is used, the amount of information acquired by each agent diminishes as the number of agents increases to infinity. As a result, the ex post efficient mechanism is also ex ante efficient when the number of agents is large.

Lastly, I study the general ex ante efficient mechanisms without restricting attention to symmetric mechanisms or symmetric equilibria. As in the symmetric case, the social planner never withholds the item in an ex ante efficient mechanism. Because characterizing optimal mechanisms in the general setting is extremely hard, I again restrict attention to the special case in which the information structures are uniformly supermodular ordered. Under this assumption, I provide conditions under which the socially optimal information choices are the same for all agents and there exists a symmetric ex ante efficient mechanism. I also provide an example in which an asymmetric mechanism generates higher expected social surplus than the optimal symmetric mechanism does when these conditions are violated.

This paper is related to the literature studying agents' incentives to acquire information given some commonly used mechanisms. Earlier papers focus on the comparison between first-price and second-price auctions. For example, Matthews (1984a) considers a first-price auction with pure common values, and examines how an increase in the number of agents affects the information acquisition. Stegeman (1996) finds that both auctions lead to identical and, more importantly, efficient incentives for information acquisition when agents' values are private and independent. In contrast, Persico (2000) finds that first-price auctions provide stronger incentive for agents to acquire information than second-price auctions do when their values are affiliated. The two most closely related papers are Bergemann and Välimäki (2002) and Bergemann et al. (2009). Both study the efficiency of information acquisition by agents when ex post efficient mechanisms is used. Instead of focusing on a particular mechanism, this paper studies the design of ex ante efficient mechanisms.

This paper is also related to papers that study the revenue-maximizing mechanisms with endogenous information acquisition. The most closely related paper is Shi (2012) who considers the sale of a single asset when buyers have independent private values and who, before the auction, can simultaneously and independently decide how much information to acquire. He finds that the optimal reserve price is always below the standard monopoly price to encourage information acquisition. Several other papers consider the case where the seller can control the timing of information acquisition (see, for example, Levin and Smith (1994), Ye (2004), and Crémer et al. (2009)).

The rest of the paper is organized as follows. Section 4.2 presents the model. Section 4.3 contains the main results. More specifically, Section 4.3.2 characterizes optimal symmetric mechanisms for each given information choice and Section 4.3.3 studies the socially optimal information choice. Section 4.4 examines ex ante efficient mechanisms without imposing symmetry restrictions. Section 4.5 concludes. All omitted proofs are relegated to appendix.

4.2. Model

There are n agents, indexed by $i \in \{1, \cdots, n\}$, who compete for a single object. Each agent i has a payoff-relevant type θ_i, which is unknown ex ante. The true value of the object to agent i is

$$v_i(\boldsymbol{\theta}) := \theta_i + \gamma \sum_{j \neq i} \theta_j,$$

where $\boldsymbol{\theta} := (\theta_1, \ldots, \theta_n)$ is the type profile, and $\gamma \in [0, 1]$ is a measure of interdependence.[2] Each agent has a quasi-linear utility. Specifically, agent i's payoff is $q_i v_i(\boldsymbol{\theta}) - t_i$ if he gets the object with probability $q_i \in [0, 1]$ and pays $t_i \in \mathbb{R}$.

Initially, agents only know that $\{\theta_i\}$ are independently distributed with a common prior distribution F and support $[\underline{\theta}, \overline{\theta}] \subset \mathbb{R}_+$. F has a positive and continuous density function f over $[\underline{\theta}, \overline{\theta}]$, and its mean is denoted by $\mu := \int_{\underline{\theta}}^{\overline{\theta}} \theta f(\theta) \mathrm{d}\theta$. Each agent i can covertly acquire a costly signal $x_i \in \mathbb{R}$ about his type θ_i by selecting a joint distribution of (x_i, θ_i) from a family of joint distributions $\{G(x_i, \theta_i | \alpha_i)\}$, indexed by $\alpha_i \in \mathbb{A} := [\underline{\alpha}, \overline{\alpha}] \subset \mathbb{R}$. For each $\alpha \in \mathbb{A}$, we also refer the joint distribution $G(\cdot, \cdot | \alpha)$ as an *information structure*. Let g denote the density function associated with G. For each $\alpha \in \mathbb{A}$, $G(\cdot, \cdot | \alpha)$ admits the same marginal distribution of θ, i.e., $\int_{\mathbb{R}} g(x, \theta | \alpha) \mathrm{d}x = f(\theta)$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$. Assume that $\mathbb{E}[\theta | x, \alpha]$ is strictly increasing in x for all $\alpha \in \mathbb{A}$. That is, a higher signal leads to a higher conditional expectation.[3] A signal with a higher α is more precise, in a sense which I formally define below. Let $C(\alpha)$ denote the cost of acquiring a signal with precision α. As is standard in the literature, assume that C is non-negative, strictly increasing, twice continuously differentiable and strictly convex. Furthermore, assume that $C(\underline{\alpha}) = C'(\underline{\alpha}) = 0$.

[2]Under this specification, the range of possible valuations for agent i depends on the number of agents. An alternative normalized specification is given by

$$v_i(\theta) := (1 - \gamma)\theta_i + \frac{\gamma}{n - 1} \sum_{j \neq i} \theta_j.$$

With this specification, the range of possible valuations for agent i does not depend on the number of agents, but now the single-crossing condition does. Most of my results hold for either specifications.

[3]For each $\alpha \in \mathbb{A}$, let $G(\theta | x, \alpha)$ denote the conditional distribution of θ given x. Then one sufficient condition for this is to assume that $G(\theta | x, \alpha)$ have the monotone likelihood ratio property.

4.2.1. Information order

Let $G(x|\alpha)$ denote the marginal distribution of signal x given precision α. I define a new signal by applying the probability integral transformation on the original signal. Let $s := G(x|\alpha)$. This transformed signal s is uniformly distributed on $[0, 1]$.[4] Clearly, the transformed signal has the same informational content as the original signal. Furthermore, because any two transformed signals have the same marginal distribution, their realizations are directly comparable. Therefore, I will henceforth work with the transformed signal directly. Let $w(s, \alpha) := \mathbb{E}[\theta|s, \alpha]$ be the conditional expectation of θ given signal s and precision α. Then, by assumption, $w(s, \alpha)$ is strictly increasing in s. For each $\alpha \in \mathbb{A}$, let $H(w|\alpha) := \mathbb{P}(w(s, \alpha) \le w)$ denote the cumulative distribution function of $w(s, \alpha)$, and $h(w|\alpha)$ denote its corresponding density function. I assume that both $H(w|\alpha)$ and $h(w|\alpha)$ are twice continuously differentiable in w and α. Throughout the paper, I assume that the information structures are *supermodular ordered*:

Definition 1 *The information structures are* supermodular ordered *if for all $\alpha \in \mathbb{A}$,*

$$-\frac{H_\alpha(w|\alpha)}{h(w|\alpha)} \text{ is strictly increasing in } w \text{ on } [w(0, \alpha), w(1, \alpha)].$$

To understand Definition 1, note that $w_\alpha(s, \alpha) = -H_\alpha(w(s, \alpha)|\alpha)/h(w(s, \alpha)|\alpha)$ which is strictly increasing in s. Hence, $w(s, \alpha)$ is supermodular in (s, α). Formally, I prove in the appendix that if the information structures are supermodular ordered, then $w(s, \alpha)$ satisfies the following property:[5]

Lemma 11 *Suppose that the information structures are supermodular ordered. Then $w(\cdot, \cdot)$ is*

[4] s is uniform on $[0, 1]$ only if $G(x)$ is continuous and strictly increasing. This can be assumed without loss of generality. If G has a discontinuity at z, where $\mathbb{P}(\tilde{x} = z) = p$, x can be transformed into x^*, which has a continuous and strictly increasing distribution function using the following construction proposed in Lehmann (1988): $x^* = x$ for $x < z$, $x^* = x + pU$ if $x = z$, where U is uniform on $(0, 1)$, and $x^* = x + p$ for $x > z$.

[5] Lemma 11 is not an equivalent definition of supermoduler ordered information structures. If $w(\cdot, \cdot)$ is strictly supermodular, i.e., satisfies inequality (4.1), then $-H_\alpha(\cdot|\alpha)/h(\cdot|\alpha)$ is non-decreasing, but not necessarily strictly increasing. I conjecture that all the results still hold if we only require that $-H_\alpha(\cdot|\alpha)/h(\cdot|\alpha)$ is non-decreasing, but the stronger assumption simplifies analysis.

strictly supermodular: for all $s, s' \in (0,1)$*,* $s > s'$ *and* $\alpha > \alpha'$

$$w(s,\alpha) - w(s',\alpha) > w(s,\alpha') - w(s',\alpha'). \tag{4.1}$$

Intuitively, if s contains little information about θ, then $w(s,\alpha)$ does not vary much as s changes and its distribution concentrates around μ. As s becomes more informative about θ, $w(s,\alpha)$ varies more as s changes and its distribution becomes more dispersed. Formally, if $\alpha > \alpha'$ then $w(s,\alpha)$ is strictly larger than $w(s,\alpha')$ in the *dispersive order*.[6] Based on this notion of dispersion, Ganuza and Penalva (2010) first introduce the information order called "supermodular precision" . Shi (2012) also assumes that the information structures are supermodular ordered for some of his results.

For some results of the paper, I further require that the information structures are *uniformly supermodular ordered*. Recall that if an information structure is more precise, then $w(s,\alpha)$ changes more dramatically as s changes, i.e., $w_s(s,\alpha)$ is larger. Hence, we can interpret $w_s(s,\alpha)$ as a local measure of the information structure's precision around s. Then, $w_{s,\alpha}(s,\alpha)/w_s(s,\alpha)$ is the percentage change of the information structure's precision around s as α increases. We say the information structures are uniformly supermodular ordered if

$$\frac{w_{s,\alpha}(s,\alpha^*)}{w_s(s,\alpha^*)} = \frac{\partial}{\partial w}\left[-\frac{H_\alpha(w(s,\alpha^*)|\alpha^*)}{h(w(s,\alpha^*)|\alpha^*)}\right]$$

is independent of s (or equivalently w). In other words, when α increases, the information structure becomes more precise "uniformly" over $[0,1]$. The formal definition is given as follows:

Definition 2 *The information structures are* uniformly supermodular ordered *if there exists a positive function* $b : \mathbb{A} \to \mathbb{R}_{++}$ *such that, for all* $\alpha \in \mathbb{A}$ *and* $w \in [w(0,\alpha), w(1,\alpha)]$*,*

$$-\frac{H_\alpha(w|\alpha)}{h(w|\alpha)} = \frac{w-\mu}{b(\alpha)}.$$

[6](See Ganuza and Penalva (2010)) Let Y and Z be two real-valued random variables with distributions F and G, respectively. We say Y is greater than Z in the *dispersive order* if for all $q, p \in (0,1)$ and $q > p$,

$$F^{-1}(q) - F^{-1}(p) \geq G^{-1}(q) - G^{-1}(p).$$

When the information structures are uniformly supermodular ordered, we can obtain a sharper and simpler characterization of the optimal mechanisms. The following two commonly used information technologies in the literature are uniformly supermodular ordered:[7]

Example 3 (Linear experiments) *Consider the following information structure, which is called "truth-or-noise" in Lewis and Sappington (1994), Johnson and Myatt (2006) and Shi (2012). With probability* $\alpha_i \in [0,1]$*,* x_i *is equal to agent* i*'s true type* θ_i*, and with probability* $1-\alpha_i$*,* x_i *is an independent draw from* F*. Since the marginal distribution of* x_i *is* F*, the transformed signal is* $s_i = F(x_i)$*. Then the posterior estimate of an agent who chooses* α_i *and receives* s_i *is* $w(s_i,\alpha_i) = \alpha_i F^{-1}(s_i) + (1-\alpha_i)\mu$*. It is easy to verify that*

$$-\frac{H_{\alpha_i}(w_i|\alpha_i)}{h(w_i|\alpha_i)} = \frac{w_i-\mu}{\alpha_i}.$$

Hence, the information structures are uniformly supermodular ordered.

Example 4 (Normal experiments) *Let* $\{\theta_i\}$ *be independently distributed with a normal distribution:* $\theta_i \overset{iid}{\sim} \mathcal{N}(\mu, 1/\beta)$*. Agent* i *can obtain a costly signal* $x_i = \theta_i + \varepsilon_i$*, where* $\varepsilon_i \overset{iid}{\sim} \mathcal{N}(0, 1/\alpha_i)$*. Since the marginal distribution of* x_i *is also normal, i.e.,* $x_i \sim \mathcal{N}(\mu, (\beta+\alpha_i)/\beta\alpha_i)$*, the transformed signal is* $s_i = \Phi\left(\sqrt{\beta\alpha_i}(x_i-\mu)/\sqrt{\beta+\alpha_i}\right)$*, where* Φ *is the distribution function of the standard normal distribution. Then the posterior estimate of an agent who chooses* α_i *and receives* s_i *is*

$$w(s_i,\alpha_i) = \mu + \frac{\sqrt{\alpha_i}\Phi^{-1}(s_i)}{\sqrt{\beta(\alpha_i+\beta)}}.$$

It is easy to verify that

$$-\frac{H_{\alpha_i}(w_i|\alpha_i)}{h(w_i|\alpha_i)} = \frac{\beta(w_i-\mu)}{2\alpha_i(\alpha_i+\beta)}.$$

Hence, the information structures are uniformly supermodular ordered.

[7] See, for example, Ganuza and Penalva (2010) and Shi (2012).

4.2.2. Timing

The game proceeds in the following way: The social planner announces a mechanism. After observing the mechanism, agents simultaneously choose their information structures $\{\alpha_i\}$ and observe the realized signals $\{s_i\}$. Both α_i and s_i are agent i's private information. Then agents simultaneously decide whether to participate in the mechanism. Each participating agent reports his private information. Based on their reports, an allocation and payments are implemented according to the announced mechanism.

I assume that the payoff structure, the timing of the game, the family of information structures and the prior distribution F are common knowledge. The solution concept is Bayesian Nash equilibrium.

4.2.3. Mechanisms

The private information of agent i is two-dimensional, including the choice of information structure α_i and the realized signal s_i. However, similar to Biais et al. (2000), Szalay (2009) and Shi (2012), the usual difficulties inherent in multi-dimensional mechanism design problem do not arise here. This is because the posterior estimate, $w(s_i, \alpha_i)$, summarizes all the private information needed to compute agent i's expected valuation of the object:

$$\mathbb{E}_{\boldsymbol{\theta}}[v_i(\boldsymbol{\theta})|\alpha_i, s_i] = w(s_i, \alpha_i) + \gamma \sum_{j \neq i} \mathbb{E}[\theta_j].$$

Furthermore, the social planner cannot screen agents with the same posterior estimate but different choices of information structures. Hence, I can appeal to the revelation principle and focus on direct mechanisms in which agents report their posterior estimates directly. For ease of notation, I use w_i to denote $w_i(s_i, \alpha_i)$ and $\boldsymbol{w} := (w_1, \dots, w_n)$ denote a vector or posterior estimates. A direct mechanism is a pair $(\boldsymbol{q}, \boldsymbol{t})$, where $\boldsymbol{q} := (q_1, \dots, q_n)$ and $\boldsymbol{t} := (t_1, \dots, t_n)$. For $i = 1, \dots, n$, $q_i : [\underline{\theta}, \overline{\theta}]^n \to [0, 1]$ is agent i's allocation rule and $t_i : [\underline{\theta}, \overline{\theta}]^n \to \mathbb{R}$ is agent i's payment rule. Specifically, given a reported vector of posterior estimates $\boldsymbol{w}$, agent i receives the object with probability $q_i(\boldsymbol{w})$, and makes a payment $t_i(\boldsymbol{w})$ to the social planner. I note here that the message space for each agent in a direct mechanism is $[\underline{\theta}, \overline{\theta}]$ because, without further knowledge on agents' choices of information structures, any $w_i \in [\underline{\theta}, \overline{\theta}]$ can arise in the game.

Given a mechanism $(\boldsymbol{q}, \boldsymbol{t})$, let $\boldsymbol{\alpha}^* := (\alpha_1^*, \ldots, \alpha_n^*)$ denote the equilibrium vector of information structures. Define agent i's interim allocation rule as

$$Q_i(w_i) := \mathbb{E}_{w_{-i}}[q_i(w_i, w_{-i})|\alpha_{-i}^*], \ \forall w_i \in [\underline{\theta}, \overline{\theta}], \tag{4.2}$$

where α_{-i}^* are his opponents' information structures. Then the interim utility of agent i who has a posterior estimate w_i and reports w_i' is

$$U_i(w_i, w_i') := w_i Q_i(w_i') + \mathbb{E}_{w_{-i}}\left[\gamma\left(\sum_{j\neq i} w_j\right) q_i(w_i', w_{-i}) - t_i(w_i', w_{-i}) \middle| \alpha_{-i}^*\right].$$

Note that $Q_i(w_i)$ and $U_i(w_i, w_i')$ also depend on α_{-i}^*, and I suppress the dependence for ease of notation.

I require that the mechanism chosen by the social planner must satisfy the following constraints. First, the mechanism must be individually rational (IR):

$$U_i(w_i) := U_i(w_i, w_i) \geq 0, \ \forall w_i \in [\underline{\theta}, \overline{\theta}], \tag{IR}$$

so that the agents are willing to participate in the mechanism. Because the social planner's goal is to maximize social surplus and transfers between agents and the social planner do not affect social surplus, we can guarantee that (IR) is satisfied by making sufficiently large lump sum transfers to agents. Hence, we can safely ignore (IR) for the remainder of the paper. Second, the mechanism must be Bayesian incentive compatible (IC):

$$U_i(w_i) \geq U_i(w_i, w_i'), \forall w_i, w_i' \in [\underline{\theta}, \overline{\theta}], \tag{IC}$$

so that truth-telling is a Bayesian Nash equilibrium. Lastly, because the information structure chosen by an agent is unobservable, the mechanism must also satisfy the information acquisition constraint

(IA). That is, no agent stands to gain by deviating from his equilibrium choice: for each agent i,

$$\alpha_i^* \in \operatorname*{argmax}_{\alpha_i} \mathbb{E}_{\boldsymbol{w}} \left[q_i(\boldsymbol{w}) \left(w_i + \gamma \sum_{j \neq i} \mathbb{E}[\theta_j] \right) - t_i(\boldsymbol{w}) \middle| \alpha_i, \alpha_j = \alpha_j^* \ \forall j \neq i \right] - C(\alpha_i). \tag{IA}$$

Then the social planner's problem is to choose a mechanism $(\boldsymbol{q}, \boldsymbol{t})$ and a vector of recommendations of information structures $\boldsymbol{\alpha}^*$ to maximize the social surplus:

$$\max_{\boldsymbol{\alpha}^*, (\boldsymbol{q}, \boldsymbol{t})} \mathbb{E} \left[\sum_i (w_i + \gamma \sum_{j \neq i} w_j) q_i(\boldsymbol{w}) \middle| \alpha_i = \alpha_i^* \ \forall i \right] - \sum_i C(\alpha_i^*),$$

subject to (IC), (IA) and the feasibility constraint (F):,

$$0 \leq q_i(\boldsymbol{w}) \leq 1, \ \sum_i q_i(\boldsymbol{w}) \leq 1, \forall \boldsymbol{w}. \tag{F}$$

We say a mechanism is *ex post efficient* if for all i, $q_i(\boldsymbol{w}) = 1$ if $w_i > \max_j w_j$ and $q_i(\boldsymbol{w}) = 0$ if $w_i < \max_j w_j$. We say a mechanism $(\boldsymbol{q}, \boldsymbol{t})$ is *ex ante efficient* or *optimal* if there exists $\boldsymbol{\alpha}^*$ such that $\boldsymbol{\alpha}^*$ and $(\boldsymbol{q}, \boldsymbol{t})$ solve the social planner's problem.

4.3. Efficient mechanisms

In this section, I restrict attention to mechanisms that treat all agents symmetrically[8] as well as symmetric equilibria in which all agents acquire the same information structure, i.e., $\alpha_i^* = \alpha^*$ for all i. This restriction significantly simplifies the analysis, but it may result in a loss of generality. Section 4.4 presents a study of ex ante efficient mechanisms without imposing this symmetry restriction, and provides conditions under which the socially optimal information choices are the same for all agents and there exists a symmetric ex ante efficient mechanism. Note that when $\boldsymbol{q}$ is symmetric and all agents acquire the same information structure, the corresponding interim allocation rule Q_i is independent of i. From here on, I drop the subscript i from Q, w and α whenever the meaning is clear.

[8]The formal definition of symmetric mechanisms can be found at the beginning of Appendix C.1.

Now the social planner's problem becomes:

$$\max_{\alpha^*,(\boldsymbol{q},\boldsymbol{t})} \mathbb{E}\left[\sum_i (w_i + \gamma \sum_{j\neq i} w_j) q_i(\boldsymbol{w}) \middle| \alpha_i = \alpha^* \ \forall i\right] - nC(\alpha^*),$$

subject to (IC), (IA) and (F).

By the standard argument (see, for example, Myerson (1981)), (IC) holds if and only if

$$Q(w) \text{ is non-decreasing in } w, \tag{MON}$$

and $U_i(w)$ is absolutely continuous and satisfies the following envelope condition

$$U_i(w) = U_i(\underline{\theta}) + \int_{\underline{\theta}}^{w} Q(\tilde{w})\mathrm{d}\tilde{w}, \ \forall w \in [\underline{\theta}, \overline{\theta}]. \tag{4.3}$$

Suppose that agent i chooses α_i, then his expected payoff is

$$\begin{aligned}
&\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} U_i(w)\mathrm{d}H(w|\alpha_i) - C(\alpha_i) \\
=&\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} \left[U_i(\underline{\theta}) + \int_{\underline{\theta}}^{w} Q(\tilde{w})\mathrm{d}\tilde{w}\right] \mathrm{d}H(w|\alpha_i) - C(\alpha_i) \\
=&U_i(\underline{\theta}) + \int_{\underline{\theta}}^{w(1,\alpha_i)} Q(w)\mathrm{d}w - \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} H(w|\alpha_i)Q(w)\mathrm{d}w - C(\alpha_i) \\
=&U_i(w(0,\alpha_i)) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} \left[1 - H(w_i|\alpha_i)\right] Q_i(w_i)\mathrm{d}w_i - C(\alpha_i),
\end{aligned}$$

where the first and the last lines hold by the envelope condition (4.3) and the second line holds by integration by parts.

Next, I replace (IA) by a one-sided first-order necessary condition. In an earlier paper, Bergemann and Välimäki (2002) show that if the social planner adopts the ex post efficient mechanism, then agents tend to acquire more information than the socially desired level. This result suggests that an ex ante efficient mechanism would distort the allocation of the object to discourage agents from acquiring information. Hence, I hypothesize that to ensure that (IA) holds in an ex ante efficient mechanism, it suffices to ensure that no agent has incentives to acquire more precise information than recommended: for all $\alpha_i > \alpha^*$,

$$\begin{aligned}
&U_i(w(0,\alpha^*)) + \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \left[1 - H(w|\alpha^*)\right] Q(w)\mathrm{d}w - C(\alpha^*) \\
&\geq U_i(w(0,\alpha_i)) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} \left[1 - H(w|\alpha_i)\right] Q(w)\mathrm{d}w - C(\alpha_i).
\end{aligned}$$

This implies the following one-sided first-order condition:[9]

$$\mathbb{E}_w \left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)} Q(w) \middle| \alpha_i = \alpha^* \right] \leq C'(\alpha^*). \tag{IA$'$}$$

The left-hand side of the above inequality is agent i's marginal benefit from acquiring α_i^*, and the right-hand side is the marginal cost. In Lemma 43 in the appendix, I show that for any non-decreasing Q_i, an agent's marginal benefit from acquiring information is non-negative. Subsequently, I consider the relaxed problem of the social planner by replacing (IA) by (IA$'$). I later show that (IA$'$) holds with equality when α^* is chosen optimally. The first-order approach is valid if the second-order condition of the agents' optimization problem is satisfied. Appendix C.1.3 provides sufficient conditions that ensure the first-order approach is valid.

Although (IA$'$) is easier to work with than (IA), it is still nonstandard and prevents me from solving the social planner's problem directly as in Myerson (1981). To overcome this difficulty, I focus on reduced form auctions. Formally,

Definition 3 *An allocation rule* $\boldsymbol{q}$ implements $Q : [\underline{\theta}, \overline{\theta}] \to [0,1]$ *and* Q *is the* reduced form *of* $\boldsymbol{q}$ *if* $\boldsymbol{q}$ *satisfies* (4.2) *and* (F) *for all* $w \in [\underline{\theta}, \overline{\theta}]$. Q *is* implementable *if* $\boldsymbol{q}$ *exists implementing* Q.

One important prior result I use in this paper is the necessary and sufficient condition of Maskin and Riley (1984a), Matthews (1984c) and Border (1991), which characterizes the interim allocation rules implementable by symmetric mechanisms. By Theorem 1 in Matthews (1984c), any non-decreasing

[9]The main reason why I consider a one-sided first-order condition here is to sign the Lagrangian multiplier associated with (IA$'$). Admittedly, this relaxation also simplifies the proof of Theorem 14. But my conjecture is that Theorem 14 can be proved even if we require the first-order condition holds with equality.

function $Q : [\underline{\theta}, \overline{\theta}] \to [0,1]$ is implementable if and only if it satisfies

$$Y(w) := \int_{w}^{\overline{\theta}} \left[H(z|\alpha^*)^{n-1} - Q(z) \right] h(z|\alpha^*) \mathrm{d}z \geq 0, \ \forall w \in [\underline{\theta}, \overline{\theta}]. \tag{F'}$$

The above condition says that the probability of assigning the object to an agent whose posterior estimate is above w, $n \int_{w}^{\overline{\theta}} Q(z) h(z|\alpha^*) \mathrm{d}z$, must not exceed the probability with which there exists an agent whose posterior estimate is above w, $1 - H(w|\alpha^*)^n = n \int_{w}^{\overline{\theta}} H(z|\alpha^*)^{n-1} h(z|\alpha^*) \mathrm{d}z$. Clearly, this is a necessary condition for Q to be implementable. If Q is non-decreasing, then Theorem 1 in Matthews (1984c) shows that it is also sufficient. Hence, given (MON), we can replace (F) by (F′). Note that the support of w is $W := [w(0, \alpha^*), w(1, \alpha^*)] \subset [\underline{\theta}, \overline{\theta}]$. Therefore, (F′) imposes no restriction on Q in outside W.

Now all three constraints ((IC), (IA) and (F)) are replaced by constraints ((MON), (IA′) and (F′)) that are expressed as functionals of the interim allocation rule Q. In order to work with reduced form auctions, we must express the social planner's objective function or the expected social surplus as a functional of Q as well. This exercise is trivial when agents have independent private values ($\gamma = 0$), because in this case an agent's expected valuation of the object and his winning probability are independent conditional on his private information. In general, this is impossible when agents' values are interdependent ($\gamma > 0$), because in this case both an agent's expected valuation of the object and his winning probability depend on other agents' private information. Nonetheless, we can still write the expected social surplus as a functional of Q if Q is the reduced form of an ex ante efficient allocation rule, which never withholds the object:

Theorem 14 *Suppose that the information structures are supermodular ordered, and α^* and $(\boldsymbol{q}, \boldsymbol{t})$ solve the relaxed problem of the social planner. Then*

$$\sum_i q_i(\boldsymbol{w}) = 1 \textit{ for almost all } \boldsymbol{w} \in W^n. \tag{4.4}$$

The proof of Theorem 14 can be found in Section 4.3.1. Using Theorem 14 and the law of iterated expectations, the social planner's objective function can be rewritten as a functional of Q:

$$\mathbb{E}_{\boldsymbol{w}}\left[\sum_i (w_i + \gamma \sum_{j\neq i} w_j) q_i(\boldsymbol{w}) \middle| \alpha_i = \alpha^* \ \forall i\right] = \sum_i \mathbb{E}_{w_i}\left[(1-\gamma) w_i Q(w_i) \middle| \alpha_i = \alpha^*\right] + n\gamma\mu.$$

Because the second term, $n\gamma\mu$, is a constant, we ignore it from here on. To summarize, the social planner's relaxed problem, denoted by ($\mathcal{P}'$), is as follows:

$$\max_{\alpha^*, Q} \mathbb{E}_w \left[(1-\gamma) w Q(w) \middle| \alpha^*\right] - C(\alpha^*), \tag{$\mathcal{P}'$}$$

subject to

$$Y(w) = \int_w^{\overline{\theta}} [H(z|\alpha^*)^{n-1} - Q(z)] h(z|\alpha^*) \mathrm{d}z \geq 0, \ \forall w \in [\underline{\theta}, \overline{\theta}]. \tag{F$'$}$$

$$Q(w) \text{ is non-decreasing in } w, \tag{MON}$$

$$\mathbb{E}_w \left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)} Q(w) \middle| \alpha^* \right] \leq C'(\alpha^*). \tag{IA$'$}$$

In addition to being instrumental in solving the social planner's problem, the result of Theorem 14 also has some inherent economic interest. Obviously, when information is exogenous, the efficient mechanism never withholds the object. This is not obvious when information is endogenous, because, by withholding the object occasionally, the social planner can discourage agents from acquiring information, which may improve efficiency ex ante. However, intuitively, whenever the social planner withholds the object, she can also allocate it randomly. By doing so, the ex post allocative efficiency improves while an agent's ex ante incentive to acquire information remains unaffected.

Though intuitive, the proof of Theorem 14 is non-trivial. This is because the resulting mechanism, by simply randomizing the object whenever it is withheld, is likely to violate (MON) or (IA$'$). To illustrate this difficulty, let A be a set of types such that $\sum_i q_i(\boldsymbol{w}) < 1$ whenever $\boldsymbol{w} \in A^n$, and suppose there exists an interval $(\underline{w}, \overline{w})$ such that $(\underline{w}, \overline{w}) \cap A = \emptyset$ and $\inf A < \underline{w} < \overline{w} < \sup A$.

If we simply redefine $\boldsymbol{q}$ such that it remains unchanged outside A^n and $\sum_i q_i(\boldsymbol{w}) = 1$ for all $\boldsymbol{w} \in A^n$, then the resulting Q remains unchanged for all $w \in (\underline{w}, \overline{w})$ but increases for all $w \in A$. If we allocate the object too often to agents whose types are in $[\underline{\theta}, \underline{w}] \cap A$, the resulting Q will no longer be non-decreasing and therefore violate (MON). If we allocate the object too often to agents whose types are in $[\overline{w}, \overline{\theta}] \cap A$, this will increase an agent's marginal benefit from acquiring information and lead to a violation of (IA′). Hence, to ensure that the new $\boldsymbol{q}$ generates a higher social surplus while satisfies all the constraints, we must adjust $\boldsymbol{q}$ not only inside A^n, but also outside A^n.

4.3.1. Proof of Theorem 14

This section contains the proof of Theorem 14. The readers who are not interested in the proof may skip this section and proceed directly to Section 4.3.2 without loss of continuity.

I prove Theorem 14 by proving Lemmas 12 and 13. Observe first that if $\alpha_i = \alpha^*$ for all i, then $Y(w(0, \alpha^*))$ is equal to 1 minus the probability of assigning the object to some agent. Clearly, (4.4) is violated if and only if $Y(w(0, \alpha^*)) > 0$. Then

Lemma 12 *Suppose that the information structures are supermodular ordered, and $\alpha_i = \alpha^*$ for all i. Let Q be any interim allocation rule satisfying* (F′)*,* (MON)*,* (IA′) *and $Y(w(0, \alpha^*)) > 0$, then there exists $\hat{Q}$ satisfying* (F′)*,* (MON) *and* (IA′) *such that*

$$\hat{Q}(w) \geq Q(w), \ \forall w \in W, \tag{4.5}$$

and strict inequalities hold for a set of w with positive measure.

The intuition behind the proof of Lemma 12 can be illustrated by Figure 12. Suppose that Q satisfies the assumptions in Lemma 12, then one can construct a $\hat{Q}$ by increasing Q at the lower end of its domain as in Figure 12. Clearly, the resulting $\hat{Q}$ is non-decreasing and implementable if the change is sufficiently small. It remains to verify that $\hat{Q}$ also satisfies (IA′). Intuitively, agents have fewer incentives to acquire information if outcomes are less sensitive to changes in their private information, which is the case when they face a less steep allocation rule. In Lemma 43 in Appendix C.1, I show that if $\hat{Q}$ is less steep than Q in the sense that it differs from Q by a non-increasing

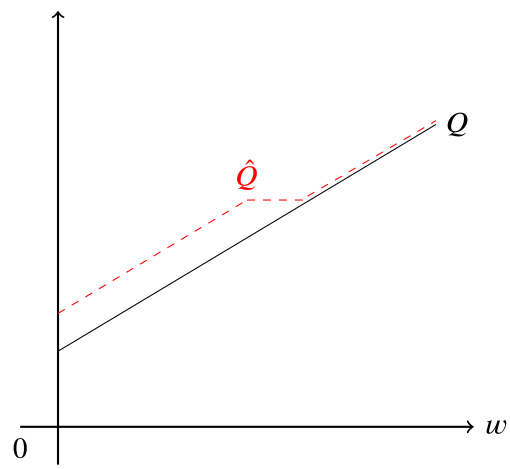


Figure 12: Proof idea of Lemma 12

function (as in Figure 12), then for any amount of information acquired (or any α), $\hat{Q}$ gives agents a smaller marginal benefit of acquiring more information. Hence, $\hat{Q}$ satisfies (IA′) as Q does.

The gap between Lemma 12 and Theorem 14 is that when $\gamma > 0$, the expected social surplus

$$\mathbb{E}_{\boldsymbol{w}}\left[\sum_i (w_i + \gamma \sum_{j \neq i} w_j) q_i(\boldsymbol{w}) \middle| \alpha_i = \alpha^* \ \forall i\right]$$

does not directly depend on Q. To prove Theorem 14, we need to show that, for any ex-post allocation rule $\boldsymbol{q}$ implementing Q, we can find a $\hat{\boldsymbol{q}}$ that implements $\hat{Q}$ and yields higher social surplus. This is the result of the following Lemma 13.

Lemma 13 *Suppose that the information structures are supermodular ordered, and $\alpha_i = \alpha^*$ for all i. Let Q and $\hat{Q}$ be two implementable allocation rules satisfying* (4.5). *Let $\boldsymbol{q}$ be an ex-post allocation rule that implements Q. Then there exists an ex-post allocation rule $\hat{\boldsymbol{q}}$ that implements $\hat{Q}$ and satisfies*

$$\mathbb{E}_{\boldsymbol{w}}\left[\sum_i (w_i + \gamma \sum_{j \neq i} w_j) \hat{q}_i(\boldsymbol{w}) \middle| \alpha_i = \alpha^* \ \forall i\right] > \mathbb{E}_{\boldsymbol{w}}\left[\sum_i (w_i + \gamma \sum_{j \neq i} w_j) q_i(\boldsymbol{w}) \middle| \alpha_i = \alpha^* \ \forall i\right].$$

The proof of Lemma 13 relies on the following technical lemma. I abuse notation a bit and let h

denote the probability measure on W corresponding to $H(w_i|\alpha^*)$, then

Lemma 14 *Let $Q : W \to [0,1]$ be an interim allocation rule and $\rho : W^n \to [0,1]$ be a symmetric measurable function. Then there exists a symmetric ex post allocation rule $\boldsymbol{q}$ that implements Q and satisfies $\sum_i q_i(\boldsymbol{w}) \geq \rho(\boldsymbol{w})$ for almost all $\boldsymbol{w} \in W^n$ if and only if for all measurable sets $A \subset W$, the following inequality holds*

$$\int_{A^n} \rho(\boldsymbol{w})\mathrm{d}h^n(\boldsymbol{w}) \leq n \int_A Q(w_i)\mathrm{d}h(w_i) \leq \int_{A^n} \mathrm{d}h^n(\boldsymbol{w}). \tag{4.6}$$

To see that inequality (4.6) is necessary, suppose that there exists an ex post allocation rule $\boldsymbol{q}$ that implements Q and satisfies $\sum_i q_i(\boldsymbol{w}) \geq \rho(\boldsymbol{w})$ for almost all $\boldsymbol{w} \in W^n$. For any measurable set $A \subset W$, the probability with which an agent whose type is in A receives the object is given by $n \int_A Q(w_i)\mathrm{d}h(w_i)$. This probability cannot exceed the probability that there exists an agent whose type is in A, $\int_{A^n} \mathrm{d}h^n(\boldsymbol{w})$; and must exceed the probability with which an agent receives the object when all agents' types are in A, $\int_{A^n} \sum_i q_i(\boldsymbol{w})\mathrm{d}h^n(\boldsymbol{w})$, which is greater than $\int_{A^n} \rho(\boldsymbol{w})\mathrm{d}h^n(\boldsymbol{w})$ by assumption. In Appendix C.1.1, I show that (4.6) is also sufficient. Note also that if $A = [w, \overline{\theta}]$, then the second inequality in (4.6) becomes (F′).

With Lemma 14 in hand, it is easy to prove Lemma 13.

Proof of Lemma 13. Consider two implementable allocation rules Q and $\hat{Q}$ satisfying (4.5). Let $\boldsymbol{q}$ be a symmetric ex-post allocation rule that implements Q. Define $\rho : W^n \to [0,1]$ by $\rho(\boldsymbol{w}) := \sum_i q_i(\boldsymbol{w})$ for all $\boldsymbol{w} \in W^n$. Then ρ is symmetric. By Lemma 14,

$$\int_{A^n} \mathrm{d}h^n(\boldsymbol{w}) \geq n \int_A \hat{Q}(w_i)\mathrm{d}h(w_i) \geq n \int_A Q(w_i)\mathrm{d}h(w_i) \geq \int_{A^n} \rho(\boldsymbol{w})\mathrm{d}h^n(\boldsymbol{w}).$$

By Lemma 14, there exists an allocation rule $\hat{\boldsymbol{q}}$ that implements $\hat{Q}$ and satisfies $\sum \hat{q}_i(\boldsymbol{w}) \geq \rho(\boldsymbol{w}) =$

$\sum_i q_i(\boldsymbol{w})$ for almost all $\boldsymbol{w} \in W^n$. Hence,

$$
\begin{aligned}
&\mathbb{E}_{\boldsymbol{w}}\left[\sum_i\left(w_i+\gamma\sum_{j\neq i}w_j\right)\hat{q}_i(\boldsymbol{w})\middle|\,\alpha_i=\alpha^*\ \forall i\right]\\
&=\sum_i\mathbb{E}_{w_i}\left[(1-\gamma)w_i\hat{Q}(w_i)\middle|\,\alpha_i=\alpha^*\right]+\mathbb{E}_{\boldsymbol{w}}\left[\left(\gamma\sum_i w_i\right)\left(\sum_i\hat{q}_i(\boldsymbol{w})\right)\middle|\,\alpha_i=\alpha^*\ \forall i\right]\\
&>\sum_i\mathbb{E}_{w_i}\left[(1-\gamma)w_iQ(w_i)\middle|\,\alpha_i=\alpha^*\right]+\mathbb{E}_{\boldsymbol{w}}\left[\left(\gamma\sum_i w_i\right)\left(\sum_i q_i(\boldsymbol{w})\right)\middle|\,\alpha_i=\alpha^*\ \forall i\right]\\
&=\mathbb{E}_{\boldsymbol{w}}\left[\sum_i\left(w_i+\gamma\sum_{j\neq i}w_j\right)q_i(\boldsymbol{w})\middle|\,\alpha_i=\alpha^*\ \forall i\right],
\end{aligned}
$$

where the strict inequality holds because Q and $\hat{Q}$ satisfies (4.5) and $\sum \hat{q}_i(\boldsymbol{w}) \geq \sum_i q_i(\boldsymbol{w})$ for almost all $\boldsymbol{w} \in W^n$. This completes the proof. ∎

4.3.2. Optimal mechanisms for fixed α^*

I solve the principal's relaxed problem ($\mathcal{P}'$) in two steps. In this subsection, I solve the following sub-problem for each $\alpha^* \in \mathbb{A}$, denoted by ($\mathcal{P}'$-α^*):

$$V(\alpha^*) := \max_{Q} \mathbb{E}_w\left[wQ(w)\middle|\,\alpha^*\right], \qquad (\mathcal{P}'\text{-}\alpha^*)$$

subject to (F′), (MON) and (IA′). In Section 4.3.3, I solve $\max_{\alpha\in\mathbb{A}}(1-\gamma)V(\alpha) - C(\alpha)$.

Fix α^*. If the principal adopts the ex post efficient mechanism, then the interim allocation rule is given by $Q(w) = H(w|\alpha^*)^{n-1}$ for all w. Clearly, if α^* is such that

$$\mathbb{E}_w\left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)}H(w|\alpha^*)^{n-1}\middle|\,\alpha^*\right] \leq C'(\alpha^*), \tag{4.7}$$

then the ex post efficient mechanism solves ($\mathcal{P}'$-α^*). Hence, in the rest of this subsection, I assume that α^* is such that (4.7) is violated. In what follows, I consider two cases in turn. In Section 4.3.2, I first solve a relaxed problem by ignoring the monotonicity constraint (MON), and then show that if the information structures are uniformly supermodular ordered then the solutions of the relaxed problem automatically satisfy (MON). In 4.3.2, I consider the general case when the information structures are supermodular ordered. In this case, the solutions of the relaxed problem violates (MON) in general. In the main text, I present an informal arugument to derive the optimal solutions of ($\mathcal{P}'$-α^*) using Myerson (1981)'s ironing procedure. The formal analysis can be found in Appendix C.1.2.

Optimal mechanisms in the regular case

If we ignore the monotonicity constraint (MON), then we can use the following Lagrangian relaxation to get an intuition of the optimal solution:

$$\mathscr{L} := \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \varphi^{\lambda_X}\left(H(w|\alpha^*),\alpha^*\right) Q(w) h(w|\alpha^*) \mathrm{d}w + \lambda_X C'(\alpha^*), \tag{4.8}$$

where $\lambda_X > 0$ is the Lagrangian multiplier associated with (IA$'$) and $\varphi^{\lambda_X}(\cdot,\alpha^*)$ is defined by[10]

$$\varphi^{\lambda_X}(t,\alpha^*) := H^{-1}(t|\alpha^*) + \lambda_X \frac{H_\alpha(H^{-1}(t|\alpha^*)|\alpha^*)}{h(H^{-1}(t|\alpha^*)|\alpha^*)},\ \forall t \in [0,1].$$

Note that because $H(\cdot|\alpha)$ is strictly increasing, $\varphi^{\lambda_X}(\cdot,\alpha^*)$ is strictly increasing (or decreasing) if and only if $\varphi^{\lambda_X}(H(\cdot|\alpha^*),\alpha^*)$ is strictly increasing (or decreasing). Here, $\varphi^{\lambda_X}(H(w|\alpha^*),\alpha^*)$ can be viewed as the *"virtual value"* associated with posterior estimate w:

$$\varphi^{\lambda_X}(H(w|\alpha^*),\alpha^*) = w + \lambda_X \frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)}.\ \forall w \in W.$$

The standard virtual value in a revenue maximization problem is defined as the difference between a type's true value and the information rents necessary to induce truthtelling. Here, as in the standard virtual value, the first term, w, is an agent's posterior estimate of his type. In the case of private values ($\gamma = 0$), this is equal to his expected value of the object. Because agents can acquire information, we must subtract, from an agent's posterior estimate, λ_X multiplied by the marginal change of an

[10] I define $\varphi^{\lambda_X}(\cdot,\alpha^*)$ as a function of percentiles rather than posterior estimates simply to make it easier to define "ironed virtual values" later when the pointwise virtual surplus maximizer violates (MON).

agent's posterior estimate if he acquires more precise information:

$$-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)} = w_\alpha(s, \alpha^*),$$

where s is such that $w(s, \alpha^*) = w$. If the information structures are supermodular ordered, then $-H_\alpha(w|\alpha^*)/h(w|\alpha^*)$ is strictly increasing in w. Hence, the virtual value function is less steeper than the identity function. In the ex post efficient mechanism, an agent is rewarded based on his posterior estimate. In contrast, in an ex ante efficient mechanism, an agent is rewarded based on his virtual value which is less sensitive to his private information than his posterior estimate does. Intuitively, this difference would discourage agents from acquiring excessive information. Finally, because the social planner's goal is to maximize the social surplus rather than her revenue, the inverse hazard rate associated with the information rents does not appear.

If λ_X is chosen optimally, then the optimal solution to ($\mathcal{P}'$-α^*) can be obtained by maximizing the virtual surplus pointwise. If, in addition, the virtual values are non-decreasing in w, then there exists a pointwise virtual surplus maximizer that is non-decreasing and therefore incentive compatible.

This method works in the simple case in which the information structures are uniformly supermodular ordered. Recall that in this case we have

$$-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)} = \frac{w-\mu}{b(\alpha^*)}, \ \forall w.$$

Hence, the virtual values are given by

$$\varphi^{\lambda_X}(H(w|\alpha^*), \alpha^*) = w - \lambda_X \frac{w-\mu}{b(\alpha^*)}, \ \forall w.$$

I first argue that the optimal λ_X is equal to $b(\alpha^*)$. Suppose that $\lambda_X < b(\alpha^*)$, then the pointwise virtual surplus maximizer is the ex post efficient allocation rule: $Q(w) = H(w|\alpha^*)^{n-1}$ for all w. However, because, by assumption, α^* is such that (4.7) is violated, (IA$'$) is violated. This is a contradiction. Hence, $\lambda_X \geq b(\alpha^*)$. Suppose that $\lambda_X > b(\alpha^*)$, then $\varphi^{\lambda_X}(H(w|\alpha^*), \alpha^*)$ is strictly decreasing. Hence, the pointwise virtual value maximizer is strictly decreasing and not incentive compatible (or violates (MON)). It is easy to see that the interim allocation rule Q that maximizes the expected virtual surplus and satisfies (MON) is constant. However, in this case (IA$'$) holds with strict inequality, which implies that $\lambda_X = 0$ and contradicts to hypothesis that $\lambda_X > b(\alpha^*) > 0$. Hence, the optimal λ_X is equal to $b(\alpha^*)$.

If $\lambda_X = b(\alpha^*)$, then the virtual values are given by $\varphi^{\lambda_X}(w, \alpha^*) = \mu$ for all w, which is constant. Hence, any feasible non-decreasing allocation rule Q satisfying condition (4.4) maximizes the expected virtual surplus. If Q also satisfies (IA$'$) with equality, then it solves ($\mathcal{P}'$-α^*). In particular, there exists $\xi \in [0, 1]$ such that the following interim allocation rule solves ($\mathcal{P}'$-α^*):

$$\hat{Q}(w) = \xi H(w|\alpha^*)^{n-1} + (1 - \xi)\frac{1}{n}, \ \forall w.$$

Recall that (4.4) holds if and only if $Y(w(0, \alpha^*)) = 0$. These arguments prove the following Proposition 9.

Proposition 9 *Suppose that the first-order approach is valid, and the information structures are uniformly supermodular ordered. Suppose, in addition, that α^* is such that* (4.7) *is violated. Then Q solves* ($\mathcal{P}'$-α^*) *if and only if Q is non-decreasing, $Y(w(0, \alpha^*)) = 0$ and Q satisfies* (IA$'$) *with equality.*

Optimal mechanisms in the general case

If the information structures are not uniformly supermodular ordered, then typically the pointwise virtual surplus maximizer is not incentive compatible (or violates (MON)), and ironing is necessary. In particular, an optimal solution can be obtained by ironing $\varphi^{\lambda_X}(\cdot, \alpha^*)$ in the following procedure first introduced by Myerson (1981). For each $t \in [0, 1]$, define

$$J^{\lambda_X}(t, \alpha^*) := \int_0^t \varphi^{\lambda_X}(\tau, \alpha^*)\mathrm{d}\tau.$$

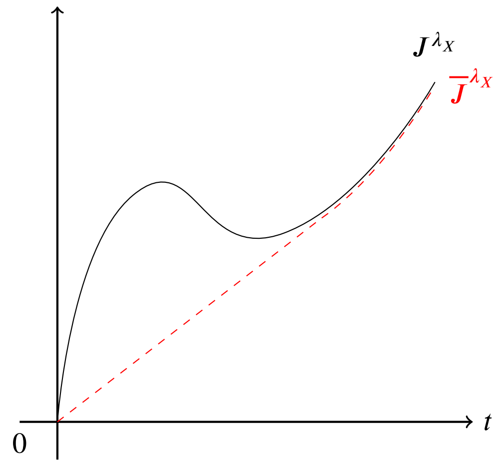


Figure 13: Ironing

Let $\overline{J}^{\lambda_X}$ denote the convex hull of J^{λ_X}, defined by

$$\overline{J}^{\lambda_X}(t,\alpha^*) := \min\left\{\beta J(t_1,\alpha^*) + (1-\beta)J(t_2,\alpha^*) | t_1, t_2 \in [0,1], \beta t_1 + (1-\beta)t_2 = t\right\}, \ \forall t \in [0,1].$$

This is illustrated by Figure 13. Because $\overline{J}^{\lambda_X}(\cdot,\alpha^*)$ is convex, it is continuously differentiable virtually everywhere. Define the *ironed virtual value* $\overline{\varphi}^{\lambda_X}(\cdot,\alpha^*)$ as follows. First, for each $t \in (0,1)$ such that $\partial \overline{J}^{\lambda_X}(t,\alpha^*)/\partial t$ exists, let $\overline{\varphi}^{\lambda_X}(t,\alpha^*) := \partial \overline{J}^{\lambda_X}(t,\alpha^*)/\partial t$. Then extend $\overline{\varphi}^{\lambda_X}(\cdot,\alpha^*)$ to $[0,1]$ by right continuity. Because $\overline{J}^{\lambda_X}(\cdot,\alpha^*)$ is convex, $\overline{\varphi}^{\lambda_X}(\cdot,\alpha^*)$ is non-decreasing. By construction, $J^{\lambda_X}(H(\cdot|\alpha^*),\alpha^*) \geq \overline{J}^{\lambda_X}(H(\cdot|\alpha^*),\alpha^*)$. If $J^{\lambda_X}(H(w|\alpha^*),\alpha^*) > \overline{J}^{\lambda_X}(H(w|\alpha^*),\alpha^*)$ for some w, then $\overline{J}^{\lambda_X}(H(\cdot|\alpha^*),\alpha^*)$ is linear and therefore $\overline{\varphi}^{\lambda_X}(H(\cdot|\alpha^*),\alpha^*)$ is constant in a neighborhood of w. (See Figure 13.)

Suppose that λ_X is chosen optimally. I argue that, any interim allocation rule Q solves ($\mathcal{P}'$-α^*) if and only if (i) Q satisfies (IA$'$) with equality; (ii) $Y(w(0,\alpha^*)) = 0$; and (iii) Q satisfies the following two *pooling properties*:

1. If $J^{\lambda_X}(H(w|\alpha^*),\alpha^*) > \overline{J}^{\lambda_X}(H(w|\alpha^*),\alpha^*)$ for all $w \in (\underline{w},\overline{w})$ and let $(\underline{w},\overline{w})$ be chosen maximally, then Q is constant on $(\underline{w},\overline{w})$.

2. If $Y(w) > 0$ for all $w \in (\underline{w},\overline{w})$ and let $(\underline{w},\overline{w})$ be chosen maximally, then $\overline{\varphi}^{\lambda_X}(H(\cdot|\alpha^*),\alpha^*)$ is

constant on $(\underline{w}, \overline{w})$.

The proof is based on Toikka (2011). For the ease of notation, I suppress the dependence of φ^{λ_X}, $\overline{\varphi}^{\lambda_X}$, J^{λ^X} and $\overline{J}^{\lambda_X}$ on λ_X. Then the expected social surplus is

$$\begin{aligned}
&\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} zQ(z)h(z|\alpha^*)\mathrm{d}z \\
\leq& \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \left[z + \lambda_X \frac{H_\alpha(z|\alpha^*)}{h(z|\alpha^*)}\right] Q(z)h(z|\alpha^*)\mathrm{d}z + \lambda_X C'(\alpha^*) && (4.9)\\
=& \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \varphi^{\lambda_X}(H(z|\alpha^*))Q(z)\mathrm{d}H(z|\alpha^*) + \lambda_X C'(\alpha^*)\\
=& \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \left[\varphi^{\lambda_X}(H(z|\alpha^*)) - \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))\right] Q(z)\mathrm{d}H(z|\alpha^*)\\
&+ \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))Q(z)\mathrm{d}H(z|\alpha^*) + \lambda_X C'(\alpha^*)\\
=& - \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \left[J^{\lambda_X}(H(z|\alpha^*)) - \overline{J}^{\lambda_X}(H(z|\alpha^*))\right] \mathrm{d}Q(z)\\
&+ \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))Q(z)\mathrm{d}H(z|\alpha^*) + \lambda_X C'(\alpha^*) && (4.10)\\
\leq& \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))Q(z)\mathrm{d}H(z|\alpha^*) + \lambda_X C'(\alpha^*) && (4.11)\\
=& \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))H(z|\alpha^*)^{n-1}\mathrm{d}H(z|\alpha^*) + \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))Y'(z)\mathrm{d}z + \lambda_X C'(\alpha^*) && (4.12)\\
\leq& \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))H(z|\alpha^*)^{n-1}\mathrm{d}H(z|\alpha^*) - \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} Y(z)\mathrm{d}\overline{\varphi}^{\lambda_X}(H(z|\alpha^*)) + \lambda_X C'(\alpha^*) && (4.13)\\
=& \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))H(z|\alpha^*)^{n-1}\mathrm{d}H(z|\alpha^*) + \lambda_X C'(\alpha^*). && (4.14)
\end{aligned}$$

Here, the inequality (4.9) holds because $\lambda_X \geq 0$ and Q satisfies (IA′); and the equality holds if and only if Q satisfies (IA′) with equality. (4.10) follows from integration by parts. The inequality (4.11) holds because $J^{\lambda_X} \geq \overline{J}^{\lambda_X}$; and the equality holds if and only if Q satisfies the first pooling property. (4.12) follows from the definition of Y. The inequality (4.13) follows from integration by parts and the fact that $Y((w(0,\alpha^*))) \geq 0$; and the equality holds if and only if $Y(w(0,\alpha^*)) = 0$. Finally, the equality (4.14) because $Y \geq 0$; and the equality holds if and only if Q satisfies the second pooling property.

It remains to determine the optimal multiplier λ_X. In order to do so, I first define the “steepest" allocation rule Q^+ and the “least steep" allocation rule Q^- satisfying conditions (ii) and (iii) given above. Define $Q^+(\cdot, \lambda_X)$ as follows. If $J^{\lambda_X}(H(w|\alpha^*), \alpha^*) > \overline{J}^{\lambda_X}(H(w|\alpha^*), \alpha^*)$ for $w \in (\underline{w}, \overline{w})$ and let $(\underline{w}, \overline{w})$ be chosen maximally, then let

$$Q^+(w, \lambda_X) := \frac{\frac{1}{n}[H(\overline{w}|\alpha^*)^n - H(\underline{w}|\alpha^*)^n]}{H(\overline{w}|\alpha^*) - H(\underline{w}|\alpha^*)}, \; \forall w \in (\overline{w}, \underline{w}).$$

Otherwise, let $Q^+(w, \lambda^X) := H(w|\alpha^*)^{n-1}$. Define $Q^-(\cdot, \lambda_X)$ as follows. If $\overline{\varphi}^{\lambda_X}(H(\cdot|\alpha^*), \alpha^*)$ is constant on $(\underline{w}, \overline{w})$ with $\underline{w} < \overline{w}$ and let $(\underline{w}, \overline{w})$ be chosen maximally, then let

$$Q^-(w, \lambda_X) := \frac{\frac{1}{n}[H(\overline{w}|\alpha^*)^n - H(\underline{w}|\alpha^*)^n]}{H(\overline{w}|\alpha^*) - H(\underline{w}|\alpha^*)}, \; \forall w \in (\overline{w}, \underline{w}).$$

Otherwise, let $Q^-(z, \lambda_X) := H(z|\alpha^*)^{n-1}$. Clearly, both Q^+ and Q^- are non-decreasing, implementable and satisfy conditions (ii) and (iii). I demonstrate Corollary 12 in Appendix C.1.2 that Q^+ is the “steepest" allocation rule and Q^- is the “least steep" allocation rule among all non-decreasing implementable Q's satisfying conditions (ii) and (iii) in the following sense: for all non-decreasing implementable Q's satisfying conditions (ii) and (iii),

$$\mathbb{E}\left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)}Q^+(w, \lambda_X)\middle|\alpha^*\right] \geq \mathbb{E}\left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)}Q(w)\middle|\alpha^*\right] \geq \mathbb{E}\left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)}Q^-(w, \lambda_X)\middle|\alpha^*\right].$$

Hence, there exists a non-decreasing implementable Q satisfying conditions (i)-(iii) if and only if λ_X is such that

$$\mathbb{E}\left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)}Q^+(w, \lambda_X)\middle|\alpha^*\right] \geq C'(\alpha^*) \geq \mathbb{E}\left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)}Q^-(w, \lambda_X)\middle|\alpha^*\right]. \tag{4.15}$$

Lemma 15 proves that such a λ_X exists and is unique, and its proof can be found in Appendix C.1.2.

Lemma 15 *Suppose that the first-order approach is valid and the information structures are supermodular ordered. Suppose, in addition, that α^* is such that* (4.7) *is violated. There exists a unique $\lambda_X > 0$ such that inequality* (4.15) *holds.*

The main result of this section is the following Theorem 15, which demonstrates that the unique λ_X given in Lemma 15 is indeed optimal, and the allocation rules I have derived above solve ($\mathcal{P}'$-α^*):

Theorem 15 *Suppose that the first-order approach is valid and the information structures are supermodular ordered. Suppose, in addition, that α^* is such that* (4.7) *is violated. Let $\lambda_X > 0$ be such that inequality* (4.15) *holds and Q be a non-decreasing implementable allocation rule. Then Q solves* ($\mathcal{P}'$-α^*) *if and only if $Y(w(0, \alpha^*)) = 0$, and Q satisfies* (IA$'$) *with equality and the two pooling properties.*

My interpretation of the optimal pooling areas is as follows: Optimally, pooling occurs where $\varphi^{\lambda_X}(H(\cdot|\alpha^*), \alpha^*)$ is not strictly increasing, i.e.,

$$\frac{w_{s,\alpha}(s, \alpha^*)}{w_s(s, \alpha^*)} = \frac{\partial}{\partial w}\left[-\frac{H_\alpha(w(s, \alpha^*)|\alpha^*)}{h(w(s, \alpha^*)|\alpha^*)}\right] \geq \frac{1}{\lambda_X}$$

Recall that if an information structure is more precise, then $w(s, \alpha)$ changes more dramatically as s changes, i.e., $w_s(s, \alpha)$ is larger. Hence, one can interpret $w_s(s, \alpha)$ as a local measure of the information structures' precision around s. Then, $w_{s,\alpha}(s, \alpha)/w_s(s, \alpha)$ is the percentage change of the information structures' precision around s as α increases. Intuitively, the most effective way to discourage agents from acquiring too much information is to introduce randomization to where the information structures' precision increases most rapidly. If the information structures are uniformly supermodular ordered, then $w_{s,\alpha}(s, \alpha)/w_s(s, \alpha)$ is a constant. In other words, when α increases, the information structure becomes more precision *uniformly* over $[0, 1]$.

Lastly, given Theorem 15, it is straightforward that there exists an optimal Q which takes the following relatively simple form:

Corollary 8 *There exists $\xi \in [0, 1]$ such that $\hat{Q}(w) := \xi Q^+(w, \lambda_X) + (1 - \xi)Q^-(w, \lambda_X)$ solves*

($\mathcal{P}'$-α^*).

Though the result is intuitive, the proof of Theorem 15 is difficult because of the presence of both the non-standard constraint (IA$'$) and (MON). In this paper, I use the following approach first proposed by Reid (1968) and later introduced into the mechanism design problem by Mierendorff (2009). I first solve ($\mathcal{P}'$-α^*) under an additional restriction, that Q is Lipschitz continuous with global Lipschitz constant K:

$$|Q(w) - Q(w')| \leq K|w - w'|, \ \forall w, w' \in W.$$

Denote the modified maximization problem by ($\mathcal{P}^K$-α^*). I show that the optimal solutions of ($\mathcal{P}^K$-α^*) converge to that of ($\mathcal{P}'$-α^*) as $K \to \infty$. Then I can obtain a characterization of the optimal solutions of ($\mathcal{P}'$-α^*) in the limit. The formal analysis can be found in Appendix C.1.2.

The proof is not a straightforward modification of Mierendorff (2009). Let Q and Q^K denote the optimal solutions to ($\mathcal{P}^K$-α^*) and ($\mathcal{P}'$-α^*), respectively. In Mierendorff (2009), Q is discontinuous at exactly one point, and it can be shown that for K sufficiently large, the slope of Q^K is equal to K only in a neighborhood around the discontinuity point. In this paper, however, Q could be discontinuous at most countably many times, at unknown points. If Q is discontinuous at w, then it is possible that every neighborhood of w contains another discontinuity point. Hence, it is non-trivial to characterize Q as the limit of Q^K.

I conclude this subsection by briefly discussing why I cannot apply control they directly. In the published version of Mierendorff (2009), Mierendorff (2016) does not use the approach described above, and directly appeals to Theorems 7 and 8 in Seierstad and Sydsæter (1987). However, the problem considered here, ($\mathcal{P}'$-α^*), is more complex for the following two reasons. First, as I have mentioned above, in Mierendorff (2016), state variable Q is discontinuous at exactly one point, while in ($\mathcal{P}'$-α^*), Q could be discontinuous at most countably many points. Second, the problem in Mierendorff (2016) can be written as a control problem without restrictions on the state variables, while ($\mathcal{P}'$-α^*) contains *pure state constraints* (constraints in which control variables do not appear). To the best of my knowledge, no existing theorem can be applied to provide necessary and sufficient conditions for the optimal solutions of ($\mathcal{P}'$-α^*).

4.3.3. *Optimal α^**

Given optimal solutions of ($\mathcal{P}'$-α^*), I can now study the optimal information choices. Let $\pi^s(\alpha) := (1-\gamma)V(\alpha) - C(\alpha)$ and $\alpha^* \in \operatorname{argmax}_{\alpha\in\mathbb{A}} \pi^s(\alpha)$. I show in Appendix C.1.2 that

$$\pi^s(\alpha) = (1-\gamma)\int_{w(0,\alpha)}^{w(1,\alpha)} \overline{\varphi}^{\lambda_X}(H(w|\alpha),\alpha)H(w|\alpha)^{n-1}h(w|\alpha)\mathrm{d}w + (1-\gamma)\lambda_X C'(\alpha) - C(\alpha),$$

where, as demonstrated in Theorem 15, λ_X depends on α in a complex way. In general, it is hard to solve the optimal α^*. In this section, I first give a condition that the optimal α^* must satisfy and then solve it when the information structures are uniformly supermodular ordered.

In Lemma 63 in Appendix C.1.3, I show that if the second-order condition of the agents' optimization problem is satisfied, then

$$\int_{w(0,\alpha)}^{w(1,\alpha)} -H_\alpha(w|\alpha)H(w|\alpha)^{n-1}\mathrm{d}w - C'(\alpha) \text{ is strictly decreasing in } \alpha. \tag{4.16}$$

Let

$$\hat{\alpha} := \inf\left\{\alpha \in \mathbb{A} | (4.7) \text{ holds for } \alpha\right\}. \tag{4.17}$$

Then $\hat{\alpha}$ is independent of γ and $\lim_{n\to\infty}\hat{\alpha} = \underline{\alpha}$. By (4.16), inequality (4.7) holds for all $\alpha > \hat{\alpha}$. I claim that the socially optimal $\alpha^* \leq \hat{\alpha}$. Note first that for all $\alpha > \hat{\alpha}$, the optimal solution to ($\mathcal{P}'$-α^*) is $Q(w) = H(w|\alpha)^{n-1}$. Hence, the average social surplus is

$$\pi^s(\alpha) = (1-\gamma)\int_{w(0,\alpha)}^{w(1,\alpha)} wH(w|\alpha)^{n-1}h(w|\alpha)\mathrm{d}w - C(\alpha).$$

Taking derivative with respect to α gives

$$\pi^{s\prime}(\alpha) = (1-\gamma)\int_{w(0,\alpha)}^{w(1,\alpha)} -H_\alpha(w|\alpha)H(w|\alpha)^{n-1}\mathrm{d}w - C'(\alpha).$$

Because $C'(\alpha)$ is strictly increasing, $\pi^{s\prime}(\alpha)$ is strictly decreasing by (4.16). By construction, $\pi^{s\prime}(\hat{\alpha}) =$

$-\gamma C'(\hat{\alpha}) \leq 0$. Hence, $\pi^{s\prime}(\alpha) < 0$ for all $\alpha > \hat{\alpha}$ and at optimum $\alpha^* \leq \hat{\alpha}$. This result is summarized in the following proposition:

Proposition 10 *Suppose that the second-order condition of the agents' optimization problem is satisfied, and the information structures are supermodular ordered. The socially optimal information choice α^* is such that $\alpha^* \leq \hat{\alpha}$, where $\hat{\alpha}$ is such that* (4.7) *holds with equality if $\alpha = \hat{\alpha}$.*

Proposition 10 states it is not optimal for the social planner to encourage agents to acquire more information than they would under the ex post efficient mechanism. This result is not surprising given the results of Bergemann et al. (2009). Proposition 10 also implies that (IA′) always holds with equality when α^* is chosen optimally.[11] Hence, it is sufficient to consider the one-sided first-order condition.

To obtain further results about the socially optimal information choice α^*, I assume that the information structures are uniformly supermodular ordered for the rest of this section. In this case, the average social surplus is

$$\pi^s(\alpha) = (1-\gamma)\left[\frac{\mu}{n} + b(\alpha)C'(\alpha)\right] - C(\alpha),\ \forall \alpha \in [\underline{\alpha}, \hat{\alpha}],$$

Hence,

$$\pi^{s\prime}(\alpha) = \left[(1-\gamma)b'(\alpha) - 1\right] C'(\alpha) + (1-\gamma)b(\alpha)C''(\alpha),\ \forall \alpha \in [\underline{\alpha}, \hat{\alpha}].$$

Hence, $\pi^{s\prime}(\alpha) \leq 0$ if and only if $1 - C'(\alpha)/\left[b'(\alpha)C'(\alpha) + b(\alpha)C''(\alpha)\right] \leq \gamma$. Assume that $C'(\alpha)/\left[b'(\alpha)C'(\alpha) + b(\alpha)C''(\alpha)\right]$ is strictly increasing. Then there exists a unique $\alpha^\circ \in \mathbb{A}$ such that $\pi^{s\prime}(\alpha) \lesseqgtr 0$ if and only if $\alpha \gtreqless \alpha^\circ$. Furthermore, α° is strictly decreasing in γ and independent of n, and $\lim_{\gamma\to 0} \alpha^\circ = \underline{\alpha}$. Therefore the socially optimal information choice is $\alpha^* = \min\{\alpha^\circ, \hat{\alpha}\}$. This result is summarized by the following proposition:

[11] If $\alpha^* = \hat{\alpha}$, then the ex post efficient mechanism is optimal and (IA′) holds with equality by the definition of $\hat{\alpha}$. Suppose that $\alpha^* < \hat{\alpha}$. Suppose, to the contrary, that (IA′) holds with strict inequality. Then the ex post efficient mechanism is optimal. However, because inequality (4.7) is violated when $\alpha^* < \hat{\alpha}$, (IA′) is violated, which is a contradiction.

Proposition 11 *Suppose that the second-order condition of the agents' optimization problem is satisfied, and the information structures are uniformly supermodular ordered. Suppose, in addition, that* $C'(\alpha)/\left[b'(\alpha)C'(\alpha) + b(\alpha)C''(\alpha)\right]$ *is strictly increasing in* α*. At optimum,* $\alpha^*(n, \gamma) = \min\{\alpha^\circ(\gamma), \hat{\alpha}(n)\}$*, where* α° *is strictly decreasing in* γ*,* $\lim_{\gamma \to 0} \alpha^\circ = \underline{\alpha}$ *and* $\lim_{n \to \infty} \hat{\alpha} = \underline{\alpha}$*.*

Proposition 11 implies that the ex post efficient mechanism is also ex ante efficient if the level of interdependency is low. As the level of interdependency increases, the socially optimal information choice decreases, and an ex ante efficient mechanism introduces more randomization into the allocation rule to discourage agents from acquiring too much information. Proposition 11 also implies that the ex post efficient mechanism is also ex ante efficient for a sufficiently large number of agents. Intuitively, when there is a large number of agents, the incentive for each agent to acquire information is already small and there is no need for the social planner to further discourage them from acquiring information by distorting the allocation rule.

Example 5 (Linear experiments) *Consider the information structures in Example 3. Assume that* $F(\theta) = \theta$ *with support* $[0, 1]$*, and the cost function (used in Persico (2000)) is of the form*

$$C(\alpha) = K\left(\alpha - \underline{\alpha}\right)^2, \ \forall \alpha \in [\underline{\alpha}, 1],$$

where $0 < \underline{\alpha} < 1$ *and* $K \geq 1/8\underline{\alpha}$*. Then, as I demonstrate in Appendix C.1.3, the first-order approach is valid. In this case* $\hat{\alpha}$ *is such that*

$$2K(\hat{\alpha} - \underline{\alpha}) = \frac{n-1}{2n(n+1)}.$$

Note that the left-hand side of the above equation is strictly increasing in $\hat{\alpha}$*; and the right-hand side is strictly decreasing in* n *for* $n \geq 2$ *and converges to* 0 *as* n *goes to infinity. Hence,* $\hat{\alpha}$ *is strictly decreasing in* n *and goes to* $\underline{\alpha}$ *as* n *goes to infinity. Finally,*

$$\pi^{s\prime}(\alpha) = 2K\left[\gamma\underline{\alpha} - (2\gamma - 1)\alpha\right].$$

If $\gamma \leq \frac{1}{2}$, then $\pi^{s\prime}(\alpha) \geq 0$ for all α and therefore $\alpha^* = \hat{\alpha}$. If $\gamma > \frac{1}{2}$, then $\pi^{s\prime}(\alpha)$ is strictly decreasing in α and therefore

$$\alpha^* = \min\left\{\frac{\gamma\underline{\alpha}}{2\gamma - 1}, \hat{\alpha}\right\}.$$

Thus, if γ sufficiently small or n sufficiently large, then $\alpha^* = \hat{\alpha}$, and the ex post efficient mechanism is also ex ante efficient. If γ sufficiently large or n sufficiently small, then the optimal α^* is strictly decreasing in γ, and goes to $\underline{\alpha}$ as γ increases to 1.

4.4. Efficient asymmetric mechanisms

In this section, I study the ex ante efficient mechanisms without restricting attention to symmetric mechanisms and symmetric equilibria. First, I show that the result in Theorem 14 is still valid for general asymmetric mechanisms, i.e., an ex ante efficient mechanism never withholds the object. Second, I derive the ex ante efficient mechanisms when the information structures are uniformly supermodular ordered, and provide conditions under which the socially optimal information choices are the same for all agents and there exists a symmetric ex ante efficient mechanism. Finally, I give an example in which an asymmetric mechanism generates higher net social surplus that the optimal symmetric mechanism.

As in the symmetric case, I consider the relaxed problem of the social planner by replacing (IA) by the one-sided first-order conditions

$$\mathbb{E}_w\left[-\frac{H_{\alpha_i}(w|\alpha_i^*)}{h(w|\alpha_i^*)}Q(w)\middle|\,\alpha_i = \alpha_i^*\right] \leq C'(\alpha_i^*),\ \forall i. \tag{AIA$'$}$$

and focus on reduced form auctions. Let $\boldsymbol{Q} := (Q_1, \ldots, Q_n)$, where $Q_i : [\underline{\theta}, \overline{\theta}] \to [0,1]$ is non-decreasing for all i. By Theorem 3 in Mierendorff (2011), $\boldsymbol{Q}$ is implementable if and only if it satisfies

$$\sum_{i=1}^{n}\int_{w_i}^{w(1,\alpha_i^*)} Q_i(z_i)\mathrm{d}H(z_i|\alpha_i^*) \leq 1 - \prod_{i=1}^{n} H(w_i|\alpha_i^*),\ \forall \boldsymbol{w} \in \prod_{i=1}^{n}[w(0,\alpha_i^*), w(1,\alpha_i^*)]. \tag{AF$'$}$$

Thus, given (MON), we can replace (F) by (AF$'$). Finally, as in the symmetric case, an ex ante efficient mechanism never withholds the object:

Theorem 16 *Suppose that the information structures are supermodular ordered, and $\boldsymbol{\alpha}^*$ and $(\boldsymbol{q}, \boldsymbol{t})$ solve the relaxed problem of the social planner. Then*

$$\sum_i q_i(\boldsymbol{w}) = 1 \textit{ for almost all } \boldsymbol{w} \in \prod_{i=1}^{n} [w(0, \alpha_i^*), w(1, \alpha_i^*)]. \tag{4.18}$$

By Theorem 16 and the Law of iterated expectations, we can rewrite the social planner's objective function as a function of $\boldsymbol{Q}$:

$$\mathbb{E}_{\boldsymbol{w}} \left[\sum_i (w_i + \gamma \sum_{j \neq i} w_j) q_i(\boldsymbol{w}) \middle| \alpha_i = \alpha_i^* \ \forall i \right] = \sum_i \mathbb{E}_{w_i} \left[(1-\gamma) w_i Q_i(w_i) \middle| \alpha_i = \alpha_i^* \right] + n\gamma\mu.$$

Because the second term, $n\gamma\mu$, is a constant, we ignore it from here on. To summarize, the social planner's relaxed problem, denoted by $(\mathcal{P}')$, becomes:

$$\max_{\boldsymbol{\alpha}^*, \boldsymbol{Q}} \mathbb{E}_{\boldsymbol{w}} \left[\sum_i (1-\gamma) w_i Q_i(w_i) \middle| \alpha_i = \alpha_i^* \ \forall i \right] - \sum_i C(\alpha_i^*),$$

subject to

$$\sum_{i=1}^{n} \int_{w_i}^{w(1,\alpha_i^*)} Q_i(z_i) \mathrm{d}H(z_i|\alpha_i^*) \leq 1 - \prod_{i=1}^{n} H(w_i|\alpha_i^*), \ \forall \boldsymbol{w} \in \prod_{i=1}^{n} [w(0, \alpha_i^*), w(1, \alpha_i^*)], \tag{AF$'$}$$

$$Q_i(w_i) \text{ is non-decreasing in } w_i, \ \forall i, \tag{MON}$$

$$\mathbb{E}_{w_i} \left[-\frac{H_{\alpha_i}(w_i|\alpha_i^*)}{h(w_i|\alpha_i^*)} Q_i(w_i) \middle| \alpha_i = \alpha_i^* \right] \leq C'(\alpha_i^*), \ \forall i. \tag{AIA$'$}$$

As in the symmetric case, I solve $(\mathcal{P}')$ in two steps. First, for each $\boldsymbol{\alpha}^* \in \mathbb{A}^n$, I solve the following sub-problem, denoted by $(\mathcal{P}'\text{-}\boldsymbol{\alpha}^*)$:

$$V(\boldsymbol{\alpha}^*) := \max_{\boldsymbol{Q}} \mathbb{E}_{\boldsymbol{w}} \left[\sum_i w_i Q_i(w_i) \middle| \boldsymbol{\alpha}^* \right] \text{ subject to (AF$'$), (MON) and (AIA$'$)},$$

Second, I solve $\max_{\boldsymbol{\alpha} \in \mathbb{A}^n} \pi^s(\boldsymbol{\alpha}) := (1-\gamma) V(\boldsymbol{\alpha}) - \sum_i C(\alpha_i)$.

Fix $\boldsymbol{\alpha}^*$. If the principal adopts the ex post efficient mechanism, then the interim allocation rule is given by $Q_i(w_i) = \prod_{j\neq i} H(w_i|\alpha_i^*)$ for all w_i and all i. Clearly, if $\boldsymbol{\alpha}^*$ is such that

$$\mathbb{E}_{w_i}\left[-\frac{H_{\alpha_i}(w_i|\alpha_i^*)}{h(w_i|\alpha_i^*)}\prod_{j\neq i} H(w_i|\alpha_j^*)\middle|\boldsymbol{\alpha}^*\right] \leq C'(\alpha_i^*), \forall i, \tag{4.19}$$

then the ex post efficient mechanism solves $(\mathcal{P}'\text{-}\boldsymbol{\alpha}^*)$. Furthermore, I show Lemma 67 in the appendix that if $\boldsymbol{\alpha}^*$ is chosen optimally, then (AIA$'$) holds with equality for all i and

$$\pi^s(\boldsymbol{\alpha}^*) = (1-\gamma)\left[\mu + \sum_{i=1}^n b(\alpha_i^*)C'(\alpha_i^*)\right] - \sum_i C(\alpha_i^*). \tag{4.20}$$

Suppose that $\boldsymbol{\alpha}^*$ is such that there exists agent i,

$$\mathbb{E}_{w_i}\left[-\frac{H_{\alpha_i}(w_i|\alpha_i^*)}{h(w_i|\alpha_i^*)}\prod_{j\neq i} H(w_i|\alpha_j^*)\middle|\boldsymbol{\alpha}^*\right] > C'(\alpha_i^*). \tag{4.21}$$

Suppose that there exists $0 < k \leq n$ such that (AIA$'$) binds for the first k agents. Then we can ignore (AIA$'$) for the last $n-k$ agents. Let λ_i denote the Lagrangian multiplier associated with (AIA$'$) for agent i ($i \leq k$). By a similar argument to that in Section 4.3.2, we have $\lambda_i = b(\alpha_i^*)$ for all $i \leq k$. Then the Lagrangian relaxation becomes

$$\begin{aligned}
\mathscr{L} &= \sum_{i\leq k}\int_{w(0,\alpha_i^*)}^{w(1,\alpha_i^*)} \mu Q_i(w_i)h(w_i|\alpha_i^*)\mathrm{d}w_i + \sum_{i>k}\int_{w(0,\alpha_i^*)}^{w(1,\alpha_i^*)} w_i Q_i(w_i)h(w_i|\alpha_i^*)\mathrm{d}w_i + \sum_{i\leq k} b(\alpha_i^*)C'(\alpha_i^*) \\
&= \int_{w(0,\alpha_1^*)}^{w(1,\alpha_1^*)}\cdots\int_{w(0,\alpha_n^*)}^{w(1,\alpha_n^*)}\left(\sum_{i\leq k}\mu q_i(\boldsymbol{w}) + \sum_{i>k} w_i q_i(\boldsymbol{w})\right)\prod_{i=1}^n h(w_i|\alpha_i^*)\mathrm{d}w_1\ldots\mathrm{d}w_n + \sum_{i\leq k} b(\alpha_i^*)C'(\alpha_i^*).
\end{aligned}$$

Suppose that $k < n$, then a pointwise virtual surplus maximizer must satisfy for all $\boldsymbol{w}$,

$$\sum_{i\leq k} q_i(\boldsymbol{w}) = \begin{cases} 1 & \text{if } \max_{j>k}\{w_j\} < \mu, \\ 0 & \text{if } \max_{j>k}\{w_j\} > \mu. \end{cases}$$

and for all $j > k$,

$$q_j(\boldsymbol{w}) = \begin{cases} 1 & \text{if } w_j > \mu \text{ and } w_j = \max_{j>k}\{w_j\}, \\ 0 & \text{if } w_j < \mu \text{ or } w_j < \max_{j>k}\{w_j\}. \end{cases}$$

Therefore, for all $i > k$, the optimal interim allocation rule is given by $Q_i(w_i) = \prod_{j>k, j\neq i} H(w_i|\alpha_j^*)$ if $w_i > \mu$ and $Q_i(w_i) = 0$ if $w_i < \mu$. Hence,

$$V(\boldsymbol{\alpha}^*) = \mu \prod_{i>k} H(\mu|\alpha_i^*) + \sum_{i>k} \int_{\mu}^{w(1,\alpha_i^*)} w_i \prod_{j>k, j\neq i} H(w_i|\alpha_j^*) h(w_i|\alpha_i^*) \mathrm{d}w_i + \sum_{i\leq k} b(\alpha_i^*) C'(\alpha_i^*).$$

Finally, (AIA′) holds for $i > k$ if and only if

$$\int_{\mu}^{w(1,\alpha_i^*)} -H_{\alpha_i}(w_i|\alpha_i^*) \prod_{j>k, j\neq i} H(w_i|\alpha_j^*) \mathrm{d}w_i \leq C'(\alpha_i^*). \tag{4.22}$$

Consider an agent i $(i > k)$. I argue that if $\boldsymbol{\alpha}^*$ is chosen optimally, then (4.22) holds with equality. Suppose, to the contrary, that (4.22) holds with strict inequality, then

$$\frac{\partial \pi^s(\boldsymbol{\alpha}^*)}{\partial \alpha_i} = -(1-\gamma) \int_{\mu}^{w(1,\alpha_i^*)} H_{\alpha_i}(w_i|\alpha_i^*) \prod_{j>k, j\neq i} H(w_i|\alpha_j^*) \mathrm{d}w_i - C'(\alpha_i^*) < -\gamma C'(\alpha_i^*) \leq 0,$$

a contradiction to the optimality of α_i^*. Hence, (4.22) holds with equality for all $i > k$. Furthermore, because the information structures are uniformly supermodular ordered, we have

$$\int_{\mu}^{w(1,\alpha_i^*)} w_i \prod_{j>k, j\neq i} H(w_i|\alpha_j^*) \mathrm{d}w_i = C'(\alpha_i^*) b(\alpha_i^*) + \int_{\mu}^{w(1,\alpha_i^*)} \mu \prod_{j>k, j\neq i} H(w_i|\alpha_j^*) \mathrm{d}w_i.$$

Substituting this into the expression of $V(\boldsymbol{\alpha}^*)$ yields

$$\begin{aligned} V(\boldsymbol{\alpha}^*) =& \mu \prod_{i>k} H(\mu|\alpha_i^*) + \sum_{i>k} \int_{\mu}^{w(1,\alpha_i^*)} \mu \prod_{j>k, j\neq i} H(w_i|\alpha_j^*) h(w_i|\alpha_i^*) \mathrm{d}w_i + \sum_{i=1}^{n} b(\alpha_i^*) C'(\alpha_i^*) \\ =& \mu \prod_{i>k} H(\mu|\alpha_i^*) + \mu \int_{\mu}^{\overline{\theta}} \mathrm{d} \prod_{j>k} H(w|\alpha_j^*) + \sum_{i=1}^{n} b(\alpha_i^*) C'(\alpha_i^*) \\ =& \mu \prod_{i>k} H(\mu|\alpha_i^*) + \mu \left[1 - \prod_{i>k} H(\mu|\alpha_i^*) \right] + \sum_{i=1}^{n} b(\alpha_i^*) C'(\alpha_i^*) = \mu + \sum_{i=1}^{n} b(\alpha_i^*) C'(\alpha_i^*). \end{aligned}$$

Hence,

$$\pi^s(\boldsymbol{\alpha}^*) = (1-\gamma)\left[\mu + \sum_{i=1}^{n} b(\alpha_i^*)C'(\alpha_i^*)\right] - \sum_i C(\alpha_i^*). \tag{4.20}$$

Suppose that $k = n$, then a pointwise virtual surplus maximizer must satisfy $\sum_i q_i(\boldsymbol{w}) = 1$ for all $\boldsymbol{w}$. Hence, (4.20) still holds in this case.

These results are summarized by the following proposition:

Proposition 12 *Suppose that the second-order condition of the agents' optimization problem is satisfied, and the information structures are uniformly supermodular ordered. Let $\boldsymbol{\alpha}^*$ be a socially optimal information choices. Then there exists agent i,*

$$\mathbb{E}_{w_i}\left[-\frac{H_{\alpha_i}(w_i|\alpha_i^*)}{h(w_i|\alpha_i^*)}\prod_{j\neq i} H(w_i|\alpha_j^*)\middle|\boldsymbol{\alpha}^*\right] = C'(\alpha_i^*).$$

and the average social surplus $\pi^s(\boldsymbol{\alpha}^*)$ is given by (4.20).

It follows immediately from Proposition 12 that if there exists $\alpha^\circ \in \arg\max_{\alpha\in\mathbb{A}}(1-\gamma)b(\alpha)C'(\alpha) - C(\alpha)$ such that $\alpha^\circ \leq \hat{\alpha}$, where $\hat{\alpha}$ is define by (4.17), then the socially optimal information choices are the same for all agents:

Proposition 13 *Suppose that the second-order condition of the agents' optimization problem is satisfied, and the information structures are uniformly supermodular ordered. Suppose, in addition, that there exists $\alpha^\circ \in \arg\max_{\alpha\in\mathbb{A}}(1-\gamma)b(\alpha)C'(\alpha) - C(\alpha)$ such that $\alpha^\circ \leq \hat{\alpha}$, where $\hat{\alpha}$ is define by* (4.17). *Then the socially optimal information choices are the same for all agents.*

I conclude this section by giving an example in which an asymmetric mechanism generates higher net social surplus that the optimal symmetric mechanism when the conditions in Proposition 13 is violated.

Example 6 (Linear experiments) *Consider Example 5. Let $n = 2$, $\underline{\alpha} = 1/2$ and $K = 3/8 \geq 1/8\underline{\alpha}$. Then, as I demonstrate in Appendix C.1.3, the first-order approach is valid. A socially optimal* *information choice* $\boldsymbol{\alpha}$ *must be such that*

$$\alpha_1 \leq \frac{9\alpha_2}{18\alpha_2 - 2} \text{ or } \alpha_2 \leq \frac{9\alpha_1}{18\alpha_1 - 2}. \tag{4.23}$$

When $\boldsymbol{\alpha}$ *satisfies* (4.23)*, the average social surplus is given by*

$$\pi^s(\boldsymbol{\alpha}) = (1-\gamma)\left[\frac{1}{2} + \frac{3}{4}\alpha_1\left(\alpha_1 - \frac{1}{2}\right) + \frac{3}{4}\alpha_2\left(\alpha_2 - \frac{1}{2}\right)\right] - \frac{3}{8}\left(\alpha_1 - \frac{1}{2}\right)^2 - \frac{3}{8}\left(\alpha_2 - \frac{1}{2}\right)^2.$$

In this case, $\hat{\alpha}$ *is such that*

$$\frac{3}{4}\left(\hat{\alpha} - \frac{1}{2}\right) = \frac{1}{12} \text{ or } \hat{\alpha} = \frac{11}{18}.$$

Furthermore, $(1-\alpha)b(\alpha)C'(\alpha) - C(\alpha) = (1-\gamma)\frac{3}{4}\alpha(\alpha - \frac{1}{2}) - \frac{3}{8}(\alpha - \frac{1}{2})^2$ *has a unique maximizer on* $[\frac{1}{2}, 1]$*. If* $\gamma \leq \frac{1}{2}$*, then* $\alpha^\circ = 1$*; and if* $\gamma > \frac{1}{2}$*, then* $\alpha^\circ = \gamma/(4\gamma - 2)$*. By Proposition 13, if* $\alpha^\circ \leq \hat{\alpha}$ *or* $\gamma \geq 11/13$*, then the socially optimal information choices are the same for all agents.*

Assume for the rest of the example that $\gamma \leq 11/13$*. In this case, the optimal symmetric mechanism is ex ante efficient, and induces the following symmetric equilibrium:* $\alpha_1 = \alpha_2 = \hat{\alpha}$*. In this case, the average social surplus is given by*

$$\pi^s(\hat{\alpha}, \hat{\alpha}) = \frac{192 - 195\gamma}{324} \approx 0.59 - 0.60\gamma.$$

Consider the following asymmetric mechanism in which

$$q_1(w_1, w_2) = \begin{cases} 0 & \text{if } \min\left\{\max\left\{\frac{7}{32}, w_1\right\}, \frac{25}{32}\right\} < w_2 \\ 1 & \text{if } \min\left\{\max\left\{\frac{7}{32}, w_1\right\}, \frac{25}{32}\right\} > w_2 \end{cases},$$

and $q_2(w_1, w_2) = 1 - q_1(w_1, w_2)$ *for all* $(w_1, w_2) \in [0,1]^2$*. Given this mechanism, the following information choices is an equilibrium* $\alpha_1^* = 9/16$ *and* $\alpha_2^* = 1$*. Let* $\pi_i(\boldsymbol{\alpha})$ *denote agent i's ex ante* *expected payoff for $i = 1, 2$. Given $\alpha_2^* = 1$, the interim allocation rule of agent 1 is given by*

$$Q_1(w_1) = \begin{cases} \frac{7}{32} & \text{if } w_1 \in \left[0, \frac{7}{32}\right] \\ w_1 & \text{if } w_1 \in \left[\frac{7}{32}, \frac{25}{32}\right] \\ \frac{25}{32} & \text{if } w_1 \in \left[\frac{25}{32}, 1\right] \end{cases}.$$

It is easy to verify that

$$\frac{\partial \pi_1(\alpha_1, \alpha_2^*)}{\partial \alpha_1} = \int_{w(0,\alpha_1)}^{w(1,\alpha_1)} -H_\alpha(w_1, \alpha_1) Q_1(w_1) \mathrm{d}w_1 - C'(\alpha_1) = -\frac{8\alpha_1}{12} + \frac{3}{8}.$$

Hence, it is optimal for agent 1 to choose $\alpha_1^* = 9/16$. Similarly, given $\alpha_1^* = 9/16$, the interim allocation rule of agent 2 is given by

$$Q_2(w_2) = \frac{16}{9} w_2 - \frac{7}{18}, \forall w_2 \in [0, 1].$$

It is easy to verify that

$$\frac{\partial \pi_2(\alpha_1^*, \alpha_2)}{\partial \alpha_2} = \int_{w(0,\alpha_2)}^{w(1,\alpha_2)} -H_\alpha(w_2, \alpha_1) Q_2(w_2) \mathrm{d}w_2 - C'(\alpha_2) = -\frac{65\alpha_2}{108} + \frac{3}{8} > 0, \ \forall \alpha_2 \in \left[\frac{1}{2}, 1\right].$$

Hence, it is optimal for agent 2 to choose $\alpha_2^* = 1$. In this case, the average social surplus is given by

$$\pi^s(\alpha_1^*, \alpha_2^*) = \frac{728 - 923\gamma}{1024} \approx 0.71 - 0.90\gamma.$$

Clearly, if $\gamma < 0.4$, then this asymmetric mechanism generates strictly higher net social surplus that the optimal symmetric mechanism does.

4.5. Conclusion

I have studied ex ante efficient mechanisms in the sale of a single object when agents have positively interdependent values, and information is independent and endogenous. Specifically, I assume agents are initially uncertain about the value of the object on sale, and they are able to pay a cost to acquire information about this value before participating in a mechanism. In an earlier paper, Bergemann and Välimäki (2002) find that using the ex post efficient mechanism will lead to ex ante over investment in information by agents. This suggests potential gains in ex ante efficiency by adjusting ex post efficient mechanisms in a way to discourage agents from gathering information.

First, I find that an ex ante efficient mechanism never withholds the object. Intuitively, whenever the object is withheld, one can instead allocate it randomly among agents. This improves the allocative efficiency without giving agents additional incentive to acquire information. Second, I fully characterize ex ante efficient mechanisms. When the interdependence is low or the number of agents is large, the ex-post efficient mechanism is also ex ante efficient. When the interdependence is high or the number of agents is small, an ex ante efficient mechanism involves randomization. Specifically, an ex ante efficient allocation rule randomizes in areas in which the accuracy of an agent's posterior estimate can be significantly improved if an additional piece of information in acquired.

In this paper, I assume all agents simultaneously acquire information prior to the auction. One interesting direction for future research is to allow for the possibility of sequential information acquisition. It is likely that the efficiency can be improved if agents are asked to acquire information in turn, and one's information acquisition decision can depend on the signals received by those who take actions earlier. Another interesting direction for future research is to consider the impact of initial private information. In this paper I only considered static mechanisms in which agents only report their private information once. In general, one can consider a dynamic mechanism in which agents report their private information both before and after acquiring information.

APPENDIX TO CHAPTER 2

A.1. Omitted proofs in Sections 2.3

A polymatroid is a polytope of type

$$P(g) := \left\{ x \in \mathbb{R}^E \,\middle|\, x \geq 0, \sum_{e \in A} x_e \leq g(A) \text{ for all } A \subset E \right\}, \tag{A.1}$$

where E is a finite set and $g : 2^E \to \mathbb{R}_+$ is a submodular function.

Lemma 16 *There exists a monotone and submodular function* $\overline{g} : 2^E \to \mathbb{R}_+$ *with* $\overline{g}(\emptyset) = 0$ *and* $P(g) = P(\overline{g})$.

Proof. Let $\overline{g}(\emptyset) := 0$ and $\overline{g}(X) := \min_{A \supset X} g(A)$ for $X \neq \emptyset$. Let $X \subset Y \subset E$. If $X = \emptyset$, then $\overline{g}(X) = 0 \leq \overline{g}(Y)$. If $X \neq \emptyset$, then $A \supset Y$ implies that $A \supset X$, and therefore we have

$$\overline{g}(X) = \min_{A \supset X} g(A) \leq \min_{A \supset Y} g(A) = \overline{g}(Y).$$

Hence, $\overline{g}$ is monotone. Let $e \in E \backslash Y$. To show that $\overline{g}$ is submodular, it suffices to show that

$$\overline{g}(Y \cup \{e\}) - \overline{g}(Y) \leq \overline{g}(X \cup \{e\}) - \overline{g}(X).$$

Because $\overline{g}(\emptyset) = 0 \leq \min_A g(A)$, it suffices to show that

$$\min_{C \supset Y \cup \{e\}} g(C) + \min_{D \supset X} g(D) \leq \min_{A \supset X \cup \{e\}} g(A) + \min_{B \supset Y} g(B).$$

Let $A^* \in \arg\min_{A \supset X \cup \{e\}} g(A)$ and $B^* \in \arg\min_{B \supset Y} g(B)$. Then $A^* \cup B^* \supset Y \cup \{e\}$ and $A^* \cap B^* \supset$

X. Hence,

$$\min_{A\supset X\cup\{e\}} g(A) + \min_{B\supset Y} g(B) = g(A^*) + g(B^*)$$
$$\geq g(A^* \cup B^*) + g(A^* \cap B^*)$$
$$\geq \min_{C\supset Y\cup\{e\}} g(C) + \min_{D\supset X} g(D),$$

where the first inequality holds because g is submodular. Hence, $\overline{g}$ is submodular. Finally, I want to show that $P(g) = P(\overline{g})$. Because $g(A) \geq \overline{g}(A)$ for all $A \subset E$, we have $P(\overline{g}) \subset P(g)$. Suppose that there exists $x \in \mathbb{R}^E$ such that $x \in P(g)$ and $x \notin P(\overline{g})$. Then there exists $A \neq \emptyset$ such that $\sum_{e\in A} x_e > \overline{g}(A)$. By construction, there exists $B \supset A$ such that $\overline{g}(A) = g(B)$. However, then we have $\sum_{e\in B} x_e \geq \sum_{e\in A} x_e > \overline{g}(A) = g(B)$, which is a contradiction to that $x \in P(g)$. Hence, $P(g) = P(\overline{g})$. ■

Proof of Lemma 1. First, because $\overline{H}(\emptyset) = 0$, and $\overline{H}$ is non-decreasing and submodular, $\hat{z}^t$ is feasible. Next, I show that $\hat{z}^t$ is optimal.

I begin the analysis by characterizing $\overline{H}$. Clearly, there exists a unique $\underline{t} \in \{1, \dots, m\}$ such that

$$\frac{1}{n}\left(\sum_{\tau=1}^{\underline{t}-1} f^\tau\right)^{n-1} < \varphi \leq \frac{1}{n}\left(\sum_{\tau=1}^{\underline{t}} f^\tau\right)^{n-1}.$$

Here, $\underline{t}$ is the minimum t such that if all agents whose values are weakly less than v^t are pooled together and ranked below any other agents with higher values, then they receive the object with probability of at least φ. It is easy to verify that[1]

$$\overline{H}(S^t) = \begin{cases} 1 - \left(\sum_{\tau=1}^{t-1} f^\tau\right)^n - n\varphi \sum_{\tau=t}^{m} f^\tau & \text{if } t > \underline{t} \\ 1 - n\varphi & \text{if } t \leq \underline{t} \end{cases}. \tag{A.2}$$

Let $\Delta(t) := \overline{H}(S^t) - n\sum_{\tau=t}^{m} \frac{c\varphi f^\tau}{1-c}$ for $t = 1, \dots, m+1$. Then $\Delta(m+1) = 0$ and $\Delta(t) = \tilde{\Delta}\left(\sum_{\tau=1}^{t-1} f^\tau\right)$,

[1]This result can be seen as a corollary of Lemmas 22 and 23 in Appendix A.3.

where $\tilde{\Delta}(x) = 1 - \frac{n\varphi}{1-c} - x^n - \frac{n\varphi x}{1-c}$ is concave in x. If $\Delta(1) = 1 - n\varphi/(1-c) \geq 0$, then let $\bar{t} := 0$; otherwise, there exists a unique $\bar{t} \in \{1, \dots, m+1\}$ such that

$$\overline{H}(S^{\bar{t}}) \leq n \sum_{\tau=\bar{t}}^{m} \frac{c\varphi f^{\tau}}{1-c} \text{ and } \overline{H}(S^{\bar{t}+1}) > n \sum_{\tau=\bar{t}+1}^{m} \frac{c\varphi f^{\tau}}{1-c}.$$

Let $\lambda := (\lambda^1, \dots, \lambda^m)$ and $\mu := (\mu^1, \dots, \mu^m)$ denote the dual variables corresponding to the upper-bounds and lower-bounds in (IC$'m1$), and $\beta := (\beta(S))_S$ denote the dual variables corresponding to (F2$m1$) in problem $(OPTm1 - \varphi)$. Consider the dual to problem $(OPTm1 - \varphi)$, denoted by $(DOPTm1 - \varphi)$,

$$\min_{\lambda,\beta,\mu} \sum_{t=1}^{m} \frac{c\varphi f^t \lambda^t}{1-c} + \sum_{S} \beta(S)\overline{H}(S) + \varphi \sum_{t=1}^{m} f^t v^t,$$

subject to

$$v^t - \frac{k}{c} - \lambda^t + \mu^t - n \sum_{S \ni t} \beta(S) \geq 0, \forall t,$$

$$\lambda \geq 0, \beta \geq 0, \mu \geq 0.$$

Let $\hat{z}$ be define in (2.7), and $(\hat{\lambda}, \hat{\beta}, \hat{\mu})$ be the corresponding dual variables. Let t^0 be such that $v^t \geq k/c$ if and only if $t \geq t^0$.

Case 1: $v^{\bar{t}} < \frac{k}{c}$ or $\bar{t} < t^0$. In this case, we have

$$\hat{z}^t := \begin{cases} \frac{c\varphi f^t}{1-c} & \text{if } t > \bar{t} \\ 0 & \text{if } t \leq \bar{t} \end{cases}.$$

Let $\hat{\beta}(S) = 0$ for all S. If $v^t < k/c$, then let $\hat{\lambda}^t = 0$ and $\hat{\mu}^t = k/c - v^t > 0$; if $v^t \geq k/c$, then let $\hat{\mu}^t = 0$ and $\hat{\lambda}^t = v^t - k/c \geq 0$. It is easy to verify that this is a feasible solution to $(DOPTm1 - \varphi)$, and the complementary slackness conditions are satisfied. Finally, the dual objective is equal to the primal objective:

$$\sum_{t=t^0}^{m} \frac{c\varphi f^t}{1-c} \left(v^t - \frac{k}{c}\right) + \varphi \sum_{t=1}^{m} f^t v^t.$$

By the duality theorem, $\hat{z}$ is an optimal solution to $(OPTm1 - \varphi)$.

Case 2: $v^{\bar{t}} \geq \frac{k}{c}$ or $\bar{t} \geq t^0$. In this case, we have

$$\hat{z}^t := \begin{cases} \frac{c\varphi f^t}{1-c} & \text{if } t > \bar{t} \\ \frac{1}{n}\overline{H}(S^{\bar{t}}) - \sum_{\tau=\bar{t}+1}^{m} \frac{c\varphi f^t}{1-c} & \text{if } t = \bar{t} \\ \frac{1}{n}\left[\overline{H}(S^t) - \overline{H}(S^{t+1})\right] & \text{if } t^0 \leq t < \bar{t} \\ 0 \text{ if } t < t^0 \end{cases},$$

Let $\hat{\beta}(S) > 0$ if $S = S^t$ for $t^0 \leq t \leq \bar{t}$; and $\hat{\beta}(S) = 0$ otherwise. If $t < t^0$, then let $\hat{\lambda}^t = 0$ and $\hat{\mu}^t = k/c - v^t \geq 0$. If $t^0 \leq t \leq \bar{t}$, then let $\hat{\lambda}^t = \hat{\mu}^t = 0$, $\hat{\beta}(S^t) = (v^t - v^{t-1})/n$ for $t > t^0$ and $\hat{\beta}(S^{t^0}) = (v^{t^0} - k/c)/n$. If $t > \bar{t}$, then let $\hat{\lambda}^t = v^t - v^{\bar{t}}$ and $\hat{\mu}^t = 0$. It is easy to verify that this is a feasible solution to $(DOPTm1 - \varphi)$, and the complementary slackness conditions are satisfied. Finally, the dual objective is equal to the primal objective:

$$\frac{1}{n}\overline{H}(S^{t^0})\left(v^{t^0} - \frac{k}{c}\right) + \sum_{t=t^0+1}^{\bar{t}} \frac{1}{n}\overline{H}(S^t)\left(v^t - v^{t-1}\right) + \sum_{t=\bar{t}+1}^{m} \frac{c\varphi f^t}{1-c}\left(v^t - \frac{k}{c}\right) + \varphi\sum_{t=1}^{m} f^t v^t.$$

By the duality theorem, $\hat{z}$ is an optimal solution to $(OPTm1 - \varphi)$. ■

Lemma 17 *An optimal solution to $(OPT - \varphi)$ exists.*

Proof. Let $\mathscr{D}$ denote the set of feasible solutions, i.e., solutions satisfying (IC′) and (F2). Consider $\mathscr{D}$ as a subset of L_2, the set of square integrable functions with respect to the probability measure corresponding to F. Topologize L_2 with its weak*, or $\sigma(L_2, L_2)$, topology. It is straightforward to verify that $\mathscr{D}$ is $\sigma(L_2, L_2)$ compact. See, for example, Border (1991).

Let $V(\varphi) := \sup_{P\in\mathscr{D}} \mathbb{E}_v\left[P(v)\left(v - \frac{k}{c}\right)\right] + \frac{\varphi k}{c}$. Let $\{P_\nu\}$ be a sequence of feasible solutions to $(OPT - \varphi)$ such that

$$\int P_\nu(v)\left(v - \frac{k}{c}\right)\mathrm{d}F(v) + \frac{\varphi k}{c} \to V(\varphi).$$

By Helly's selection theorem, after taking subsequences, I can assume that there exists P such that

$\{P_\nu\}$ converges pointwise to P. Because $\mathscr{D}$ is $\sigma(L_2, L_2)$ compact, after taking subsequences again, I can assume that there exists $P \in \mathscr{D}$ such that $\{P_\nu\}$ converges to P in $\sigma(L_2, L_2)$ topology. Because $v - k/c \in L_2$, the weak convergence of $\{P_\nu\}$ implies that

$$\int P(v)\left(v - \frac{k}{c}\right) \mathrm{d}F(v) + \frac{\varphi k}{c} = V(\varphi).$$

■

Proof of Theorem 1. Let $\{P_m\}$ be the sequence of optimal solutions to $(OPTm - \varphi)$ defined in Corollary 1. Let $\overline{P}_m^t := \overline{z}^t / f^t + \varphi$ for all t. Then

$$P_m^t := \begin{cases} \overline{P}_m^t & \text{if } v^t > \frac{k}{c} \\ \varphi & \text{if } v^t < \frac{k}{c} \end{cases}.$$

Recall that $\overline{H}$ is given by (A.2). Thus, there are three cases. If $\bar{t} > \underline{t}$, then

$$\overline{P}_m^t = \begin{cases} \frac{\varphi}{1-c} & \text{if } t > \bar{t} \\ \frac{\frac{1}{n} - \frac{1}{n}\left(\sum_{\tau=1}^{\bar{t}} f^\tau\right)^n - \sum_{\tau=\bar{t}+1}^{m} \frac{\varphi f^t}{1-c}}{f^{\bar{t}}} & \text{if } t = \bar{t} \\ \frac{\frac{1}{n}\left(\sum_{\tau=1}^{t} f^\tau\right)^n - \frac{1}{n}\left(\sum_{\tau=1}^{t-1} f^\tau\right)^n}{f^t} & \text{if } \underline{t} < t < \bar{t} \\ \frac{\frac{1}{n}\left(\sum_{\tau=1}^{\underline{t}} f^\tau\right)^n - \varphi \sum_{\tau=1}^{\underline{t}-1} f^\tau}{f^t} & \text{if } t = \underline{t} \\ \varphi & \text{if } t < \underline{t} \end{cases}.$$

If $\bar{t} = \underline{t}$, then

$$\overline{P}_m^t = \begin{cases} \frac{\varphi}{1-c} & \text{if } t > \bar{t} \\ \frac{\frac{1}{n} - \varphi \sum_{t=1}^{t-1} f^t - \sum_{\tau=t+1}^{m} \frac{\varphi f^t}{1-c}}{f^{\bar{t}}} & \text{if } t = \bar{t} \\ \varphi & \text{if } t < \bar{t} \end{cases}.$$

If $\bar{t} < \underline{t}$, then

$$\overline{P}_m^t = \begin{cases} \frac{\varphi}{1-c} & \text{if } t > \bar{t} \\ \varphi & \text{if } t < \bar{t} \end{cases} .$$

I can extend P_m to V by setting

$$P_m(v) := P_m^t \text{ for } v \in \left[\underline{v} + \frac{(t-1)(\overline{v}-\underline{v})}{m}, \underline{v} + \frac{t(\overline{v}-\underline{v})}{m}\right], t = 1, \ldots, m.$$

Extend $\overline{P}_m$ to V in a similar fashion. Compare $\overline{P}_m$ and $\overline{P}_\varphi$. It is easy to see that $\{\overline{P}_m\}$ converges pointwise to $\overline{P}_\varphi$. Hence, $\{P_m\}$ converges pointwise to P_φ^*, which is a feasible solution to $(OPT - \varphi)$.

To show the optimality of P_φ^*, let $\hat{P}$ be an optimal solution to $(OPT - \varphi)$, which exists by Lemma 17 in the appendix. Define $\hat{P}_m$ be such that

$$\hat{P}_m^t := \frac{1}{f^t} \int_{\underline{v}+\frac{(t-1)(\overline{v}-\underline{v})}{m}}^{\underline{v}+\frac{t(\overline{v}-\underline{v})}{m}} \hat{P}(v) \mathrm{d}F(v) \text{ for } t = 1, \ldots, m,$$

and it can be extended to V by setting

$$\hat{P}_m(v) := \hat{P}_m^t \text{ for } v \in \left[\underline{v} + \frac{(t-1)(\overline{v}-\underline{v})}{m}, \underline{v} + \frac{t(\overline{v}-\underline{v})}{m}\right], t = 1, \ldots, m.$$

By the Lebesgue differentiation theorem, $\{\hat{P}_m\}$ converges pointwise to $\hat{P}$. It is easy to verify that $\hat{P}_m$ defined on $\{v^1, \ldots, v^m\}$ is a feasible solution to $(OPTm - \varphi)$. Hence

$$\sum_{t=1}^{m} f^t \hat{P}_m^t \left(v^t - \frac{k}{c}\right) + \frac{\varphi k}{c} \leq \sum_{t=1}^{m} f^t P_m^t \left(v^t - \frac{k}{c}\right) + \frac{\varphi k}{c}$$

By the dominated convergence theorem,

$$\sum_{t=1}^{m} f^t \hat{P}_m^t \left(v^t - \frac{k}{c}\right) = \int_V \hat{P}_m(v) \left(v - \frac{k}{c}\right) \mathrm{d}F(v) \to \int_V \hat{P}(v) \left(v - \frac{k}{c}\right) \mathrm{d}F(v),$$

and

$$\sum_{t=1}^{m} f^t P_m^t \left(v^t - \frac{k}{c}\right) = \int_V P_m(v) \left(v - \frac{k}{c}\right) \mathrm{d}F(v) \to \int_V P_\varphi^*(v) \left(v - \frac{k}{c}\right) \mathrm{d}F(v).$$

Hence,

$$\int_V P_\varphi^*(v)\left(v-\frac{k}{c}\right)\mathrm{d}F(v)=\int_V \hat{P}(v)\left(v-\frac{k}{c}\right)\mathrm{d}F(v),$$

which implies that P_φ^* is optimal. ■

Lemma 18 *Suppose* $(1-c)/n \le \varphi \le \min\{1/n, 1-c\}$. *Then* $v^l \ge v^u$ *if and only if* $v^l \le v^\natural$, *where* $v^\natural$ *is defined by* (2.14). *Furthermore, if* $n(1-c) < 1$ *then* $v^\natural$ *is strictly increasing in* n *and strictly decreasing in* c.

Proof. Because $(1-c)/n \le \varphi \le \min\{1/n, 1-c\}$, v^l and v^u satisfies:

$$\frac{1-F(v^u)^n}{1-F(v^u)}=\frac{F(v^l)^{n-1}}{1-c}. \tag{A.3}$$

Define

$$\Delta(v) := \frac{F(v)^{n-1}(1-F(v))}{1-c}-1+F(v)^n.$$

Then $\Delta(\underline{v})=-1<0$ and $\Delta(\overline{v})=0$. Then

$$\Delta'(v)=\frac{F(v)^{n-2}f(v)}{1-c}\left[-cnF(v)+n-1\right].$$

Clearly, the term in the brackets is strictly decreasing in v. Moreover, $\Delta'(\underline{v}) = n-1 > 0$ and $\Delta'(\overline{v}) = n(1-c)-1$.

If $n(1-c)\ge 1$, then $\Delta'(v)\ge 0$ for all v. Hence, $\Delta(v)$ is non-decreasing, and therefore $\Delta(v)\le 0$ for all v. Hence,

$$\frac{1-F(v^u)^n}{1-F(v^u)}=\frac{F(v^l)^{n-1}}{1-c}\le\frac{1-F(v^l)^n}{1-F(v^l)},$$

which implies $v^l \ge v^u$.

If $n(1-c)<1$, then there exists $v^\flat$ such that $\Delta'(v)>0$ for $v\in[\underline{v}, v^\flat]$ and $\Delta'(v)<0$ for $v\in[v^\flat,\overline{v}]$. Hence, $\Delta(v)$ in strictly increasing in $[\underline{v}, v^\flat]$, and strictly decreasing in $[v^\flat,\overline{v}]$. Hence, there exists a unique $v^{\natural} \in (\underline{v}, \overline{v})$ such that $\Delta(v) \leq 0$ if and only if $v \leq v^{\natural}$. By (A.3), this implies that $v^l \geq v^u$ if and only if $v^l \leq v^{\natural}$. Finally, for any v, $\Delta(v)$ is strictly decreasing in n, and strictly increasing in c. Hence, $v^{\natural}$ is strictly increasing in n, and strictly decreasing in c. ■

Proof of Theorem 3. First, if $\varphi \leq (1-c)/n$, then $v^u = \hat{v} = \underline{v}$, and

$$P^*_{\varphi}(v) := \begin{cases} \frac{\varphi}{1-c} & \text{if } v \geq \frac{k}{c} \\ \varphi & \text{if } v < \frac{k}{c} \end{cases}.$$

The principal's objective becomes

$$\frac{c\varphi}{1-c}\int_{\frac{k}{c}}^{\overline{v}}\left(v-\frac{k}{c}\right)\mathrm{d}F(v) + \varphi\int_{\underline{v}}^{\overline{v}} v\mathrm{d}F(v),$$

which is strictly increasing in φ. Hence, in optimum, $\varphi \geq (1-c)/n$.

Given φ, let $Z(\varphi)$ denote the principal's optimal payoff. Suppose that $\varphi \geq 1-c$ or equivalently $F(v^l)^{n-1} \geq n(1-c)$. Then $v^u = \overline{v}$, and the principal's payoff is $Z(\varphi) = Z_1(v^l(\varphi))$, where

$$\begin{aligned} Z_1(v^l) := & \int_{\max\left\{v^l, \frac{k}{c}\right\}}^{\overline{v}}\left(v-\frac{k}{c}\right)F(v)^{n-1}\mathrm{d}F(v) \\ & + \frac{1}{n}F(v^l)^{n-1}\int_{\underline{v}}^{\max\left\{v^l, \frac{k}{c}\right\}}\left(v-\frac{k}{c}\right)\mathrm{d}F(v) + \frac{1}{n}F(v^l)^{n-1}\frac{k}{c}. \end{aligned}$$

If $v^l < k/c$, then $Z_1(v^l)$ is strictly increasing in v^l. If $v^l \geq k/c$, then

$$Z_1'(v^l) = \frac{n-1}{n}F(v^l)^{n-2}f(v^l)\left\{\mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^l\}] + \frac{k}{c}\right\}.$$

Clearly, the term inside the braces is strictly decreasing in v^l. Recall that $v^* \geq k/c$ is defined by (2.12). Hence, $Z_1'(v^l) \geq 0$ if and only if $v^l \leq v^*$, and Z_1 achieves its maximum at $v^l = v^*$. I show in Lemma 19 in the appendix that for any φ and the corresponding v^l, we have $Z(\varphi) \leq Z_1(v^l(\varphi))$.

Hence,

$$Z(\varphi) \leq Z_1(v^l(\varphi)) \leq Z_1(v^*).$$

Thus, if $F(v^*)^{n-1} \geq n(1-c)$, then it is optimal to set $\varphi^* = F(v^*)^{n-1}/n$ and $v^l = v^*$. This proves the first part of Theorem 3.

Suppose that $F(v^*)^{n-1} < n(1-c)$. Then in optimum $\varphi \leq 1-c$. Because $(1-c)/n \leq \varphi \leq 1/n$, there is a one-to-one correspondence between $\hat{v}$ and φ. Given φ, $\hat{v}(\varphi)$ is uniquely pinned down by

$$1 - n\varphi F(\hat{v}) - \frac{n\varphi}{1-c}\left[1 - F(\hat{v})\right] = 0.$$

If φ is such that $v^l \geq v^u$, then $Z(\varphi) = Z_2(\hat{v}(\varphi))$, where

$$\begin{aligned} Z_2(\hat{v}) :=& \frac{1-c}{n(1-cF(\hat{v}))} \int_{\underline{v}}^{\max\left\{\hat{v}, \frac{k}{c}\right\}} \left(v - \frac{k}{c}\right) \mathrm{d}F(v) \\ &+ \frac{1}{n(1-cF(\hat{v}))} \int_{\max\left\{\hat{v}, \frac{k}{c}\right\}}^{\overline{v}} \left(v - \frac{k}{c}\right) \mathrm{d}F(v) + \frac{1-c}{n(1-cF(\hat{v}))} \frac{k}{c}. \end{aligned}$$

If $\hat{v} < k/c$, then $Z_2(\hat{v})$ is strictly increasing in $\hat{v}$. If $\hat{v} \geq k/c$, then

$$Z_2'(\hat{v}) = \frac{cf(\hat{v})}{n(1-cF(\hat{v}))^2} \left\{ \mathbb{E}_v[v] - \mathbb{E}_v[\min\{v, \hat{v}\}] + (1-c)\left[\mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, \hat{v}\}] + \frac{k}{c}\right] \right\}.$$

Clearly, the term inside the braces is strictly decreasing in $\hat{v}$. Recall that $v^{**} > v^* \geq k/c$ is defined by (2.13). Hence, $Z_2'(\hat{v}) \geq 0$ if and only if $\hat{v} \leq v^{**}$, and Z_2 achieves its maximum at $\hat{v} = v^{**}$. I show in Lemma 20 in the appendix that for any $\varphi \leq 1-c$ and the corresponding $\hat{v}$, we have $Z(\varphi) \leq Z_2(\hat{v}(\varphi))$. Hence,

$$Z(\varphi) \leq Z_2(\hat{v}(\varphi)) \leq Z_2(v^{**}).$$

Finally, by Lemma 18, $v^l \geq v^u$ if and only if $v^l \leq v^\natural$. Thus, if $v^{**} \leq v^\natural$, then it is optimal to set $\varphi^* = F(v^{**})^{n-1}/n$ and $v^l = v^{**}$. This proves the second part of Theorem 3.

Suppose that $F(v^*)^{n-1} < n(1-c)$ and $v^{**} > v^\natural$. Then

$$Z(\varphi) = \varphi \int_{\underline{v}}^{\max\left\{v^l, \frac{k}{c}\right\}} \left(v - \frac{k}{c}\right) \mathrm{d}F(v) + \int_{\max\left\{v^l, \frac{k}{c}\right\}}^{\max\left\{v^u, \frac{k}{c}\right\}} \left(v - \frac{k}{c}\right) F(v)^{n-1} \mathrm{d}F(v)$$
$$+ \frac{\varphi}{1-c} \int_{\max\left\{v^u, \frac{k}{c}\right\}}^{\overline{v}} \left(v - \frac{k}{c}\right) \mathrm{d}F(v) + \frac{\varphi k}{c}.$$

If φ is such that $v^l < k/c$, then $Z(\varphi)$ is strictly increasing in φ. If $v^l \geq k/c$, then

$$Z'(\varphi) = \frac{1}{1-c} \left[\mathbb{E}_v[v] - \mathbb{E}_v[\min\{v, v^u\}]\right] + \mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^l\}] + \frac{k}{c}.$$

Because both v^l and v^u are strictly increasing in φ, $Z'(\varphi)$ is strictly decreasing in φ. Let φ^* be such that

$$\mathbb{E}_v[v] - \mathbb{E}_v[\min\{v, v^u(\varphi^*)\}] + (1-c)\left[\mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^l(\varphi^*)\}] + \frac{k}{c}\right] = 0. \tag{2.15}$$

Compare (2.15) with (2.13) and (2.12), and it is easy to see that $v^u(\varphi^*) > v^{**} > v^l(\varphi^*) > v^*$. Hence, $Z'(\varphi) \geq 0$ if and only if $\varphi \leq \varphi^*$, and Z achieves its maximum at $\varphi = \varphi^*$. This proves the third part of Theorem 3. ■

Lemma 19 *Let Z and Z_1 be defined as in the proof of Theorem 3. Then $Z(\varphi) \leq Z_1(v^l(\varphi))$.*

Proof. Fix φ and the corresponding v^l. Note that $Z_1(v^l)$ is attained by the following allocation rule

$$P_1(v) := \begin{cases} F(v)^{n-1} & \text{if } v \geq \max\left\{v^l, \frac{k}{c}\right\} \\ \varphi & \text{if } v < \max\left\{v^l, \frac{k}{c}\right\} \end{cases}.$$

It is easy to see that $P_1 - P^*_\varphi$ is non-decreasing, and

$$\int_{\underline{v}}^{\overline{v}} P_1(v) \mathrm{d}F(v) = \int_{\underline{v}}^{\overline{v}} P^*_\varphi(v) \mathrm{d}F(v) = \frac{1}{n}.$$

Moreover, $v - k/c$ is non-decreasing in v. Hence, by Lemma 1 in Persico (2000), $Z(\varphi) \leq Z_1(v^l(\varphi))$.

■

Lemma 20 *Let Z and Z_2 be defined as in the proof of Theorem 3. If $\varphi \leq 1-c$, then $Z(\varphi) \leq Z_2(\hat{v}(\varphi))$.*

Proof. Fix φ and the corresponding $\hat{v}$. Note that $Z_2(\hat{v})$ is attained by the following allocation rule

$$P_2(v) := \begin{cases} \frac{\varphi}{1-c} & \text{if } v \geq \max\left\{\hat{v}, \frac{k}{c}\right\} \\ \varphi & \text{if } v < \max\left\{\hat{v}, \frac{k}{c}\right\} \end{cases}.$$

It is easy to see that $P_2 - P^*_\varphi$ is non-decreasing, and

$$\int_{\underline{v}}^{\overline{v}} P_2(v)\mathrm{d}F(v) = \int_{\underline{v}}^{\overline{v}} P^*_\varphi(v)\mathrm{d}F(v) = \frac{1}{n},$$

Moreover, $v - k/c$ is non-decreasing in v. Hence, by Lemma 1 in Persico (2000), $Z(\varphi) \leq Z_1(v_2(\hat{\varphi}))$.

■

Proof of Theorem 2. Let Z and Z_1 be defined as in the proof of Theorem 3. If $\overline{v} - k/c \leq \mathbb{E}_v[v]$, then $Z_1(v)$ is strictly increasing in v^l and achieves its maximum when $v^l = \overline{v}$. By Lemma 19, $Z(\varphi) \leq Z_1(v^l(\varphi)) \leq Z_1(\overline{v})$. Note that $Z_1(\overline{v})$ can be achieved via pure randomization. This completes the proof. ■

A.2. Omitted proofs in Section 2.4

Proof of Corollary 2. The analysis in Section 2.4 has proved most results of Corollary 2. What is left to prove is that if $n^*(\rho, c) < n < n^{**}(\rho, c)$, then $v^l(n, \rho, c)$ is strictly increasing in n, ρ and c and $v^u(n, \rho, c)$ is strictly decreasing in n and strictly decreasing in ρ and c. If $n^*(\rho, c) < n < n^{**}(\rho, c)$, then v^l and v^u satisfy (A.3). It is easy to see that v^u is strictly increasing in v^l and vice versa.

To prove the properties of v^l, let

$$\Delta_l(v^l, n, \rho, c) := \mathbb{E}_v[v] - \mathbb{E}_v[\min\{v, v^u\}] + (1-c)\left[\mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^l\}] + \frac{k}{c}\right],$$

where v^u is a function of v^l, n and c defined by (A.3). Then $\Delta_l(v^l, n, \rho, c) \equiv 0$ by (2.15). Furthermore, we have

$$\begin{aligned}
\frac{\partial \Delta_l}{\partial v^l} &= -[1 - F(v^u)]\frac{\partial v^u}{\partial v^l} - (1-c)F(v^l) < 0,\\
\frac{\partial \Delta_l}{\partial n} &= -[1 - F(v^u)]\frac{\partial v^u}{\partial n} > 0,\\
\frac{\partial \Delta_l}{\partial \rho} &= 1 - c > 0,\\
\frac{\partial \Delta_l}{\partial c} &= -\left[\mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^l\}] + \frac{k}{c}\right] > 0.
\end{aligned} \tag{A.4}$$

Hence, by the implicit function theorem, we have $\partial v^l/\partial n > 0$, $\partial v^l/\partial \rho > 0$ and $\partial v^l/\partial c > 0$. To see that $\partial v^u/\partial n < 0$ in the second line in (A.4), let

$$\Delta(v^u, v^l, n) := \frac{F(v^l)^{n-1}(1 - F(v^u))}{1-c} - 1 + F(v^u)^n.$$

Then $\Delta(v^u, v^l, n) \equiv 0$ by (A.3). Furthermore, we have

$$\begin{aligned}
\frac{\partial \Delta}{\partial v^u} &= \left[-\frac{F(v^l)^{n-1}}{1-c} + nF(v^u)^{n-1}\right] f(v^u) = \left[-\frac{1 - F(v^u)^n}{1 - F(v^u)} + nF(v^u)^{n-1}\right] f(v^u) < 0,\\
\frac{\partial \Delta}{\partial n} &= \frac{F(v^l)^{n-1}[1 - F(v^u)] \log F(v^l)}{1-c} + F(v^u)^n \log F(v^u) < 0.
\end{aligned}$$

Hence, by the implicit function theorem, $\partial v^u/\partial n = -(\partial \Delta/\partial n)/(\partial \Delta/\partial v^u) < 0$.

To prove the properties of v^u, let

$$\Delta_u(v^u, n, \rho, c) := \mathbb{E}_v[v] - \mathbb{E}_v[\min\{v, v^u\}] + (1-c)\left[\mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^l\}] + \frac{k}{c}\right],$$

where v^l is a function of v^u, n and c defined by (A.3). Then $\Delta_u(v^u, n, \rho, c) \equiv 0$ by (A.3). Furthermore, we have

$$\begin{aligned}
\frac{\partial \Delta_u}{\partial v^u} &= -[1 - F(v^u)] - (1-c)F(v^l)\frac{\partial v^l}{\partial v^u} < 0,\\
\frac{\partial \Delta_u}{\partial n} &= -(1-c)F(v^l)\frac{\partial v^l}{\partial n} < 0,\\
\frac{\partial \Delta_u}{\partial \rho} &= 1 - c > 0,\\
\frac{\partial \Delta_u}{\partial c} &= -\left[\mathbb{E}_v[v] - \mathbb{E}_v[\max\{v, v^l\}] + \frac{k}{c}\right] > 0.
\end{aligned} \tag{A.5}$$

Hence, by the implicit function theorem, we have $\partial v^u/\partial n < 0$, $\partial v^u/\partial \rho > 0$ and $\partial v^u/\partial c > 0$. To see that $\partial v^l/\partial n > 0$ in the second line in (A.5), note that

$$\frac{\partial \Delta}{\partial v^l} = \frac{(n-1)F(v^l)^{n-2}f(v^l)[1-F(v^u)]}{1-c} > 0.$$

Hence, by the implicit function theorem, $\partial v^l/\partial n = -(\partial \Delta/\partial n)/(\partial \Delta/\partial v^l) > 0$. ■

A.3. Asymmetric environment

A.3.1. Finite case

Let $\mathcal{D} := \cup_i[\underline{v}_i - k_i/c_i, \overline{v}_i - k_i/c_i]$. Let $\underline{d} := \inf \mathcal{D}$ and $\overline{d} := \sup \mathcal{D}$. Fix an integer $m \geq 2$. For $t = 1, \ldots m$, let

$$\begin{aligned} d^t &:= \underline{d} + \frac{(2t-1)(\overline{d}-\underline{d})}{2m}, \\ f_i^t &:= F_i\left(\underline{d} + \frac{t(\overline{d}-\underline{d})}{m} + \frac{k_i}{c_i}\right) - F_i\left(\underline{d} + \frac{(t-1)(\overline{d}-\underline{d})}{m} + \frac{k_i}{c_i}\right), i = 1, \ldots, n. \end{aligned}$$

Consider the finite model in which, for each agent i, $v_i - k_i/c_i$ can take only m possible different values, i.e., $v_i - k_i/c_i \in \{d^1, \ldots, d^m\}$ and the probability mass function satisfies $f_i(d^t) = f_i^t$ for $t = 1, \ldots, m$. It is possible that $f_i^t = 0$ for some t. The corresponding problem of $(OPTA - \varphi)$ in the finite model, denoted by $(OPTAm - \varphi)$, is given by:

$$\max_P \sum_{i=1}^n \left[\sum_{t=1}^m f_i^t P_i^t d^t + \frac{\varphi_i k_i}{c_i}\right],$$

subject to

$$\varphi_i \leq P_i^t \leq \frac{\varphi_i}{1-c_i}, \forall t, \tag{AIC$'m$}$$

$$\sum_{i=1}^n \sum_{t\in S_i} f_i^t P_i^t \leq 1 - \prod_{i=1}^n \sum_{t\notin S_i} f_i^t, \forall S_i \subset \{1, \ldots, m\}. \tag{AF2m}$$

Define $H(\mathcal{S}) := 1-\prod_{i=1}^n \sum_{t\notin S_i} f_i^t - \sum_{i=1}^n \sum_{t\in S_i} \varphi_i f_i^t$ for all $\mathcal{S} := (S_1, \dots, S_n)$ and $S_i \subset \{1, \dots, m\}$ for all i. Define $\overline{H}(\mathcal{S}) := \min_{\mathcal{S}'\supset\mathcal{S}} H(\mathcal{S}')$ for all $\mathcal{S}$. Let $z_i^t := f_i^t \left(P_i^t - \varphi_i\right)$ for all i and t. By Lemma 16, $(OPTAm - \varphi)$ can be rewritten as $(OPTAm1 - \varphi)$

$$\max_z \sum_{i=1}^n \sum_{t=1}^m z_i^t d^t + \sum_{i=1}^n \varphi_i \left(\sum_{t=1}^m f_i^t d^t + \frac{k_i}{c_i} \right),$$

subject to

$$0 \le z_i^t \le \frac{c_i \varphi_i f_i^t}{1-c_i}, \forall i, \forall t, \tag{AIC$'m1$}$$

$$\sum_{i=1}^n \sum_{t\in S_i} z_i^t \le \overline{H}(\mathcal{S}), \forall \mathcal{S} \subset \{1, \dots, m\}^n. \tag{AF2$m1$}$$

Note that if $f_i^t = 0$, then $z_i^t = 0$ by definition and therefore satisfies (AIC$'m1$) automatically.

Algorithm 1 below describes an algorithm that finds a feasible solution to $(OPTAm - \varphi)$. I start by giving a verbal overview of the algorithm. It is in the spirit of greedy algorithms. It begins by assigning values to $\{z_i^m\}_i$ who have the largest weight d^m in the objective function. Let set $\mathcal{I}_0^m$ collect all the agents whose highest net values are below d^m. If $i \in \mathcal{I}_0^m$, then $f_i^m = 0$ by definition and $z_i^m = 0$ by (AIC$'m1$). Next check whether there exists some agent $i \notin \mathcal{I}_0^m$ such that if z_i^m is assigned the highest value allowed by (AF2$m1$), then the upper-bound on z_i^m in (AIC$'m1$) is respected. If so, assign z_i^m this highest value. Continue until no such agent can be found. Then, among all the agents whose z_i^m have not been assign values yet, check whether there exists a pair of agents, a triple of agents and etc. until there does not exist a group of agents $\mathcal{I}'$ such that if assign $\sum_{i\in\mathcal{I}'} z_i^m$ the highest value allowed by (AF2$m1$), then the the upper-bounds in (AIC$'m1$) can be respected. If now there still exists an agent i whose z_i^m has not been assigned a value yet, then I conjecture that the upper-bound on z_i^m in (AIC$'m1$) binds, and let $z_i^m = c_i\varphi_i f_i^t/(1-c_i)$. Let set $\mathcal{I}_1^m$ collect all the agents not in $\mathcal{I}_0^m$ and for whom the upper-bounds on z_i^m in (AIC$'m1$) do not bind. Continue to assign values to $\{z_i^{m-1}\}_i, \{z_i^{m-2}\}_i, \dots \{z_i^1\}_i$ in the same fashion.

In order to define the algorithm formally, I introduce some notations. Let $S_i^t := \{t, \dots, m\}$ and

$S_i^{m+1} := \emptyset$ for all i and t, $\mathcal{S}^t := \{t, \ldots, m\}^n$ for all t and $\mathcal{S}^{m+1} := \emptyset$. Define $\mathcal{S} + (t,i) := (S_1, \ldots, S_{i-1}, S_i \cup \{t\}, S_{i+1} \ldots, S_n)$ and $\mathcal{S} - (t,i) := (S_1, \ldots, S_{i-1}, S_i \backslash \{t\}, S_{i+1} \ldots, S_n)$.

Algorithm 1 *Let $\mathcal{I}_0^m := \left\{ i \,\middle|\, f_i^m = 0 \right\}$ and $\overline{z}_i^m := 0$ for all $i \in \mathcal{I}_0^m$. Define $\mathcal{I}_1^m \subset \mathcal{I} \backslash \mathcal{I}_0^m$, n^m, $\left\{ \pi^{m,1}, \ldots, \pi^{m,n^m} \right\}$, $\left\{ \mathcal{S}^{m,1}, \ldots, \mathcal{S}^{m,n^m} \right\}$ and $\overline{z}_i^m$ for all $i \notin \mathcal{I}_0^m$ recursively as follows.*

1. *Let $\mathcal{I}_1^m = \emptyset$ and $\nu = 1$.*

2. *If $\mathcal{I}_1^m = \mathcal{I} \backslash \mathcal{I}_0^m$, then go to step 5. Otherwise, let $\iota = 1$ and go to step 2.*

3. *If there exists $\mathcal{I}' \neq \emptyset$ such that $|\mathcal{I}'| = \iota$, $\mathcal{I}' \cap \left(\mathcal{I}_0^m \cup \mathcal{I}_1^m \right) = \emptyset$ and*

$$\overline{H}\left(\mathcal{S} + \sum_{i \in \mathcal{I}'} (m,i) \right) - \overline{H}(\mathcal{S}) \leq \sum_{i \in \mathcal{I}'} \frac{c_i \varphi_i f_i^m}{1 - c_i},$$

   where $S_j = S_j^m$ if $j \in \mathcal{I}_1^m$ and $S_j = S_j^{m+1}$ otherwise, then let $\overline{z}_i^m \leq c_i \varphi_i f_i^m / (1 - c_i)$ for $i \in \mathcal{I}'$ be such that

$$\sum_{i \in \mathcal{I}'} \overline{z}_i^m = \overline{H}\left(\mathcal{S} + \sum_{i \in \mathcal{I}'} (m,i) \right) - \overline{H}(\mathcal{S}).$$

   Let $\pi^{m,\nu} := \mathcal{I}'$ and $\mathcal{S}^{m,\nu} := \mathcal{S}$. Redefine ν as $\nu + 1$ and $\mathcal{I}_1^m$ as $\mathcal{I}' \cup \mathcal{I}_1^m$, and go to step 2. If there does not exist such an $\mathcal{I}'$, go to step 4.

4. *If $\iota < n - |\mathcal{I}_0^m \cup \mathcal{I}_1^m|$, then redefine ι as $\iota + 1$ and go to step 3. If $\iota = n - |\mathcal{I}_0^m \cup \mathcal{I}_1^m|$, then go to step 5.*

5. *Let $n^m := \nu - 1$ and $\overline{z}_i^m := c_i \varphi_i f_i^m / (1 - c_i)$ for all $i \in \mathcal{I} \backslash \left(\mathcal{I}_0^m \cup \mathcal{I}_1^m \right)$.*

Let $1 \leq t \leq m - 1$. Suppose that we have defined $\mathcal{I}_0^\tau$, $\mathcal{I}_1^\tau$, n^τ, $\left\{ \pi^{\tau,1}, \ldots, \pi^{\tau,n^\tau} \right\}$, $\left\{ \mathcal{S}^{\tau,1}, \ldots, \mathcal{S}^{\tau,n^\tau} \right\}$ and $\{\overline{z}_i^\tau\}_i$ for all $\tau \geq t + 1$. Let $\mathcal{I}_0^t := \left\{ i \middle| f_i^t = 0 \right\}$ and $\overline{z}_i^t := 0$ for all $i \in \mathcal{I}_0^t$. Define $\mathcal{I}_1^t \subset \mathcal{I} \backslash \mathcal{I}_0^t$, $\left\{ \pi^{t,1}, \ldots, \pi^{t,n^t} \right\}$, $\left\{ \mathcal{S}^{t,1}, \ldots, \mathcal{S}^{t,n^t} \right\}$ and $\overline{z}_i^t$ for all $i \notin \mathcal{I}_0^t$ recursively as follows.

1. *Let $\mathcal{I}_1^t := \emptyset$ and $\nu = 1$.*

2. *If* $\mathcal{I}_1^t = \mathcal{I} \backslash \mathcal{I}_0^t$*, then go to step 5. Otherwise, let* $\imath = 1$ *and go to step 2.*

3. *If there exists* $\mathcal{I}' \neq \emptyset$ *such that* $|\mathcal{I}'| = \imath$, $\mathcal{I}' \cap \left(\mathcal{I}_0^t \cup \mathcal{I}_1^t\right) = \emptyset$ *and*

$$\min_{\mathcal{S}} \overline{H}\left(\mathcal{S} + \sum_{i \in \mathcal{I}'} (t, i)\right) - \sum_{j=1}^{n} \sum_{\tau \in S_j} \overline{z}_j^\tau \leq \sum_{i \in \mathcal{I}'} \frac{c_i \varphi_i f_i^t}{1 - c_i},$$

where $\mathcal{S} = (S_1^{t_1}, \ldots, S_n^{t_n})$ *with* $t_j \geq t$ *if* $j \in \pi^{t,1} \cup \cdots \cup \pi^{t,\nu-1}$, $t_j = t+1$ *if* $j \in \mathcal{I}'$ *and* $t_j \geq t+1$ *otherwise, then let* $\overline{z}_i^t \leq c_i \varphi_i f_i^t / (1 - c_i)$ *for* $i \in \mathcal{I}'$ *be such that*

$$\sum_{i \in \mathcal{I}'} \overline{z}_i^t = \min_{\mathcal{S}} \overline{H}\left(\mathcal{S} + \sum_{i \in \mathcal{I}'} (m, i)\right) - \sum_{i=1}^{n} \sum_{\tau \in S_i} \overline{z}_i^\tau.$$

Let $\pi^{t,\nu} := \mathcal{I}'$ *and* $\mathcal{S}^{t,\nu}$ *as a minimizer of the right-hand side of the above equation such that there is no* $\mathcal{S} \supsetneq \mathcal{S}^{t,\nu}$ *which is also a minimizer. Redefine* ν *as* $\nu + 1$ *and* $\mathcal{I}_1^t$ *as* $\mathcal{I}' \cup \mathcal{I}_1^t$*, and go to step 2. If there does not exist such an* $\mathcal{I}'$*, then go to step 4.*

4. *If* $\imath < n - |\mathcal{I}_0^t \cup \mathcal{I}_1^t|$*, then redefine* $\imath$ *as* $\imath + 1$ *and go to step 3. If* $\imath = n - |\mathcal{I}_0^t \cup \mathcal{I}_1^t|$*, then go to step 5.*

5. *Let* $n^t := \nu - 1$ *and* $\overline{z}_i^t := c_i \varphi_i f_i^t / (1 - c_i)$ *for all* $i \in \mathcal{I} \backslash \left(\mathcal{I}_0^t \cup \mathcal{I}_1^t\right)$.

Note that $\{\mathcal{S}^{t,\nu} + \sum_{i \in \pi^{t,\nu}} (t, i)\}$ is the collection of sets for which (AF2m1) bind.

Let $\overline{z}$ be a solution found by Algorithm 1. I first prove that $\overline{z}$ is a feasible solution to $(OPTAm1-\varphi)$. For each i and t, let $\overline{P}_i^t := \overline{z}_i^t / f_i^t + \varphi_i$ if $f_i^t > 0$ and $\overline{P}_i^t := 0$ otherwise. Then $\overline{z}$ is a feasible solution to $(OPTAm1-\varphi)$ if and only if $\overline{P}$ is a feasible solution to $(OPTAm-\varphi)$. Lemma 25 below proves that $\overline{P}$ is non-decreasing. By Theorem 2 in Mierendorff (2011), $\overline{P}$ is a feasible solution to $(OPTAm-\varphi)$ if and only if for all $t_1, \ldots, t_n \in \{1, \ldots, m\}$

$$\sum_{i=1}^{n} \sum_{t \in S_i} \overline{P}_i^t \leq 1 - \prod_{i=1}^{n} \sum_{t \notin S_i^{t_i}} f_i^t.$$

By construction, this is true if and only if for all $t_1, \ldots, t_n \in \{1, \ldots, m\}$,

$$\sum_{i=1}^{n} \sum_{t \in S_i} \overline{z}_i^t \leq \overline{H}(\mathcal{S}), \tag{A.6}$$

where $\mathcal{S} = (S_1^{t_1}, \ldots, S_n^{t_n})$, which is proved in Lemma 24.

Hence, $\overline{\boldsymbol{P}}$ is a feasible solution to $(OPTAm{-}\varphi)$, or equivalently, $\overline{z}$ is a feasible solution to $(OPTAm1-\varphi)$. Let

$$\hat{z}_i^t := \begin{cases} \overline{z}_i^t & \text{if } d^t \geq 0 \\ 0 & \text{if } d^t < 0 \end{cases}. \tag{A.7}$$

Clearly, $\hat{z}$ is also a feasible solution to $(OPTAm1 - \varphi)$. Furthermore,one can verify that $\hat{z}$ is an optimal solution to $(OPTAm1 - \varphi)$ by the duality theorem:

Lemma 21 *$\hat{z}$ define in* (A.7) *is an optimal solution to* $(OPTAm1 - \varphi)$*.*

Finally, let $\boldsymbol{P}^m := (P_i^{m,t})_{i,t}$, where

$$P_i^{m,t} := \begin{cases} \overline{P}_i^{m,t} & \text{if } d^t \geq 0 \\ \varphi_i & \text{if } d^t < 0 \end{cases}. \tag{A.8}$$

The following corollary directly follows from Lemma 21:

Corollary 9 *$\boldsymbol{P}^m$ defined in* (A.8) *is an optimal solution to* $(OPTAm - \varphi)$*.*

The rest of this subsection is organized as follows. In Appendix A.3.1, I prove two technical lemmas on H and $\overline{H}$, which are useful for later proofs. In Appendix A.3.1, I prove that $\overline{z}$ is a feasible solution to $(OPTAm - \varphi)$ by proving Lemmas 24 and 25. In Appendix A.3.1, I prove that $\hat{z}$ is an optimal solution to $(OPTAm - \varphi)$. In Appendix A.3.1, I prove two technical lemmas that are useful in characterizing the limit of $\{\boldsymbol{P}^m\}$.

Properties of H and $\overline{H}$

Here, I introduce two technical lemmas on H and $\overline{H}$. Lemma 22 proves a useful property of H. Lemma 23 characterizes $\overline{H}$.

Lemma 22 *If $H(\mathcal{S}) < 1 - \sum_{i=1}^n \varphi_i$ and $\mathcal{S}' \subset \mathcal{S}$, then $H(\mathcal{S}') \leq H(\mathcal{S})$.*

Proof. Consider $\mathcal{S} = (S_1, \dots, S_n)$. We have

$$H(\mathcal{S}) - 1 + \sum_{i=1}^n \varphi_i = \sum_{i=1}^n \varphi_i \sum_{\tau \notin S_i} f_i^\tau - \prod_{i=1}^n \sum_{\tau \notin S_i} f_i^\tau .$$

Let $S_i^{supp} := \{t | f_i^t > 0\}$. If $S_i = S_i^{supp}$ for some i, then $\sum_{\tau \notin S_i} f_i^\tau = 0$ and therefore $H(\mathcal{S}) \geq 1 - \sum_{i=1}^n \varphi_i$. Hence, $H(\mathcal{S}) < 1 - \sum_{i=1}^n \varphi_i$ implies that $S_i \neq S_i^{supp}$ or $\sum_{\tau \notin S_i} f_i^\tau > 0$ for all i. Thus, $\varphi_i \leq \prod_{j \neq i} \sum_{\tau \notin S_j} f_j^\tau$ for all i. Let $\mathcal{S}' := (S_1, \dots, S_{i-1}, S_i \backslash \{t\}, S_{i+1} \dots, S_n)$. Then

$$H(\mathcal{S}) - H(\mathcal{S}') = f_i^t \left(\prod_{j \neq i} \sum_{\tau \notin S_j} f_j^\tau - \varphi_i \right) \geq 0.$$

Hence, $H(\mathcal{S}') \leq H(\mathcal{S})$. By induction, $H(\mathcal{S}') \leq H(\mathcal{S})$ for all $\mathcal{S}' \subset \mathcal{S}$. ■

Lemma 23 $\overline{H}(\mathcal{S}) = \min \{H(\mathcal{S}), 1 - \sum_{i=1}^n \varphi_i\}$.

Proof. Recall that $\overline{H}(\mathcal{S}) = \min_{\mathcal{S}'' \supset \mathcal{S}} H(\mathcal{S})$. Recall that $\mathcal{S}^1 := \{1, \dots, m\}^n$. Because $\mathcal{S}^1 \supset \mathcal{S}$ and $H(\mathcal{S}^1) = 1 - \sum_{i=1}^n \varphi_i$, we have $\overline{H}(\mathcal{S}) \leq 1 - \sum_{i=1}^n \varphi_i$.

Suppose that $H(\mathcal{S}) \leq 1 - \sum_{i=1}^n \varphi_i$. Let $\mathcal{S}'' \supset \mathcal{S}$. If $H(\mathcal{S}'') \geq 1 - \sum_{i=1}^n \varphi_i$, then $H(\mathcal{S}) \leq 1 - \sum_{i=1}^n \varphi_i \leq H(\mathcal{S}'')$. If $H(\mathcal{S}'') < 1 - \sum_{i=1}^n \varphi_i$, then $H(\mathcal{S}) \leq H(\mathcal{S}'')$ by Lemma 22. Hence, $\overline{H}(\mathcal{S}) = H(\mathcal{S})$.

Suppose that $H(\mathcal{S}) > 1 - \sum_{i=1}^n \varphi_i$. I claim that $\overline{H}(\mathcal{S}) = 1 - \sum_{i=1}^n \varphi_i$. Suppose not, then there exists $\mathcal{S}'' \supset \mathcal{S}$ such that $H(\mathcal{S}'') < 1 - \sum_{i=1}^n \varphi_i$. Then, by Lemma 22, $H(\mathcal{S}) \leq H(\mathcal{S}'') < 1 - \sum_{i=1}^n \varphi_i$, which is a contradiction to the fact that $H(\mathcal{S}) > 1 - \sum_{i=1}^n \varphi_i$. Hence, $\overline{H}(\mathcal{S}) = 1 - \sum_{i=1}^n \varphi_i$. ■

Proofs of feasibility

Lemma 24 *For all* $t_1, \dots, t_n \in \{1, \dots, m\}$,

$$\sum_{i=1}^{n} \sum_{t \in S_i} \overline{z}_i^t \le \overline{H}(\mathcal{S}), \tag{A.6}$$

where $\mathcal{S} = (S_1^{t_1}, \dots, S_n^{t_n})$.

Proof. For each t, let $\pi^{t,0} := \emptyset$ and $\pi^{t,n^t+1} := \mathcal{I} \backslash (\mathcal{I}_0^t \cup \mathcal{I}_1^t)$. Suppose that $\mathcal{S} \subset \mathcal{S}^m$, i.e., $t_i \ge m$ for all i. By Algorithm 1, we have

$$\sum_{i \in \pi^{m,1}} \overline{z}_i^m = \overline{H}\left(\emptyset + \sum_{i \in \pi^{m,1}} (m,i)\right),$$
$$\sum_{i \in \mathcal{I}''} \overline{z}_i^m \le \sum_{i \in \mathcal{I}''} \frac{c_i \varphi_i f_i^m}{1 - c_i} \le \overline{H}\left(\emptyset + \sum_{i \in \mathcal{I}''} (m,i)\right), \forall \emptyset \ne \mathcal{I}'' \subsetneq \pi^{m,1},$$

where the second inequality in the second line holds because otherwise $|\pi^{m,1}| \le |\mathcal{I}''|$ by Algorithm 1, which is a contradiction to $\mathcal{I}'' \subsetneq \pi^{m,1}$. Thus, (A.6) holds if $t_i = m+1$ for all $i \notin \pi^{m,1}$. Suppose that we have shown that (A.6) holds if $t_i = m+1$ for all $i \notin \pi^{m,1} \cup \cdots \cup \pi^{m,\nu-1}$ and $\nu \ge 2$. Suppose that $t_i = m+1$ for all $i \notin \pi^{m,1} \cup \cdots \cup \pi^{m,\nu}$. Let $\mathcal{S}' := \emptyset + \sum_{i \in \pi^{m,1} \cup \cdots \cup \pi^{m,\nu-1}} (m,i)$ and $\mathcal{I}'' := \left\{ i \in \pi^{m,\nu} | t_i = m \right\} \subset \pi^{m,\nu}$. By Algorithm 1, we have

$$\sum_{i \in \mathcal{I}''} \overline{z}_i^m = \overline{H}\left(\mathcal{S}' + \sum_{i \in \mathcal{I}'} (m,i)\right) - \overline{H}(\mathcal{S}') \text{ if } \mathcal{I}'' = \pi^{m,\nu} \text{ and } \nu \le n^m,$$
$$\sum_{i \in \mathcal{I}''} \overline{z}_i^m \le \sum_{i \in \mathcal{I}''} \frac{c_i \varphi_i f_i^m}{1 - c_i} \le \overline{H}\left(\mathcal{S}' + \sum_{i \in \mathcal{I}''} (m,i)\right) - \overline{H}(\mathcal{S}') \text{ if } \mathcal{I}'' \subsetneq \pi^{m,\nu} \text{ or } \nu = n^m + 1.$$

Because $\mathcal{S} - \sum_{i \in \pi^{m,\nu}} (m, i) \subset \mathcal{S}'$, we have

$$\begin{aligned}
\sum_{i=1}^{n} \sum_{t \in S_i' \setminus S_i} \overline{z}_i^t &= \sum_{i=1}^{n} \sum_{t \in S_i'} \overline{z}_i^t - \sum_{i \notin \pi^{m,\nu}} \sum_{t \in S_i} \overline{z}_i^t \\
&\geq \overline{H}(\mathcal{S}') - \overline{H}\left(\mathcal{S} - \sum_{i \in \pi^{m,\nu}} (m, i) \right) \\
&\geq \overline{H}\left(\mathcal{S}' + \sum_{i \in \mathcal{I}'} (m, i) \right) - \overline{H}(\mathcal{S}),
\end{aligned}$$

where the last inequality holds because $\overline{H}$ is submodular. Hence,

$$\begin{aligned}
\sum_{i=1}^{n} \sum_{t \in S_i} \overline{z}_i^t &\leq \sum_{i=1}^{n} \sum_{t \in S_i'} \overline{z}_i^t - \sum_{i=1}^{n} \sum_{t \in S_i' \setminus S_i} \overline{z}_i^t + \overline{H}\left(\mathcal{S}' + \sum_{i \in \mathcal{I}'} (m, i) \right) - \overline{H}(\mathcal{S}') \\
&\leq \overline{H}(\mathcal{S}') - \overline{H}\left(\mathcal{S}' + \sum_{i \in \mathcal{I}'} (m, i) \right) + \overline{H}(\mathcal{S}) + \overline{H}\left(\mathcal{S}' + \sum_{i \in \mathcal{I}'} (m, i) \right) - \overline{H}(\mathcal{S}') \\
&= \overline{H}(\mathcal{S}).
\end{aligned}$$

By induction, (A.6) holds for all $\mathcal{S} \subset \mathcal{S}^m$.

Suppose that $\mathcal{S} \subset \mathcal{S}^{t+1} + \sum_{i \in \pi^{t,1} \cup \cdots \cup \pi^{t,\nu}} (t, i)$ for $t \leq m-1$ and $1 \leq \nu \leq n^t + 1$. Let $\mathcal{I}' := \{ i \in \pi^{t,\nu} | t_i = t \}$ and $\mathcal{S}' := \mathcal{S} - \sum_{i \in \mathcal{I}'} (t, i)$. Suppose, w.l.o.g., that $\mathcal{I}' \neq \emptyset$. If $\mathcal{I}' = \pi^{t,\nu}$, then, by Algorithm 1, we have

$$\sum_{i \in \mathcal{I}'} \overline{z}_i^\tau \leq \overline{H}\left(\mathcal{S}' + \sum_{i \in \mathcal{I}'} (t, i) \right) - \sum_{i=1}^{n} \sum_{\tau \in S_i'} \overline{z}_i^\tau = \overline{H}\,(\mathcal{S}) - \sum_{i=1}^{n} \sum_{\tau \in S_i'} \overline{z}_i^\tau.$$

If $\mathcal{I}' \subsetneq \pi^{t,\nu}$, then, by Algorithm 1, we have

$$\sum_{i \in \mathcal{I}'} \overline{z}_i^\tau \leq \sum_{i \in \mathcal{I}'} \frac{c_i \varphi_i f_i^t}{1 - c_i} \leq \overline{H}\left(\mathcal{S}' + \sum_{i \in \mathcal{I}'} (t, i) \right) - \sum_{i=1}^{n} \sum_{\tau \in S_i'} \overline{z}_i^\tau = \overline{H}\,(\mathcal{S}) - \sum_{i=1}^{n} \sum_{\tau \in S_i'} \overline{z}_i^\tau,$$

where the second inequality holds because otherwise $|\pi^{t,\nu}| \leq |\mathcal{I}'|$ by Algorithm 1, which is a contradiction to $\mathcal{I}' \subsetneq \pi^{t,\nu}$. Hence, (A.6) holds for $\mathcal{S}$. ■

Lemma 25 $\overline{P}_i^t$ *is non-decreasing in t on* $\{t \,|\, f_i^t > 0\}$.

To prove Lemma 25, I first prove the following lemma which says that if the upper-bound in (AIC$'m$1) does not bind for z_i^{t+1}, then it does not bind for z_i^t.

Lemma 26 *Suppose that* $f_i^t, f_i^{t+1} > 0$. *Then* $\overline{z}_i^{t+1} \in \mathcal{I}_1^{t+1}$ *implies that* $\overline{z}_i^t \in \mathcal{I}_1^t$.

Proof. Suppose that $f_i^t, f_i^{t+1} > 0$ and $\overline{z}_i^{t+1} \in \mathcal{I}_1^{t+1}$. Then, by Algorithm 1, there exists $\mathcal{S}$ with $S_j = S_j^{t_j} \subset S_j^{t+1}$ for all $j \neq i$ and $S_i = S_i^{t+1}$ such that

$$\sum_{j=1}^{n} \sum_{\tau \in S_j} \overline{z}_j^\tau = \overline{H}(\mathcal{S}).$$

Suppose that $H(\mathcal{S}) < 1 - \sum_{j=1}^n \varphi_j$. Because, by Lemma 24,

$$\sum_{j \neq i} \sum_{\tau \in S_j} \overline{z}_j^\tau + \sum_{\tau \in S_i \setminus \{t+1\}} \overline{z}_i^\tau \leq \overline{H}(\mathcal{S} - (t+1, i)),$$

we have

$$\frac{c_i \varphi_i f_i^{t+1}}{1 - c_i} \geq \overline{z}_i^{t+1} \geq \overline{H}(\mathcal{S}) - \overline{H}(\mathcal{S} - (t+1, i)) = f_i^{t+1} \left(\prod_{j \neq i} \sum_{\tau \in S_j} f_j^\tau - \varphi_i \right),$$

where the last equality holds by Lemmas 22 and 23. This implies that $\prod_{j \neq i} \sum_{\tau \in S_j} f_j^\tau \leq \frac{\varphi_i}{1 - c_i}$. Hence,

$$\begin{aligned} \overline{z}_i^t \leq& \overline{H}(\mathcal{S} + (t, i)) - \sum_{j=1}^{n} \sum_{\tau \in S_j} \overline{z}_j^\tau \\ \leq& H(\mathcal{S} + (t, i)) - \overline{H}(\mathcal{S}) \\ =& f_i^t \left(\prod_{j \neq i} \sum_{\tau \in S_j} f_j^\tau - \varphi_i \right) \leq \frac{c_i \varphi_i f_i^t}{1 - c_i}, \end{aligned}$$

where the equality holds by Lemmas 22 and 23.

Suppose that $H(\mathcal{S}) \geq 1 - \sum_{j=1}^n \varphi_j$, then by Lemmas 22 and 23, $\overline{H}(\mathcal{S}) = \overline{H}(\mathcal{S} + (t, i)) = 1 - \sum_{j=1}^n \varphi_j$.

Hence,

$$
\begin{aligned}
\overline{z}_i^t \leq & \overline{H}(\mathcal{S} + (t,i)) - \sum_{j=1}^{n} \sum_{\tau \in S_j} \overline{z}_j^\tau \\
\leq & \overline{H}(\mathcal{S} + (t,i)) - \overline{H}(\mathcal{S}) \\
= & 0 \leq \frac{c_i \varphi_i f_i^t}{1 - c_i}.
\end{aligned}
$$

Hence, $z_i^t \in \mathcal{I}_1^t$. ∎

Proof of Lemma 25. Suppose that $f_i^t, f_i^{t+1} > 0$. Recall that $\overline{P}_i^t = \overline{z}_i^t / f_i^t + \varphi_i$ if $f_i^t > 0$. Suppose that $\overline{z}_i^{t+1} \notin \mathcal{I}_1^{t+1}$, then $\overline{P}_i^t \leq \frac{\varphi_i}{1-c_i} = \overline{P}_i^{t+1}$. Suppose that $\overline{z}_i^{t+1} \in \mathcal{I}_1^{t+1}$. Then there exists $\mathcal{S}$ with $S_j = S_j^{t_j} \subset S_j^{t+1}$ for all $j \neq i$ and $S_i = S_i^{t+1}$ such that

$$
\sum_{j=1}^{n} \sum_{\tau \in S_j} \overline{z}_j^\tau = \overline{H}(\mathcal{S}).
$$

Suppose that $H(\mathcal{S}) < 1 - \sum_{j=1}^{n} \varphi_j$. In the proof of Lemma 26, we have shown that $\overline{z}_i^{t+1} \geq f_i^{t+1} \left(\prod_{j \neq i} \sum_{\tau \in S_j} f_j^\tau - \varphi_i \right)$ and $\overline{z}_i^t \leq f_i^t \left(\prod_{j \neq i} \sum_{\tau \in S_j} f_j^\tau - \varphi_i \right)$. Hence,

$$
\overline{P}_i^t \leq \prod_{j \neq i} \sum_{\tau \in S_j} f_j^\tau \leq \overline{P}_i^{t+1}.
$$

Suppose that $H(\mathcal{S}) \geq 1 - \sum_{j=1}^{n} \varphi_j$. By the proof of Lemma 26, we have $\overline{P}_i^t = \varphi_i \leq \overline{P}_i^{t+1}$. ∎

Proofs of optimality

Before proving Lemma 21, I first prove some useful properties of $\mathcal{S}^{t,\nu}$ and $\overline{z}$. Recall that $\{\mathcal{S}^{t,\nu} + \sum_{i \in \pi^{t,\nu}} (t,i)\}$ is the collection of sets for which (AF2m1) bind. The result in Lemma 27 implies that this collection is a *nested sequence of sets*. In fact, Lemma 27 proves a stronger statement.

Lemma 27 $\mathcal{S}^{t,1} \supset \mathcal{S}^{t+1,n^{t+1}} + \sum_{i \in \pi^{t+1,n^{t+1}}} (t+1,i)$ *for* $1 \leq t \leq m-1$*; and* $\mathcal{S}^{t,\nu+1} \supset \mathcal{S}^{t,\nu} + \sum_{i \in \pi^{t,\nu}} (t,i)$

for $1 \leq t \leq m$.

Proof. By Algorithm 1, $\mathcal{S}^{m,\nu+1} \supset \mathcal{S}^{m,\nu} + \sum_{i\in\pi^{m,\nu}}(m,i)$. Let $t \leq m-1$ and $\mathcal{I}' = \pi^{t,1}$. Let $\mathcal{S} := \mathcal{S}^{t+1,n^{t+1}} + \sum_{i\in\pi^{t+1,n^{t+1}}}(t+1,i)$. Then $\sum_{j=1}^{n}\sum_{\tau\in S_j}\overline{z}_j^\tau = \overline{H}(\mathcal{S})$. Suppose $\mathcal{S}^{t,1} \not\supset \mathcal{S}$. Let $\mathcal{S}' := \mathcal{S}\cup\mathcal{S}^{t,1}$. Then $S'_j = S_j^{t_j}$ for some $t_j \geq t+1$ for all j. By Lemma 24, we have

$$\begin{aligned}
&\sum_{j=1}^{n}\sum_{\tau\in S'_j\setminus S_j^{t,1}}\overline{z}_j^\tau \\
=&\sum_{j=1}^{n}\sum_{\tau\in S_j}\overline{z}_j^\tau - \sum_{j=1}^{n}\sum_{\tau\in S_j^{t,1}\cap S_j}\overline{z}_j^\tau \\
\geq&\overline{H}(\mathcal{S}) - \overline{H}(\mathcal{S}\cap\mathcal{S}^{t,1}).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\overline{H}\left(\mathcal{S}' + \sum_{i\in\mathcal{I}'}(t,i)\right) - \sum_{j=1}^{n}\sum_{\tau\in S'_j}\overline{z}_j^\tau - \overline{H}\left(\mathcal{S}^{t,1} + \sum_{i\in\mathcal{I}'}(t,i)\right) + \sum_{j=1}^{n}\sum_{\tau\in S_j^{t,1}}\overline{z}_j^\tau \\
=&\overline{H}\left(\mathcal{S}' + \sum_{i\in\mathcal{I}'}(t,i)\right) - \overline{H}\left(\mathcal{S}^{t,1} + \sum_{i\in\mathcal{I}'}(t,i)\right) - \sum_{j=1}^{n}\sum_{\tau\in S'_j\setminus S_j^{t,1}}\overline{z}_j^\tau \\
\leq&\left[\overline{H}\left(\mathcal{S}' + \sum_{i\in\mathcal{I}'}(t,i)\right) - \overline{H}(\mathcal{S})\right] - \left[\overline{H}\left(\mathcal{S}^{t,1} + \sum_{i\in\mathcal{I}'}(t,i)\right) - \overline{H}(\mathcal{S}\cap\mathcal{S}^{t,1})\right] \\
\leq&0,
\end{aligned}$$

where the last inequality holds because $\overline{H}$ is submodular, which is a contradiction to the definition of $\mathcal{S}^{t,\nu}$. Hence, $\mathcal{S}^{t,1} \supset \mathcal{S}^{t+1,n^{t+1}} + \sum_{i\in\pi^{t+1,n^{t+1}}}(t+1,i)$. By a similar argument, one can show that $\mathcal{S}^{t,\nu+1} \supset \mathcal{S}^{t,\nu} + \sum_{i\in\pi^{t,\nu}}(t,i)$ for all $t \leq m-1$. ■

By Lemmas 22, 23 and 27, there exists $\underline{t}$ and $\overline{\nu}$ such that

$$\overline{H}\left(\mathcal{S}^{t,\nu} + \sum_{i\in\pi^{t,\nu}}(t,i)\right) = \begin{cases} 1 - \sum_i \varphi_i & \text{if } t < \underline{t} \text{ or } \nu \geq \overline{\nu}, t = \underline{t} \\ H\left(\mathcal{S}^{t,\nu} + \sum_{i\in\pi^{t,\nu}}(t,i)\right) < 1 - \sum_i \varphi_i & \text{otherwise} \end{cases}. \quad \text{(A.9)}$$

The definition of $\underline{t}$ is analogous to that in the symmetric case. By a similar argument to that in Lemma 26, we have

Lemma 28 *If $t < \underline{t}$, or $t = \underline{t}$ and $i \notin \pi^{t,1} \cup \cdots \cup \pi^{t,\bar{\nu}}$, then $\overline{z}_i^t = 0$.*

Proof of Lemma 21. Consider the dual to problem $(OPTAm1-\varphi)$, which is denoted by $(DOPTAm1-\varphi)$,

$$\min_{\lambda,\beta,\mu} \sum_{i=1}^{n} \sum_{t=1}^{m} \frac{\lambda_i^t c_i \varphi_i f_i^t}{1-c_i} + \sum_{\mathcal{S}} \beta(\mathcal{S}) \overline{H}(\mathcal{S}) + \sum_{i=1}^{n} \varphi_i \left(\sum_{t=1}^{m} f_i^t d^t + \frac{k_i}{c_i} \right),$$

subject to

$$d^t - \lambda_i^t + \mu_i^t - \sum_{\mathcal{S}_i \ni t} \beta(\mathcal{S}) \geq 0 \text{ if } f_i^t > 0, \forall i, \forall t,$$

$$\lambda \geq 0, \mu \geq 0, \beta \geq 0.$$

Let $\hat{z}$ be defined by (A.7) and $(\hat{\beta}, \hat{\lambda}, \hat{\mu})$ be the corresponding dual variables. Let t^0 be such that $d^{t^0} \geq 0$ if and only if $t \geq t^0$.

Let $\hat{\beta}\left(\mathcal{S}^{t,n^t} + \sum_{i \in \pi^{t,n^t}} (t,i)\right) \geq 0$ for $t \geq \max\{t^0, \underline{t}\}$ and $\hat{\beta}(\mathcal{S}) = 0$ otherwise. (i) If $t < \max\{t^0, \underline{t}\}$, then let $\hat{\lambda}_i^t = 0$ and $\hat{\mu}_i^t = -d^t \geq 0$. (ii) If $t = \max\{t^0, \underline{t}\}$, then let $\hat{\beta}\left(\mathcal{S}^{t,n^t} + \sum_{j \in \pi^{t,n^t}} (t,j)\right) = d^t \geq 0$ and $\hat{\mu}_i^t = 0$. If $i \in \mathcal{I}_1^t$, then let $\hat{\lambda}_i^t = 0$. If $i \notin \mathcal{I}_0^t \cup \mathcal{I}_1^t$, let $\hat{\lambda}_i^t = d^t \geq 0$. (iii) If $t > \max\{t^0, \underline{t}\}$, let $\hat{\beta}\left(\mathcal{S}^{t,n^t} + \sum_{j \in \pi^{t,n^t}} (t,j)\right) = d^t - d^{t-1} \geq 0$ and $\hat{\mu}_i^t = 0$. If $i \in \mathcal{I}_1^t$, then let $\hat{\lambda}_i^t = 0$. If $i \notin \mathcal{I}_0^t \cup \mathcal{I}_1^t$ and $i \in \mathcal{I}_1^{\max\{t^0, \underline{t}\}}$, then let $\hat{\lambda}_i^t = d^t - d^{t^*} \geq 0$ where $t^* = \min\{t' \geq \max\{t^0, \underline{t}\} | \mathcal{I}_1^{t'} \ni i\}$. If $i \notin \mathcal{I}_1^{\max\{t^0, \underline{t}\}}$, then let $\hat{\lambda}_i^t = 0$. Hence, $(\hat{\lambda}, \hat{\mu}, \hat{\beta})$ is a feasible solution to $(DOPTAm1 - \varphi)$ and the complementary slackness conditions are satisfied. Finally, it is easy to verify that the dual objective is equal to the primal objective. By the duality theorem, $\hat{z}$ is an optimal solution to $(OPTAm1 - \varphi)$. ■

Properties of $S^{t,\nu}$

Before moving on to the continuum case, I prove the following two lemmas which are useful in characterizing the limit of $\{P^m\}$.

Lemma 29 *Suppose* $\mathcal{S}^{t,\nu} = (S_1^{t_1^*}, \dots, S_n^{t_n^*})$. *Then* $t_i^* = t$ *if* $i \in \pi^{t,1} \cup \cdots \cup \pi^{t,\nu-1}$, $t_i^* = t+1$ *if* $i \in \mathcal{I}_1^{t+1} \cup \pi^{t,\nu} \backslash \left(\pi^{t,1} \cup \cdots \cup \pi^{t,\nu-1}\right)$, *and* $t_i^* \in \{t+1, m+1\}$ *otherwise. Furthermore, for* $h \notin \mathcal{I}_1^{t+1} \cup \pi^{t,1} \cup \cdots \cup \pi^{t,\nu}$, *we have*

1. *If* $\frac{\varphi_h}{1-c_h} - \prod_{i \neq h} \sum_{\tau=1}^{t_i^*-1} f_i^\tau \geq 0$, *then* $t_h^* = t+1$.

2. *If* $\frac{\varphi_h}{1-c_h} - \prod_{i \neq h} \sum_{\tau=1}^{t_i^*-1} f_i^\tau < 0$ *and* $\overline{H}\left(\mathcal{S}^{t,\nu} + \sum_{i \in \pi^{t,\nu}} (t,i)\right) < 1 - \sum_{i=1}^n \varphi_i$, *then* $t_h^* = m+1$.

Proof. By Algorithm 1, $t_i^* = t+1$ if $i \in \pi^{t,\nu}$. By Lemma 27, $t_i^* = t$ if $i \in \pi^{t,1} \cup \cdots \cup \pi^{t,\nu-1}$ and $t_i^* = t+1$ if $i \in \mathcal{I}_1^{t+1} \backslash \left(\pi^{t,1} \cup \cdots \cup \pi^{t,\nu-1}\right)$. If $t = m$, then, by Algorithm 1, $t_i^* = m+1$ for $i \notin \pi^{t,1} \cup \cdots \cup \pi^{t,\nu-1}$.

Let $t \leq m-1$. For the ease of notation, let $\mathcal{I}' = \pi^{t,\nu}$ and $\mathcal{S} = (S_1^{t_1}, \dots, S_n^{t_n})$ be such that $t_i = t$ if $i \in \pi^{t,1} \cup \cdots \cup \pi^{t,\nu-1}$, $t_i = t+1$ if $i \in \mathcal{I}_1^{t+1} \cup \pi^{t,\nu} \backslash \left(\pi^{t,1} \cup \cdots \cup \pi^{t,\nu-1}\right)$ and $t_i \geq t+1$ otherwise. Fix $h \notin \mathcal{I}_1^{t+1} \cup \pi^{t,1} \cup \cdots \cup \pi^{t,\nu}$ and t_i for all $i \neq h$. Define

$$\Delta(t_h) := \overline{H}\left(\mathcal{S} + \sum_{i \in \mathcal{I}'} (t,i)\right) - \sum_{i=1}^n \sum_{\tau \in S_i} \overline{z}_i^\tau.$$

By Lemma 28 and the fact that $h \notin \mathcal{I}_1^{t+1}$, there exists $t \leq t^* \leq m+1$ such that if $t+1 \leq t_h \leq t^*$, then $\Delta(t_h) = 1 - \sum_{i=1}^n \varphi_i - \sum_{i=1}^n \sum_{\tau \in S_i} \overline{z}_i^\tau$; and if $t^* < t_h \leq m+1$, then

$$\Delta(t_h) = 1 - \left(\prod_{i \notin \mathcal{I}'} \sum_{\tau=1}^{t_i-1} f_i^\tau\right)\left(\prod_{i \in \mathcal{I}'} \sum_{\tau=1}^{t-1} f_i^\tau\right) - \sum_{i \notin \mathcal{I}'} \sum_{\tau=t_i}^m f_i^\tau \varphi_i - \sum_{i \in \mathcal{I}'} \sum_{\tau=t}^m f_i^\tau \varphi_i - \sum_{i=1}^n \sum_{\tau=t_i}^m \overline{z}_i^\tau.$$

Because $h \notin \mathcal{I}_1^{t+1}$, we have $\overline{z}_h^{t_h} = c_h \varphi_h f_h^{t_h}/(1-c_h)$ for all $t_h \geq t+1$. If $t_h < t^*$, then we have $\Delta(t_h+1) - \Delta(t_h) = c_h \varphi_h f_h^{t_h}/(1-c_h) \geq 0$. Hence, $\Delta(t+1) \leq \Delta(t_h)$ for all $t_h \leq t^*$. If $t_h > t^*$, we have

$$\begin{aligned}
&\Delta(t_h+1) - \Delta(t_h) \\
=& f_h^{t_h}\left(\frac{\varphi_h}{1-c_h} - \left(\prod_{i \notin \mathcal{I}', i \neq h} \sum_{\tau=1}^{t_i-1} f_i^\tau\right)\left(\prod_{i \in \mathcal{I}'} \sum_{\tau=1}^{t-1} f_i^\tau\right)\right).
\end{aligned}$$

If $t_h = t^*$, we have

$$\begin{aligned}&\Delta(t_h+1)-\Delta(t_h)\\ \geq& f_h^{t_h}\left(\frac{\varphi_h}{1-c_h}-\left(\prod_{i\notin\mathcal{I}',i\neq h}\sum_{\tau=1}^{t_i-1}f_i^\tau\right)\left(\prod_{i\in\mathcal{I}'}\sum_{\tau=1}^{t-1}f_i^\tau\right)\right).\end{aligned}$$

Hence, if $\frac{\varphi_h}{1-c_h}-\left(\prod_{i\notin\mathcal{I}',i\neq h}\sum_{\tau=1}^{t_i-1}f_i^\tau\right)\left(\prod_{i\in\mathcal{I}'}\sum_{\tau=1}^{t-1}f_i^\tau\right)\geq 0$, then $\Delta(t_h+1)\geq\Delta(t_h)$ for all $t_h\geq t^*$. Furthermore, because $\Delta(t+1)\leq\Delta(t_h)$ for all $t_h\leq t^*$, we have $\Delta(t+1)\leq\Delta(t_h)$ for all $t_h\geq t+1$, hence $t_h^*=t+1$.

If $\frac{\varphi_h}{1-c_h}-\left(\prod_{i\notin\mathcal{I}',i\neq h}\sum_{\tau=1}^{t_i-1}f_i^\tau\right)\left(\prod_{i\in\mathcal{I}'}\sum_{\tau=1}^{t-1}f_i^\tau\right) < 0$, then $\Delta(t_h+1)\leq\Delta(t_h)$ for all $t_h > t^*$. Hence, $\Delta(m+1)\leq\Delta(t_h)$ for all $t_h > t^*$. Recall that $\Delta(t+1)\leq\Delta(t_h)$ for all $t_h\leq t^*$. Hence, $t_h^*\in\arg\min\{\Delta(t+1),\Delta(m+1)\}$. If $\overline{H}\left(\mathcal{S}^{t,\nu}+\sum_{i\in\pi^{t,\nu}}(t,i)\right) < 1-\sum_{i=1}^n\varphi_i$, then $t^*=t$ by definition, which implies that $t_h^*=m+1$. ■

Lemma 30 *Suppose $\mathcal{S}^{t,\nu}=(S_1^{t_1^*},\dots,S_n^{t_n^*})$ and $h\notin\mathcal{I}_1^{t+1}\cup\pi^{t,1}\cup\cdots\cup\pi^{t,\nu}$, then $t_h^*=t+1$ implies that $h\in\mathcal{I}_1^t$.*

Proof. Suppose $\overline{H}\left(\mathcal{S}^{t,\nu}+\sum_{i\in\pi^{t,\nu}}(t,i)\right) = 1-\sum_{i=1}^n\varphi_i$, then by Lemma 28, $h\in\mathcal{I}_1^t$. Suppose $\overline{H}\left(\mathcal{S}^{t,\nu}+\sum_{i\in\pi^{t,\nu}}(t,i)\right) < 1-\sum_{i=1}^n\varphi_i$. By Lemma 29, $\frac{\varphi_h}{1-c_h}\geq\prod_{i\neq h}\sum_{\tau=1}^{t_i^*-1}f_i^\tau$. Hence,

$$\begin{aligned}&\overline{H}\left(\mathcal{S}^{t,\nu}+\sum_{i\in\pi^{t,\nu}}(t,i)+(t,h)\right)-\overline{H}\left(\mathcal{S}^{t,\nu}+\sum_{i\in\pi^{t,\nu}}(t,i)\right)\\ \leq& f_h^t\left(\prod_{i\neq h}\sum_{\tau=1}^{t_i^*-1}f_i^\tau-\varphi_h\right)\\ \leq& f_h^t\left(\frac{\varphi_h}{1-c_h}-\varphi_h\right)=\frac{c_h\varphi_h f_h^t}{1-c_h}.\end{aligned}$$

By Algorithm 1, $h\in\mathcal{I}_1^t$. ■

A.3.2. Continuum case

I characterize an optimal solution in the continuum case by taking m to infinity. Let $\mathcal{I}_1^{m,t}$ denote $\mathcal{I}_1^t$ and $\underline{t}^m$ be defined by (A.9) when $\mathcal{D}$ is discretized by m grid points. Clearly, if $i \in \mathcal{I}_1^{m,t}$ then $i \in \mathcal{I}_1^{2m,2t-1}$. Let $\bar{t}_i^m := \max\left\{t | i \in \mathcal{I}_1^{m,t}\right\}$ and $\bar{d}_i^m := \underline{d} + \frac{(\bar{t}_i^m - 1)(\bar{d} - \underline{d})}{m}$. Then the sequence of $\left\{\bar{d}_i^{2^\kappa}\right\}_\kappa$ is non-decreasing and bounded from above by $\bar{d}$. Hence, the sequence converges and let $d_i^u := \lim_{\kappa\to\infty} \bar{d}_i^{2^\kappa}$ denote its limit. For each κ, let $\underline{d}^{2^\kappa} := \underline{d} + \frac{(\underline{t}^{2^\kappa} - 1)(\bar{d} - \underline{d})}{2^\kappa}$, which is bounded. After taking subsequences, we can assume $\left\{\underline{d}^{2^\kappa}\right\}_\kappa$ converges and let $d^l := \lim_{\kappa\to\infty} \underline{d}^{2^\kappa}$ denote its limit. Let

$$\bar{P}_i(v_i) := \begin{cases} \frac{\varphi_i}{1-c_i} & \text{if } v_i > d_i^u + \frac{k_i}{c_i} \\ \prod_{j\neq i, d_j^u \geq v_i - \frac{k_i}{c_i}} F_j\left(v_i - \frac{k_i}{c_i} + \frac{k_j}{c_j}\right) & \text{if } d^l + \frac{k_i}{c_i} < v_i < d_i^u + \frac{k_i}{c_i} \\ \varphi_i & \text{if } v_i < d^l + \frac{k_i}{c_i} \end{cases}.$$

Finally, let $\boldsymbol{P}^* := (P_i^*)_i$ where

$$P_i^*(v_i) := \begin{cases} \bar{P}_i(v_i) & \text{if } v_i > \frac{k_i}{c_i} \\ \varphi_i & \text{if } v_i < \frac{k_i}{c_i} \end{cases}. \tag{2.19}$$

We are now ready to prove Theorem 4.

Proof of Theorem 4. We can extend $\bar{P}_i^m$ (P_i^m) to $[\underline{v}_i, \bar{v}_i]$ by setting, for each $t = 1, \ldots, m$,

$$\bar{P}_i^m(v_i) := \bar{P}_i^{m,t} (P_i^m(v_i) := P_i^{m,t}) \text{ for } v_i \in \left[\underline{d} + \frac{(t-1)(\bar{d} - \underline{d})}{m} + \frac{k_i}{c_i}, \underline{d} + \frac{t(\bar{d} - \underline{d})}{m} + \frac{k_i}{c_i}\right].$$

I show that, after taking subsequences, $\bar{P}_i^m$ converges to $\bar{P}_i$ pointwise.

First, by construction and Lemma 28, $\bar{P}_i^{2^\kappa}(v_i) = \varphi_i$ for all $v_i < \underline{d}^{2^\kappa} + \frac{k_i}{c_i}$, we have $\lim_{\kappa\to\infty} \bar{P}_i^{2^\kappa}(v_i) = \bar{P}_i(v_i)$ for all $v_i < d^l + \frac{k_i}{c_i}$. Similarly, by construction, $\bar{P}_i^{2^\kappa}(v_i) = \frac{\varphi_i}{1-c_i}$ for all $v_i > \bar{d}_i^{2^\kappa} + \frac{k_i}{c_i}$, we have $\lim_{\kappa\to\infty} \bar{P}_i^{2^\kappa}(v_i) = \bar{P}_i(v_i)$ for all $v_i > d_i^u + \frac{k_i}{c_i}$.

Suppose $d^l < v_i - \frac{k_i}{c_i} < d_i^u$. Assume without loss of generality that $d_1^u \geq \cdots \geq d_n^u \geq d^l$. If $d_i^u = d^l$, then we are done. Assume for the rest of the proof that $d_i^u > d^l$. Let $d_{n+1}^u := d^l$. Consider v_i such that $d_i^u \geq d_j^u > v_i - \frac{k_i}{c_i} > d_{j+1}^u$ for some $j \geq i$. For m sufficiently large, there exists t such that

$$d_{j+1}^u < \underline{d} + \frac{(t-1)(\overline{d} - \underline{d})}{m} < \underline{d} + \frac{t(\overline{d} - \underline{d})}{m} < v_i - \frac{k_i}{c_i} < \underline{d} + \frac{(t+1)(\overline{d} - \underline{d})}{m} < d_j^u \leq d_i^u.$$

Hence, by construction, we have $\mathcal{I}_1^{m,t} = \mathcal{I}_1^{m,t+1} = \{1, \dots, j\}$. By Lemmas 29 and 30, there exists $\mathcal{S} = (S_1^{t_1}, \dots, S_n^{t_n})$ such that $t_i = t+1$, $t_h \in \{t, t+1\}$ if $h \leq j$ and $h \neq i$, $t_h = m+1$ if $h > j$, and

$$f_i^t\left(\overline{P}_i^{m,t} - \varphi_i\right) = \overline{z}_i^{m,t} = \overline{H}(\mathcal{S} + (t,i)) - \overline{H}(\mathcal{S}).$$

Because $\overline{H}$ is submodular, we have

$$\begin{aligned} f_i^t\left(\overline{P}_i^{m,t} - \varphi_i\right) &\leq \overline{H}(\mathcal{S}' + (t,i)) - \overline{H}(\mathcal{S}') \\ &= f_i^t\left(\prod_{h \leq j, h \neq i} \sum_{\tau=1}^{t} f_h^\tau - \varphi_i\right), \end{aligned}$$

where $\mathcal{S}' = (S_1^{t+1}, \dots, S_j^{t+1}, S_{j+1}^{m+1}, \dots, S_n^{m+1})$; and

$$\begin{aligned} f_i^t\left(\overline{P}_i^{m,t} - \varphi_i\right) &\geq \overline{H}(\mathcal{S}'' - (t,i)) - \overline{H}(\mathcal{S}') \\ &= f_i^t\left(\prod_{h \leq j, h \neq i} \sum_{\tau=1}^{t-1} f_h^\tau - \varphi_i\right), \end{aligned}$$

where $\mathcal{S}'' = (S_1^t, \dots, S_j^t, S_{j+1}^{m+1}, \dots, S_n^{m+1})$. Hence,

$$\prod_{h \leq j, h \neq i} \sum_{\tau=1}^{t-1} f_h^\tau \leq \overline{P}_i^{m,t} \leq \prod_{h \leq j, h \neq i} \sum_{\tau=1}^{t} f_h^\tau.$$

Take $m = 2^\kappa$ to infinity and we have $\lim_{\kappa \to \infty} \overline{P}_i^{2^\kappa}(v_i) = \overline{P}_i(v_i)$.

It follows that, after taking subsequences, $\boldsymbol{P}_i^m$ converges to $\boldsymbol{P}_i^*$ pointwise. $\boldsymbol{P}^*$ is feasible by a similar argument to that in the proof of Lemma 17, and optimal by a similar argument to that in the proof of Theorem 1. ■

A.3.3. Optimal one-threshold mechanism

By a similar argument to that in the proof of Theorem 2, we can show that pure randomization is optimal if verification is sufficiently costly or the principal's ability to punish an agent is sufficiently limited, i.e., $\overline{v}_i - k_i/c_i \leq \mathbb{E}_{v_i}[v_i]$ for all i. To make the problem interesting, in what follows, I assume:

Assumption 5 $\overline{v}_i - k_i/c_i > \mathbb{E}_{v_i}[v_i]$ *for some* i.

If $d_i^u = \overline{v}_i - \frac{k_i}{c_i}$ for all i, then $d^l \geq \max_j \left\{ \underline{v}_j - k_j/c_j \right\}$ satisfies that

$$\sum_{i=1}^{n} \varphi_i F_i \left(d^l + \frac{k_i}{c_i} \right) = \prod_{i=1}^{n} F_i \left(d^l + \frac{k_i}{c_i} \right).$$

Lemma 31 below shows that there exists a unique d^l satisfying the above equation. Note that unless $\varphi_i = 0$ for all i, we have $d^l > \max_j \left\{ \underline{v}_j - k_j/c_j \right\}$. Clearly, in optimum, $\varphi_i > 0$ for some i. Hence, $d^l > \max_j \left\{ \underline{v}_j - k_j/c_j \right\}$. Let d_i^* $(i = 1, \ldots, n)$ be defined by

$$\mathbb{E}_{v_i}[v_i] - \mathbb{E}_{v_i} \left[\max \left\{ v_i, d_i^* + \frac{k_i}{c_i} \right\} \right] + \frac{k_i}{c_i} = 0, \tag{A.10}$$

and $d^{l*} := \max_i d_i^*$. Now we are ready to state the main result in this subsection which characterizes the set of optimal $\boldsymbol{\varphi}$:

Theorem 17 *Suppose that Assumption 5 holds. If*

$$\sum_{i=1}^{n} (1 - c_i) F_i \left(d^{l*} + \frac{k_i}{c_i} \right) \prod_{j \neq i} F_j \left(\overline{v}_j - \frac{k_i}{c_i} + \frac{k_j}{c_j} \right) \leq \prod_{i=1}^{n} F_i \left(d^{l*} + \frac{k_i}{c_i} \right),$$

then the set of optimal $\boldsymbol{\varphi}$ *is the convex hull of*

$$\left\{ \boldsymbol{\varphi} \;\middle|\; \begin{array}{c} i^* \in \arg\max_i d_i^*, \varphi_i = (1 - c_i) \prod_{j \neq i} F_j \left(\overline{v}_j - \frac{k_i}{c_i} + \frac{k_j}{c_j} \right) \ \forall i \neq i^*, \\ \varphi_{i^*} = \frac{\prod_{i=1}^{n} F_i \left(d^{l*} + \frac{k_i}{c_i} \right) - \sum_{i \neq i^*} (1 - c_i) \prod_{j \neq i} F_j \left(\overline{v}_j - \frac{k_i}{c_i} + \frac{k_j}{c_j} \right) F_i \left(d^* + \frac{k_i}{c_i} \right)}{F_{i^*} \left(d^{l*} + \frac{k_{i^*}}{c_{i^*}} \right)} \end{array} \right\}.$$

For each optimal $\boldsymbol{\varphi}^*$, the following allocation rule is optimal:

$$P_i^{**}(v_i) := \begin{cases} \prod_{j\neq i} F_j\left(v_i - \frac{k_i}{c_i} + \frac{k_j}{c_j}\right) & \text{if } v_i \geq d^l + \frac{k_i}{c_i} \\ \varphi_i^* & \text{if } v_i < d^l + \frac{k_i}{c_i} \end{cases}.$$

Proof. Let $\Phi(d^l, d_1^u, \ldots, d_n^u) \subset \left\{\boldsymbol{\varphi} \mid \sum \varphi_i \leq 1\right\}$ denote the feasible set of $\boldsymbol{\varphi}$ given d^l and $d_1^u, \ldots, d_n^u$. I often abuse notation and use Φ to denote the feasible set when its meaning is clear. Fix $d^l > \max_j \left\{\underline{v}_j - k_j/c_j\right\}$ and $d_i^u = \overline{v}_i - \frac{k_i}{c_i}$ for all i. Then $\boldsymbol{\varphi}$ is feasible if and only if

$$\sum_{i=1}^n \varphi_i F_i\left(d^l + \frac{k_i}{c_i}\right) = \prod_{i=1}^n F_i\left(d^l + \frac{k_i}{c_i}\right),$$
$$\prod_{j\neq i} F_j\left(\overline{v}_i - \frac{k_i}{c_i} + \frac{k_j}{c_j}\right) \leq \frac{\varphi_i}{1 - c_i}, \forall i.$$

Hence, Φ is non-empty if and only if

$$\sum_{i=1}^n (1 - c_i) F_i\left(d^l + \frac{k_i}{c_i}\right) \prod_{j\neq i} F_j\left(\overline{v}_i - \frac{k_i}{c_i} + \frac{k_j}{c_j}\right) \leq \prod_{i=1}^n F_i\left(d^l + \frac{k_i}{c_i}\right).$$

Suppose that Φ is non-empty. It is not hard to see that Φ is convex. Because the objective function is linear in $\boldsymbol{\varphi}$ and the feasible set is convex, there is an optimal $\boldsymbol{\varphi}$ which is an extreme point.

Clearly, $\boldsymbol{\varphi}$ is an extreme point of Φ if and only if there exists i^* such that

$$\sum_{i=1}^n \varphi_i F_i\left(d^l + \frac{k_i}{c_i}\right) = \prod_{i=1}^n F_i\left(d^l + \frac{k_i}{c_i}\right),$$
$$\varphi_j = (1 - c_j) \prod_{i\neq j} F_i\left(\overline{v}_j - \frac{k_j}{c_j} + \frac{k_i}{c_i}\right), \forall j \neq i^*.$$

In this case, denote the principal's payoff by $Z_{1,i^*}(d^l)$. For ease of notation, let $i^* = 1$. Let $\bar{\varphi}_j :=$

$(1-c_j)\prod_{i\neq j} F_i\left(\overline{v}_j - \frac{k_j}{c_j} + \frac{k_i}{c_i}\right)$ for all j. Then the principal's payoff is given as follows:

$$\begin{aligned}
Z_{1,1}(d^l) :=& \sum_{i=1}^{n} \int_{\max\left\{d^l+\frac{k_i}{c_i},\frac{k_i}{c_i}\right\}}^{\overline{v}_i} \left(v_i - \frac{k_i}{c_i}\right) \prod_{j\neq i} F_j\left(v_i - \frac{k_i}{c_i} + \frac{k_j}{c_j}\right) \mathrm{d}F_i(v_i) \\
&+ \sum_{i\neq 1} \int_{\underline{v}_i}^{\max\left\{d^l+\frac{k_i}{c_i},\frac{k_i}{c_i}\right\}} \left(v_i - \frac{k_i}{c_i}\right) \overline{\varphi}_i \mathrm{d}F_i(v_i) \\
&+ \int_{\underline{v}_1}^{\max\left\{d^l+\frac{k_1}{c_1},\frac{k_1}{c_1}\right\}} \left(v_1 - \frac{k_1}{c_1}\right) \frac{\prod_{i=1}^n F_i\left(d^l + \frac{k_i}{c_i}\right) - \sum_{i\neq 1} F_i\left(d^l + \frac{k_i}{c_i}\right)\overline{\varphi}_i}{F_1\left(d^l + \frac{k_1}{c_1}\right)} \mathrm{d}F_1(v_1) \\
&+ \sum_{i\neq 1} \frac{\overline{\varphi}_i k_i}{c_i} + \frac{k_1}{c_1} \frac{\prod_{i=1}^n F_i\left(d^l + \frac{k_i}{c_i}\right) - \sum_{i\neq 1} F_i\left(d^l + \frac{k_i}{c_i}\right)\overline{\varphi}_i}{F_1\left(d^l + \frac{k_1}{c_1}\right)}.
\end{aligned}$$

If $d^l < 0$, then it is not hard to show that $Z_{1,1}$ is strictly increasing in d^l. If $d^l \geq 0$, then, after some algebra, we have

$$\begin{aligned}
&Z'_{1,1}(d^l) \\
=&\left\{\sum_{i\neq 1}\left[f_i\left(d^l + \frac{k_i}{c_i}\right)\prod_{j\neq i,1} F_j\left(d^l + \frac{k_j}{c_j}\right) - \overline{\varphi}_i \frac{f_i\left(d^l + \frac{k_i}{c_i}\right)F_1\left(d^l + \frac{k_1}{c_1}\right) - F_i\left(d^l + \frac{k_i}{c_i}\right) f_1\left(d^l + \frac{k_1}{c_1}\right)}{F_1^2\left(d^l + \frac{k_1}{c_1}\right)}\right]\right\} \\
&\cdot\left[\int_{\underline{v}_1}^{d^l+\frac{k_1}{c_1}} \left(v_1 - d^l - \frac{k_1}{c_1}\right)\mathrm{d}F_1(v_1) + \frac{k_1}{c_1}\right].
\end{aligned}$$

Because $\overline{\varphi}_i \leq \prod_{j\neq i} F_j\left(d^l + \frac{k_j}{c_j}\right)$, the first-term in the above equation is strictly positive. The second-term is strictly decreasing in d^l. Let d_1^* be such that

$$\int_{\underline{v}_1}^{d_1^*+\frac{k_1}{c_1}} \left(v_1 - d_1^* - \frac{k_1}{c_1}\right)\mathrm{d}F_1(v_1) + \frac{k_1}{c_1} = 0. \tag{A.11}$$

Then $Z'_{1,1}(d^l) > 0$ if $d^l < d_1^*$ and $Z'_{1,1}(d^l) < 0$ if $d^l > d_1^*$. Hence, $Z_{1,1}(d^l)$ achieves its maximum at $d^l = d_1^*$.

Define d_i^* for all $i \geq 2$ as in (A.11). Suppose that $d_1^* \geq d_2^*$. By a similar argument to that in Lemma 31, $\Phi\left(d_2^*, \overline{v}_1 - \frac{k_1}{c_1}, \ldots, \overline{v}_n - \frac{k_n}{c_n}\right) \neq \emptyset$ implies that $\Phi\left(d_1^*, \overline{v}_1 - \frac{k_1}{c_1}, \ldots, \overline{v}_n - \frac{k_n}{c_n}\right) \neq \emptyset$. Suppose that both $\Phi\left(d_2^*, \overline{v}_1 - \frac{k_1}{c_1}, \ldots, \overline{v}_n - \frac{k_n}{c_n}\right)$ and $\Phi\left(d_1^*, \overline{v}_1 - \frac{k_1}{c_1}, \ldots, \overline{v}_n - \frac{k_n}{c_n}\right)$ are non-empty. Then

$$\begin{aligned}
& Z_{1,1}(d^l) - Z_{1,2}(d^l) \\
&= \left[\prod_{i=1}^n F_i\left(d^l + \frac{k_i}{c_i}\right) - \sum_{i=1} F_i\left(d^l + \frac{k_i}{c_i}\right)\overline{\varphi}_i\right] \\
&\quad \cdot \left\{\frac{1}{F_1\left(d^l + \frac{k_1}{c_1}\right)}\left[\int_{\underline{v}_1}^{d^l + \frac{k_1}{c_1}}\left(v_1 - \frac{k_1}{c_1}\right)\mathrm{d}F_1(v_1) + \frac{k_1}{c_1}\right]\right. \\
&\qquad \left. - \frac{1}{F_2\left(d^l + \frac{k_2}{c_2}\right)}\left[\int_{\underline{v}_2}^{d^l + \frac{k_2}{c_2}}\left(v_1 - \frac{k_2}{c_2}\right)\mathrm{d}F_2(v_2) + \frac{k_2}{c_2}\right]\right\}
\end{aligned}$$

If $d^l = d_2^*$, then by definition we have

$$\begin{aligned}
& Z_{1,1}(d_2^*) - Z_{1,2}(d_2^*) \\
&= \left[\prod_{i=1}^n F_i\left(d_2^* + \frac{k_i}{c_i}\right) - \sum_{i=1} F_i\left(d_2^* + \frac{k_i}{c_i}\right)\overline{\varphi}_i\right] \\
&\quad \cdot \left\{\frac{1}{F_1\left(d_2^* + \frac{k_1}{c_1}\right)}\left[\int_{\underline{v}_1}^{d_2^* + \frac{k_1}{c_1}}\left(v_1 - \frac{k_1}{c_1}\right)\mathrm{d}F_1(v_1) + \frac{k_1}{c_1}\right] - d_2^*\right\} \\
&\geq \left[\prod_{i=1}^n F_i\left(d_2^* + \frac{k_i}{c_i}\right) - \sum_{i=1} F_i\left(d_2^* + \frac{k_i}{c_i}\right)\overline{\varphi}_i\right]\left(d_2^* - d_2^*\right) = 0,
\end{aligned}$$

where the last inequality holds because $d_1^* \geq d_2^*$, and the inequality holds strictly if $d_1^* > d_2^*$. Hence, $Z_{1,1}(d_1^*) \geq Z_{1,2}(d_2^*)$ and the inequality holds strictly if $d_1^* > d_2^*$.

Let $d^{l*} := \max_i d_i^*$. If

$$\sum_{i=1}^n (1 - c_i)\prod_{j \neq i} F_j\left(\overline{v}_j - \frac{k_i}{c_i} + \frac{k_j}{c_j}\right) F_i\left(d^* + \frac{k_i}{c_i}\right) \leq \prod_{i=1}^n F_i\left(d^{l*} + \frac{k_i}{c_i}\right),$$

then $\Phi\left(d^{l*}, \overline{v}_1 - \frac{k_1}{c_1}, \ldots, \overline{v}_n - \frac{k_n}{c_n}\right)$ is feasible. This completes the proof. ■

Lemma 31 *There exists a unique* $d^l \geq \max_j \left\{ \underline{v}_j - k_j/c_j \right\}$ *such that*

$$\sum_{i=1}^{n} \varphi_i F_i \left(d_l + \frac{k_i}{c_i} \right) = \prod_{i=1}^{n} F_i \left(d^l + \frac{k_i}{c_i} \right). \tag{A.12}$$

Proof. If $\varphi_i = 0$ for all i, then $d^l = \max_j \left\{ \underline{v}_j - k_j/c_j \right\}$ is the unique solution to (A.12). Assume, for the rest of the proof, that $\varphi_i > 0$ for some i. Let

$$\Delta(d^l) := \sum_{i=1}^{n} \varphi_i F_i \left(d^l + \frac{k_i}{c_i} \right) - \prod_{i=1}^{n} F_i \left(d^l + \frac{k_i}{c_i} \right).$$

Then

$$\Delta'(d^l) = \sum_{i=1}^{n} f_i \left(d^l + \frac{k_i}{c_i} \right) \left[\varphi_i - \prod_{j \neq i} F_j \left(d^l + \frac{k_j}{c_j} \right) \right].$$

Because $\Delta_l \left(\max_j \left\{ \underline{v}_j - k_j/c_j \right\} \right) > 0$, a solution to (A.12) must satisfy that $d^l > \max_j \left\{ \underline{v}_j - k_j/c_j \right\}$. Assume, for the rest of the proof, that $d^l > \max_j \left\{ \underline{v}_j - k_j/c_j \right\}$. Then $F_i \left(d^l + \frac{k_i}{c_i} \right) > 0$ for all i. If $\Delta(d^l) \leq 0$, then $\varphi_i \leq \prod_{j \neq i} F_j \left(d^l + \frac{k_j}{c_j} \right)$ for all i, and the strict inequality holds for some i, which implies that $\Delta'(d^l) < 0$. Hence, $\Delta(d^l)$ crosses zero at most once, in which case it does so from above. Because $\Delta_l \left(\max_j \left\{ \underline{v}_j - k_j/c_j \right\} \right) > 0$ and $\Delta_l \left(\max_j \left\{ \overline{v}_j - k_j/c_j \right\} \right) = \sum_i \varphi_i - 1 \leq 0$, there exists a unique d^l satisfying (A.12). ■

A.3.4. Symmetric environment revisited

Fix $\boldsymbol{\varphi}$ and let d_i^u and d^l be the associated optimal thresholds. Assume, without loss of generality, that $d_1^u \geq \cdots \geq d_n^u \geq d^l$. Let $1 \leq \xi_1 < \cdots < \xi_L \leq n$ be such that $d_1^u = \cdots = d_{\xi_1}^u$, $d_{\xi_\iota}^u > d_{\xi_\iota+1}^u = \cdots = d_{\xi_{\iota+1}}^u$ for $\iota = 1, \ldots, L-1$ and $d_{\xi_L}^u > d_{\xi_L+1}^u = \cdots = d_n^u = d^l$. Note that in the symmetric environment $d_i^u \geq d_j^u$ only if $\varphi_i \geq \varphi_j$. The proof of Theorem 5 uses the following properties of d^l and d_i^u:

Lemma 32 *If* $\frac{\varphi_i}{1-c} \geq 1$ *for all* $i \leq \xi_1$, *then* $d_{\xi_1}^u = \overline{v} - \frac{k}{c}$; *otherwise* $\frac{\varphi_i}{1-c} < 1$ *for all* $i \leq \xi_1$ *and* $d_{\xi_1}^u$

satisfies

$$1 - F\left(d^u_{\xi_1} + \frac{k}{c}\right)^{\xi_1} = \sum_{i=1}^{\xi_1} \frac{\varphi_i}{1-c_i}\left[1 - F\left(d^u_{\xi_1} + \frac{k}{c}\right)\right],$$

$$1 - F(v)^{\xi_1} \lesseqgtr \sum_{i=1}^{\xi_1} \frac{\varphi_i}{1-c}\left[1 - F(v)\right] \textit{ if } v \lesseqgtr d^u_{\xi_1} + \frac{k}{c},$$

$$\frac{\varphi_i}{1-c} \geq F\left(d^u_{\xi_1} + \frac{k}{c}\right)^{\xi_1 - 1}, \forall i = 1, \dots, \xi_1.$$

For $\iota = 1, \dots, L-1$, $\frac{\varphi_i}{1-c} < 1$ *for* $\xi_\iota + 1 \leq i \leq \xi_{\iota+1}$ *and* $d^u_{\xi_{\iota+1}}$ *satisfies*

$$F\left(d_{\xi_{\iota+1}} + \frac{k}{c}\right)^{\xi_\iota} - F\left(d_{\xi_{\iota+1}} + \frac{k}{c}\right)^{\xi_{\iota+1}} = \sum_{i=\xi_\iota+1}^{\xi_{\iota+1}} \frac{\varphi_i}{1-c_i}\left[1 - F\left(d_{\xi_{\iota+1}} + \frac{k}{c}\right)\right],$$

$$F(v)^{\xi_\iota} - F(v)^{\xi_{\iota+1}} \lesseqgtr \sum_{i=\xi_\iota+1}^{\xi_{\iota+1}} \frac{\varphi_i}{1-c}\left[1 - F(v)\right] \textit{ if } v \lesseqgtr d_{\xi_{\iota+1}} + \frac{k}{c},$$

$$\frac{\varphi_i}{1-c} \geq F\left(d^u_{\xi_{\iota+1}} + \frac{k}{c}\right)^{\xi_{\iota+1} - 1}, \forall i = \xi_\iota + 1, \dots, \xi_{\iota+1}.$$

Finally, d^l *satisfies*

$$F\left(d^l + \frac{k}{c}\right)^{\xi_L} = \sum_{i=1}^{n} \varphi_i F\left(d^l + \frac{k}{c}\right) + \sum_{i=\xi_L+1}^{n} \frac{\varphi_i}{1-c}\left[1 - F\left(d^l + \frac{k}{c}\right)\right],$$

$$F(v)^{\xi_L} \lesseqgtr \sum_{i=1}^{n} \varphi_i F(v) + \sum_{i=\xi_L+1}^{n} \frac{\varphi_i}{1-c}\left[1 - F(v)\right] \textit{ if } v \lesseqgtr d^l + \frac{k}{c}.$$

The arguments used to prove Lemma 32 are similar to that used to show that $\overline{P}_i^m$ converges to $\overline{P}_i$ if $d^l < v_i - \frac{k_i}{c_i} < d^u_i$, and are neglected here.

Proof of Theorem 5. The first part of the theorem directly follows from Theorem 17. Assume, for the rest of the proof, that $F(v^*)^{n-1} < n(1-c)$. Consider an optimal $\boldsymbol{\varphi}$, and let d^u_i and d^l be the associated optimal thresholds. Assume, without loss of generality, that $d^u_1 \geq \cdots \geq d^u_n \geq d^l$. Let ξ_ι ($\iota = 1, \dots, L$) be defined as in the beginning of this subsection.

First, I show that $L = 1$. Suppose, to the contrary, that $L \geq 2$. Suppose that $d^u_{\xi_2} < 0$, then the principal's objective function is strictly increasing in φ_i for $i > \xi_2$. Hence, in optimum, it must be that $d^u_{\xi_2} \geq 0$. Construct a new $\boldsymbol{\varphi}^*$ as follows: Let

$$\varphi^*_i = \frac{1}{\xi_2} \sum_{j=1}^{\xi_2} \varphi_j, \text{ for all } i = 1, \ldots, \xi_2,$$

and $\varphi^*_i = \varphi_i$ for all $i > \xi_2$. Let d^{u*}_i and d^{l*} be the optimal thresholds associated with $\boldsymbol{\varphi}^*$. Then $d^{u*}_1 = \ldots d^{u*}_{\xi_2}$ and $d^{u*}_i = d^u_i$ for all $i > \xi_2$. There are two cases: (1) $\varphi_i < 1 - c$ for all $i \leq \xi_1$ and (2) $\varphi_i \geq 1 - c$ for all $i \leq \xi_1$.

Case 1: $\varphi_i < 1 - c$ for all $i \leq \xi_1$. In this case, $\varphi^*_1 < 1 - c$. Then $d^{u*}_{\xi_2}$ is defined by

$$\left[1 - F\left(d^{u*}_{\xi_2} + \frac{k}{c}\right)\right] \sum_{i=1}^{\xi_2} \frac{\varphi^*_i}{1 - c} = 1 - F\left(d^{u^*}_{\xi_2} + \frac{k}{c}\right)^{\xi_2}. \tag{A.13}$$

Hence, $d^{u}_{\xi_2} < d^{u^*}_{\xi_2} < d^{u}_{\xi_1}$. Let $Z(\boldsymbol{\varphi})$ denote the principal's payoff given $\boldsymbol{\varphi}$. Then

$$
\begin{aligned}
& Z(\boldsymbol{\varphi}^*) - Z(\boldsymbol{\varphi}) \\
=& \sum_{i=1}^{\xi_2} \left[\int_{d^{u*}_{\xi_2}+\frac{k}{c}}^{\overline{v}} \left(v_i - \frac{k}{c}\right) \frac{\varphi_i^*}{1-c} \mathrm{d}F(v_i) + \int_{d^{u}_{\xi_2}+\frac{k}{c}}^{d^{u^*}_{\xi_2}+\frac{k}{c}} \left(v_i - \frac{k}{c}\right) F(v_i)^{\xi_2 - 1} \mathrm{d}F(v_i) \right] \\
& - \sum_{i=\xi_1+1}^{\xi_2} \int_{d^{u}_{\xi_2}+\frac{k}{c}}^{\overline{v}} \left(v_i - \frac{k}{c}\right) \frac{\varphi_i}{1-c} \mathrm{d}F(v_i) \\
& - \sum_{i=1}^{\xi_1} \left[\int_{d^{u}_{\xi_2}+\frac{k}{c}}^{d^{u}_{\xi_1}+\frac{k}{c}} \left(v_i - \frac{k}{c}\right) F(v_i)^{\xi_1 - 1} \mathrm{d}F(v_i) + \int_{d^{u}_{\xi_1}+\frac{k}{c}}^{\overline{v}} \left(v_i - \frac{k}{c}\right) \frac{\varphi_i}{1-c} \mathrm{d}F(v_i) \right] \\
=& \int_{d^{u}_{\xi_2}+\frac{k}{c}}^{d^{u*}_{\xi_2}+\frac{k}{c}} \left(v - \frac{k}{c}\right) \xi_2 F(v)^{\xi_2 - 1} \mathrm{d}F(v) + \int_{d^{u*}_{\xi_2}+\frac{k}{c}}^{\overline{v}} \left(v - \frac{k}{c}\right) \sum_{i=1}^{\xi_2} \frac{\varphi_i^*}{1-c} \mathrm{d}F(v) \\
& - \int_{d^{u}_{\xi_2}+\frac{k}{c}}^{d^{u}_{\xi_1}+\frac{k}{c}} \left(v - \frac{k}{c}\right) \left(\sum_{i=\xi_1+1}^{\xi_2} \frac{\varphi_i}{1-c} + \xi_1 F(v)^{\xi_1 - 1} \right) \mathrm{d}F(v) - \int_{d^{u}_{\xi_1}+\frac{k}{c}}^{\overline{u}} \left(v - \frac{k}{c}\right) \sum_{i=1}^{\xi_2} \frac{\varphi_i}{1-c} \mathrm{d}F(v) \\
=& \int_{d^{u}_{\xi_2}+\frac{k}{c}}^{d^{u*}_{\xi_2}+\frac{k}{c}} \left(v - \frac{k}{c}\right) \xi_2 F(v)^{\xi_2 - 1} \mathrm{d}F(v) + \int_{d^{u*}_{\xi_2}+\frac{k}{c}}^{d^{u}_{\xi_1}+\frac{k}{c}} \left(v - \frac{k}{c}\right) \sum_{i=1}^{\xi_2} \frac{\varphi_i}{1-c} \mathrm{d}F(v) \\
& - \int_{d^{u}_{\xi_2}+\frac{k}{c}}^{d^{u}_{\xi_1}+\frac{k}{c}} \left(v - \frac{k}{c}\right) \left(\sum_{i=\xi_1+1}^{\xi_2} \frac{\varphi_i}{1-c} + \xi_1 F(v)^{\xi_1 - 1} \right) \mathrm{d}F(v) \\
=& d^{u}_{\xi_1} \left[F\left(d^{u}_{\xi_1} + \frac{k}{c}\right) \sum_{i=1}^{\xi_1} \frac{\varphi_i}{1-c} - F\left(d^{u}_{\xi_1} + \frac{k}{c}\right)^{\xi_1} \right] + d^{u*}_{\xi_2} \left[F\left(d^{u*}_{\xi_2} + \frac{k}{c}\right)^{\xi_2} - F\left(d^{u*}_{\xi_2} + \frac{k}{c}\right) \sum_{i=\xi_1+1}^{\xi_2} \frac{\varphi_i}{1-c} \right] \\
& - d^{u}_{\xi_2} \left[F\left(d^{u}_{\xi_2} + \frac{k}{c}\right)^{\xi_2} - F\left(d^{u}_{\xi_2} + \frac{k}{c}\right) \sum_{i=\xi_1+1}^{\xi_2} \frac{\varphi_i}{1-c} - F\left(d^{u}_{\xi_2} + \frac{k}{c}\right)^{\xi_1} \right] \\
& - \int_{d^{u}_{\xi_2}+\frac{k}{c}}^{d^{u*}_{\xi_2}+\frac{k}{c}} \left[F(v)^{\xi_2} - F(v) \sum_{i=\xi_1}^{\xi_2} \frac{\varphi_i}{1-c} - F(v)^{\xi_1} \right] \mathrm{d}v - \int_{d^{u*}_{\xi_2}+\frac{k}{c}}^{d^{u}_{\xi_1}+\frac{k}{c}} \left[F(v) \sum_{i=1}^{\xi_1+1} \frac{\varphi_i}{1-c} - F(v)^{\xi_1} \right] \mathrm{d}v,
\end{aligned}
$$

where the third equality holds because $\sum_{i=1}^{\xi_2} \varphi_i = \sum_{i=1}^{\xi_2} \varphi_i^*$, and the last equality holds by integration by parts. Because $d^{u*}_{\xi_2}$ satisfies (A.13), $d^{u}_{\xi_1}$ satisfies that

$$
\begin{aligned}
1 - F\left(d^{u}_{\xi_1} + \frac{k}{c}\right)^{\xi_1} &= \left[1 - F\left(d^{u}_{\xi_1} + \frac{k}{c}\right)\right] \sum_{i=1}^{\xi_1} \frac{\varphi_i}{1-c} \\
1 - F(v)^{\xi_1} &< [1 - F(v)] \sum_{i=1}^{\xi_1} \frac{\varphi_i}{1-c}, \forall v < d^{u}_{\xi_1} + \frac{k}{c},
\end{aligned}
$$

and $d^u_{\xi_2}$ satisfies that

$$1 - F\left(d^u_{\xi_2} + \frac{k}{c}\right)^{\xi_2} = \left[1 - F\left(d^u_{\xi_2} + \frac{k}{c}\right)\right] \sum_{i=\xi_1+1}^{\xi_2} \frac{\varphi_i}{1-c} + 1 - F\left(d^u_{\xi_2} + \frac{k}{c}\right)^{\xi_1}$$

$$1 - F(v)^{\xi_2} > [1 - F(v)] \sum_{i=\xi_1+1}^{\xi_2} \frac{\varphi_i}{1-c} + 1 - F(v)^{\xi_1}, \forall v > d^u_{\xi_2} + \frac{k}{c},$$

we have

$$\begin{aligned}
&Z(\boldsymbol{\varphi}^*) - Z(\boldsymbol{\varphi}) \\
&> \left(d_{\xi_1} - d^{u^*}_{\xi_2}\right)\left(\sum_{i=1}^{\xi_1} \frac{\varphi_i}{1-c} - 1\right) - \left(d^{u*}_{\xi_2} - d_{\xi_2}\right) \sum_{\xi_1+1}^{\xi_2} \frac{\varphi_i}{1-c} \\
&+ \int_{d^u_{\xi_2}+\frac{k}{c}}^{d^{u*}_{\xi_2}+\frac{k}{c}} \sum_{i=\xi_1+1}^{\xi_2} \frac{\varphi_i}{1-c} \mathrm{d}v - \int_{d^{u*}_{\xi_2}+\frac{k}{c}}^{d^u_{\xi_1}+\frac{k}{c}} \left(\sum_{i=1}^{\xi_1} \frac{\varphi_i}{1-c} - 1\right) \mathrm{d}v = 0,
\end{aligned}$$

which is a contradiction to the optimality of $\boldsymbol{\varphi}$.

Case 2: $\varphi_i \geq 1 - c$ for all $i \leq \xi_1$. If $\varphi_1^* \geq 1 - c$, then $d^{u*}_{\xi_2} = \overline{d}$. In this case, we have

$$\begin{aligned}
&Z(\boldsymbol{\varphi}^*) - Z(\boldsymbol{\varphi}) \\
&= \sum_{i=1}^{\xi_2} \int_{d^u_{\xi_2}+\frac{k}{c}}^{\overline{v}} \left(v_i - \frac{k}{c}\right) F(v_i)^{\xi_2 - 1} \mathrm{d}F(v_i) - \sum_{i=\xi_1+1}^{\xi_2} \int_{d^u_{\xi_2}+\frac{k}{c}}^{\overline{v}} \left(v_i - \frac{k}{c}\right) \frac{\varphi_i}{1-c} \mathrm{d}F(v_i) \\
&- \sum_{i=1}^{\xi_1} \int_{d^u_{\xi_2}+\frac{k}{c}}^{\overline{v}} \left(v_i - \frac{k}{c}\right) F(v_i)^{\xi_1 - 1} \mathrm{d}F(v_i) \\
&= \int_{d^u_{\xi_2}+\frac{k}{c}}^{\overline{v}} \left(v - \frac{k}{c}\right) \xi_2 F(v)^{\xi_2 - 1} \mathrm{d}F(v) - \int_{d^u_{\xi_2}+\frac{k}{c}}^{\overline{v}} \left(v - \frac{k}{c}\right) \left(\sum_{i=\xi_1+1}^{\xi_2} \frac{\varphi_i}{1-c} + \xi_1 F(v)^{\xi_1 - 1}\right) \mathrm{d}F(v) \\
&= \left(v - \frac{k}{c}\right) \left[F(v)^{\xi_2} - F(v) \sum_{i=\xi_1+1}^{\xi_2} \frac{\varphi_i}{1-c} - F(v)^{\xi_1}\right] \Bigg|_{d^u_{\xi_2}+\frac{k}{c}}^{\overline{v}} \\
&- \int_{d^u_{\xi_2}+\frac{k}{c}}^{\overline{v}} \left[F(v)^{\xi_2} - F(v) \sum_{i=\xi_1+1}^{\xi_2} \frac{\varphi_i}{1-c} - F(v)^{\xi_1}\right] \mathrm{d}v,
\end{aligned}$$

where the last equality holds by integration by parts. Because $d^u_{\xi_2}$ satisfies that

$$\left[1-F\left(d^u_{\xi_2}+\frac{k}{c}\right)\right]\sum_{i=\xi_1+1}^{\xi_2}\frac{\varphi_i}{1-c}+1-F\left(d^u_{\xi_2}+\frac{k}{c}\right)^{\xi_1}=1-F\left(d^u_{\xi_2}+\frac{k}{c}\right)^{\xi_2},$$
$$[1-F(v)]\sum_{i=\xi_1+1}^{\xi_2}\frac{\varphi_i}{1-c}+1-F(v)^{\xi_1}<1-F(v)^{\xi_2},\forall v>d^u_{\xi_2}+\frac{k}{c},$$

we have

$$Z(\boldsymbol{\varphi}^*)-Z(\boldsymbol{\varphi})>-\left(\overline{v}-\frac{k}{c}-d^u_{\xi_2}\right)\sum_{i=1}^{\xi_2}\frac{\varphi_i}{1-c}+\int_{d^u_{\xi_2}+\frac{k}{c}}^{\overline{v}}\sum_{i=\xi_1+1}^{\xi_2}\frac{\varphi_i}{1-c}\mathrm{d}v=0,$$

which is a contradiction to the optimality of $\boldsymbol{\varphi}$. If $\varphi^*_1<1-c$, then let $d^{u*}_1=\cdots=d^{u*}_{\xi_2}$ be defined by (A.13). Note that if $\xi_1=1$ and $\varphi_1/(1-c)=1$, then the new mechanism using $\boldsymbol{\varphi}^*$ coincides with the old mechanism using $\boldsymbol{\varphi}$. In this case, we can redefine $d^u_1:=d^u_{\xi_2}$ without changing the mechanism. Except for this case, we can show, by a similar argument to that in Case 1, that $Z(\boldsymbol{\varphi}^*)-Z(\boldsymbol{\varphi})>0$, which is a contradiction to the optimality of $\boldsymbol{\varphi}$.

Hence, by induction, we have $L=1$. For ease of notation, let $j:=\xi_1$. Next, we show that $j=0$ or n. Suppose, to the contrary, that $0<j<n$. Suppose that $d^l<0$, then the principal's objective function is strictly increasing in φ_i for $i>j$. Hence, in optimum, it must be that $d^l\geq 0$. Construct a new $\boldsymbol{\varphi}^*$ as follows: Let

$$\varphi^*_i=\frac{1}{n}\sum_{j=1}^{n}\varphi_j,\ \text{for all } i=1,\dots,n.$$

Case 1: $\varphi_i<1-c$ **for all** $i\leq j$**.** In this case, $\varphi^*_1<1-c$. Let $d^{u*}_1=\cdots=d^{u*}_n$ be such that

$$1-F\left(d^{u*}_j+\frac{k}{c}\right)^n=\sum_{i=1}^{n}\frac{\varphi^*_i}{1-c}\left[1-F\left(d^{u*}_j+\frac{k}{c}\right)\right].\tag{A.14}$$

Then $d^{u*}_j<d^u_j$. Let d^{l*} be such that

$$F\left(d^{l*}+\frac{k}{c}\right)^n=\sum_{i=1}^{n}\varphi^*_iF\left(d^{l*}+\frac{k}{c}\right).\tag{A.15}$$

Then $d^l < d^{l*}$. There are two subcases to consider: (i) $d^{l*} \le d_j^{u*}$ and (ii) $d^{l*} > d_j^{u*}$.

(i) Suppose that $d^{l*} \le d_j^{u*}$. Then

$$
\begin{aligned}
&Z(\boldsymbol{\varphi}^*) - Z(\boldsymbol{\varphi}) \\
=&\sum_{i=1}^{n}\left[\int_{\underline{v}}^{d^{l^*}+\frac{k}{c}}\left(v_i-\frac{k}{c}\right)\varphi_i^*\mathrm{d}F(v_i)+\int_{d^{l*}+\frac{k}{c}}^{d_j^{u*}+\frac{k}{c}}\left(v_i-\frac{k}{c}\right)F(v_i)^{n-1}\mathrm{d}F(v_i)+\int_{d_j^{u*}+\frac{k}{c}}^{\overline{v}}\left(v_i-\frac{k}{c}\right)\frac{\varphi_i^*}{1-c}\mathrm{d}F(v_i)\right] \\
&-\sum_{i=1}^{n}\int_{\underline{v}}^{d^l+\frac{k}{c}}\left(v_i-\frac{k}{c}\right)\varphi_i\mathrm{d}F(v_i)-\sum_{i=j+1}^{n}\int_{d^l+\frac{k}{c}}^{\overline{v}}\left(v_i-\frac{k}{c}\right)\frac{\varphi_i}{1-c}\mathrm{d}F(v_i) \\
&-\sum_{i=1}^{j}\left[\int_{d^l+\frac{k}{c}}^{d_j^u+\frac{k}{c}}\left(v_i-\frac{k}{c}\right)F(v_i)^{j-1}\mathrm{d}F(v_i)+\int_{d_j^u+\frac{k}{c}}^{\overline{v}}\left(v_i-\frac{k}{c}\right)\frac{\varphi_i}{1-c}\mathrm{d}F(v_i)\right] \\
=&\int_{d^l+\frac{k}{c}}^{d^{l^*}+\frac{k}{c}}\left(v-\frac{k}{c}\right)\sum_{i=1}^{n}\varphi_i\mathrm{d}F(v)+\int_{d^{l*}+\frac{k}{c}}^{d_j^{u*}+\frac{k}{c}}\left(v-\frac{k}{c}\right)nF(v)^{n-1}\mathrm{d}F(v)+\int_{d_j^{u*}+\frac{k}{c}}^{d_j^u+\frac{k}{c}}\left(v-\frac{k}{c}\right)\sum_{i=1}^{n}\frac{\varphi_i}{1-c}\mathrm{d}F(v) \\
&-\int_{d^l+\frac{k}{c}}^{d_j^u+\frac{k}{c}}\left(v-\frac{k}{c}\right)\sum_{i=j+1}^{n}\frac{\varphi_i}{1-c}\mathrm{d}F(v)-\int_{d^l+\frac{k}{c}}^{d_j^u+\frac{k}{c}}\left(v-\frac{k}{c}\right)jF(v)^{j-1}\mathrm{d}F(v) \\
=&d_j^u\left[F\left(d_j^u+\frac{k}{c}\right)\sum_{i=1}^{j}\frac{\varphi_i}{1-c}-F\left(d_j^u+\frac{k}{c}\right)^j\right]+d_j^{u*}\left[F\left(d_j^{u*}+\frac{k}{c}\right)^n-F\left(d_j^{u*}+\frac{k}{c}\right)\sum_{i=1}^{n}\frac{\varphi_i}{1-c}\right] \\
&+d^{l*}\left[F\left(d^{l*}+\frac{k}{c}\right)\sum_{i=1}^{n}\varphi_i-F\left(d^{l*}+\frac{k}{c}\right)^n\right] \\
&+d^l\left[-F\left(d^l+\frac{k}{c}\right)\sum_{i=1}^{n}\varphi_i+F\left(d^l+\frac{k}{c}\right)\sum_{i=j+1}^{n}\frac{\varphi_i}{1-c}+F\left(d^l+\frac{k}{c}\right)^j\right] \\
&-\int_{d^l+\frac{k}{c}}^{d^{l*}+\frac{k}{c}}F(v)\sum_{i=1}^{n}\varphi_i\mathrm{d}v-\int_{d^{l*}+\frac{k}{c}}^{d_j^{u*}+\frac{k}{c}}F(v)^n\mathrm{d}v-\int_{d_j^u+\frac{k}{c}}^{d_j^{u*}+\frac{k}{c}}F(v)\sum_{i=1}^{n}\frac{\varphi_i}{1-c}\mathrm{d}v \\
&+\int_{d^l+\frac{k}{c}}^{d_j^u+\frac{k}{c}}\left[F(v)\sum_{i=j+1}^{n}\frac{\varphi_i}{1-c}+F(v)^j\right]\mathrm{d}v,
\end{aligned}
$$

where the second equality holds because $\sum_{i=1}^n \varphi_i^* = \sum_{i=1}^n \varphi_i$ and the last equality holds by integration by parts. Because d_j^{u*} satisfies (A.14), d^{l*} satisfies (A.15), d_j^u satisfies that

$$
\begin{aligned}
1-F\left(d_j^u+\frac{k}{c}\right)^j &= \sum_{i=j+1}^{n}\frac{\varphi_i}{1-c}\left[1-F\left(d_j^u+\frac{k}{c}\right)\right], \\
1-F(v)^j &< \sum_{i=j+1}^{n}\frac{\varphi_i}{1-c}\left[1-F(v)\right], \forall v < d_j^u,
\end{aligned}
$$

and d^l satisfies that

$$1 - F\left(d^l + \frac{k}{c}\right)^j + \sum_{i=j+1}^{n} \frac{\varphi_i}{1-c}\left[1 - F\left(d^l + \frac{k}{c}\right)\right] + \sum_{i=1}^{n} \varphi_i F\left(d^l + \frac{k}{c}\right) = 1,$$

$$1 - F(v)^j + \sum_{i=j+1}^{n} \frac{\varphi_i}{1-c}\left[1 - F(v)\right] + \sum_{i=1}^{n} \varphi_i F(v) < 1, \forall v > d^l$$

$$F(v)^j - F(v)^n > [1 - F(v)] \sum_{i=j+1}^{n} \frac{\varphi_i}{1-c}, \forall v > d^l = d^u_{j+1},$$

we have

$$\begin{aligned}
&Z(\boldsymbol{\varphi}^*) - Z(\boldsymbol{\varphi}) \\
>&d^u_j \left(\sum_{i=1}^{j} \frac{\varphi_i}{1-c} - 1\right) + d^{u*}_j \left(1 - \sum_{i=1}^{n} \frac{\varphi_i}{1-c}\right) + d^l \sum_{i=j+1}^{n} \frac{\varphi_i}{1-c} \\
&- \int_{d^l+\frac{k}{c}}^{d^{l*}+\frac{k}{c}} \sum_{i=j+1}^{n} \frac{\varphi_i}{1-c} \mathrm{d}v - \int_{d^{l*}+\frac{k}{c}}^{d^{u*}_j+\frac{k}{c}} \sum_{i=j+1}^{n} \frac{\varphi_i}{1-c} \mathrm{d}v - \int_{d^u_j+\frac{k}{c}}^{d^{u*}_j+\frac{k}{c}} \left(\sum_{i=1}^{j} \frac{\varphi_i}{1-c} - 1\right) \mathrm{d}v = 0,
\end{aligned}$$

which is a contradiction to the optimality of $\boldsymbol{\varphi}$.

(ii) Suppose that $d^{l*} > d^{u*}_j$. In this case, redefine $d^{l*} = d^{u*}_j$ such that

$$\sum_{i=1}^{n} \left[\varphi^*_i F\left(d^{l*} + \frac{k}{c}\right) + \frac{\varphi^*_i}{1-c}\left(1 - F\left(d^{l*} + \frac{k}{c}\right)\right)\right] = 1.$$

Then $d^l < d^{l*} = d^{u*}_j < d^u_j$. By a similar argument to that in case (ii), we can show that $Z(\boldsymbol{\varphi}^*) - Z(\boldsymbol{\varphi}) > 0$, which is a contradiction to the optimality of $\boldsymbol{\varphi}$.

Case 2: $\varphi_i \geq 1 - c$ **for all** $i \leq j$**.** If $\varphi^*_1 \geq 1 - c$, then let $d^{u*}_1 = \ldots d^{u*}_n = \overline{d}$ and d^{l*} be defined by (A.15). By a similar argument to that in Case 1, we can show that $Z(\boldsymbol{\varphi}^*) - Z(\boldsymbol{\varphi}) > 0$, which is a contradiction to the optimality of $\boldsymbol{\varphi}$.

If $\varphi^*_1 < 1-c$, then let $d^{u*}_1 = \ldots d^{u*}_n = \overline{d}$ be defined by (A.14) and d^{l*} be defined by (A.15). Note that if $j = 1$ and $\varphi_1/(1-c) = 1$, then the new mechanism using $\boldsymbol{\varphi}^*$ coincides with the old mechanism using $\boldsymbol{\varphi}$. In this case, we can redefine $d^u_1 := d^l$ without changing the mechanism. Except for this case, we can show, by a similar argument to that in Case 1, that $Z(\boldsymbol{\varphi}^*) - Z(\boldsymbol{\varphi}) > 0$, which is a contradiction to the optimality of $\boldsymbol{\varphi}$.

Hence, $j = 0$ or n.

Case 1: $j = 0$**.** In this case, for all i,

$$P_i^*(v_i) = \begin{cases} \frac{\varphi_i}{1-c} & \text{if } v_i \geq d^l + \frac{k}{c} \\ \varphi_i & \text{if } v_i < d^l + \frac{k}{c} \end{cases}.$$

is an optimal mechanism given $\boldsymbol{\varphi}$. Furthermore, $\boldsymbol{\varphi}$ and d^l must satisfy

$$\left[1 - F\left(d^l + \frac{k}{c}\right)\right] \sum_{j=1}^{i} \frac{\varphi_j}{1-c} \leq 1 - F\left(d^l + \frac{k}{c}\right)^i, \forall i \leq n, \tag{A.16}$$

$$F\left(d^l + \frac{k}{c}\right) \sum_{i=1}^{n} \varphi_i + \left[1 - F\left(d^l + \frac{k}{c}\right)\right] \sum_{i=1}^{n} \frac{\varphi_i}{1-c} = 1. \tag{A.17}$$

In particular, (A.16) holds for $i = n$, which implies

$$\sum_{i=1}^{n} \frac{\varphi_i}{1-c} \leq \frac{1 - F\left(d^l + \frac{k}{c}\right)^n}{1 - F\left(d^l + \frac{k}{c}\right)}.$$

Substituting this into (A.17) yields

$$F\left(d^l + \frac{k}{c}\right)^{n-1} \leq \sum_{i=1}^{n} \varphi_i \leq \frac{(1-c)\left[1 - F\left(d^l + \frac{k}{c}\right)^n\right]}{1 - F\left(d^l + \frac{k}{c}\right)}.$$

By the proof of the second part in Theorem 3, $j = 0$ is optimal if $v^{**} \leq v^{\natural}$, in which case the optimal $d^l = d_1^u = \cdots = d_n^u = v^{**} - \frac{k}{c}$. The set of optimal $\boldsymbol{\varphi}$ is given by $\Phi(d^l, d_1^u, \ldots, d_n^u)$. Clearly, $\boldsymbol{\varphi} \in \Phi$ if and only if $\boldsymbol{\varphi}$ satisfies conditions (A.16) and (A.17). Because $v^{**} \leq v^{\natural}$ implies that

$$1 \leq \frac{1}{1 - cF(v^{**})} \leq \frac{1 - F(v^{**})^n}{1 - F(v^{**})},$$

there exists $1 \le h \le n$ such that

$$\frac{1 - F(v^{**})^{h-1}}{1 - F(v^{**})} \le \frac{1}{1 - cF(v^{**})} < \frac{1 - F(v^{**})^{h}}{1 - F(v^{**})}.$$

Hence, for all $i > h$, (A.16) holds if (A.17) holds. Given this, it is easy to see that the set of optimal $\boldsymbol{\varphi}$ is the convex hull of

$$\left\{ \boldsymbol{\varphi} \;\middle|\; \begin{array}{l} \varphi_{i_j} = (1-c)F(v^{**})^{j-1} \text{ if } j \le h-1, \varphi_{i_h} = \frac{1-c}{1-cF(v^{**})} - \sum_{j=1}^{h-1}(1-c)F(v^{**})^{j-1}, \\ \varphi_{i_j} = 0 \text{ if } j \ge h+1 \text{ and } (i_1, \dots, i_n) \text{ is a permutation of } (1, \dots, n) \end{array} \right\}.$$

Case 2: $j = n$**.** In this case, let $d^u := d_1^u = \cdots = d_n^u$, and

$$P_i^*(v_i) = \begin{cases} \frac{\varphi_i}{1-c} & \text{if } v_i \ge d^u + \frac{k}{c} \\ F(v)^{n-1} & \text{if } d^l + \frac{k}{c} < v_i < d^u + \frac{k}{c} \\ \varphi_i & \text{if } v_i \le d^l + \frac{k}{c} \end{cases}.$$

Furthermore, $\boldsymbol{\varphi}$, d^l and d^u must satisfy that

$$\left[1 - F\left(d^u + \frac{k}{c}\right)\right] \sum_{j=1}^{i} \frac{\varphi_j}{1-c} \le 1 - F\left(d^u + \frac{k}{c}\right)^i, \forall i \le n-1, \tag{A.18}$$

$$\left[1 - F\left(d^u + \frac{k}{c}\right)\right] \sum_{i=1}^{n} \frac{\varphi_i}{1-c} = 1 - F\left(d^u + \frac{k}{c}\right)^n, \tag{A.19}$$

$$F\left(d^l + \frac{k}{c}\right) \sum_{i=1}^{n} \varphi_i = F\left(d^l + \frac{k}{c}\right)^n. \tag{A.20}$$

(A.19) and (A.20) imply that d^l and d^u satisfy that

$$\frac{1 - F\left(d^u + \frac{k}{c}\right)^n}{1 - F\left(d^u + \frac{k}{c}\right)} = \frac{F\left(d^l + \frac{k}{c}\right)^{n-1}}{1-c}.$$

By the proof of the third part in Theorem 3, $j = n$ is optimal if $v^{**} > v^{\natural}$, in which case the optimal $d^l = v^l(\varphi^*) - \frac{k}{c}$ and the optimal $d_1^u = \cdots = d_n^u = v^u(\varphi^*) - \frac{k}{c}$. The set of optimal $\boldsymbol{\varphi}$ is given by

$\Phi(d^l, d_1^u, \ldots, d_n^u)$. Clearly, $\boldsymbol{\varphi} \in \Phi$ if and only if $\boldsymbol{\varphi}$ satisfies conditions (A.18)-(A.20). It is easy to see that Φ is the convex hull of

$$\left\{ \boldsymbol{\varphi} \,\middle|\, \varphi_{i_j} = (1-c)F(v^u(\varphi^*))^{j-1} \ \forall j \text{ and } (i_1, \ldots, i_n) \text{ is a permutation of } (1, \ldots, n) \right\}.$$

■

APPENDIX TO CHAPTER 3

B.1. The revelation principle

Consider a general mechanism that consists of a message space $\mathscr{M}$ and a quadruplet (a, p, q, θ), where $a : \mathscr{M} \to [0, 1]$ denotes the probability an agent obtains the good, $p : \mathscr{M} \to [0, 1]$ denotes the payment an agent must make, $q : \mathscr{M} \to [0, 1]$ denotes the probability of inspecting and $\theta : \mathscr{M} \times \{n, b_1, b_2\} \to [0, 1]$ denotes the probability an agent is penalized. In particular, $\theta(m, n)$ denotes the probability an agent is penalized if he is not inspected and $\theta(m, b)$ denotes the probability an agent is penalized if he is inspected and his budget is revealed to be b.

Given a mechanism, an agent of type $t = (v, b)$ chooses $m \in \mathscr{M}$ to maximize

$$a(m)v - p(m) - (1 - q(m))\theta(m, n)c - q(m)\theta(m, b)c$$

subject to the constraint that $p(m) \leq b$. Let $m^*(t)$ denote the solution to the agent's problem. For simplicity, I assume $m^*(t)$ is deterministic, but it is easy to accommodate mixed strategies. If the agent's problem has multiple solutions, then some deterministic selection rule is used. Consider a new mechanism with message space T. Let $a^*(t) = a(m^*(t))$, $p^*(t) = p(m^*(t))$, $q^*(t) = a(m^*(t))$ and $\theta^*(t, \cdot) = \theta(m^*(t), \cdot)$. Then the new mechanism is incentive compatible. Clearly, an agent has no incentive to report $\hat{t}$ such that $p^*(\hat{t}) > b$. For $\hat{t}$ such that $p^*(\hat{t}) \leq b$, we have

$$\begin{aligned} &a(m^*(t))v - p(m^*(t)) - (1 - q(m^*(t)))\theta(m^*(t), n)c - q(m^*(t))\theta(m^*(t), b)c \\ \geq &a(m^*(\hat{t}))v - p(m^*(\hat{t})) - (1 - q(m^*(\hat{t})))\theta(m^*(\hat{t}), n)c - q(m^*(\hat{t}))\theta(m^*(\hat{t}), b)c. \end{aligned}$$

The inequality simply follows from the fact that $m^*(t)$ is the solution to a type t agent's problem in the original mechanism. Clearly, the principal's payoff in the truthtelling equilibrium is as same as that in the original mechanism.

B.2. Common knowledge budgets

Proof of Theorem 7. Let (a, p) be a feasible mechanism. For each $b \in B$, $a(\cdot, b)$ is non-decreasing and $p(v, b) = va(v, b) - \int_{\underline{v}}^{v} a(\nu, b)\mathrm{d}\nu - u(\underline{v}, b)$. Consider another mechanism (a^*, p^*). Let $a^*(\cdot, b)$ be defined by

$$a^*(v, b) = \begin{cases} a(\overline{v}, b) & \text{if } v \geq v_b^* \\ 0 & \text{otherwise} \end{cases},$$

where v_b^* is such that

$$\int_{\underline{v}}^{\overline{v}} a(v, b) f(v)\mathrm{d}v = a(\overline{v}, b)(1 - F(v_b^*)). \tag{B.1}$$

Let $p^*(v, b) = va^*(v, b) - \int_{\underline{v}}^{v} a^*(\nu, b)\mathrm{d}\nu - u(\underline{v}, b)$. Clearly, (a^*, p^*) satisfies constraints (IR), (IC) and (S) and improves welfare. The revenue obtained by (a^*, p^*) is

$$\mathbb{E}_t[p^*(t)] = -(1 - \pi)u(\underline{v}, b_1) - \pi u(\underline{v}, b_2) + \int_{\underline{v}}^{\overline{v}} \left[v - \frac{1 - F(v)}{f(v)}\right] \left[(1 - \pi)a^*(v, b_1) + \pi a^*(v, b_2)\right] \mathrm{d}v.$$

By Assumption 2, $v - [1 - F(v)]/f(v)$ is strictly increasing. Thus, (a^*, p^*) also improves revenue, and therefore satisfies the (BB) constraint. Finally, we show that the (BC) constraint holds:

$$\begin{aligned} p^*(\overline{v}, b) =& \overline{v}a(\overline{v}, b) - \int_{\underline{v}}^{\overline{v}} a^*(v, b)\mathrm{d}v - u(\underline{v}, b) \\ \leq& \overline{v}a(\overline{v}, b) - \int_{\underline{v}}^{\overline{v}} a(v, b)\mathrm{d}v - u(\underline{v}, b) \leq b. \end{aligned}$$

The inequality holds if and only if

$$\begin{aligned} & \int_{\underline{v}}^{\overline{v}} [a^*(v, b) - a(v, b)]\mathrm{d}v \geq 0, \\ \Longleftrightarrow & \int_{v_b^*}^{\overline{v}} [a^*(v, b) - a(v, b)]\mathrm{d}v \geq \int_{\underline{v}}^{v_b^*} [a(v, b) - a^*(v, b)]\mathrm{d}v. \end{aligned}$$

The inequality holds since

$$\begin{aligned}\int_{v_b^*}^{\overline{v}}[a^*(v,b)-a(v,b)]\mathrm{d}v &= \int_{v_b^*}^{\overline{v}}[a^*(v,b)-a(v,b)]f(v)\frac{1}{f(v)}\mathrm{d}v\\ &\geq \int_{v_b^*}^{\overline{v}}[a^*(v,b)-a(v,b)]f(v)\frac{1}{f(v_b^*)}\mathrm{d}v\\ &= \int_{\underline{v}}^{v_b^*}[a(v,b)-a^*(v,b)]f(v)\frac{1}{f(v_b^*)}\mathrm{d}v\\ &\geq \int_{\underline{v}}^{v_b^*}[a(v,b)-a^*(v,b)]f(v)\frac{1}{f(v)}\mathrm{d}v\\ &= \int_{\underline{v}}^{v_b^*}[a(v,b)-a^*(v,b)]\mathrm{d}v,\end{aligned}$$

where the second and fourth line holds since f is non-increasing by Assumption 3 and the third line holds by (B.1). Hence there exists v_1^* and v_2^* such that the optimal allocation rule satisfies $a(v,b_1)=\chi_{\{v\geq v_1^*\}}(v)\min\left\{\frac{u(\underline{v},b_1)+b_1}{v_1^*},1\right\}$ and $a(v,b_2)=\chi_{\{v\geq v_2^*\}}(v)$. ■

B.3. Privately known budgets

B.3.1. No verification

Proof of Lemma 2. If $v=\underline{v}$, (3.3) reduces to $u(\underline{v},b_2)\geq u(\underline{v},b_1)$. Suppose $u(\underline{v},b_2)>u(\underline{v},b_1)$. Let

$$u^*(\underline{v},b_1)=u^*(\underline{v},b_2)=(1-\pi)u(\underline{v},b_1)+\pi u(\underline{v},b_2).$$

Let v^* be such that

$$v^* := \sup\left\{v\ \middle|\ \begin{array}{c}\int_{\underline{v}}^{v}a(\nu,b_1)\mathrm{d}\nu+u(\underline{v},b_1)-\int_{\underline{v}}^{v}\min\{a(\nu,b_1),a(\nu,b_2)\}\mathrm{d}\nu\\ -(1-\pi)u(\underline{v},b_1)-\pi u(\underline{v},b_2)\leq 0\end{array}\right\}.$$

Let $v^- := \sup\{v\leq v^*|a(v,b_2)\geq a(v,b_1)\}$ and $v^+ := \inf\{v\geq v^*|a(v,b_2)\geq a(v,b_1)\}$. Note that if $v^*=\overline{v}$, then $v^+=v^*=\overline{v}$. Note also that if $a(v^*,b_2)\geq a(v^*,b_1)$, then $v^+=v^-=v^*$. Clearly,

$a(v,b_1) > a(v,b_2)$ for all $v \in (v^-, v^+)$. There exists $\alpha \in (0,1)$ such that

$$\int_{\underline{v}}^{v^+} a(\nu,b_1)\mathrm{d}\nu + u(\underline{v},b_1) - \int_{\underline{v}}^{v^-} \min\{a(\nu,b_1),a(\nu,b_2)\}\mathrm{d}\nu$$
$$-\int_{v^-}^{v^+} [\alpha a(\nu,b_1) + (1-\alpha)a(\nu,b_2)]\mathrm{d}\nu - (1-\pi)u(\underline{v},b_1) - \pi u(\underline{v},b_2) = 0.$$

Assume without loss of generality that $a(v^-,b_1) = a(v^-,b_2)$ and $a(v^+,b_1) = a(v^+,b_2)$. Let

$$a^*(v,b_1) = \begin{cases} \min\{a(v,b_1),a(v,b_2)\} & \text{if } v < v^- \\ \alpha a(v,b_1) + (1-\alpha)a(v,b_2) & \text{if } v^- < v < v^+ \\ a(v,b_1) & \text{if } v > v^+ \end{cases},$$

and

$$a^*(v,b_2) = \begin{cases} \frac{(1-\pi)\left[a(v,b_1) - \min\{a(v,b_1),a(v,b_2)\}\right]}{\pi} + a(v,b_2) & \text{if } v < v^- \\ \frac{(1-\pi)\left[a(v,b_1) - \alpha a(v,b_1) - (1-\alpha)a(v,b_2)\right]}{\pi} + a(v,b_2) & \text{if } v^- < v < v^+ \\ a(v,b_2) & \text{if } v > v^+ \end{cases}.$$

Clearly, $a^*(v,b)$ is feasible and non-decreasing in v. By construction, we have $(1-\pi)u(\underline{v},b_1) + \pi u(\underline{v},b_2) = (1-\pi)u^*(\underline{v},b_1) + \pi u^*(\underline{v},b_2)$, $(1-\pi)a(v,b_1) + \pi a(v,b_2) = (1-\pi)a^*(v,b_1) + \pi a^*(v,b_2)$ for all v, and

$$u(\underline{v},b_1) + \int_{\underline{v}}^{v^+} a(v,b_1)\mathrm{d}v = u^*(\underline{v},b_1) + \int_{\underline{v}}^{v^+} a^*(v,b_1)\mathrm{d}v. \tag{B.2}$$

Hence

$$u(\underline{v},b_2) + \int_{\underline{v}}^{v^+} a(v,b_2)\mathrm{d}v = u^*(\underline{v},b_2) + \int_{\underline{v}}^{v^+} a^*(v,b_2)\mathrm{d}v. \tag{B.3}$$

Let $p^*(v,b) = va^*(v,b) - \int_{\underline{v}}^{v} a(\nu,b)\mathrm{d}\nu - u^*(\underline{v},b)$. Then

$$\begin{aligned}
p^*(\overline{v},b_1) &= \overline{v}a^*(\overline{v},b_1) - \int_{\underline{v}}^{\overline{v}} a^*(v,b_1)\mathrm{d}v - u^*(\underline{v},b_1) \\
&= \overline{v}a(\overline{v},b_1) - \int_{\underline{v}}^{v^+} a^*(v,b_1)\mathrm{d}v - \int_{v^+}^{\overline{v}} a(v,b_1)\mathrm{d}v - u^*(\underline{v},b_1) \\
&= \overline{v}a(\overline{v},b_1) - \int_{0}^{v^+} a(v,b_1)\mathrm{d}v - \int_{v^+}^{\overline{v}} a(v,b_1)\mathrm{d}v - u(\underline{v},b_1) \le b_1,
\end{aligned}$$

where the third line follows from (B.2). Hence the (BC) constraint holds. For $v < v^-$, we have $a^*(v,b_2) \ge a^*(v,b_1)$ and $u^*(\underline{v},b_1) = u^*(\underline{v},b_2)$. Hence (3.3) holds. For $v > v^+$, we have

$$\begin{aligned}
u^*(\underline{v},b_1) + \int_{\underline{v}}^{v} a^*(\nu,b_1)\mathrm{d}\nu &= u(\underline{v},b_1) + \int_{\underline{v}}^{v} a(\nu,b_1)\mathrm{d}\nu \\
&\le u(\underline{v},b_2) + \int_{\underline{v}}^{v} a(\nu,b_2)\mathrm{d}\nu \\
&= u^*(\underline{v},b_2) + \int_{\underline{v}}^{v} a^*(\nu,b_2)\mathrm{d}\nu,
\end{aligned}$$

where the first line follows from (B.2) and the third line follows from (B.3). Finally, consider $v \in [v^-, v^+]$. Suppose $\alpha \le 1-\pi$, then $a^*(v,b_1) \le a^*(v,b_2)$ for $v \in (v^-, v^+)$ and we have

$$\begin{aligned}
u^*(\underline{v},b_1) + \int_{\underline{v}}^{v} a^*(\nu,b_1)\mathrm{d}\nu &= u^*(\underline{v},b_1) + \int_{\underline{v}}^{v^-} a^*(\nu,b_1)\mathrm{d}\nu + \int_{v^-}^{v} a^*(\nu,b_1)\mathrm{d}\nu \\
&\le u^*(\underline{v},b_2) + \int_{\underline{v}}^{v^-} a^*(\nu,b_2)\mathrm{d}\nu + \int_{v^-}^{v} a^*(\nu,b_2)\mathrm{d}\nu \\
&= u^*(\underline{v},b_2) + \int_{\underline{v}}^{v} a^*(\nu,b_2)\mathrm{d}\nu.
\end{aligned}$$

Suppose $\alpha > \pi$, then $a^*(v,b_1) > a^*(v,b_2)$ for $v \in [v^-, v^+]$ and we have

$$\begin{aligned}
u^*(\underline{v},b_1) + \int_{\underline{v}}^{v} a^*(\nu,b_1)\mathrm{d}\nu &= u^*(\underline{v},b_1) + \int_{\underline{v}}^{v^+} a^*(\nu,b_1)\mathrm{d}\nu - \int_{v}^{v^+} a^*(\nu,b_1)\mathrm{d}\nu \\
&\le u^*(\underline{v},b_2) + \int_{\underline{v}}^{v^+} a^*(\nu,b_2)\mathrm{d}\nu - \int_{v}^{v^+} a^*(\nu,b_2)\mathrm{d}\nu \\
&= u^*(\underline{v},b_2) + \int_{\underline{v}}^{v} a^*(\nu,b_2)\mathrm{d}\nu.
\end{aligned}$$

Hence the (IC-b) constraint holds. Clearly, (a^*, p^*) also satisfies constraints (IR), (IC-v), (S) and (BB), and does not change welfare. ■

Proof of Lemma 3. Given Lemma 2, (3.3) becomes

$$\int_{\underline{v}}^{v} a(\nu, b_2)\mathrm{d}\nu \geq \int_{\underline{v}}^{v} a(\nu, b_1)\mathrm{d}\nu, \ \forall v. \tag{B.4}$$

For each $b \in \boldsymbol{B}$, we have

$$\begin{aligned}\int_{\underline{v}}^{v} a(\nu, b)f(\nu)\mathrm{d}\nu &= \int_{\underline{v}}^{v} f(\nu)\mathrm{d}\int_{\underline{v}}^{\nu} a(\nu', b)\mathrm{d}\nu' \\ &= f(v)\int_{\underline{v}}^{v} a(\nu', b)\mathrm{d}\nu' - \int_{\underline{v}}^{v}\left[\int_{\underline{v}}^{\nu} a(\nu', b)\mathrm{d}\nu'\right] f'(\nu)\mathrm{d}\nu.\end{aligned}$$

Since $f \geq 0$ and $-f' \geq 0$, (3.4) follows from (B.4). ■

Proof of Theorem 8. We first solve the optimal mechanism of $\mathcal{P}'$ and then verify that the optimal mechanism also satisfies the (IC) constraint of low-budget agents. Let (a, p) be a feasible mechanism. For each $b \in \boldsymbol{B}$, $a(\cdot, b)$ is non-decreasing and $p(v, b) = va(v, b) - \int_{\underline{v}}^{v} a(\nu, b)\mathrm{d}\nu - u(\underline{v}, b)$. Consider another mechanism (a^*, p^*).

Let $\hat{v} := \inf\{v | a(v, b_2) \geq a(\overline{v}, b_1)\}$. Note that $\hat{v} = \overline{v}$ if $a(\overline{v}, b_1) > a(\overline{v}, b_2)$ and $\hat{v} = \underline{v}$ if $a(\overline{v}, b_1) \leq a(\underline{v}, b_2)$. Let a^* be defined by

$$a^*(v, b_1) = \begin{cases} a(\overline{v}, b_1) & \text{if } v \geq v_1^*, \\ 0 & \text{otherwise,} \end{cases}$$

where v_1^* satisfies $a(\overline{v}, b_1)[1 - F(v_1^*)] = \int_{\underline{v}}^{\overline{v}} a(v, b_1) f(v) \mathrm{d}v$, and

$$a^*(v, b_2) = \begin{cases} 1 & \text{if } v \geq v_2^{**}, \\ a(\overline{v}, b_1) & \text{if } v_2^* \leq v < v_2^{**}, \\ 0 & \text{otherwise,} \end{cases}$$

where $v_2^* \leq \hat{v}$ satisfies $a(\overline{v}, b_1)[F(\hat{v}) - F(v_2^*)] = \int_{\underline{v}}^{\hat{v}} a(v, b_2) f(v) \mathrm{d}v$ and $v_2^{**} \geq \hat{v}$ satisfies $1 - F(v_2^{**}) + a(\overline{v}, b_1)[F(v_2^{**}) - F(\hat{v})] = \int_{\hat{v}}^{\overline{v}} a(v, b_2) f(v) \mathrm{d}v$. Let $p^*(v, b) = va^*(v, b) - \int_{\underline{v}}^{v} a^*(\nu, b) \mathrm{d}\nu - u(\underline{v}, b)$.

We show that $v_1^* \geq v_2^*$. If $v_1^* \geq \hat{v}$, then $v_1^* \geq v_2^*$. If $v_1^* < \hat{v}$, then

$$\begin{aligned} a(\overline{v}, b_1)[F(\hat{v}) - F(v_1^*)] &= \int_{\underline{v}}^{\hat{v}} a(v, b_1) f(v) \mathrm{d}v + \int_{\hat{v}}^{\overline{v}} [a(v, b_1) - a(\overline{v}, b_1)] f(v) \mathrm{d}v \\ &\leq \int_{\underline{v}}^{\hat{v}} a(v, b_1) f(v) \mathrm{d}v \\ &\leq \int_{\underline{v}}^{\hat{v}} a(v, b_2) f(v) \mathrm{d}v \\ &= a(\overline{v}, b_1)[F(\hat{v}) - F(v_2^*)], \end{aligned}$$

where the third line holds by Lemma 3. In this case, it must be that $a(\overline{v}, b_1) > 0$ since otherwise $a(\overline{v}, b_1) = 0 \leq a(0, b_2)$, which implies $\hat{v} = \underline{v} \leq v_1^*$. Hence, $v_2^* \leq v_1^*$. Thus, (a^*, p^*) satisfies the (IC-b) constraint.

Clearly, (a^*, p^*) also satisfies constraints (BC), (IR), (IC-v), (S) and (BB) and strictly improves welfare. Suppose $v_2^* < v_1^*$, then it is welfare improving to increase v_2^* and reduce v_1^* without affecting any constraint. Hence, it is optimal to set $v_1^* = v_2^* = v^*$. Let $u^* = u(\underline{v}, b_1) = u(\underline{v}, b_2)$. Then the optimal allocation rule satisfies $a(v, b_1) = \chi_{\{v \geq v^*\}} \min\left\{\frac{u^* + b_1}{v^*}, 1\right\}$ and $a(v, b_2) = \chi_{\{v \geq v^*\}} \min\left\{\frac{u^* + b_1}{v^*}, 1\right\} + \chi_{\{v \geq v_2^{**}\}} \left(1 - \min\left\{\frac{u^* + b_1}{v^*}, 1\right\}\right)$.

Clearly, if $u^* + b_1 > v^*$, we can reduce u^* such that $u^* + b_1 = v^*$ without affecting any constraint or the principal's objective function. This completes the characterization of the optimal mechanism of $\mathcal{P}'$. Finally, it is easy to see that the (IC) constraint of low-budget types is satisfied. This completes the proof. ■

B.3.2. The general case

Proof of Lemma 4. Suppose not. Then one can construct another feasible mechanism (a^*, p^*, q^*), which strictly improves welfare.

Let $\hat{v}_2^m = \inf\{v | a(v, b_2) \geq a^m\}$ for $m = 1, \dots, M$, $\hat{v}_2^0 = 0$ and $\hat{v}_2^{M+1} = \overline{v}$. Given a, the optimal verification rule satisfies $q(v, b_1) = q^m$ if $v \in (v_1^{m-1}, v_1^m)$ for $m = 1, \dots, M$, where

$$q^m = \frac{1}{c} \max \left\{ 0, u(\underline{v}, b_1) - u(\underline{v}, b_2) - \int_{\underline{v}}^{\hat{v}_2^m} a(\nu, b_2) \mathrm{d}\nu + a^m(\hat{v}_2^m - v_1^{m-1}) + \int_{\underline{v}}^{v_1^{m-1}} a(\nu, b_1) \mathrm{d}\nu \right\}.$$

For each $m = 1, \dots, M+1$, there exists $v_2^{m-1} \in [\hat{v}_2^{m-1}, \hat{v}_2^m]$ such that

$$\int_{\hat{v}_2^{m-1}}^{\hat{v}_2^m} a(v, b_2) f(v) \mathrm{d}v = a^{m-1}[F(v_2^{m-1}) - F(\hat{v}_2^{m-1})] + a^m[F(\hat{v}_2^m) - F(v_2^{m-1})]. \tag{B.5}$$

Consider $a^*(v, b_2)$ such that $a^*(v, b_2) = a^m$ if $v \in (v_2^{m-1}, v_2^m)$ for $m = 1, \dots, M$, $a^*(v, b_2) = 0$ if $v < v_2^0$ and $a^*(v, b_2) = 1$ if $v > v_2^M$. Note that if $a^1 = 0$, then $v_2^0 = \underline{v}$. If $a^M = 1$, then v_2^M is in-determined and we assume $v_2^M = v_2^{M-1}$. Let $a^*(v, b_1) = a(v, b_1)$.

Let $p^*(v, b) = va^*(v, b) - \int_{\underline{v}}^{v} a^*(\nu, b) \mathrm{d}\nu - u(\underline{v}, b)$. Let $q^*(v, b_1) = q(v, b_1)$. We show that the (IC-b) constraint is satisfied. That is, for $m = 1, \dots, M$,

$$q^m c \geq u(\underline{v}, b_1) - u(\underline{v}, b_2) - \int_{\underline{v}}^{v_2^{m-1}} a^*(\nu, b_2) \mathrm{d}\nu + a^m(v_2^{m-1} - v_1^{m-1}) + \int_{\underline{v}}^{v_1^{m-1}} a(\nu, b_1) \mathrm{d}\nu.$$

Since $a^*(v, b_2) = a^m$ for $v \in (v_2^{m-1}, \hat{v}_2^m)$, we have

$$\begin{aligned}
&u(\underline{v}, b_1) - u(\underline{v}, b_2) - \int_{\underline{v}}^{v_2^{m-1}} a^*(\nu, b_2)\mathrm{d}\nu + a^m(v_2^{m-1} - v_1^{m-1}) + \int_{\underline{v}}^{v_1^{m-1}} a(\nu, b_1)\mathrm{d}\nu \\
=&u(\underline{v}, b_1) - u(\underline{v}, b_2) - \int_{\underline{v}}^{\hat{v}_2^{m}} a^*(\nu, b_2)\mathrm{d}\nu + a^m(\hat{v}_2^{m} - v_1^{m-1}) + \int_{\underline{v}}^{v_1^{m-1}} a(\nu, b_1)\mathrm{d}\nu \\
\leq&u(\underline{v}, b_1) - u(\underline{v}, b_2) - \int_{\underline{v}}^{\hat{v}_2^{m}} a(\nu, b_2)\mathrm{d}\nu + a^m(\hat{v}_2^{m} - v_1^{m-1}) + \int_{\underline{v}}^{v_1^{m-1}} a(\nu, b_1)\mathrm{d}\nu,
\end{aligned}$$

where the last inequality holds if and only if

$$\int_{\underline{v}}^{\hat{v}_2^m} [a^*(\nu, b_2) - a(\nu, b_2)]\mathrm{d}\nu \geq 0.$$

To prove this, we prove that for $m = 1, \dots, M$

$$\int_{\hat{v}_2^{m-1}}^{\hat{v}_2^m} [a^*(\nu, b_2) - a(\nu, b_2)]\mathrm{d}\nu \geq 0. \tag{B.6}$$

(B.6) holds if and only if

$$\int_{v_2^{m-1}}^{\hat{v}_2^m} [a^*(\nu, b_2) - a(\nu, b_2)]\mathrm{d}\nu \geq \int_{\hat{v}_2^{m-1}}^{v_2^{m-1}} [a(\nu, b_2) - a^*(\nu, b_2)]\mathrm{d}\nu. \tag{B.7}$$

(B.7) follows from the construction of a^* and Assumption 3:

$$\begin{aligned}
\int_{v_2^{m-1}}^{\hat{v}_2^m} [a^*(\nu, b_2) - a(\nu, b_2)]\mathrm{d}\nu &\geq \int_{v_2^{m-1}}^{\hat{v}_2^m} [a^*(\nu, b_2) - a(\nu, b_2)] f(\nu) \frac{1}{f(v_2^{m-1})}\mathrm{d}\nu \\
&= \int_{\hat{v}_2^{m-1}}^{v_2^{m-1}} [a(\nu, b_2) - a^*(\nu, b_2)] f(\nu) \frac{1}{f(v_2^{m-1})}\mathrm{d}\nu \\
&\geq \int_{\hat{v}_2^{m-1}}^{v_2^{m-1}} [a(\nu, b_2) - a^*(\nu, b_2)]\mathrm{d}\nu.
\end{aligned}$$

By Assumption 2, $\mathbb{E}_t[p^*(t)] \geq \mathbb{E}_t[p(t)]$. Hence, constraint (BB) is satisfied. It is also clear that (a^*, p^*, q^*) satisfies constraints (IR), (IC-v), (BC) and (S), and strictly improves welfare. ■

Proof of Lemma 5. The proof is by contradiction. Let (a, p, q) be a feasible mechanism, where a is a M-step allocation rule, p satisfies the envelope condition and q is given by (3.7). Suppose (a, p, q) satisfies neither (C1) nor (C2). I show that one can construct another feasible mechanism (a^*, p^*, q^*), which strictly improves welfare and satisfies one of the two conditions. Furthermore, a^* is a M'-step function for some $M' \leq M$. I break the proof into three steps.

Step 1. Suppose $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1 \left(v_2^0 - v_1^0\right) < 0$. Let $m > 1$ be such that $v_2^{m'-1} - v_1^{m'-1} \leq 0$ for all $m' < m$ and $v_2^{m-1} - v_1^{m-1} > 0$. If there is no such m, then (a, p, q) satisfies (C2). Let $\hat{v}$ be defined by $F(\hat{v}) = \pi F(v_2^{m-1}) + (1 - \pi)F(v_1^{m-1})$ if $F(v_1^m) > \pi F(v_2^{m-1}) + (1 - \pi)F(v_1^{m-1})$ and $\hat{v} = v_1^m$ otherwise. Consider two different cases.

Case 1

Suppose $(a^m - a^{m-1})(\hat{v} - v_1^{m-1}) \geq \pi[u(\underline{v}, b_2) - u(\underline{v}, b_1) - a^1 \left(v_2^0 - v_1^0\right)]$, let $\tilde{v}_1^{m-1} \in [v_1^{m-1}, \hat{v}]$ be such that

$$(a^m - a^{m-1})(\tilde{v}_1^{m-1} - v_1^{m-1}) = \pi[u(\underline{v}, b_2) - u(\underline{v}, b_1) - a^1 \left(v_2^0 - v_1^0\right)].$$

Let $\tilde{v}_2^{m-1} \in [\hat{v}, v_2^{m-1}]$ be such that $\pi[F(v_2^{m-1}) - F(\tilde{v}_2^{m-1})] = (1 - \pi)[F(\tilde{v}_1^{m-1}) - F(v_1^{m-1})]$. Let $\tilde{v}_i^{m'} = v_i^{m'}$ for $i = 1, 2$ and $m' \neq m - 1$. Let $a^*(v, b_1) = a^{m-1}$ if $v \in (v_1^{m-1}, \tilde{v}_1^{m-1})$ and $a^*(v, b_1) = a(v, b_1)$ otherwise. Let $a^*(v, b_2) = a^m$ if $v \in (\tilde{v}_2^{m-1}, v_2^{m-1})$ and $a^*(v, b_2) = a(v, b_2)$ otherwise. Let $u^*(\underline{v}, b_1) = (1 - \pi)u(\underline{v}, b_1) + \pi u(\underline{v}, b_2) - \pi a^1 \left(v_2^0 - v_1^0\right)$ and $u^*(\underline{v}, b_2) = (1 - \pi)u(\underline{v}, b_1) + \pi u(\underline{v}, b_2) + (1 - \pi)a^1 \left(v_2^0 - v_1^0\right)$. Let $p^*(v, b) = va^*(v, b) - \int_{\underline{v}}^{v} a^*(\nu, b)\mathrm{d}\nu - u^*(\underline{v}, b)$. By construction, $p^*(\overline{v}, b_1) = p(\overline{v}, b_1) \leq b_1$. Hence, the (BC) constraint is satisfied. Let $q^*(v, b_1) = q(v, b_1)$. By Assumption 2, the (BB) constraint holds. For $v \in (\tilde{v}_1^{m'-1}, \tilde{v}_1^{m'})$, $m' = 1, \ldots, m - 1$, (IC-b) holds since

$$u^*(\underline{v}, b_1) - u^*(\underline{v}, b_2) + \sum_{j=1}^{m'} (a^j - a^{j-1})(\tilde{v}_2^{j-1} - \tilde{v}_1^{j-1}) \leq 0 \leq q^*(v, b_1)c.$$

For $v \in (\tilde{v}_1^{m'-1}, \tilde{v}_1^{m'})$, $m' = m, \dots, M$, we have $q^*(v, b_1) = q^m$. Then (IC-b) holds since

$$\begin{aligned}
& u^*(\underline{v}, b_1) - u^*(\underline{v}, b_2) + \sum_{j=1}^{m'} (a^j - a^{j-1})(\tilde{v}_2^{j-1} - \tilde{v}_1^{j-1}) \\
=& \sum_{j=1}^{m'} (a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) + (a^m - a^{m-1})(\tilde{v}_2^{m-1} - \tilde{v}_1^{m-1} - v_2^{m-1} + v_1^{m-1}) - a^1 \left(v_2^0 - v_1^0\right) \\
\leq& \sum_{j=1}^{m'} (a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) + \frac{(a^m - a^{m-1})(v_1^{m-1} - \tilde{v}_1^{m-1})}{\pi} - a^1 \left(v_2^0 - v_1^0\right) \\
=& \sum_{j=1}^{m'} (a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) + u(\underline{v}, b_1) - u(\underline{v}, b_2) \\
=& q^{m'} c,
\end{aligned}$$

where the third line holds since by Assumption 3

$$\begin{aligned}
v_2^{m-1} - \tilde{v}_2^{m-1} \geq& \frac{1}{f(\tilde{v}_2^{m-1})} [F(v_2^{m-1}) - F(\tilde{v}_2^{m-1})] \\
\geq& \frac{1-\pi}{\pi} \frac{1}{f(\tilde{v}_1^{m-1})} [F(\tilde{v}_1^{m-1}) - F(v_1^{m-1})] \\
\geq& \frac{1-\pi}{\pi} \left(\tilde{v}_1^{m-1} - v_1^{m-1}\right).
\end{aligned}$$

Clearly, (a^*, p^*, q^*) also satisfies constraints (IR), (IC-v) and (S) and strictly increases welfare. Note also that the new mechanism satisfies $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1 \left(v_2^0 - v_1^0\right) = 0$.

Suppose $\tilde{v}_1^{m-1} < v_1^m$, then continue with the argument in step 2.

Suppose $\tilde{v}_1^{m-1} = v_1^m < \tilde{v}_2^{m-1}$, then by the arguments in Lemma 4, we can construct a new mechanism which is feasible and strictly increases welfare, and whose allocation rule is a $(M-1)$-step allocation rule. Continue with the argument in step 2.

Case 2

Suppose $(a^m - a^{m-1})(\hat{v} - v_1^{m-1}) < \pi[u(\underline{v}, b_2) - u(\underline{v}, b_1) + a^1 \left(v_2^0 - v_1^0\right)]$. Let $\tilde{v}_1^{m-1} = \hat{v}$. Let $\tilde{v}_2^{m-1} \in [\hat{v}, v_2^{m-1}]$ be such that $\pi[F(v_2^{m-1}) - F(\tilde{v}_2^{m-1})] = (1-\pi)[F(\tilde{v}_1^{m-1}) - F(v_1^{m-1})]$. Let

$\tilde{v}_i^{m'} = v_i^{m'}$ for $i = 1,2$ and $m' \neq m-1$. Let $a^*(v,b_1) = a^{m-1}$ if $v \in (v_1^{m-1}, \tilde{v}_1^{m-1})$, and $a^*(v,b_1) = a(v,b_1)$ otherwise. Let $a^*(v,b_2) = a^m$ if $v \in (\tilde{v}_2^{m-1}, v_2^{m-1})$ and $a^*(v,b_2) = a(v,b_2)$ otherwise. Let $u^*(\underline{v},b_1) = u(\underline{v},b_1)+(a^m-a^{m-1})(\hat{v}-v_1^{m-1})$ and $u^*(\underline{v},b_2) = u(\underline{v},b_2)-(1-\pi)(a^m-a^{m-1})(\hat{v}-v_1^{m-1})/\pi$. Then $u^*(\underline{v},b_2) > u^*(\underline{v},b_1) + a^1\left(v_2^0 - v_1^0\right) \geq 0$. Let $p^*(v,b) = va^*(v,b) - \int_{\underline{v}}^{v} a^*(\nu,b)\mathrm{d}\nu - u^*(\underline{v},b)$. By construction, $p^*(\overline{v},b_1) = p(\overline{v},b_1) \leq b_1$. Hence, the (BC) constraint is satisfied. Let $q^*(v,b_1) = q(v,b_1)$. By Assumption 2, the (BB) constraint is satisfied. By the same argument in Case 1, the (IC-b) constraint is satisfied. Clearly, (a^*,p^*,q^*) also satisfies constraints (IR), (IC-v) and (S) and strictly increases welfare.

In this case, by construction, we have $\tilde{v}_1^{m-1} = \min\left\{\tilde{v}_2^{m-1}, v_1^m\right\}$.

Suppose $\tilde{v}_1^{m-1} = \tilde{v}_2^{m-1} < v_1^m$, then let $m^* > m$ be such that $\tilde{v}_2^{m'-1} - \tilde{v}_1^{m'-1} \leq 0$ for all $m' < m^*$ and $\tilde{v}_2^{m^*-1} - \tilde{v}_1^{m^*-1} > 0$. If there is no such m^*, (a^*,p^*,q^*) then satisfies (C2). Otherwise repeat the argument in step 1 for m^*.

Suppose $\tilde{v}_1^{m-1} = v_1^m \leq \tilde{v}_2^{m-1}$, then by the argument in Lemma 4, we can construct a new mechanism which is feasible and strictly increases welfare, and whose allocation rule is a $(M-1)$-step allocation rule. Repeat the arguments in step 1 for m.

Since M is finite, in finite steps we can construct a feasible mechanism (a,p,q) that either satisfies (C2) or $u(0,b_1) - u(0,b_2) + a^1 v_2^0 \geq 0$. In the latter case, continue with the argument in step 2.

Step 2. Suppose $u(\underline{v},b_1) - u(\underline{v},b_2) + a^1\left(v_2^0 - v_1^0\right) \geq 0$. Consider $m \geq 2$. Suppose (3.8) holds for all $m' < m$ and

$$u(\underline{v},b_1) - u(\underline{v},b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) < 0.$$

If there is no such m, then (a,p,q) satisfies (C1). It must be the case that $v_2^{m-1} < v_1^{m-1}$. Suppose $v_2^{m-1} < v_2^M$. Let $m^* \geq m$ be the smallest m' such that $v_2^{m'} > v_2^{m-1}$. That is, $v_2^{m^*} > v_2^{m-1}$ and

$v_2^{m'} = v_2^{m-1}$ for $m' = m, \dots, m^* - 1$. Let $\hat{v} \in [v_2^{m-1}, v_1^{m-1}]$ be such that

$$u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m-1} (a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) + (a^m - a^{m-1})(\hat{v} - v_1^{m-1}) = 0.$$

We consider two different cases.

Case 1

Suppose $(a^{m^*} - a^{m-1})[F(\hat{v}) - F(v_2^{m-1})] \leq (a^{m^*+1} - a^{m^*})[F(v_2^{m^*}) - F(\hat{v})]$. Let $\tilde{v}_2^{m^*} \in [\hat{v}, v_2^{m^*})$ be such that

$$(a^{m^*} - a^{m-1})[F(\tilde{v}_2^{m-1}) - F(v_2^{m-1})] = (a^{m^*+1} - a^{m^*})[F(v_2^{m^*}) - F(\tilde{v}_2^{m^*})]. \tag{B.8}$$

Let $\tilde{v}_2^{m'} = \hat{v}$ for $m' = m-1, \dots, m^* - 1$ and $\tilde{v}_2^{m'} = v_2^{m'}$ if $m' < m-1$ or $m' > m^*$. Let $\tilde{v}_1^{m'} = v_1^{m'}$ for all m'. Let $a^*(v, b_1) = a(v, b_1)$. Let $a^*(v, b_2) = a^{m-1}$ if $v \in (v_2^{m-1}, \tilde{v}_2^{m-1})$, $a^*(v, b_2) = a^{m^*+1}$ if $v \in (\tilde{v}_2^{m^*}, v_2^{m^*})$ and $a^*(v, b_2) = a(v, b_2)$ otherwise. Let $p^*(v, b) = va^*(v, b) - \int_{\underline{v}}^{v} a^*(\nu, b)\mathrm{d}\nu - u(\underline{v}, b)$. Clearly, $p^*(\overline{v}, b_1) = p(\overline{v}, b_1) \leq b_1$. Hence, the (BC) constraint is satisfied. Let $q^*(v, b_1) = q(v, b_1)$. By Assumption 2, the (BB) constraint holds.

Finally, we show that (a^*, p^*, q^*) satisfies the (IC-b) constraint. That is, for $m' = 1, \dots, M$

$$u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m'} (a^j - a^{j-1})(\tilde{v}_2^{j-1} - \tilde{v}_1^{j-1}) \leq q^{m'} c.$$

This is trivial for $m' \leq m$. For $m' = m+1, \dots, m^*$, we have $\tilde{v}_2^{m'-1} = \tilde{v}_2^{m-1} \leq v_1^{m-1} < v_1^{m'-1}$. Hence

$$u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m} (a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) + \sum_{j=m+1}^{m'} (a^j - a^{j-1})(\tilde{v}_2^{j-1} - v_1^{j-1}) < 0 \leq q^{m'} c.$$

Finally, consider $m' \geq m^* + 1$. It suffices to show that

$$\begin{aligned} &u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m^*+1} (a^j - a^{j-1})(\tilde{v}_2^{j-1} - \tilde{v}_1^{j-1}) \\ \leq &u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m^*+1} (a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}), \end{aligned}$$

which holds if and only if

$$(a^{m^*} - a^{m-1})(\tilde{v}_2^{m-1} - v_2^{m-1}) \leq (a^{m^*+1} - a^{m^*})(v_2^{m^*} - \tilde{v}_2^{m^*}).$$

The last inequality holds by (B.8) and Assumption 3. Clearly, (a^*, p^*, q^*) also satisfies constraints (IR), (IC-v) and (S) and strictly increases welfare. Let $m'' > m$ be such that (3.8) holds for all $m' < m''$ and is violated for m''. If there is no such m'', then (a^*, p^*, q^*) satisfies (C1). Otherwise repeat the argument in step 2 for m''.

Case 2

Suppose $(a^{m^*} - a^{m-1})[F(\hat{v}) - F(v_2^{m-1})] > (a^{m^*+1} - a^{m^*})[F(v_2^{m^*}) - F(\hat{v})]$. Let $\tilde{v}_2^{m-1}$ be such that

$$(a^{m^*} - a^{m-1})[F(\tilde{v}_2^{m-1}) - F(v_2^{m-1})] = (a^{m^*+1} - a^{m^*})[F(v_2^{m^*}) - F(\tilde{v}_2^{m-1})].$$

Let $\tilde{v}_2^{m'} = \tilde{v}_2^{m-1}$ for $m' = m, \dots, m^*$ and $\tilde{v}_2^{m'} = v_2^{m'}$ if $m' < m - 1$ or $m' > m^*$. Let $\tilde{v}_1^{m'} = v_1^{m'}$ for all m'. Let $a^*(v, b_1) = a(v, b_1)$. Let $a^*(\cdot, b_2)$ such that $a^*(v, b_2) = a^{m-1}$ if $v \in (v_2^{m-1}, \tilde{v}_2^{m-1})$, $a^*(v, b_2) = a^{m^*+1}$ if $v \in (\tilde{v}_2^{m-1}, v_2^{m^*})$ and $a^*(v, b_2) = a(v, b_2)$ otherwise. Let $p^*(v, b) = va^*(v, b) - \int_{\underline{v}}^{v} a^*(\nu, b)\mathrm{d}\nu - u(\underline{v}, b)$. Clearly, $p^*(\overline{v}, b_1) = p(\overline{v}, b_1) \leq b_1$. Hence, the (BC) constraint is satisfied. Let $q^*(v, b_1) = q(v, b_1)$. By Assumption 2, the (BB) constraint holds. By the same argument in Case 1, the (IC-b) constraint is satisfied. Clearly, (a^*, p^*, q^*) also satisfies constraints (IR), (IC-v) and (S) and strictly increases welfare. Note that for (a^*, p^*, q^*) we have $\tilde{v}_2^{m^*} = \tilde{v}_2^{m-1}$. Repeat the argument in step 2 for m with m^* replaced by $m^* + 1$.

Since M is finite, in finite steps we can construct a feasible mechanism (a, p, q) that either satisfies

(C1), or $v_2^M = v_2^{m-1} < v_1^{m-1}$ and

$$u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) < 0.$$

In the latter case, continue with the argument in step 3.

Step 3. Suppose $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1 \left(v_2^0 - v_1^0\right) \geq 0$,

$$u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m'}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) \geq 0$$

for all $m' = 1, \ldots, m-1$, $v_2^M = v_2^{m-1} < v_1^{m-1}$, and $u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) < 0$.

Let $\tilde{v}_1^{m-1} = v_1^{m-1} - \varepsilon$ for some $\varepsilon > 0$ and $\tilde{v}_2^{m'} = v_2^{m-1} + \delta$ for $m' = m-1, \ldots, M$, where $\delta > 0$ is such that

$$(1-\pi)(a^m - a^{m-1})\left[F(v_1^{m-1}) - F(\tilde{v}_1^{m-1})\right] = \pi(1 - a^{m-1})\left[F(v_2^{m-1}) - F(\tilde{v}_2^{m-1})\right]. \tag{B.9}$$

Let $\tilde{v}_i^{m'} = v_i^{m'}$ if $m' \neq m-1$ for $i = 1, 2$. Let $\varepsilon > 0$ be such that

$$\min\left\{\tilde{v}_1^{m-1} - v_1^{m-2}, u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(\tilde{v}_2^{j-1} - \tilde{v}_1^{j-1})\right\} = 0. \tag{B.10}$$

Since $\sum_{j=1}^{m-1}(a^j - a^{j-1})(\tilde{v}_2^{j-1} - \tilde{v}_1^{j-1}) \geq 0$, we have $\tilde{v}_2^{m'} \leq \tilde{v}_1^{m'}$ for all $m' \geq m-1$. Let $a^*(v, b_i) = a^m$ if $v \in (\tilde{v}_i^{m-1}, \tilde{v}_i^m)$ for $i = 1, 2$ and $m = 1, \ldots, M$, $a^*(v, b_2) = 0$ if $v < \tilde{v}_2^0$ and $a^*(v, b_2) = 1$ if $v > \tilde{v}_2^M$. Let $p^*(v, b) = va^*(v, b) - \int_{\underline{v}}^{v} a^*(\nu, b)\mathrm{d}\nu - u(\underline{v}, b)$. Since $a^*(\overline{v}, b_1) = a(\overline{v}, b_1)$ and $a^*(v, b_1) \geq a(v, b_1)$ for all v, we have $p^*(\overline{v}, b_1) \leq p(\overline{v}, b_1) \leq b_1$. Hence, the (BC) constraint is satisfied. Let $q^*(v, b_1) = q^m$ if $v \in (\tilde{v}_1^{m-1}, \tilde{v}_1^m)$ for $m = 1, \ldots, M$. Then the change of the verification cost is

$$k(q^m - q^{m-1})[F(v_1^{m-1}) - F(\tilde{v}_1^{m-1})].$$

Since $q^m = 0 \leq q^{m-1}$, the verification cost is reduced. Furthermore, by Assumption 2, the revenue increases. Hence, the (BB) constraint holds. Finally, we show that the (IC-b) constraint is satisfied. That is, for $m' = 1, \ldots, M$

$$u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m'} (a^j - a^{j-1})(\tilde{v}_2^{j-1} - \tilde{v}_1^{j-1}) \leq q^{m'} c.$$

This is trivial for $m' < m$. For $m' \geq m$, this holds since

$$u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m'} (a^j - a^{j-1})(\tilde{v}_2^{j-1} - \tilde{v}_1^{j-1}) \leq 0 = q^{m'} c.$$

Clearly, (a^*, p^*, q^*) also satisfies constraints (IR), (IC-v) and (S) and strictly increases welfare.

If the first term of (B.10) reaches zero first, then by the argument in Lemma 4, we can construct a new mechanism which is feasible and strictly increases welfare, and whose allocation rule is a $(M-1)$-step allocation rule. Then repeat the argument in step 3 for $m-1$. If the second term of (B.10) reaches zero first and $m < M$, then repeat the argument in step 3 for $m+1$. If the second term of (B.10) reaches zero first and $m = M$, then (a^*, p^*, q^*) satisfies (C1).

Since M is finite, in finite steps we can construct a feasible mechanism (a^*, p^*, q^*), which strictly improves welfare and satisfies (C1). Furthermore, a^* is a M'-step allocation rule for some $M' \leq M$.

■

Lemma 33 *Suppose Assumptions 2 and 3 hold. An optimal mechanism of $\mathcal{P}'(M, d)$ satisfies that $v_2^1 \geq v_1^1$.*

Proof of Lemma 33. Assume without loss of generality that $a^2 > a^1$. Suppose, on the contrary, that $v_2^1 < v_1^1$. Since (3.8) holds for $m = 2$, it must be that $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_0^2 - v_0^1) > 0$. Hence, it is either (i) $u(\underline{v}, b_1) > u(\underline{v}, b_2) \geq 0$, or (ii) $a^1 > 0$ and $v_2^0 > v_0^1$.

Suppose $u(\underline{v}, b_1) > u(\underline{v}, b_2) \geq 0$.

We construct another feasible mechanism (a^*, p^*, q^*), which strictly improves welfare. Let $\varepsilon > 0$ be sufficiently small. Let $\tilde{v}_1^1 = v_1^1 - \pi\varepsilon/(1-\pi)$ and $\tilde{v}_2^1 > v_2^1$ be such that $(1 -$

$\pi)\left[F(v_1^1) - F(\tilde{v}_1^1)\right] = \pi\left[F(\tilde{v}_2^1) - F(v_2^1)\right]$. By Assumption 3,

$$\begin{aligned}\tilde{v}_2^1 - v_2^1 &\leq \left[F(\tilde{v}_2^1) - F(v_2^1)\right]\frac{1}{f(\tilde{v}_2^1)}\\ &\leq \frac{1-\pi}{\pi}\left[F(v_1^1) - F(\tilde{v}_1^1)\right]\frac{1}{f(\tilde{v}_1^1)}\\ &\leq \frac{1-\pi}{\pi}(v_1^1 - \tilde{v}_1^1) = \varepsilon.\end{aligned}$$

For $\varepsilon > 0$ sufficiently small, $\tilde{v}_2^1 \leq \tilde{v}_1^1$. Let $\tilde{v}_i^m = v_i^m$ for $i = 1, 2$ and $m \neq 1$. Let $u^*(\underline{v}, b_2) = u(\underline{v}, b_2) + (a^2 - a^1)\varepsilon$ and $u^*(\underline{v}, b_1) = u(\underline{v}, b_1) - \pi(a^2 - a^1)\varepsilon/(1-\pi)$. For $\varepsilon > 0$ sufficiently small, $u^*(\underline{v}, b_1) \geq u^*(\underline{v}, b_2) > 0$. Let $a^*(v, b_i) = a^m$ for $v \in (\tilde{v}_i^{m-1}, \tilde{v}_i^m)$ for $i = 1, 2$ and $m = 1, \ldots, M$, $a^*(v, b_2) = 0$ if $v < \tilde{v}_2^0$ and $a^*(v, b_1) = 1$ if $v > \tilde{v}_2^M$. Let $p^*(v, b) = va^*(v, b) - \int_{\underline{v}}^{v} a^*(\nu, b)\mathrm{d}\nu - u^*(\underline{v}, b)$. By construction, $p^*(\overline{v}, b_1) = p(\overline{v}, b_1)$. Hence, the (BC) constraint is satisfied. Let $q^*(v, b) = q(v, b)$. Clearly, (a^*, p^*, q^*) satisfies constraints (IR), (IC-v) and (S), and strictly improves welfare. (a^*, p^*, q^*) satisfies (BB) by Assumption 2.

Finally, we show that (a^*, p^*, q^*) satisfies the (IC-b) constraint. If $v < \tilde{v}_1^1$, then

$$u^*(\underline{v}, b_1) - u^*(\underline{v}, b_2) + a^1(\tilde{v}_2^0 - \tilde{v}_1^0) = u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) - \frac{(a^2 - a^1)\varepsilon}{1-\pi} \leq q^1 c.$$

If $v \in (\tilde{v}_1^1, v_1^2)$, then

$$\begin{aligned}&u^*(\underline{v}, b_1) - u^*(\underline{v}, b_2) + a^1(\tilde{v}_2^0 - \tilde{v}_1^0) + (a^2 - a^1)(\tilde{v}_2^1 - \tilde{v}_1^1)\\ =&u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) + (a^2 - a^1)(v_2^1 - v_1^1)\\ &- \frac{(a^2 - a^1)\varepsilon}{1-\pi} + (a^2 - a^1)(\tilde{v}_2^1 - v_2^1 + v_1^1 - \tilde{v}_1^1)\\ \leq&u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) + (a^2 - a^1)(v_2^1 - v_1^1) - \frac{(a^2 - a^1)\varepsilon}{1-\pi} + (a^2 - a^1)\left(\varepsilon + \frac{\pi\varepsilon}{1-\pi}\right)\\ =&u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) + (a^2 - a^1)(v_2^1 - v_1^1)\\ \leq&\min\{q^2 c, q^1 c\},\end{aligned}$$

where the first inequality holds since $\tilde{v}_2^1 - v_2^1 \leq \varepsilon$ and the last inequality holds since $v_2^1 < v_1^1$.

If $v \in (v_1^{m-1}, v_1^m)$ for $m \geq 3$, then

$$\begin{aligned}
&u^*(\underline{v}, b_1) - u^*(\underline{v}, b_2) + a^1(\tilde{v}_2^0 - \tilde{v}_1^0) + (a^2 - a^1)(\tilde{v}_2^1 - \tilde{v}_1^1) + \sum_{j=3}^{m}(a^j - a^{j-1})(\tilde{v}_2^{j-1} - \tilde{v}_1^{j-1}) \\
\leq &u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) + (a^2 - a^1)(v_2^1 - v_1^1) + \sum_{j=3}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) \\
\leq &q^m c.
\end{aligned}$$

Hence, (IC-b) constraint is satisfied. This contradicts to that (a, p, q) is optimal.

Suppose $a^1 > 0$ and $v_2^0 > v_1^0$.

We construct another feasible mechanism (a^*, p^*, q^*), which strictly improves welfare. Let $\varepsilon \in (0, a^1]$ be sufficiently small. Let

$$\tilde{v}_1^1 = \frac{(a^2 - a^1)v_1^1 + \varepsilon v_1^0}{a^2 - a^1 + \varepsilon} < v_1^1.$$

By Assumption 3, we have

$$\begin{aligned}
&(a^2 - a^1)\left[F(v_1^1) - F(\tilde{v}_1^1)\right] \\
\leq &(a^2 - a^1)(v_1^1 - \tilde{v}_1^1)f(\tilde{v}_1^1) \\
= &\varepsilon(\tilde{v}_1^1 - v_1^0)f(\tilde{v}_1^1) \\
\leq &\varepsilon\left[F(\tilde{v}_1^1) - F(v_1^0)\right].
\end{aligned}$$

Let $\Delta := (a^2 - a^1)\left[F(v_1^1) - F(\tilde{v}_1^1)\right] - \varepsilon\left[F(\tilde{v}_1^1) - F(v_1^0)\right] \geq 0$. If $v_2^1 > v_2^0$, then let $\tilde{v}_2^0 = v_2^0$ and $\tilde{v}_2^1$ be such that

$$\pi(a^2 - a^1)\left[F(v_2^1) - F(\tilde{v}_2^1)\right] = \pi\varepsilon\left[F(\tilde{v}_2^1) - F(v_2^0)\right] + (1 - \pi)\Delta.$$

For $\varepsilon > 0$ sufficiently small, $\tilde{v}_2^1 \geq \tilde{v}_2^0 \geq v_1^0$. If $v_2^1 = v_2^0$, then let $\tilde{v}_2^1 = \tilde{v}_2^0$ be such that

$$\pi(a^2 - a^1)\left[F(v_2^1) - F(\tilde{v}_2^1)\right] = (1-\pi)\Delta.$$

For $\varepsilon > 0$ sufficiently small, $\tilde{v}_2^1 = \tilde{v}_2^0 \geq v_1^0$. Let $\tilde{v}_i^m = v_i^m$ for $i = 1, 2$ and $m \geq 2$. For $i = 1, 2$, let $a^*(v, b_i) = a^1 - \varepsilon$ if $v \in (\tilde{v}_i^0, \tilde{v}_i^1)$, $a^*(v, b_i) = a^2$ if $v \in (\tilde{v}_i^1, v_i^1)$, and $a^*(v, b_i) = a(v, b_i)$ otherwise. Let $u^*(\underline{v}, b) = u(\underline{v}, b)$ and $p^*(v, b) = va^*(v, b) - \int_{\underline{v}}^{v} a^*(\nu, b)\mathrm{d}\nu - u^*(\underline{v}, b)$. By construction, $p^*(\overline{v}, b_1) = p(\overline{v}, b_1)$. Hence, the (BC) constraint is satisfied. Let $q^*(v, b) = q(v, b)$. Clearly, (a^*, p^*, q^*) satisfies constraints (IR), (IC-v) and (S), and strictly improves welfare. (a^*, p^*, q^*) satisfies (BB) by Assumption 2.

Finally, we show that (a^*, p^*, q^*) satisfies the (IC-b) constraint. Suppose $v_2^1 > v_2^0$. If $v < \tilde{v}_1^1$, then

$$u^*(\underline{v}, b_1) - u^*(\underline{v}, b_2) + (a^1 - \varepsilon)(v_2^0 - v_1^0) < u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) = q^1 c.$$

If $v \in (\tilde{v}_1^1, \tilde{v}_1^2)$, then

$$\begin{aligned}
&u^*(\underline{v}, b_1) - u^*(\underline{v}, b_2) + (a^1 - \varepsilon)(v_2^0 - v_1^0) + (a^2 - a^1 + \varepsilon)(\tilde{v}_2^1 - \tilde{v}_1^1) \\
=&u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) + (a^2 - a^1)(v_2^1 - v_1^1) \\
&+ (a^2 - a^1 + \varepsilon)(\tilde{v}_2^1 - \tilde{v}_1^1) - \varepsilon(v_2^0 - v_1^0) - (a^2 - a^1)(v_2^1 - v_1^1) \\
=&u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) + (a^2 - a^1)(v_2^1 - v_1^1) \\
&+ (a^2 - a^1 + \varepsilon)\tilde{v}_2^1 - \varepsilon v_2^0 - (a^2 - a^1)v_2^1 \\
\leq&u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) + (a^2 - a^1)(v_2^1 - v_1^1) \leq \min\{q^1 c, q^2 c\},
\end{aligned}$$

where the last inequality holds since $v_2^1 < v_1^1$. To see that the first inequality holds, note that by Assumption 3,

$$\begin{aligned}(a^2 - a^1)(v_2^1 - \tilde{v}_2^1) &\geq (a^2 - a^1)\left[F(v_2^1) - F(\tilde{v}_2^1)\right]\frac{1}{f(\tilde{v}_2^1)}\\ &\geq \varepsilon\left[F(\tilde{v}_2^1) - F(v_2^0)\right]\\ &\geq \varepsilon(\tilde{v}_2^1 - v_2^0).\end{aligned}$$

Hence, $(a^2 - a^1 + \varepsilon)\tilde{v}_2^1 \leq (a^2 - a^1)v_2^1 + \varepsilon v_2^0$. Furthermore, $\tilde{v}_i^m = v_i^m$ for $i = 1, 2$ and $m \geq 2$. Hence, the (IC-b) constraint is satisfied. Suppose $v_2^0 = v_2^1$. If $v < \tilde{v}_1^1$, then

$$u^*(\underline{v}, b_1) - u^*(\underline{v}, b_2) + (a^1 - \varepsilon)(v_2^0 - v_1^0) < u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) = q^1 c.$$

If $v \in (\tilde{v}_1^1, \tilde{v}_1^2)$, then

$$\begin{aligned}&u^*(\underline{v}, b_1) - u^*(\underline{v}, b_2) + (a^1 - \varepsilon)(v_2^0 - v_1^0) + (a^2 - a^1 + \varepsilon)(\tilde{v}_2^1 - \tilde{v}_1^1)\\ =&u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) + (a^2 - a^1)(v_2^1 - v_1^1) + a^2(\tilde{v}_2^1 - v_2^1)\\ \leq&u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) + (a^2 - a^1)(v_2^1 - v_1^1) \leq \min\{q^1 c, q^2 c\},\end{aligned}$$

where the first inequality holds since $\tilde{v}_2^1 \leq v_2^1$ and the second inequality holds since $v_2^1 < v_1^1$. Furthermore, $\tilde{v}_i^m = v_i^m$ for $i = 1, 2$ and $m \geq 2$. Hence, the (IC-b) constraint is satisfied. This contradicts to that (a, p, q) is optimal.

Hence, $v_2^1 \geq v_1^1$. ■

Let $M \geq 3$ be an integer. We note that if a mechanism is a feasible solution to $\mathcal{P}'(M-1, d)$, then it is also a feasible solution to $\mathcal{P}'(M, d)$. Clearly, $V(M-1, d) \leq V(M, d)$. Suppose $V(M-1, d) < V(M, d)$, then in an optimal solution to $\mathcal{P}'(M, d)$ the allocation rule must be a M-step allocation rule, i.e.,

$$0 = a^0 \leq a^1 < a^2 < \cdots < a^M \leq a^{M+1} = 1,$$
$$\underline{v} = v_1^0 < v_1^1 < \cdots < v_1^M = \overline{v}.$$

Hence $\alpha^2 = \cdots = \alpha^M = 0$ and $\gamma_1^1 = \cdots = \gamma_1^M = 0$. Let $\rho := k/c$. Then the first-order conditions of

$\mathcal{P}'(M,d)$ are

$$\pi\left[\int_{v_2^{m-1}}^{v_2^m}\left[(1+\lambda)v-\lambda\frac{1-F(v)}{f(v)}\right]\mathrm{d}v-\beta[F(v_2^m)-F(v_2^{m-1})]\right]$$
$$+(1-\pi)\left[\int_{v_1^{m-1}}^{v_1^m}\left[(1+\lambda)v-\lambda\frac{1-F(v)}{f(v)}\right]f(v)\mathrm{d}v\right]$$
$$-(1-\pi)(1+\lambda)\rho(v_2^{m-1}-v_1^{m-1})[F(v_1^m)-F(v_1^{m-1})]-(1-\pi)\beta[F(v_1^m)-F(v_1^{m-1})]$$
$$+(1-\pi)(1+\lambda)\rho(v_2^m-v_1^m-v_2^{m-1}+v_1^{m-1})[1-F(v_1^m)]+\eta(v_1^m-v_1^{m-1})+\mu^m(v_2^{m-1}-v_1^{m-1})$$
$$-(v_2^m-v_1^m-v_2^{m-1}+v_1^{m-1})\sum_{j=m+1}^{M}\mu^j+\alpha^m-\alpha^{m+1}=0, \qquad (a^m, 1\le m\le M-1)$$
$$\pi\left[\int_{v_2^{M-1}}^{v_2^M}\left[(1+\lambda)v-\lambda\frac{1-F(v)}{f(v)}\right]f(v)\mathrm{d}v-\beta[F(v_2^M)-F(v_2^{M-1})]\right]$$
$$+(1-\pi)\int_{v_1^{M-1}}^{v_1^M}\left[(1+\lambda)v-\lambda\frac{1-F(v)}{f(v)}\right]f(v)\mathrm{d}v$$
$$-(1-\pi)(1+\lambda)\rho(v_2^{M-1}-v_1^{M-1})[F(v_1^M)-F(v_1^{M-1})]-(1-\pi)\beta[F(v_1^M)-F(v_1^{M-1})]$$
$$-\eta v_1^{M-1}+\mu^M(v_2^{M-1}-v_1^{M-1})+\alpha^M-\alpha^{M+1}=0, \qquad (a^M)$$
$$(a^{m+1}-a^m)\left\{(1-\pi)\left[(\beta-(1+\lambda)v_1^m)f(v_1^m)+(\lambda+\rho+\lambda\rho)[1-F(v_1^m)]+(1+\lambda)\rho(v_2^m-v_1^m)f(v_1^m)\right]\right.$$
$$\left.-\sum_{j=m+1}^{M}\mu^j-\eta\right\}=0, \qquad (v_1^m, 1\le m\le M-1)$$
$$a^1\left\{\pi\left[(\beta-(1+\lambda)v_2^0)f(v_2^0)+\lambda[1-F(v_2^0)]\right]-(1-\pi)(1+\lambda)\rho+\sum_{j=1}^{M}\mu^j\right\}+\gamma_2^0-\gamma_2^1=0, \qquad (v_2^0)$$
$$(a^{m+1}-a^m)\left\{\pi\left[(\beta-(1+\lambda)v_2^m)f(v_2^m)+\lambda[1-F(v_2^m)]\right]-(1-\pi)(1+\lambda)\rho[1-F(v_1^m)]+\sum_{j=m+1}^{M}\mu^j\right\}$$
$$+\gamma_2^m-\gamma_2^{m+1}=0, \qquad (v_2^m, 1\le m\le M-1)$$
$$\pi(a^{M+1}-a^M)\left[(\beta-(1+\lambda)v_2^M)f(v_2^M)+\lambda[1-F(v_2^M)]\right]+\gamma_2^M-\gamma_2^{M+1}=0, \qquad (v_2^M)$$
$$\eta+\sum_{m=1}^{M}\mu^m-(1-\pi)(\lambda+\rho+\lambda\rho)+\xi_1=0, \qquad (u(\underline{v},b_1))$$
$$-\sum_{m=1}^{M}\mu^m-\pi\lambda+(1-\pi)(1+\lambda)\rho+\xi_2=0. \qquad (u(\underline{v},b_2))$$

The variables in the parentheses denote with respect to which variables the first-order conditions are taken.

Lemma 34 *Suppose Assumptions 2 and 3 hold and $V(M,d) > V(M-1,d)$ for some $M \geq 3$. An optimal mechanism of $\mathcal{P}'(M,d)$ satisfies that $v_2^m - v_1^m$ is strictly increasing in $m = 1,\dots M-1$.*

Proof of Lemma 34. Since $a^{m+1} > a^m$ for $m = 1,\dots M-1$, the FOCs of v_1^m become

$$(1-\pi)\left[(\beta-(1+\lambda)v_1^m)f(v_1^m) + (\lambda+\rho+\lambda\rho)\left[1-F(v_1^m)\right] + (1+\lambda)\rho(v_2^m - v_1^m)f(v_1^m)\right] - \sum_{j=m+1}^{M} \mu^j - \eta = 0,$$

for $m = 1,\dots,M-1$. Then for $m = 1,\dots,M-1$

$$v_2^m - v_1^m = \frac{1}{\rho}v_1^m - \frac{\lambda+\rho+\lambda\rho}{(1+\lambda)\rho}\frac{1-F(v_1^m)}{f(v_1^m)} - \frac{\beta}{(1+\lambda)\rho} + \frac{\eta + \sum_{j=m+1}^{M}\mu^j}{(1-\pi)(1+\lambda)\rho f(v_1^m)},$$

which is strictly increasing in v_1^m by Assumptions 2 and 3. If $\mu^{m+1} = 0$, then $v_2^{m+1} - v_1^{m+1} > v_2^m - v_1^m$ since $v_1^{m+1} > v_1^m$.

If $\mu^{m+1} > 0$, then $v_2^{m+1} \geq v_1^{m+1} > v_1^m \geq v_2^m$ since (3.8) holds for m and $m+2$ and (3.8) holds with equality for $m+1$. Hence, $v_2^{m+1} - v_1^{m+1} \geq 0 \geq v_2^m - v_1^m$. By Lemma 33, $v_2^1 \geq v_1^1$. Hence, if there exists $m \geq 1$ such that $v_2^{m+1} - v_1^{m+1} \geq 0 \geq v_2^m - v_1^m$, then it must be the case that $v_2^{m+1} - v_1^{m+1} = v_2^m - v_1^m = \cdots = v_2^1 - v_1^1 = 0$. In particular, $v_2^2 - v_1^2 = v_2^1 - v_1^1$. Then we can construct another feasible mechanism (a^*, p^*, q^*), which strictly improves welfare. Let $\hat{v} \in (v_1^1, v_1^2)$ be such that

$$(a^3 - a^2)\left[F(v_1^2) - F(\hat{v})\right] = (a^2 - a^1)\left[F(\hat{v}) - F(v_1^1)\right].$$

Let $a^*(v,b) = a^1$ if $v \in (v_1^1, \hat{v})$, $a^*(v,b) = a^3$ if $v \in (\hat{v}, v_1^2)$ and $a^*(v,b) = a(v,b)$ otherwise. Let $p^*(v,b) = va^*(v,b) - \int_{\underline{v}}^{v} a^*(\nu,b)\mathrm{d}\nu - u(\underline{v},b)$. Then the (BC) constraint is satisfied by Assumption 3. Let $q^*(v,b_1) = q(v,b_1)$. Clearly, (a^*, p^*, q^*) satisfies constraints (IR), (IC-v), (IC-b), (S) and (BB), and strictly improves welfare. A contradiction. ■

Lemma 35 *Suppose Assumptions 2 and 3 hold. Then $V(M,d) = V(5,d)$ for all $M \geq 5$ and $d \geq 0$.*

Proof of Lemma 35. Fix $d \geq 0$ and $M \geq 6$ be an integer. We show that $V(M-1,d) = V(M,d)$. Suppose, on the contrary, that $V(M-1,d) < V(M,d)$, then in an optimal solution to $\mathcal{P}'(M,d)$ the allocation rule must be a M-step allocation rule. In particular, $0 = a^0 \leq a^1 < a^2 < \cdots < a^M \leq a^{M+1} = 1$. By Lemmas 33 and 34, an optimal solution to $\mathcal{P}'(M,d)$ must satisfies

$$v_2^{M-1} - v_1^{M-1} > v_2^{M-2} - v_1^{M-2} > \cdots > v_2^1 - v_1^1 \geq 0. \tag{B.11}$$

Fix $\underline{v} = v_1^0 < v_1^1 < \cdots < v_1^M = \overline{v}$ and $0 \geq v_2^0 \leq v_2^1 \leq \cdots \leq v_2^M \leq v_2^{M+1} \leq \overline{v}$ such that (B.11) holds. Then $\mathcal{P}'(M,d)$ is linear in $u(\underline{v},b_1)$, $u(\underline{v},b_2)$ and a^m for $m = 1,\ldots,M$. Then an optimal solution can be obtained at an extreme point of the feasible region. By (B.11), inequalities corresponding to μ^m for $m = 2,\ldots M$ holds if the inequality corresponding to μ^1 holds. Hence, the feasible set is characterized by (S), (BC), (BB) and the following inequalities:

$$\begin{aligned} &u(\underline{v},b_1) \geq 0, u(\underline{v},b_2) \geq 0, \\ &u(\underline{v},b_1) - u(\underline{v},b_2) + a^1(v_2^0 - v_1^0) \geq 0, \\ &0 \leq a^0 \leq a^1 \leq \cdots \leq a^M \leq a^{M+1} = 1. \end{aligned} \tag{B.12}$$

Note that if $a^1 = 0$, then $u(\underline{v},b_1) \geq 0$ is redundant. Hence, in addition to (S), (BC), (BB) and $a^M \leq 1$, at most three of the following four inequalities are active at the same time: $u(\underline{v},b_1) \geq 0$, $u(\underline{v},b_2) \geq 0$, $a^1 \geq 0$ and (B.12). Since $M \geq 6$, at least one of the following constraints hold with equality: $a^1 \leq a^2 \cdots \leq a^{M-1} \leq a^M$, a contradiction. ■

Lemma 36 *Suppose Assumptions 2 and 3 hold. For any $d > 0$, there exists $\overline{M}(d)$ such that for all $M > \overline{M}(d)$,*

$$V - V(M,d) \leq (1-\pi)\left(1 + \frac{k}{c}\right)\frac{\mathbb{E}[v]}{M}.$$

Proof of Lemma 36. Let (a,p,q) denote an optimal mechanism of $\mathcal{P}'$. Then $p(v,b) = va(v,b) - \int_{\underline{v}}^{v} a(v,b)\mathrm{d}v - u(\underline{v},b)$ for all $(v,b) \in T$ and q is defined by (3.6). Fix $M \geq 2$. Let $a^0 = 0$, $a^{M+1} = 1$

and $a^m = (m-1)a(\overline{v}, b_1)/M$ for $m = 1, \dots, M$. Let $v_1^0 = \underline{v}$, $v_1^M = \overline{v}$ and for $m = 0, \dots, M-1$

$$v_1^m = \inf \left\{ v \,\middle|\, a(v, b_1) \geq a^{m+1} \right\}.$$

Then $\underline{v} = v_1^0 \leq v_1^1 \leq \cdots \leq v_1^M = \overline{v}$ and $0 = a^0 \leq a^1 < a^2 < \cdots < a^M \leq a^{M+1} = 1$. Let $a^*(v, b_1) = a^m$ if $v \in (v_1^{m-1}, v_1^m)$ for $m = 1, \dots, M$. Then $a(v, b_1) - 1/M \leq a^*(v, b_1) \leq a(v, b_1)$. Let $\hat{v}_2^m = \inf\{v | a(v, b_2) \geq a^m\}$ for $m = 1, \dots, M$, $\hat{v}_2^0 = 0$ and $\hat{v}_2^{M+1} = \overline{v}$. For each $m = 1, \dots, M+1$, there exists $v_2^{m-1} \in [\hat{v}_2^{m-1}, \hat{v}_2^m]$ such that

$$\int_{\hat{v}_2^{m-1}}^{\hat{v}_2^m} a(v, b_2) f(v) \mathrm{d}v = a^{m-1}[F(v_2^{m-1}) - F(\hat{v}_2^{m-1})] + a^m[F(\hat{v}_2^m) - F(v_2^{m-1})]. \tag{B.13}$$

Consider $a^*(v, b_2)$ such that $a^*(v, b_2) = a^m$ if $v \in (v_2^{m-1}, v_2^m)$ for $m = 1, \dots, M$, $a^*(v, b_2) = 0$ if $v < v_2^0$ and $a^*(v, b_2) = 1$ if $v > v_2^M$. Note that since $a^1 = 0$, we have $v_2^0 = \underline{v}$. Clearly, a^* satisfies constraint (S). Let $p^*(v, b) = v a^*(v, b) - \int_{\underline{v}}^{v} a^*(\nu, b) \mathrm{d}\nu - u(\underline{v}, b)$ for $b \in B$. Let q^* be such that

$$c q^*(v, b_1) = c q(v, b_1) + \frac{v}{M}.$$

We show that the (IC-b) constraint is satisfied, i.e., for all $v \in (v_1^{m-1}, v_1^m)$, $m = 1, \dots, M$,

$$c q^*(v, b_1) \geq u(\underline{v}, b_1) - u(\underline{v}, b_2) - \int_{\underline{v}}^{v_2^{m-1}} a(\nu, b_2) \mathrm{d}\nu + a^m(v_2^{m-1} - v) + \int_{\underline{v}}^{v} a^*(\nu, b_1) \mathrm{d}\nu.$$

Recall that for $v \in (v_1^{m-1}, v_1^m)$, we have

$$c q(v, b_1) \geq u(\underline{v}, b_1) - u(\underline{v}, b_2) - \int_{\underline{v}}^{\hat{v}_2^m} a(\nu, b_2) \mathrm{d}\nu + a(v, b_1)(\hat{v}_2^m - v) + \int_{\underline{v}}^{v} a(\nu, b_1) \mathrm{d}\nu.$$

Then for $v \in (v_1^{m-1}, v_1^m)$

$$
\begin{aligned}
cq^*(v, b_1) &= cq(v, b_1) + \frac{v}{M} \\
&\geq u(\underline{v}, b_1) - u(\underline{v}, b_2) - \int_{\underline{v}}^{\hat{v}_2^m} a(\nu, b_2)\mathrm{d}\nu + a(v, b_1)\hat{v}_2^m - \left(a(v, b) - \frac{1}{M}\right) v + \int_{\underline{v}}^{v} a(\nu, b_1)\mathrm{d}\nu \\
&\geq u(\underline{v}, b_1) - u(\underline{v}, b_2) - \int_{\underline{v}}^{\hat{v}_2^m} a(\nu, b_2)\mathrm{d}\nu + a^m(\hat{v}_2^m - v) + \int_{\underline{v}}^{v} a(\nu, b_1)\mathrm{d}\nu \\
&\geq u(\underline{v}, b_1) - u(\underline{v}, b_2) - \int_{\underline{v}}^{\hat{v}_2^m} a^*(\nu, b_2)\mathrm{d}\nu + a^m(\hat{v}_2^m - v) + \int_{\underline{v}}^{v} a^*(\nu, b_1)\mathrm{d}\nu \\
&= u(\underline{v}, b_1) - u(\underline{v}, b_2) - \int_{\underline{v}}^{v_2^{m-1}} a^*(\nu, b_2)\mathrm{d}\nu + a^m(v_2^{m-1} - v) + \int_{\underline{v}}^{v} a^*(\nu, b_1)\mathrm{d}\nu,
\end{aligned}
$$

where the third line holds since $a(v, b) - 1/M \leq a^*(v, b) \leq a(v, b)$ and the fourth line holds by the same argument in the proof of Lemma 4. Then

$$
\begin{aligned}
&\mathbb{E}_t[p^*(t) - q^*(t)k] - \mathbb{E}_t[p(t) - q(t)k] \\
=&\pi \int_{\underline{v}}^{\overline{v}} \left[v - \frac{1 - F(v)}{f(v)}\right] [a^*(v, b_2) - a(v, b_2)] f(v)\mathrm{d}v \\
&+ (1 - \pi) \int_{\underline{v}}^{\overline{v}} \left[v - \frac{1 - F(v)}{f(v)}\right] [a^*(v, b_1) - a(v, b_1)] f(v)\mathrm{d}v \\
&- (1 - \pi) \int_{\underline{v}}^{\overline{v}} k[q^*(v, b_1) - q(v, b_1)] f(v)\mathrm{d}v \\
\geq& - \frac{\mathbb{E}[v]}{M} - (1 - \pi)\frac{\mathbb{E}[v]}{M}\frac{k}{c}.
\end{aligned}
$$

For any $d > 0$, there exists $\overline{M}(d)$ such that for all $M > \overline{M}(d)$, we have $\frac{\mathbb{E}[v]}{M} + (1 - \pi)\frac{\mathbb{E}[v]}{M}\frac{k}{c} < d$. Then (a^*, p^*, q^*) is a feasible solution to $\mathcal{P}'(M, d)$ for $M > \overline{M}(d)$. Hence,

$$
\begin{aligned}
&V - V(M, d) \\
\leq&(1 - \pi) \left[\int_{\underline{v}}^{\overline{v}} v[a(v, b_1) - a^*(v, b_1)] f(v)\mathrm{d}v - \int_{\underline{v}}^{\overline{v}} \left[q(v, b_1) - q^*(v, b_1)\right] k f(v)\mathrm{d}v\right] \\
\leq&(1 - \pi) \left(1 + \frac{k}{c}\right) \frac{\mathbb{E}[v]}{M}.
\end{aligned}
$$

■

Proof of Theorem 9. By Lemmas 7 and 36, we have

$$V - V(2,d) = V - V(M,d) \leq (1-\pi)\left(1+\frac{k}{c}\right)\frac{\mathbb{E}[v]}{M}.$$

Let M goes to infinity and we have $V(2,0) \leq V \leq V(2,d)$ for all $d > 0$. By Lemma 37, $\lim_{d\to 0} V(2,d) = V(2,0)$. Hence, $V = V(2,0)$.

Hence, there exists $u(\underline{v}, b_1) \geq 0$, $u(\underline{v}, b_2) \geq 0$, $\underline{v} \leq v_1^1 \leq \overline{v}$, $\underline{v} \leq v_2^0 \leq v_2^1 \leq v_2^2 \leq \overline{v}$ and $0 \leq a^1 \leq a^2 \leq \overline{v}$ the optimal mechanism of $\mathcal{P}'$ is given by

$$\begin{aligned}
a(v,b_1) &= a^1 + \chi_{\{v \geq v_1^1\}}\left(a^2 - a^1\right),\\
a(v,b_2) &= \chi_{\{v \geq v_2^0\}} a^1 + \chi_{\{v \geq v_2^1\}}\left(a^2 - a^1\right) + \chi_{\{v \geq v_2^2\}}(1-a^2),\\
p(v,b_1) &= -u(\underline{v}, b_1) + \chi_{\{v \geq v_1^1\}}\left(a^2 - a^1\right) v_1^1,\\
p(v,b_2) &= -u(\underline{v}, b_2) + \chi_{\{v \geq v_2^0\}} a^1 v_2^0 + \chi_{\{v \geq v_2^1\}}\left(a^2 - a^1\right) v_2^1 + \chi_{\{v \geq v_2^2\}}(1-a^2) v_2^2,\\
q(v,b_1) &= \frac{1}{c}\left[u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1\left(v_2^0 - v_1^0\right) + \chi_{\{v \geq v_1^1\}}\left(a^2 - a^1\right)(v_2^1 - v_1^1)\right],\\
q(v,b_2) &= 0.
\end{aligned}$$

By Lemma 33, $v_2^1 \geq v_1^1$. We show below that $v_2^0 = \underline{v}$ and $a^1 = 0$.

First, we show that $v_2^0 = \underline{v}$. We consider two different cases: $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1\left(v_2^0 - v_1^0\right) = 0$ and $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1\left(v_2^0 - v_1^0\right) > 0$.

Suppose $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1\left(v_2^0 - v_1^0\right) = 0$.

Suppose to the contradiction that $v_2^0 > \underline{v}$. Then we can construct another feasible mechanism (a^*, p^*, q^*), which strictly improves welfare. Since $v_2^0 > \underline{v} = v_1^0$, we have $u(\underline{v}, b_2) > u(\underline{v}, b_1)$ and, by construction, $a^1 > 0$ and $v_1^1 > \underline{v}$.

Let $\varepsilon > 0$ be sufficiently small. Let $\tilde{v}_1^0 = \underline{v} + \varepsilon$ and $\tilde{v}_2^0 < v_2^0$ be such that $\pi[F(v_2^0) - F(\tilde{v}_2^0)] =$

$(1-\pi)F(\underline{v}+\varepsilon)$. For $\varepsilon > 0$ sufficiently small, $\tilde{v}_1^0 < \min\{v_1^1, \tilde{v}_2^0\}$. Let $\tilde{v}_i^1 = v_i^1$ for $i = 1,2$. Let $u^*(\underline{v}, b_1) = u(\underline{v}, b_1) + a^1\varepsilon$ and $u^*(\underline{v}, b_2) = u(\underline{v}, b_2) - (1-\pi)a^1\varepsilon/\pi$. For $\varepsilon > 0$ sufficiently small, $u^*(\underline{v}, b_2) \geq u^*(\underline{v}, b_1) > 0$. Let $a^*(v, b_1) = 0$ if $v < \tilde{v}_1^0$ and $a^*(v, b_1) = a(v, b_1)$ otherwise. Let $a^*(v, b_2) = a^1$ if $v \in (\tilde{v}_2^0, v_2^0)$ and $a^*(v, b_2) = a(v, b_2)$ otherwise. Let $p^*(v, b) = va^*(v, b) - \int_{\underline{v}}^{v} a^*(\nu, b)\mathrm{d}\nu - u^*(\underline{v}, b)$. Since $u^*(\underline{v}, b_1) - a^1\tilde{v}_1^0 = u(\underline{v}, b_1) - a^1 v_1^0$, we have $p^*(\overline{v}, b_1) = p(\overline{v}, b_1) \leq b_1$. Hence, the (BC) constraint is satisfied. Let $q^*(v, b_1) = q(v, b_1)$. Clearly, (a^*, p^*, q^*) satisfies constraints (IR), (IC-v) and (S), and strictly improves welfare. (a^*, p^*, q^*) satisfies (BB) by Assumption 2.

Finally, we show that (a^*, p^*, q^*) satisfies the (IC-b) constraint. First, for $v < \underline{v} + \varepsilon$, we have $u^*(\underline{v}, b_1) - u^*(\underline{v}, b_2) \leq 0 \leq q^*(v, b_1)c$. Next, we show that for $m = 1,2$

$$q^m c \geq u^*(\underline{v}, b_1) - u^*(\underline{v}, b_2) + \sum_{j=1}^{m} (a^j - a^{j-1})(\tilde{v}_2^{j-1} - \tilde{v}_1^{j-1}).$$

Since

$$\begin{aligned}
v_2^0 - \tilde{v}_2^0 &= \int_{\tilde{v}_2^0}^{v_2^0} f(v)\frac{1}{f(v)}\mathrm{d}v \\
&\geq \frac{1}{f(\tilde{v}_2^0)}[F(v_2^0) - F(\tilde{v}_2^0)] \\
&\geq \frac{1-\pi}{\pi}\frac{F(\underline{v}+\varepsilon)}{f(\underline{v}+\varepsilon)} \\
&\geq \frac{1-\pi}{\pi}\varepsilon,
\end{aligned}$$

where the inequalities hold by Assumption 3, we have

$$\begin{aligned}
&u^*(\underline{v}, b_1) - u^*(\underline{v}, b_2) + \sum_{j=1}^{1} (a^j - a^{j-1})(\tilde{v}_2^{j-1} - \tilde{v}_1^{j-1}) \\
=&u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1 v_2^0 + \frac{a^1\varepsilon}{\pi} + a^1(\tilde{v}_2^0 - v_2^0) - a^1(v_1^0 + \varepsilon) \\
\leq&u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1\left(v_2^0 - v_1^0\right).
\end{aligned}$$

Furthermore, $\tilde{v}_i^m = v_i^m$ for $i = 1, 2$ and $m \geq 1$. Hence, the (IC-b) constraint is satisfied. This contradicts to that (a, p, q) is optimal. Hence $v_2^0 = \underline{v}$.

Suppose $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1 \left(v_2^0 - v_1^0\right) > 0$**.**

Suppose to the contradiction that $v_2^0 > \underline{v}$. In this case, $\gamma_2^0 = 0$. By construction, we have $a^1 > 0$ and $v_1^1 > \underline{v}$. Hence, $\alpha_1 = 0$. Furthermore, since $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1 \left(v_2^0 - v_1^0\right) > 0$ and $v_2^1 \geq v_1^1$, we have $\mu^1 = \mu^2 = 0$. Then v_2^0 satisfies

$$\pi \left[(\beta - (1+\lambda)v_2^0) f(v_2^0) + \lambda[1 - F(v_2^0)]\right] - (1-\pi)(1+\lambda)\rho[1 - F(v_1^0)] = 0, \tag{B.14}$$

$$\begin{aligned} &\pi \int_{v_2^0}^{v_2^M} [v + \lambda\varphi(v) - \beta]\, f(v)\mathrm{d}v \\ &+ (1-\pi)\left[\int_{v_1^0}^{v_1^M} [v + \lambda\varphi(v) - \beta]\, f(v)\mathrm{d}v - (1+\lambda)\rho(v_2^0 - v_1^0)[1 - F(v_1^0)]\right] \\ &- \eta v_1^0 - \alpha^{M+1} = 0. \end{aligned} \tag{B.15}$$

Since $v_2^0 \geq v_1^0$, it follows from Claims 3 and (B.15) that $\int_{v_1^0}^{v_1^M} [v + \lambda\varphi(v)]\, f(v)\mathrm{d}v \geq \beta[1 - F(v_1^0)]$, i.e., $\hat{v}(\beta) = \underline{v}$.

Given β, η and λ, (B.22) and (B.23) define v_2^1 as functions of v_1^1, denoted by g_1 and g_2, respectively. By a similar argument in Claim 6, $g_1'(v) > 1$, and $g_2'(v) < 1$ if $v > \hat{v}(\beta)$ and $g_2(v) \geq v$. Let Δ_3 denote the left-hand side of (B.18) or (B.15), then

$$\begin{aligned} \frac{\partial \Delta_3}{\partial v_1^1} &= (1-\pi)\left[(\beta - v_1^1 - \lambda\varphi(v_1^1)) f(v_1^1) + (1+\lambda)\rho(v_2^1 - v_1^1) f(v_1^1) + (1+\lambda)\rho[1 - F(v_1^1)]\right] - \eta, \\ \frac{\partial \Delta_3}{\partial v_2^1} &= \pi(\beta - v_2^1 - \lambda\varphi(v_2^1)) f(v_2^1) - (1-\pi)(1+\lambda)\rho[1 - F(v_1^1)]. \end{aligned}$$

Clearly, $\partial\Delta_3(v_1, g_2(v_1))/\partial v_2 = 0$ by (B.21). Since $v_2^1 \geq v_1^1$, then $g_2(v) > g_1(v)$ for all $v < v_1^1$. Then $\partial\Delta_3(v_1, g_2(v_1))/\partial v_1 > \Delta_3(v_1, g_1(v_1))/\partial v_1 = 0$ for all $v_1 < v_1^1$. Then

$$0 = \Delta_3(v_1^1, v_2^1) = \Delta_3(v_1^0, v_2^0) + \int_{v_1^0}^{v_1^1} \frac{\partial \Delta_3(v_1, g_2(v_1))}{\partial v_1}\mathrm{d}v_1 > \Delta_3(v_1^0, v_2^0) = 0,$$

a contradiction. Hence, $v_2^0 = \underline{v}$.

Next, we show that $a^1 = 0$. Suppose $a^1 > 0$, then $\alpha^1 = 0$. Then v_2^0 satisfies

$$a^1 \left\{ \pi \left[\beta - v_2^0 - \lambda\varphi(v_2^0)\right] f(v_2^0) - (1-\pi)(1+\lambda)\rho + \sum_{j=1}^{2} \mu^j \right\} + \gamma_2^0 = 0, \tag{B.16}$$

$$\pi \int_{v_2^0}^{v_2^M} [v + \lambda\varphi(v) - \beta] f(v)\mathrm{d}v + (1-\pi) \int_{v_1^0}^{v_1^M} [v + \lambda\varphi(v) - \beta] f(v)\mathrm{d}v - \eta v_1^0 - \alpha^{M+1} = 0. \tag{B.17}$$

By Claims 3, it follows from (B.17) that $\int_{v_1^0}^{v_1^M} [v + \lambda\varphi(v)] f(v)\mathrm{d}v - \beta \geq 0$, i.e., $\hat{v}(\beta) = \underline{v}$. Since $g_2'(v) \leq 1$ if $v \geq \hat{v}(\beta)$ and $g_2(v) \geq v$, and $g_2(v_1^1) = v_2^1 \geq v_1^1$, we have $v_2^0 = g_2(v_1^0) > v_1^0 = \underline{v}$, a contradiction. Hence, $a^1 = 0$.

Let $a^* = a^2$, $v_1^* = v_1^1$, $v_2^* = v_2^1$ and $v_2^{**} = v_2^2$. Let $u_1^* = u(\underline{v}, b_1)$ and $u_2^* = u(\underline{v}, b_2)$. This completes the proof. ■

Proof of Corollary 6. This results holds trivially if the first-best can be achieved. For the rest of the proof, I assume that the first-best can be achieved. Suppose there are two optimal mechanisms (a, p, q) and $(\hat{a}, \hat{p}, \hat{q})$. By Theorem 9, there exist $(u_1^*, u_2^*, a^*, v_1^*, v_2^*, v_2^{**})$ and $(\hat{u}_1^*, \hat{u}_2^*, \hat{a}^*, \hat{v}_1^*, \hat{v}_2^*, \hat{v}_2^{**})$ that characterize the two different optimal mechanisms, respectively.

First, I show that the convex combination of the two mechanisms $(\kappa a + (1-\kappa)\hat{a}, \kappa p + (1-\kappa)\hat{p}, \kappa + (1-\kappa)\hat{q})$, where $\kappa \in (0, 1)$, is also optimal. Clearly, it satisfies (IR), (BC), (BB) and (S):

$$[\kappa a(t) + (1-\kappa)\hat{a}(t)]v - [\kappa p(t) + (1-\kappa)\hat{p}(t)] = \kappa[a(t)v - p(t)] + (1-\kappa)[\hat{a}(t)v - \hat{p}(t)] \geq 0,$$

$$\kappa p(t) + (1-\kappa)\hat{p}(t) \leq b,$$

$$\mathbb{E}_t\left[\kappa p(t) + (1-\kappa)\hat{p}(t) - [\kappa q(t) + (1-\kappa)\hat{q}(t)]k\right] = \kappa\mathbb{E}_t[p(t) - q(t)k] + (1-\kappa)\mathbb{E}_t[\hat{p}(t) - \hat{q}(t)k] \geq 0,$$

$$\mathbb{E}_t[\kappa a(t) + (1-\kappa)\hat{a}(t)] = \kappa\mathbb{E}_t[a(t)] + (1-\kappa)\mathbb{E}[\hat{a}(t)] \leq S.$$

It satisfies (IC-v) since $\kappa a(v,b)+(1-\kappa)\hat{a}(v,b)$ is non-decreasing in v and

$$
\begin{aligned}
&\kappa p(v,b)+(1-\kappa)\hat{p}(v,b)\\
=&\kappa\left[a(v,b)v-\int_{\underline{v}}^{v}a(\nu,b)\mathrm{d}\nu-u(\underline{v},b)\right]+(1-\kappa)\left[\hat{a}(v,b)v-\int_{\underline{v}}^{v}\hat{a}(\nu,b)\mathrm{d}\nu-\hat{u}(\underline{v},b)\right]\\
=&\left[\kappa a(v,b)+(1-\kappa)\right]v-\int_{\underline{v}}^{v}[\kappa a(\nu,b)+(1-\kappa)\hat{a}(\nu,b)]\mathrm{d}\nu-[\kappa u(\underline{v},b)+(1-\kappa)\hat{u}(\underline{v},b)].
\end{aligned}
$$

Finally, it satisfies (IC-b) since

$$
\begin{aligned}
&[\kappa a(v,b_2)+(1-\kappa)\hat{a}(v,b_2)]v-[\kappa p(v,b_2)+(1-\kappa)\hat{p}(v,b_2)]\\
=&\kappa[a(v,b_2)v-p(v,b_2)]+(1-\kappa)[\hat{a}(v,b_2)v-\hat{p}(v,b_2)]\\
\geq&\kappa[a(\hat{v},b_1)v-p(\hat{v},b_1)-q(\hat{v},b_1)c]+(1-\kappa)[\hat{a}(\hat{v},b_1)v-\hat{p}(\hat{v},b_1)-\hat{q}(\hat{v},b_1)c]\\
=&[\kappa a(v,b_1)+(1-\kappa)\hat{a}(v,b_1)]v-[\kappa p(\hat{v},b_1)+(1-\kappa)\hat{p}(\hat{v},b_1)]-[\kappa q(\hat{v},b_1)c+(1-\kappa)q(\hat{v},b_1)]c.
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
&\mathbb{E}_t\left[[\kappa a(t)+(1-\kappa)\hat{a}(t)]v-[\kappa q(t)+(1-\kappa)\hat{q}(t)]k\right]\\
=&\kappa\mathbb{E}_t[a(t)v-q(t)k]+(1-\kappa)\mathbb{E}_t[\hat{a}(t)v-\hat{q}(t)k]\\
=&V.
\end{aligned}
$$

Hence, $(\kappa a+(1-\kappa)\hat{a},\kappa p+(1-\kappa)\hat{p},\kappa+(1-\kappa)\hat{q})$ is an optimal mechanism of $\mathcal{P}$.

Second, I show that $v_1^*=\hat{v}_1^*$. Suppose, on the contrary, that $v_1^*<\hat{v}_1^*$. Then

$$
\kappa a(v,b_1)+(1-\kappa)\hat{a}(v,b_1)=\chi_{\{v\geq v_1^*\}}\kappa a^*+\chi_{v\geq\hat{v}_1^*}(1-\kappa)\hat{a}^*,
$$

which is a 3-step function, a contradiction.

Third, I show that $v_2^*=\hat{v}_2^*$, $v_2^{**}=\hat{v}_2^{**}$ and $a^*=\hat{a}^*$. Suppose $a^*=\hat{a}^*=1$. By Proposition 2, (S) holds with equality in an optimal mechanism. Hence, $v_2^*=v_2^{**}=\hat{v}_2^*=\hat{v}_2^{**}$.

Suppose $a^* < 1$ and $\hat{a}^* = 1$. Since (S) holds with equality in both mechanisms, it must be that $v_2^* < \hat{v}_2^*$. In this case, $a(v, b_1) = \chi_{\{v \geq v_1^*\}}[\kappa a^* + (1-\kappa)]$. If $v \in \left(v_2^*, \min\{v_2^{**}, \hat{v}_2^{**}\}\right)$, then $a(v, b_2) = \kappa a^* < \kappa a^* + (1-\kappa)$, which is a contradiction to Lemma 4.

Suppose $a^* < 1$ and $\hat{a}^* < 1$. In this case, $a(v, b_1) = \chi_{\{v \geq v_1^*\}}[\kappa a^* + (1-\kappa)\hat{a}^*]$. Suppose, on the contrary, that $v_2^* < \hat{v}_2^*$. If $v \in \left(v_2^*, \min\{v_2^{**}, \hat{v}_2^*\}\right)$, then $a(v, b_2) = \kappa a^* < \kappa a^* + (1-\kappa)\hat{a}^*$, which is a contradiction to Lemma 4. Hence, $v_2^* = \hat{v}_2^*$. Suppose, on the contrary, that $v_2^{**} < \hat{v}_2^{**}$. If $v \in \left(v_2^{**}, \hat{v}_2^{**}\right)$, then $a(v, b_2) = \kappa + (1-\kappa)\hat{a}^* > \kappa a^* + (1-\kappa)\hat{a}^*$, a contradiction to Lemma 4. Hence, $v_2^{**} = \hat{v}_2^{**}$. Finally, since (S) holds with equality in both mechanisms, it must be the case $a^* = \hat{a}^*$.

Lastly, I show that $u_i^* = \hat{u}_i^*$ for $i = 1, 2$. Proposition 14 shows that if the first-best cannot be achieved then both (BC) and (BB) hold with equality in an optimal mechanism. Hence, $u_1^* = a^* v_1^* - b_1 = \hat{a}^* \hat{v}_1^* - b_1 = \hat{u}_1^*$. If $\rho \geq \pi/(1-\pi)$, then by Proposition 4 $u_2^* = u_1^* = \hat{u}_1^* = \hat{u}_2^*$. If $\rho < \pi/(1-\pi)$, then $u_2^* = \hat{u}_2^*$ by (BB). ■

B.3.3. Proof of Lemma 7

Let $M \geq 3$ be an integer. We want to show that $V(M-1, d) = V(M, d)$. Suppose to the contradiction that $V(M-1, d) < V(M, d)$, then an optimal solution to $\mathcal{P}'(M, d)$ satisfies the first-order conditions given before the proof of Lemma 34 in Appendix 3.4.2.

For later use, we note here that the summation of FOCs of $a^{m'}$, $m+1 \leq m' \leq M$, $m = 0, \ldots, M-1$, gives:

$$\pi \left[\int_{v_2^m}^{v_2^M} \left[(1+\lambda)v - \lambda \frac{1-F(v)}{f(v)} \right] f(v) \mathrm{d}v - \beta [F(v_2^M) - F(v_2^m)] \right]$$
$$+ (1-\pi) \left[\int_{v_1^m}^{v_1^M} \left[(1+\lambda)v - \lambda \frac{1-F(v)}{f(v)} \right] f(v) \mathrm{d}v - (1+\lambda)\rho(v_2^m - v_1^m)[1 - F(v_1^m)] - \beta[1 - F(v_1^m)] \right]$$
$$- \eta v_1^m + (v_2^m - v_1^m) \sum_{j=m+1}^{M} \mu^j + \alpha^{m+1} - \alpha^{M+1} = 0. \qquad \text{(B.18)}$$

Recall that $\alpha^2 = \cdots = \alpha^M = 0$. We break the proof into several claims. In all claims, we assume, without explicitly repeating this, that Assumptions 2 and 3 hold, $u(\underline{v}, b_1)$, $u(\underline{v}, b_2)$, $\{a^m\}_{m=1}^M$, $\{v_1^m\}_{m=1}^{M-1}$ and $\{v_2^m\}_{m=0}^M\}$ define an optimal mechanism of $\mathcal{P}'(M, d)$ and β, η, λ, ξ_1, ξ_2, $\{\mu^m\}_{m=1}^M$, $\{\alpha^m\}_{m=1}^{M+1}$, $\{\gamma_1^m\}_{m=1}^M$ and $\{\gamma_2^m\}_{m=0}^{M+1}$ are the associated Lagrangian multipliers.

Claim 1 $\gamma_2^m = 0$ *for* $m = 2, \ldots, M-1$.

Proof. Since $a^{m+1} > a^m$ for $m = 1, \ldots M-1$, the FOCs of v_1^m become

$$\begin{aligned}&(1-\pi)\left[(\beta - (1+\lambda)v_1^m)f(v_1^m) + (\lambda + \rho + \lambda\rho)\left[1 - F(v_1^m)\right] + (1+\lambda)\rho(v_2^m - v_1^m)f(v_1^m)\right] \\ &- \sum_{j=m+1}^{M} \mu^j - \eta = 0,\end{aligned}$$

for $m = 1, \ldots, M-1$. Then for $m = 1, \ldots, M-1$

$$v_2^m = \frac{1+\rho}{\rho}v_1^m - \frac{\lambda+\rho+\lambda\rho}{(1+\lambda)\rho}\frac{1-F(v_1^m)}{f(v_1^m)} - \frac{\beta}{(1+\lambda)\rho} + \frac{\eta + \sum_{j=m+1}^{M}\mu^j}{(1-\pi)(1+\lambda)\rho f(v_1^m)}, \tag{B.19}$$

which is strictly increasing in v_1^m by Assumptions 2 and 3. Let $m = 1, \ldots, M-2$. If $\mu^{m+1} = 0$, then $v_2^{m+1} > v_2^m$ since $v_1^{m+1} > v_1^m$ and (B.19). If $\mu^{m+1} > 0$, then $v_2^{m+1} \geq v_1^{m+1} > v_1^m \geq v_2^m$ since (3.8) holds for m and $m+2$ and (3.8) holds with equality for $m+1$. Hence, $\gamma_2^m = 0$ for $m = 2, \ldots, M-1$. ■

Let

$$\varphi(v) := v - \frac{1-F(v)}{f(v)},$$

denote the "virtual" value, which is strictly increasing in v by Assumption 2. By Lemmas 33 and 34, we have $v_2^{M-1} > v_1^{M-1}$. In this case, $\mu^M = 0$.

Claim 2 *Suppose* $v_2^{M-1} > v_1^{M-1}$*, then* $\overline{v} + \lambda\varphi(\overline{v}) > \beta \geq v_2^{M-1} + \lambda\varphi(v_2^{M-1})$.

Proof. Since $\mu^M = 0$, the FOC of v_2^{M-1} implies that $\beta \geq v_2^{M-1} + \lambda\varphi(v_2^{M-1})$. Since $v_2^{M-1} > v_1^{M-1}$ and $\mu^M = 0$, the FOC of a^M implies that

$$\pi\int_{v_2^{M-1}}^{v_2^M}\left[v + \lambda\varphi(v) - \beta\right]f(v)\mathrm{d}v + (1-\pi)\int_{v_1^{M-1}}^{v_1^M}\left[v + \lambda\varphi(v) - \beta\right]f(v)\mathrm{d}v \geq 0.$$

Hence, it must be the case that $\beta < \overline{v} + \lambda\varphi(\overline{v})$. ■

Claim 3 *Suppose $v_2^{M-1} > v_1^{M-1}$, then $\gamma_2^M = \gamma_2^{M+1} = 0$ and $v_2^M + \lambda\varphi(v_2^M) \leq \beta$.*

Proof. Suppose $v_2^M + \lambda\varphi(v_2^M) > \beta \geq v_2^{M-1} + \lambda\varphi(v_2^{M-1})$, then $v_2^M > v_2^{M-1}$ and therefore $\gamma_2^M = 0$. Suppose $v_2^M + \lambda\varphi(v_2^M) \leq \beta < \overline{v} + \lambda\varphi(\overline{v})$, then $v_2^M < \overline{v}$ and therefore $\gamma_2^{M+1} = 0$. Since $\gamma_2^{M+1} = 0$ and $v_2^M + \lambda\varphi(v_2^M) \leq \beta$, the FOC of v_2^M implies that $\gamma_2^M = 0$. Hence, $\gamma_2^M = 0$.

Suppose $a^{M+1} > a^M$, then the FOC of v_2^M implies that $\beta \geq v_2^M + \lambda\varphi(v_2^M)$. Suppose $a^{M+1} = a^M$, then by construction $v_2^M = v_2^{M-1}$ and therefore $v_2^M + \lambda\varphi(v_2^M) \leq \beta$. Hence, $v_2^M + \lambda\varphi(v_2^M) \leq \beta < \overline{v} + \lambda\varphi(\overline{v})$, which implies that $v_2^M < \overline{v}$ and therefore $\gamma_2^{M+1} = 0$. ■

In what follows, we consider two cases: $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1\left(v_2^0 - v_1^0\right) > 0$ and $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1\left(v_2^0 - v_1^0\right) = 0$.

Case 1. $v_2^{M-1} > v_1^{M-1}$ and $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1\left(v_2^0 - v_1^0\right) = 0$.

Claim 4 *Suppose $v_2^{M-1} > v_1^{M-1}$ and $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1\left(v_2^0 - v_1^0\right) = 0$, then $\gamma_2^1 = 0$.*

Proof. Suppose $\gamma_2^0 > 0$, then $v_2^0 = \underline{v}$. Since (3.8) holds for $m = 2$, we have $v_2^1 \geq v_1^1 > \underline{v} = v_2^0$. Hence, $\gamma_2^1 = 0$.

Suppose $\gamma_2^0 = 0$. Suppose $a^1 = 0$, then the FOC of v_2^0 implies that $\gamma_2^1 = 0$. Suppose $a^1 > 0$. Suppose to the contradiction that $\gamma_2^1 > 0$, then we can construct another feasible mechanism (a^*, p^*, q^*), which strictly improves welfare. Since $\gamma_2^1 > 0$, we have $v_2^0 = v_2^1 \geq v_1^1$. We consider two different cases: (1) $v_2^0 = v_2^1 = v_1^1$ and (2) $v_2^0 = v_2^1 > v_1^1$.

Suppose $v_2^0 = v_2^1 = v_1^1$.

Let $\tilde{v}_1^1$ be such that $a^2(v_1^1 - \tilde{v}_1^1) = a^1(v_1^1 - \underline{v})$. Then, by Assumption 3, we have

$$\begin{aligned} a^2\left[F(v_1^1) - F(\tilde{v}_1^1)\right] &= (a^2 - a^1 + a^1)\left[F(v_1^1) - F(\tilde{v}_1^1)\right] \\ &\leq a^1\left[F(v_1^1) - F(\tilde{v}_1^1)\right] + (a^2 - a^1) f(\tilde{v}_1^1)(v_1^1 - \tilde{v}_1^1) \\ &= a^1\left[F(v_1^1) - F(\tilde{v}_1^1)\right] + a^1 f(\tilde{v}_1^1)\tilde{v}_1^1 \\ &\leq a^1 F(v_1^1). \end{aligned}$$

Let $\tilde{v}_2^0 = \underline{v}$ and $\tilde{v}_2^1$ be such that $\pi\left[F(v_2^1) - F(\tilde{v}_2^1)\right] = (1-\pi)\left[a^1 F(v_1^1) - a^2\left[F(v_1^1) - F(\tilde{v}_1^1)\right]\right]$. Let $\tilde{v}_1^m = v_1^m$ and $\tilde{v}_2^m = v_2^m$ for all $m \geq 1$. Let $a^*(v, b_i) = a^m$ if $v \in (\tilde{v}_i^{m-1}, \tilde{v}_i^m)$ for $m \geq 2$ and $i = 1, 2$ and $a^*(v, b_i) = 0$ if $v \in (\underline{v}, \tilde{v}_i^1)$ for $i = 1, 2$. Let $p^*(v, b) = v a^*(v, b) - \int_{\underline{v}}^{v} a^*(\nu, b)\mathrm{d}\nu - u(\underline{v}, b)$. Then, by construction, $p^*(\overline{v}, b_1) = p(\overline{v}, b_1) \leq b_1$. Hence, the (BC) constraint is satisfied. Let $q^*(v, b_1) = q(v, b_1)$. By Assumption 2, (a^*, p^*, q^*) improves revenue and therefore satisfies the (BB) constraint. Clearly, (a^*, p^*, q^*) satisfies constraints (IR), (IC-v) and (S), and strictly improves welfare.

Finally, we show that (a^*, p^*, q^*) satisfies the (IC-b) constraint. For $v \in (\underline{v}, \tilde{v}_1^1)$, we have

$$u(\underline{v}, b_1) - u(\underline{v}, b_2) < u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1\left(v_2^0 - v_1^0\right) = q^1 c.$$

For $v \in (\tilde{v}_1^1, v_1^1)$, we have

$$u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^2(\tilde{v}_2^1 - \tilde{v}_1^1) \leq u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1\left(v_2^0 - v_1^0\right) = q^1 c.$$

The first inequality holds since $a^2(\tilde{v}_2^1 - \tilde{v}_1^1) \leq a^2(v_1^1 - \tilde{v}_1^1) = a^1(v_1^1 - \underline{v}) = a^1\left(v_2^0 - v_1^0\right)$. For

$v \in (v_1^{m-1}, v_1^{m-2})$, $m \geq 2$, we have

$$\begin{aligned}
&u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^2(\tilde{v}_2^1 - \tilde{v}_1^1) + \sum_{j=3}^{m}(a^j - a^{j-1})(\tilde{v}_2^{j-1} - \tilde{v}_1^{j-1}) \\
=&u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) + a^2(\tilde{v}_2^1 - \tilde{v}_1^1) \\
&- (a^2 - a^1)(v_2^1 - v_1^1) - a^1\left(v_2^0 - v_1^0\right) \\
=&u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) + a^2\tilde{v}_2^1 - a^2\tilde{v}_1^1 - a^2 v_2^1 + (a^2 - a^1)v_1^1 + a^1 v_1^0 \\
\leq&u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) = q^m c,
\end{aligned}$$

where the last inequality holds by construction. Hence, the (IC-b) constraint is satisfied. Thus, (a^*, p^*, q^*) is feasible. However, this contradicts to that (a, p, q) is optimal.

Suppose $v_2^0 = v_2^1 > v_1^1$.

Let $a^*(v, b_1) = a^1 - \varepsilon$ for some $\varepsilon > 0$ sufficiently small if $v < v_1^1$ and $a^*(v, b_1) = a(v, b_1)$ otherwise. Let $\tilde{v}_2^0 < v_2^0$ be such that $\pi(a^1 - \varepsilon)\left[F(v_2^0) - F(\tilde{v}_2^0)\right] = (1 - \pi)\varepsilon F(v_1^1)$. For $\varepsilon > 0$ sufficiently small, $v_1^1 < \tilde{v}_2^0$. Let $\tilde{v}_2^m = v_2^m$ for $m \geq 1$. Let $a^*(v, b_2) = a^1 - \varepsilon$ if $v \in (\tilde{v}_2^0, \tilde{v}_2^1)$ and $a^*(v, b_2) = a(v, b_2)$ otherwise. Let $u^*(\underline{v}, b_1) = u(\underline{v}, b_1) + \varepsilon(v_1^1 - v_1^0)$ and $u^*(\underline{v}, b_2) = u(\underline{v}, b_2) - (1 - \pi)\varepsilon(v_1^1 - v_1^0)/\pi$. For $\varepsilon > 0$ sufficiently small, $u^*(\underline{v}, b_2) \geq u^*(\underline{v}, b_1) > 0$. Let $p^*(v, b) = va^*(v, b) - \int_{\underline{v}}^{v} a(\nu, b)\mathrm{d}\nu - u(\underline{v}, b)$. Then, by construction, we have $p^*(\overline{v}, b_1) = p(\overline{v}, b_1) \leq b_1$. Hence, the (BC) constraint is satisfied. Let $q^*(v, b_1) = q(v, b_1)$. Then (a^*, p^*, q^*) satisfies (BB) by Assumption 2. Clearly, (a^*, p^*, q^*) satisfies constraints (IR), (IC-v) and (S), and strictly improves welfare.

Finally, we show that (a^*, p^*, q^*) satisfies the (IC-b) constraint. Note that, by Assumption 3, we have

$$
\begin{aligned}
(a^1-\varepsilon)(v_2^0-\tilde{v}_2^0) &= (a^1-\varepsilon)\int_{\tilde{v}_2^0}^{v_2^0} f(v)\frac{1}{f(v)}\mathrm{d}v \\
&\geq (a^1-\varepsilon)\frac{1}{f(\tilde{v}_2^0)}[F(v_2^0)-F(\tilde{v}_2^0)] \\
&\geq \frac{1-\pi}{\pi}\varepsilon\frac{1}{f(v_1^1)}F(v_1^1) \\
&\geq \frac{1-\pi}{\pi}\varepsilon(v_1^1-v_1^0).
\end{aligned}
$$

Then, for $v < v_1^1$, we have

$$
\begin{aligned}
&u^*(\underline{v},b_1)-u^*(\underline{v},b_2)+(a^1-\varepsilon)(\tilde{v}_2^0-v_1^0) \\
=&u(\underline{v},b_1)-u(\underline{v},b_2)+a^1v_2^0+\frac{\varepsilon(v_1^1-v_1^0)}{\pi}+(a^1-\varepsilon)(\tilde{v}_2^0-v_2^0)-\varepsilon v_2^0-(a^1-\varepsilon)v_1^0 \\
\leq&u(\underline{v},b_1)-u(\underline{v},b_2)+a^1v_2^0+\frac{\varepsilon(v_1^1-v_1^0)}{\pi}-\frac{(1-\pi)\varepsilon(v_1^1-v_1^0)}{\pi}-\varepsilon v_2^0-(a^1-\varepsilon)v_1^0 \\
\leq&u(\underline{v},b_1)-u(\underline{v},b_2)+a^1\left(v_2^0-v_1^0\right)+\varepsilon(v_1^1-v_2^0) \\
<&u(\underline{v},b_1)-u(\underline{v},b_2)+a^1\left(v_2^0-v_1^0\right)=q^1c.
\end{aligned}
$$

For $v \in (v_1^{m-1}, v_1^m)$ for $m = 2, \ldots, M$, we have

$$
\begin{aligned}
&u^*(\underline{v},b_1)-u^*(\underline{v},b_2)+(a^1-\varepsilon)(\tilde{v}_2^0-v_1^0)+(a^2-a^1+\varepsilon)(v_2^1-v_1^1) \\
&+\sum_{j=3}^{m}(a^j-a^{j-1})(\tilde{v}_2^{j-1}-\tilde{v}_1^{j-1}) \\
\leq&u(\underline{v},b_1)-u(\underline{v},b_2)+\sum_{j=1}^{m}(a^j-a^{j-1})(v_2^{j-1}-v_1^{j-1})+\varepsilon(v_1^1-v_2^0)+\varepsilon(v_2^1-v_1^1), \\
=&u(\underline{v},b_1)-u(\underline{v},b_2)+\sum_{j=1}^{m}(a^j-a^{j-1})(v_2^{j-1}-v_1^{j-1})=q^mc.
\end{aligned}
$$

Hence, the (IC-b) constraint is satisfied. Thus, (a^*, p^*, q^*) is feasible. However, this contradicts to that (a, p, q) is optimal.

Hence, it must be that $\gamma_2^1 = 0$. ■

By Claims 1, 3 and 4, we have $\gamma_2^m = 0$ for $m = 1, \ldots, M+1$. Thus, for $m = 1, \ldots, M-1$, v_1^m and v_2^m satisfy

$$(1-\pi)\left[(\beta - v_1^m - \lambda\varphi(v_1^m))f(v_1^m) + (1+\lambda)\rho[1-F(v_1^m)] + (1+\lambda)\rho(v_2^m - v_1^m)f(v_1^m)\right] \\ - \sum_{j=m+1}^{M}\mu^j - \eta = 0, \tag{B.20}$$

$$\pi(\beta - v_2^m - \lambda\varphi(v_2^m))f(v_2^m) - (1-\pi)(1+\lambda)\rho[1-F(v_1^m)] + \sum_{j=m+1}^{M}\mu^j = 0. \tag{B.21}$$

Recall that (B.20) and (B.21) are the first-order conditions of v_1^m and v_2^m, respectively, for $m = 1, \ldots, M-1$.

Claim 5 *Suppose* $v_2^{M-1} > v_1^{M-1}$ *and* $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1\left(v_2^0 - v_1^0\right) = 0$, *then* $\mu^m = 0$ *for* $m = 3, \ldots, M$.

Proof. The result follows directly from Lemmas 33 and 34. ■

Claim 6 *Suppose* $v_2^{M-1} > v_1^{M-1}$ *and* $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1\left(v_2^0 - v_1^0\right) = 0$, *then* $M \leq 2$.

Proof. Let $\hat{m} = 1$ if $\mu^2 = 0$ and $\hat{m} = 2$ if $\mu^2 > 0$. For $m = \hat{m}, \ldots, M-1$, (B.20) and (B.21) become

$$(1-\pi)\left[(\beta - v_1^m - \lambda\varphi(v_1^m))f(v_1^m) + (1+\lambda)\rho[1-F(v_1^m)] + (1+\lambda)\rho(v_2^m - v_1^m)f(v_1^m)\right] - \eta = 0, \tag{B.22}$$

$$\pi(\beta - v_2^m - \lambda\varphi(v_2^m))f(v_2^m) - (1-\pi)(1+\lambda)\rho[1-F(v_1^m)] = 0, \tag{B.23}$$

Given β, η and λ, (B.22) and (B.23) define v_2^m as functions of v_1^m, denoted by g_1 and g_2, respectively. Clearly, by Assumptions 2 and 3, $g_1'(v_1^m) > 1$. Since $\mu^{\hat{m}} > 0$, (3.8) holds by equality for $\hat{m}$, which implies that $v_2^{\hat{m}} \geq v_1^{\hat{m}}$. Furthermore, since $g_1'(v_1^m) > 1$, $v_2^m \geq v_1^m$ for all $\hat{m} \leq m \leq M-1$. Since $v + \lambda\varphi(v) < \beta$ for all $v < v_2^M$, $v_2^m \geq v_1^m \geq \underline{v} \geq 0$, $\sum_{j=m+1}^{M}\mu^j = 0$, $\eta \geq 0$, $\alpha^{m+1} = 0$ and $\alpha^{M+1} \geq 0$, (B.18) implies that

$$\int_{v^m}^{\overline{v}} [v + \lambda\varphi(v) - \beta] f(v)\mathrm{d}v \geq 0,$$

which holds if and only if $v^m \geq \hat{v}(\beta)$, where

$$\hat{v}(\beta) := \inf \left\{ \hat{v} \,\middle|\, \int_{v^m}^{\overline{v}} [v + \lambda\varphi(v) - \beta] f(v)\mathrm{d}v \geq 0 \right\}.$$

By the implicit function theorem, we have

$$g_2'(v_1^m) = \frac{1-\pi}{\pi} \frac{(1+\lambda)\rho f(v_1^m)}{-(\beta - (1+\lambda)v_2^m) f'(v_2^m) + (1+2\lambda) f(v_2^m)} > 0. \tag{B.24}$$

To see that the last inequality holds, note that $(\beta - v - \lambda\varphi(v))f(v)$ is strictly decreasing in v for $v < v_2^M$. Taking derivative with respect to v yields $(\beta - (1+\lambda)v)f'(v) - (1+2\lambda)f(v) < 0$ for $v < v_2^M$. Note that Assumption 2 implies that for all $v \geq v_1^m$, we have

$$f(v) \geq f(v_1^m) \frac{1 - F(v)}{1 - F(v_1^m)}. \tag{B.25}$$

Then for $v_1^m \geq \hat{v}(\beta)$ we have

$$\begin{aligned}
1 - F(v_1^m) &\geq \frac{f(v_1^m)}{1 - F(v_1^m)} \int_{v_1^m}^{\overline{v}} (1 - F(v))\mathrm{d}v \\
&= \frac{f(v_1^m)}{1 - F(v_1^m)} \left[(1+\lambda) \int_{v_1^m}^{\overline{v}} (1 - F(v))\mathrm{d}v - \lambda \int_{v_1^m}^{\overline{v}} (1 - F(v))\mathrm{d}v \right] \\
&= \frac{f(v_1^m)}{1 - F(v_1^m)} \left[-(1+\lambda)v_1^m [1 - F(v_1^m)] + \int_{v_1^m}^{\overline{v}} \left[(1+\lambda)v - \lambda \frac{1 - F(v)}{f(v)} \right] f(v)\mathrm{d}v \right] \\
&\geq (\beta - (1+\lambda)v_1^m) f(v_1^m),
\end{aligned}$$

where the first line holds by (B.25), the third line holds by integration by parts, and the last line holds since $v_1^m \geq \hat{v}(\beta)$. Combining this and (B.23) yields

$$\begin{aligned}(\beta - v_2^m - \lambda\varphi(v_2^m))f(v_2^m) &= \frac{1-\pi}{\pi}(1+\lambda)\rho[1-F(v_1^m)] \\ &= \frac{1-\pi}{\pi}\rho\left[[1-F(v_1^m)] + \lambda[1-F(v_1^m)]\right] \\ &\geq \frac{1-\pi}{\pi}\rho\left[(\beta-(1+\lambda)v_1^m)f(v_1^m) + \lambda[1-F(v_1^m)]\right] \\ &= \frac{1-\pi}{\pi}\rho\left[\beta - v_1^m - \lambda\varphi(v_1^m)\right]f(v_1^m).\end{aligned}$$

Hence,

$$\frac{\rho f(v_1^m)}{f(v_2^m)} \leq \frac{\pi}{1-\pi}\frac{\beta - v_2^m - \lambda\varphi(v_2^m)}{\beta - v_1^m - \lambda\varphi(v_1^m)}.$$

Furthermore,

$$\begin{aligned}&-(\beta-(1+\lambda)v_2^m)f'(v_2^m) + (1+2\lambda)f(v_2^m) \\ =&-(\beta - v_2^m - \lambda\varphi(v_2^m))f'(v_2^m) + \lambda\left\{\frac{[1-F(v_2^m)]f'(v_2^m)}{f(v_2^m)} + f(v_2^m)\right\} + (1+\lambda)f(v_2^m) \\ \geq&(1+\lambda)f(v_2^m),\end{aligned}$$

where the last inequality holds since $\beta - v_2^m - \lambda\varphi(v_2^m) > 0$, $f' \leq 0$ by Assumption 3 and $[1 - F(v_2^m)]f'(v_2^m) + f^2(v_2^m) \geq 0$ by Assumption 2. Finally, since $v_2^m \geq v_1^m \geq \hat{v}(\beta)$, we have

$$g_2'(v_1^m) = \frac{1-\pi}{\pi}\frac{(1+\lambda)\rho f(v_1^m)}{-(\beta-(1+\lambda)v_2^m)f'(v_2^m) + (1+2\lambda)f(v_2^m)} \leq \frac{\beta - v_2^m - \lambda\varphi(v_2^m)}{\beta - v_1^m - \lambda\varphi(v_1^m)} \leq 1.$$

Note that $g_2'(v_1^m) < 1$ if $v_1^m > \hat{v}(\beta)$ or $v_1^m < v_2^m$.

Thus, there exists at most one $v_1^m \geq \hat{v}(\beta)$ such that $g_1(v_1^m) = g_2(v_1^m) \geq v_1^m$, i.e., (B.22) and (B.23) has at most one solution such that $v_2^m \geq v_1^m \geq \hat{v}(\beta)$. Hence, $M \leq \hat{m} + 1 \leq 3$.

Suppose $M = 3$. By Claim 3, $v + \lambda\varphi(v) < \beta$ for all $v \leq v_2^M$. Furthermore, $\eta \geq 0$ and $\alpha^{M+1} \geq 0$.

Hence, it follows from (B.18) that

$$\int_{v^1}^{\overline{v}} [v + \lambda\varphi(v) - \beta]\, f(v)\mathrm{d}v \geq 0,$$

i.e., $v^1 \geq \hat{v}(\beta)$. Then we have $v_1^{\hat{m}} > v^{\hat{m}-1} \geq \hat{v}(\beta)$, and $g_2(v_1^{\hat{m}}) = v_2^{\hat{m}} \geq v_1^{\hat{m}}$ since $\mu^{\hat{m}} > 0$. Since $g_2'(v) < 1$ if $v > \hat{v}(\beta)$ and $g_2(v) \geq v$, we have $g_2(v) > v$ for all $v < v_1^{\hat{m}}$. Hence, $v^{\hat{m}-1} = g_2(v^{\hat{m}-1}) > v^{\hat{m}-1}$, a contradiction. Hence, $M = 2$ and $v_2^1 \geq v_1^1$. ■

Case 2. $v_2^{M-1} > v_1^{M-1}$ **and** $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1 \left(v_2^0 - v_1^0\right) > 0$**.** In this case, by Lemmas 33 and 34, $\mu^m = 0$ for $m = 1, \ldots, M$.

Claim 7 *Suppose* $v_2^{M-1} > v_1^{M-1}$ *and* $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1 \left(v_2^0 - v_1^0\right) > 0$*, then* $\gamma_2^1 = 0$.

Proof. Since $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1 \left(v_2^0 - v_1^0\right) > 0$, $\mu^1 = 0$. Suppose, on the contrary, that $\gamma_2^1 > 0$. Then $v_2^1 = v_2^0$.

Suppose $\gamma_2^0 > 0$, then $v_2^1 = v_2^0 = \underline{v} = v_1^0$. Hence, $u(\underline{v}, b_1) > u(\underline{v}, b_2)$. Let $\tilde{v}_2^1 = \underline{v} + \varepsilon$ for some $\varepsilon > 0$ sufficiently small. Let $\tilde{v}_1^1$ be such that $\pi F(\varepsilon) = (1-\pi)\left[F(v_1^1) - F(\tilde{v}_1^1)\right]$. For $\varepsilon > 0$ sufficiently small, $\tilde{v}_2^1 < \tilde{v}_1^1$. Let $\tilde{v}_i^m = v_i^m$ and for $i = 1, 2$ and $m \neq 1$. Let $a^*(v, b_2) = a^1$ for all $v \in (\underline{v}, \tilde{v}_2^1)$ and $a^*(v, b_2) = a(v, b_2)$ otherwise. Let $a^*(v, b_1) = a^2$ for $v \in (\tilde{v}_1^1, v_1^1)$ and $a^*(v, b_1) = a(v, b_1)$ otherwise. Let $u^*(\underline{v}, b_1) = u(\underline{v}, b_1) - (a^2 - a^1)(v_1^1 - \tilde{v}_1^1)$ and $u^*(\underline{v}, b_2) = u(\underline{v}, b_2) + \frac{1-\pi}{\pi}(a^2 - a^1)(v_1^1 - \tilde{v}_1^1)$. For $\varepsilon > 0$ sufficiently small, $u^*(\underline{v}, b_1) \geq u^*(\underline{v}, b_2) > 0$. Let $p^*(v, b) = va^*(v, b) - \int_{\underline{v}}^{v} a^*(\nu, b)\mathrm{d}\nu - u^*(\underline{v}, b)$. By construction, $p^*(\overline{v}, b_1) = p(\overline{v}, b_1) \leq b_1$. Hence, the (BC) constraint holds. Let $q^*(v, b_1) = q(v, b_1)$. By Assumption 2, the (BB) constraint holds. For $v \in (\underline{v}, \tilde{v}_1^1)$, (IC-b) holds since

$$u^*(\underline{v}, b_1) - u^*(\underline{v}, b_2) + a^1(\tilde{v}_2^0 - \tilde{v}_1^0) = u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) - \frac{(a^2 - a^1)(v_1^1 - \tilde{v}_1^1)}{\pi} \leq q^1 c.$$

For $v \in (\tilde{v}_1^1, v_1^1)$, (IC-b) holds since

$$
\begin{aligned}
&u^*(\underline{v}, b_1) - u^*(\underline{v}, b_2) + a^1(\tilde{v}_2^0 - \tilde{v}_1^0) + (a^2 - a^1)(\tilde{v}_2^1 - \tilde{v}_1^1) \\
=&u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{2}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) \\
&- \frac{(a^2 - a^1)(v_1^1 - \tilde{v}_1^1)}{\pi} + (a^2 - a^1)(v_1^1 - \tilde{v}_1^1 + \tilde{v}_2^1 - v_2^1) \\
=&u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{2}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) + (a^2 - a^1)(\tilde{v}_2^1 - v_2^1) - \frac{(1-\pi)(a^2 - a^1)(v_1^1 - \tilde{v}_1^1)}{\pi} \\
\leq&u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{2}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) \\
\leq&u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1(v_2^0 - v_1^0) = q^1 c,
\end{aligned}
$$

where the last inequality holds since $v_2^1 = \underline{v} < v_1^1$, and the first inequality holds since by Assumption 3 we have

$$
\begin{aligned}
\tilde{v}_2^1 - v_2^1 &\leq \frac{F(\varepsilon)}{f(\tilde{v}_2^1)} \\
&\leq \frac{1}{f(\tilde{v}_1^1)} \frac{1-\pi}{\pi} \left[F(v_1^1) - F(\tilde{v}_1^1)\right] \\
&\leq \frac{(1-\pi)(v_1^1 - \tilde{v}_1^1)}{\pi}.
\end{aligned}
$$

For $v \in (v_1^{m-1}, v_1^m)$, $m = 2, \dots, M$, (IC-b) holds since

$$
\begin{aligned}
&u^*(\underline{v}, b_1) - u^*(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(\tilde{v}_2^{j-1} - \tilde{v}_1^{j-1}) \\
=&u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) + (a^2 - a^1)(\tilde{v}_2^1 - v_2^1) - \frac{(1-\pi)(a^2 - a^1)(v_1^1 - \tilde{v}_1^1)}{\pi} \\
\leq&u(\underline{v}, b_1) - u(\underline{v}, b_2) + \sum_{j=1}^{m}(a^j - a^{j-1})(v_2^{j-1} - v_1^{j-1}) = q^m c.
\end{aligned}
$$

Clearly, (a^*, p^*, q^*) also satisfies constraints (IR), (IC-v) and (S), and strictly improves welfare. This contradicts to the optimality of (a, p, q). Hence, $\gamma_2^1 = 0$.

Suppose $\gamma_2^0 = 0$. Suppose $a^1 = 0$, then the FOC of v_2^0 implies that $\gamma_2^1 = 0$. Suppose $a^1 > 0$. Then

$$\begin{aligned}
&\pi(\beta - (1+\lambda)v_2^0)f(v_2^0) + \pi\lambda[1 - F(v_2^0)] \\
\geq &(1-\pi)(1+\lambda)\rho - \sum_{j=2}^{M} \mu^j \\
> &(1-\pi)(1+\lambda)\rho[1 - F(v_1^1)] - \sum_{j=2}^{M} \mu^j \\
\geq &\pi(\beta - (1+\lambda)v_2^1)f(v_2^1) + \pi\lambda[1 - F(v_2^1)].
\end{aligned}$$

Since $(\beta - (1+\lambda)v)f(v) + \lambda[1 - F(v)]$ is strictly decreasing in v when $v + \lambda\varphi(v) < \beta$, we have $v_2^1 > v_2^0$ and therefore $\gamma_2^1 = 0$. ■

By Claims 1, 3 and 7, we have $\gamma_2^m = 0$ for $m = 1, \ldots, M$. Thus, for $m = 1, \ldots, M-1$, v_1^m and v_2^m satisfies (B.22), (B.23) and (B.18).

Claim 8 *Suppose $v_2^{M-1} > v_1^{M-1}$ and $u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^1\left(v_2^0 - v_1^0\right) > 0$, then $M \leq 2$.*

Proof. Suppose, on the contrary, that $M \geq 3$. Then there exists $1 \leq m < M-1$ such that $v_2^m \geq v_1^m$. It follows from (B.18) that $\int_{v_1^m}^{v_1^M} [v + \lambda\varphi(v)]\, f(v)\mathrm{d}v \geq \beta[1 - F(v_1^m)]$, i.e., $v_1^m \geq \hat{v}(\beta)$. Both (v_1^m, v_2^m) and (v_1^{M-1}, v_2^{M-1}) are solutions to (B.22) and (B.23), and satisfy $v_2 \geq v_1 \geq \hat{v}(\beta)$. However, by a similar argument in Claim 6, (B.22) and (B.23) have at most one solution satisfying $v_2 \geq v_1 \geq \hat{v}(\beta)$, a contradiction. Hence, it must be $M \leq 2$. ■

To summarize, we have shown in both cases that $M \leq 2$. However, this contradicts to the assumption that $M \geq 3$. Hence, it must be that $V(M, d) = V(M-1, d)$ for all $M \geq 3$. This completes the proof of Lemma 7.

B.3.4. Continuity

Let $\rho = k/c$. I abuse notation and let $\mathcal{P}'(2, \rho, \pi, S, b_1, d)$ denote the principal's problem $\mathcal{P}'(2, d)$ when verification cost is k, punishment is c, the percentage of high-budget agents is π, supply is S and low-budget agent's budget is b_1. Define $V : R_+ \times (0,1)^2 \times [\underline{v}, \overline{v}] \times \mathbb{R}_+ \to \mathbb{R}_+$ and Γ^* :

$R_+ \times (0,1)^2 \times [\underline{v}, \overline{v}] \times R_+ \to R_+^2 \times [0,1] \times [\underline{v}, \overline{v}]^3$ as follows. Let $V(\rho, \pi, S, b_1, d)$ denote the value of $\mathcal{P}'(2, \rho, \pi, S, b_1, d)$ and $\Gamma^*(\rho, \pi, S, b_1, d)$ the set of optimal solutions.

Lemma 37 *Suppose Assumption 2 holds. Then $V(\rho, \pi, S, b_1, d)$ is continuous and $\Gamma^*(\rho, \pi, S, b_1, d)$ is upper hemicontinuous.*

Proof of Lemma 37. Let correspondence $\Gamma : R_+ \times (0,1)^2 \times [\underline{v}, \overline{v}] \times R_+ \to R_+^2 \times [0,1] \times [\underline{v}, \overline{v}]^3$ be defined as follows. For each (ρ, π, S, b_1, d), let $(u(\underline{v}, b_1), u(\underline{v}, b_2), a^2, v_1^1, v_2^1, v_2^2) \in \Gamma(\rho, \pi, S, b_1, d)$ if and only if it is a feasible solution to $\mathcal{P}(2, d)$. To simplify notation, let $u_1 = u(\underline{v}, b_1)$ and $u_2 = u_2(\underline{v}, b_2)$. Clearly, Γ is compact-valued and upper hemicontinuous. I show that it is also lower hemicontinuous.

Fix (ρ, π, S, b_1, d), $(u_1, u_2, a^2, v_1^1, v_2^1, v_2^2) \in \Gamma(\rho, \pi, S, b_1, d)$ and a sequence

$$(\rho(n), \pi(n), S(n), b_1(n), d(n)) \to (\rho, \pi, S, b_1, d)$$

as $n \to \infty$. Let $\varphi(v) := v - \frac{1-F(v)}{f(v)}$ and r be such that $\varphi(r) = 0$. I show that after taking a subsequence there exist $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \in \Gamma\left(\rho(n), \pi(n), S(n), b_1(n), d(n)\right)$ for all n and $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \to (u_1, u_2, a^2, v_1^1, v_2^1, v_2^2)$ as $n \to \infty$.

Case 1: Suppose $a^2 = 0$. Then (S) and (BB) become:

$$\pi\left[1 - F(v_2^2)\right] \le S,$$
$$-u_1 - [(1-\pi)\rho - \pi](u_1 - u_2) + \pi \int_{v_2^2}^{\overline{v}} \varphi(v) f(v) \mathrm{d}v \ge -d.$$

Case 1.1: Suppose $v_2^2 < r$.

After taking a subsequence, I can assume that for all n, $F^{-1}\left(\frac{\pi(n)-S(n)}{\pi(n)}\right) < r$ and

$$-(1-\pi(n))(1+\rho(n))u_1 + [(1-\pi(n))\rho(n) - \pi(n)]\, u_2 + \pi(n) \int_r^{\overline{v}} \varphi(v) f(v) \mathrm{d}v \ge -d(n).$$

Let $u_1(n) = u_1$, $u_2(n) = u_2$, $a^2(n) = a^2$, $v_1^1(n) = v_1^1$, $v_2^1(n) = v_2^1$ and

$$v_2^2(n) = \inf \left\{ \begin{array}{c} v \geq \max\left\{ v_2^2, F^{-1}\left(\frac{\pi(n)-S(n)}{\pi(n)}\right)\right\}, \\ -(1-\pi(n))(1+\rho(n))u_1 + [(1-\pi(n))\rho(n) - \pi(n)]\, u_2 \\ +\pi(n) \int_v^{\overline{v}} \varphi(\nu) f(\nu) \mathrm{d}\nu \geq -d(n) \end{array} \right\}.$$

Clearly, $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \in \Gamma\left(\rho(n), \pi(n), S(n), b_1(n), d(n)\right)$ for all n and $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \to (u_1, u_2, a^2, v_1^1, v_2^1, v_2^2)$ as $n \to \infty$.

Case 1.2

Suppose $v_2^2 \geq r$. Suppose $v_2^2 = \overline{v}$, then (BB) implies that $u_1 = u_2 = 0$. Let $u_1(n) = u_1$, $u_2(n) = u_2$, $a^2(n) = a^2$, $v_1^1(n) = v_1^1$, $v_2^1(n) = v_2^1$ and $v_2^2(n) = v_2^2$. Clearly,

$$\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \in \Gamma\left(\rho(n), \pi(n), S(n), b_1(n), d(n)\right)$$

for all n and $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \to (u_1, u_2, a^2, v_1^1, v_2^1, v_2^2)$ as $n \to \infty$.

Suppose $v_2^2 < \overline{v}$. After taking a subsequence, I can assume that for all n, $F^{-1}\left(\frac{\pi(n)-S(n)}{\pi(n)}\right) < \overline{v}$. Let $a^2(n) = a^2$, $v_1^1(n) = v_1^1$, $v_2^1(n) = v_2^1$ and

$$v_2^2(n) = \max\left\{ v_2^2, F^{-1}\left(\frac{\pi(n) - S(n)}{\pi(n)}\right)\right\}.$$

For n sufficiently large, $\int_{v_2^2(n)}^{\overline{v}} \varphi(v) f(v) \mathrm{d}v > 0$. If $u_2 > 0$, then let $u_1(n) = u_1 + \min\{\Delta(n), 0\}$ and $u_2(n) = u_2 + \min\{\Delta(n), 0\}$; otherwise let $u_1(n) = u_1 + \min\left\{\frac{\Delta(n)}{(1-\pi(n))(1+\rho(n))}, 0\right\}$ and $u_2(n) = u_2$, where

$$\Delta(n) = -(1-\pi(n))(1+\rho(n))u_1 + [(1-\pi(n))\rho(n) - \pi(n)]\, u_2 + \pi(n) \int_{v_2^2(n)}^{\overline{v}} \varphi(\nu) f(\nu) \mathrm{d}\nu + d(n).$$

Clearly, $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \in \Gamma\left(\rho(n), \pi(n), S(n), b_1(n), d(n)\right)$ for all n and $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \to (u_1, u_2, a^2, v_1^1, v_2^1, v_2^2)$ as $n \to \infty$.

Case 2: Suppose $a^2 > 0$.

Case 2.1: Suppose $v_2^2 < r$.

Let $a^2(n) = \min\left\{\frac{b_1(n)+u_1}{v_1^1}, a^2\right\}$. After taking a subsequence, I can assume that for all n,

$$F^{-1}\left(\frac{\pi(n) - S(n) + (1-\pi)a^2(n)\left[1 - F(v_1^1)\right] - \pi a^2 F(v_2^1)}{1 - \pi(n)a^2(n)}\right) < r$$

and

$$\begin{aligned}
&-(1-\pi(n))(1+\rho(n))u_1 + \left[(1-\pi(n))\rho(n) - \pi(n)\right]u_2 + (1-\pi(n))\int_{v_1^1}^{\overline{v}} a^2(n)\varphi(v)f(v)\mathrm{d}v \\
&-(1-\pi(n))\rho(n)a^2(n)(v_2^1 - v_1^1)\left[1 - F(v_1^1)\right] \\
&+\pi(n)\int_{v_2^1}^{r} a^2(n)\varphi(v)f(v)\mathrm{d}v + \pi(n)\int_{r}^{\overline{v}} \varphi(v)f(v)\mathrm{d}v \geq -d(n).
\end{aligned}$$

Let $u_1(n) = u_1$, $u_2(n) = u_2$, $v_1^1(n) = v_1^1$, $v_2^1(n) = v_2^1$ and

$$v_2^2(n) = \inf\left\{\begin{array}{c}
v \geq \max\left\{v_2^2, F^{-1}\left(\frac{\pi(n)-S(n)+(1-\pi)a^2(n)\left[1-F(v_1^1)\right]-\pi a^2 F(v_2^1)}{1-\pi(n)a^2(n)}\right)\right\}, \\
-(1-\pi(n))(1+\rho(n))u_1 + \left[(1-\pi(n))\rho(n) - \pi(n)\right]u_2 \\
+(1-\pi(n))\int_{v_1^1}^{\overline{v}} a^2(n)\varphi(v)f(v)\mathrm{d}v - (1-\pi(n))\rho(n)a^2(n)(v_2^1 - v_1^1)\left[1 - F(v_1^1)\right] \\
+\pi(n)\int_{v_2^1}^{v} a^2(n)\varphi(v)f(v)\mathrm{d}v + \pi(n)\int_{v}^{\overline{v}} \varphi(v)f(v)\mathrm{d}v \geq -d(n)
\end{array}\right\}.$$

Clearly, $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \in \Gamma\left(\rho(n), \pi(n), S(n), b_1(n), d(n)\right)$ for all n and $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \to (u_1, u_2, a^2, v_1^1, v_2^1, v_2^2)$ as $n \to \infty$.

Case 2.2: Suppose $v_2^2 \geq r$.

Suppose $v_1^1 = \overline{v}$ and $v_2^1 = v_2^2$, then the proof follows that of Case 1.2. Assume for the rest of the proof that $v_1^1 < \overline{v}$ or $v_2^1 < v_2^2$. Let

$$A := (1-\pi)\int_{v_1^1}^{v_1^2} \varphi(v)f(v)\mathrm{d}v - (1-\pi)\rho\int_{v_1^1}^{\overline{v}} (v_2^1 - v_1^1)f(v)\mathrm{d}v + \pi\int_{v_2^1}^{v_2^2} \varphi(v)f(v)\mathrm{d}v.$$

and

$$A(n) :=(1-\pi(n))\int_{v_1^1}^{v_1^2}\varphi(v)f(v)\mathrm{d}v-(1-\pi(n))\rho(n)\int_{v_1^1}^{\overline{v}}(v_2^1-v_1^1)f(v)\mathrm{d}v$$
$$+\pi(n)\int_{v_2^1}^{v_2^2}\varphi(v)f(v)\mathrm{d}v.$$

Suppose $A<0$**.** After taking a subsequence, I can assume that for all n, $S(n)-\pi(n)\left[1-F(v_2^2)\right]>0$, $A(n)<0$ and

$$-(1-\pi(n))(1+\rho(n))u_1+[(1-\pi(n))\rho(n)-\pi(n)]\,u_2+\pi(n)\int_{v_2^2}^{\overline{v}}\varphi(v)f(v)\mathrm{d}v>-d(n).$$

Let $u_1(n)=u_1$, $u_2(n)=u_2$, $v_1^1(n)=v_1^1$, $v_2^1(n)=v_2^1$, $v_2^2(n)=v_2^2$ and

$$a^2(n)=\min\left\{\begin{array}{c} a^2,\frac{S(n)-\pi(n)\left[1-F(v_2^2)\right]}{\pi(n)[F(v_2^2)-F(v_2^1)]+(1-\pi(n))[1-F(v_1^1)]},\frac{b_1(n)+u_1}{v_1^1},\\ \frac{-(1-\pi(n))(1+\rho(n))u_1+[(1-\pi(n))\rho(n)-\pi(n)]u_2+\pi(n)\int_{v_2^2}^{\overline{v}}\varphi(v)f(v)\mathrm{d}v+d(n)}{-A(n)}\end{array}\right\}.$$

Clearly, $\left(u_1(n),u_2(n),a^2(n),v_1^1(n),v_2^1(n),v_2^2(n)\right)\in\Gamma\left(\rho(n),\pi(n),S(n),b_1(n),d(n)\right)$ for all n and $\left(u_1(n),u_2(n),a^2(n),v_1^1(n),v_2^1(n),v_2^2(n)\right)\to(u_1,u_2,a^2,v_1^1,v_2^1,v_2^2)$ as $n\to\infty$.

Suppose $A=0$**.**

Suppose $d=0$ and $v_2^2=\overline{v}$. Then $u_1=u_2=0$ and $v_2^1\geq v_1^1$. Define

$$g_n(v):=\begin{cases}(1-\pi(n))\int_v^{\overline{v}}\varphi(v)f(v)\mathrm{d}v-(1-\pi(n))\rho(n)(v_2^1-v)[1-F(v)]+\pi(n)\int_{v_2^1}^{\overline{v}}\varphi(v)f(v)\mathrm{d}v & \text{if } v<v_2^1\\ \int_v^{\overline{v}}\varphi(v)f(v)\mathrm{d}v & \text{if } v\geq v_2^1\end{cases}.$$

Then

$$g_n'(v):=\begin{cases}-(1-\pi(n))\varphi(v)f(v)+(1-\pi(n))\rho(n)(v_2^1-v)f(v)+(1-\pi(n))\rho(n)[1-F(v)] & \text{if } v<v_2^1\\ -\varphi(v)f(v) & \text{if } v\geq v_2^1\end{cases}.$$

Let g_∞ and g_∞' denote the case in which $\pi(n)=\pi$ and $\rho(n)=\rho$.

Suppose $v_1^1 < v_2^1$. Then

$$g_n(v_2^1) = \int_{v_2^1}^{\overline{v}} \varphi(v) f(v) \mathrm{d}v > 0.$$

Let $u_1(n) = u_1$, $u_2(n) = u_2$, $v_2^1(n) = v_2^1$, $v_2^2(n) = v_2^2$,

$$\begin{aligned} v_1^1(n) &= \inf \left\{ v \geq v_1^1 \,\middle|\, g_n(v) \geq 0 \right\} < v_2^1, \\ \text{and } a^2(n) &= \min \left\{ a^2, \frac{S(n) - \pi(n)\left[1 - F(v_2^2)\right]}{\pi(n)[F(v_2^2) - F(v_2^1)] + (1 - \pi(n))[1 - F(v_1^1(n))]}, \frac{b_1(n)}{v_1^1(n)} \right\}. \end{aligned}$$

If $v \in (v_1^1, v_2^1)$, then $g_n'(v)/f(v)$ is strictly decreasing. Since $g_\infty(v_1^1) = A = 0$ and $g_\infty(v_2^1) > 0$, $g_\infty(v) > 0$ for all $v \in (v_1^1, v_2^1)$. Hence, $v_1^1(n) \to v_1^1$ as $n \to \infty$. Clearly,

$$\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \in \Gamma\left(\rho(n), \pi(n), S(n), b_1(n), d(n)\right)$$

for all n and $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \to (u_1, u_2, a^2, v_1^1, v_2^1, v_2^2)$ as $n \to \infty$.

Suppose $v_1^1 = v_2^1 < \overline{v}$. Let $u_1(n) = u_1$, $u_2(n) = u_2$, $v_1^1(n) = v_1^1$, $v_2^1(n) = v_2^1$, $v_2^2(n) = v_2^2$ and

$$a^2(n) = \min \left\{ a^2, \frac{S(n) - \pi(n)\left[1 - F(v_2^2)\right]}{\pi(n)[F(v_2^2) - F(v_2^1)] + (1 - \pi(n))[1 - F(v_1^1(n))]}, \frac{b_1(n)}{v_1^1(n)} \right\}.$$

Clearly, $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \in \Gamma\left(\rho(n), \pi(n), S(n), b_1(n), d(n)\right)$ for all n and $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \to (u_1, u_2, a^2, v_1^1, v_2^1, v_2^2)$ as $n \to \infty$.

Suppose $v_2^2 < \overline{v}$ or $d > 0$. After taking a subsequence, I can assume there exists $\varepsilon > 0$ such that for all n, $S(n) - \pi(n)[1 - F(n)] > \varepsilon$, $b_1(n) > \varepsilon$ and

$$\int_{v_2^2}^{\overline{v}} \varphi(v) f(v) \mathrm{d}v + d(n) > \varepsilon.$$

Note that $(0,0,0,v_1^1,v_2^1,v_2^2) \in \Gamma(\rho(n),\pi(n),S(n),b_1(n),d(n))$. Define $\kappa(n) \in (0,1]$ as follows:

$$\kappa(n) = \sup\left\{ \kappa \leq 1 \;\middle|\; \begin{array}{c} \pi(n)\kappa a^2[F(v_2^2) - F(v_2^1)] + \pi(n)[1 - F(v_2^2)] \\ +(1-\pi(n))\kappa a^2[1 - F(v_1^1)] \geq S(n), \\ \kappa a^2 v_1^1 - \kappa u_1 \leq b_1(n), \\ -(1-\pi(n))(1+\rho(n))\kappa u_1 + [(1-\pi(n))\rho(n) - \pi(n)]\,\kappa u_2 \\ +(1-\pi(n))\int_{v_1^1}^{\overline{v}} \kappa a^2 \varphi(v) f(v)\mathrm{d}v \\ -(1-\pi(n))\rho(n)\kappa a^2(v_2^1 - v_1^1)\left[1 - F(v_1^1)\right] \\ +\pi(n)\int_{v_2^1}^{v_2^2} \kappa a^2 \varphi(v) f(v)\mathrm{d}v + \pi(n)\int_{v_2^2}^{\overline{v}} \varphi(v) f(v)\mathrm{d}v \geq -d(n) \end{array} \right\}.$$

Since at $(0,0,0,v_1^1,v_2^1,v_2^2)$ constraints (S), (BC) and (BB) hold with strict inequality by a gap at least ε, it is not hard to see that $\kappa(n) \to 1$ as $n \to \infty$. Let $u_1(n) = \kappa(n)u_1$, $u_2(n) = \kappa(n)u_2$, $a^2(n) = \kappa(n)a^2$, $v_1^1(n) = v_1^1$, $v_2^1(n) = v_2^1$ and $v_2^2(n) = v_2^2$. Clearly,

$$\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \in \Gamma\left(\rho(n), \pi(n), S(n), b_1(n), d(n)\right)$$

for all n, and $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \to (u_1, u_2, a^2, v_1^1, v_2^1, v_2^2)$ as $n \to \infty$.

Suppose $A > 0$. After taking a subsequence, I can assume that there exists $\varepsilon > 0$ such that for all n, $S(n) - \pi(n)\left[1 - F(v_2^2)\right] > \varepsilon$, $b_1(n) > \varepsilon$ and $A(n) > \varepsilon$.

Suppose $v_2^2 = \overline{v}$ and $d = 0$. Let

$$\hat{a}^2 = \min\left\{ a^2, \frac{\varepsilon}{2\left[F(v_2^2) - F(v_2^1) + 1 - F(v_1^1)\right]}, \frac{\varepsilon}{2v_1^1} \right\} > 0.$$

Then $(0, 0, \hat{a}^2, v_1^1, v_2^1, v_2^2) \in \Gamma(\rho(n), \pi(n), S(n), b_1(n), d(n))$. Define $\kappa(n) \in (0, 1]$ as follows:

$$\kappa(n) = \sup\left\{ \kappa \leq 1 \;\middle|\; \begin{array}{c} \pi(n)[\kappa a^2 + (1-\kappa)\hat{a}^2][F(v_2^2) - F(v_2^1)] + \pi(n)[1 - F(v_2^2)] \\ +(1-\pi(n))[\kappa a^2 + (1-\kappa)\hat{a}^2][1 - F(v_1^1)] \geq S(n), \\ [\kappa a^2 + (1-\kappa)\hat{a}^2]v_1^1 - \kappa u_1 \leq b_1(n), \\ -(1-\pi(n))(1+\rho(n))\kappa u_1 + [(1-\pi(n))\rho(n) - \pi(n)]\,\kappa u_2 \\ +(1-\pi(n)) \int_{v_1^1}^{\overline{v}} [\kappa a^2 + (1-\kappa)\hat{a}^2]\varphi(v) f(v) \mathrm{d}v \\ -(1-\pi(n))\rho(n)[\kappa a^2 + (1-\kappa)\hat{a}^2](v_2^1 - v_1^1)\left[1 - F(v_1^1)\right] \\ +\pi(n) \int_{v_2^1}^{v_2^2} [\kappa a^2 + (1-\kappa)\hat{a}^2]\varphi(v) f(v) \mathrm{d}v \\ +\pi(n) \int_{v_2^2}^{\overline{v}} \varphi(v) f(v) \mathrm{d}v \geq -d(n) \end{array} \right\}.$$

Since at $(0, 0, \hat{a}^2, v_1^1, v_2^1, v_2^2)$ constraints (S), (BC) and (BB) hold with strict inequality by a gap at least $\min\{\hat{a}^2\varepsilon, \varepsilon/2\}$, it is not hard to see that $\kappa(n) \to 1$ as $n \to \infty$. Let $u_1(n) = \kappa(n)u_1$, $u_2(n) = \kappa(n)u_2$, $a^2(n) = \kappa(n)a^2$, $v_1^1(n) = v_1^1$, $v_2^1(n) = v_2^1$ and $v_2^2(n) = v_2^2$. Clearly, $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \in \Gamma\left(\rho(n), \pi(n), S(n), b_1(n), d(n)\right)$ for all n, and $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \to (u_1, u_2, a^2, v_1^1, v_2^1, v_2^2)$ as $n \to \infty$.

Suppose $v_2^2 < \overline{v}$ or $d > 0$. Then by a similar argument to that of $A = 0$, there exist $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \in \Gamma\left(\rho(n), \pi(n), S(n), b_1(n), d(n)\right)$ for all n, and $\left(u_1(n), u_2(n), a^2(n), v_1^1(n), v_2^1(n), v_2^2(n)\right) \to (u_1, u_2, a^2, v_1^1, v_2^1, v_2^2)$ as $n \to \infty$.

Hence, Γ is hemicontinuous. By Berge's Maximum Theorem, V is continuous and Γ^* is upper hemicontinuous. ■

B.4. Properties of the optimal mechanism

Let $a^* = a^2$, $v_1^* = v_1^1$, $v_2^* = v_2^1$, $v_2^{**} = v_2^2$, $u_1^* = u(\underline{v}, b_1)$ and $u_2^* = u(\underline{v}, b_2)$ denote an solution to $\mathcal{P}'(2, 0)$. Let $\beta, \eta, \lambda, \mu^1, \mu^2, \alpha^3, \xi_1$ and ξ_2 denote the corresponding Lagrangian multipliers.

Proof of Proposition 3. First-best is achieved if the allocation rule satisfies $v^* := v_1^* = v_2^* =$

$F^{-1}(1-S)$ and $a^* = 1$, and verification is zero. Hence, $u_1^* = u_2^* = v^* - b_1$ and (BB) holds if and only if

$$b_1 - v^* F(v^*) \geq 0. \tag{B.26}$$

Since $v^* = F^{-1}(1-S)$, there exists $\hat{S}(b_1) < 1$ such that (B.26) holds if and only if $S \geq \hat{S}(b_1)$. Clearly, $\hat{S}(b_1)$ is strictly decreasing in b_1. ■

Proof of Proposition 2. Let $S' := (1-\pi)a^* \left[1 - F(v_1^*)\right] + \pi a^*[F(v_2^{**}) - F(v_2^*)] + \pi[1 - F(v_2^{**})]$. Suppose to the contradiction that $S' < S$. Let $\kappa \in (0,1)$ be such that $\kappa + (1-\kappa)S' = S$. Consider a new mechanism (a^*, p^*, q^*). Let $a^*(v,b) = \kappa + (1-\kappa)a(v,b)$ and $p^*(v,b) = va^*(v,b) - \int_{\underline{v}}^{v} a^*(\nu, b)\mathrm{d}\nu - (1-\kappa)u(\underline{v}, b)$ for all v and b. Finally, let $q(v, b_2) = 0$ for all v, $q(v, b_1) = (1-\kappa)\left[u(\underline{v}, b_1) - u(\underline{v}, b_2)\right]/c$ if $v < v_1^*$ and $q(v, b_1) = (1-\kappa)\left[u(\underline{v}, b_1) - u(\underline{v}, b_2) + a^*(v_2^* - v_1^*)\right]/c$ if $v > v_1^*$. Clearly, (a^*, p^*, q^*) strictly improves welfare upon (a, p, q). Now we show that (a^*, p^*, q^*) is also feasible. By construction, (IR) and (IC-v) hold. Note that

$$\begin{aligned}
p^*(v,b) &= va^*(v,b) - \int_{\underline{v}}^{v} a^*(\nu, b)\mathrm{d}\nu - (1-\kappa)u(\underline{v}, b) \\
&= (1-\kappa)va(v,b) + \kappa v - \int_{\underline{v}}^{v} \left[\kappa + (1-\kappa)a(\nu, b)\right]\mathrm{d}\nu - (1-\kappa)u(\underline{v}, b) \\
&= (1-\kappa)va(v,b) - (1-\kappa)\int_{\underline{v}}^{v} a(\nu, b)\mathrm{d}\nu - (1-\kappa)u(\underline{v}, b) \\
&= (1-\kappa)p(v,b).
\end{aligned}$$

Hence, $\mathbb{E}\left[p^*(v,b) - kq^*(v,b)\right] = (1-\kappa)\mathbb{E}\left[p(v,b) - kq(v,b)\right] \geq 0$. That is, (BB) holds. Since $p^*(\overline{v}, b_1) = (1-\kappa)p(\overline{v}, b_1) \leq b_1$, (BC) holds. Since $\mathbb{E}[a^*(v,b)] = \kappa + (1-\kappa)\mathbb{E}[a(v,b)] = \kappa + (1-\kappa)S' = S$, (S) holds. Finally, we show that (IC-b) holds. If $v \leq v_1^*$, then

$$(1-\kappa)\left[u(\underline{v}, b_1) - u(\underline{v}, b_2)\right] + \kappa(\underline{v} - \underline{v}) \leq q(v, b_1)c.$$

If $v > v_1^*$, then

$$(1-\kappa)\left[u(\underline{v}, b_1) - u(\underline{v}, b_2)\right] + \kappa(\underline{v} - \underline{v}) + (\kappa + (1-\kappa)a^* - \kappa)(v_2^* - v_1^*) \leq q(v, b_1)c.$$

Thus, we can conclude that (a^*, p^*, q^*) is feasible. However, this contradicts to that (a, p, q) is optimal. Hence, (S) holds with equality. ■

Proposition 14 *Suppose Assumptions 2 and 3 hold. Suppose also that $S < \hat{S}(b_1)$, i.e., the first-best cannot be achieved. In an optimal mechanism of $\mathcal{P}$, (S), (BB) and (BC) hold with equality.*

Proof of Proposition 14. First, it follows from Proposition 2 that (S) holds with equality. Second, we show that (BC) holds with equality. Suppose to the contradiction that (BC) holds with strict inequality. We consider four different cases: (1) $v_2^* > v_1^*$, (2) $v_2^{**} > v_2^* = v_1^*$, (3) $v_2^{**} = v_2^* = v_1^*$ and $a^* < 1$ and (4) $v_2^{**} = v_2^* = v_1^*$ and $a^* = 1$.

Suppose $v_2^* > v_1^*$.

Let $\varepsilon > 0$ and $\delta > 0$ be such that $(1-\pi)\left[F(v_1^* + \varepsilon) - F(v_1^*)\right] = \pi[F(v_2^*) - F(v_2^* - \delta)]$. For $\varepsilon > 0$ sufficiently small, we have $v_2^* - v_1^* - \varepsilon - \delta \geq 0$. Consider a new mechanism (a^*, p^*, q^*) that satisfies

$$\begin{aligned}
a^*(v, b_1) &= \chi_{\{v \geq v_1^* + \varepsilon\}} a^*, \ p(v, b_1) = \chi_{\{v \geq v_1^*\}} a^*(v_1^* + \varepsilon) - u_1^*, \\
q(v, b_1) &= \frac{1}{c}\left[\chi_{\{v \geq v_1^*\}} a^*(v_2^* - v_1^* - \varepsilon - \delta) + u_1^* - u_2^*\right], \\
a(v, b_2) &= \chi_{\{v \geq v_2^* - \delta\}} a^* + \chi_{\{v \geq v_2^{**}\}}\left(1 - a^*\right), \\
p(v, b_2) &= \chi_{\{v \geq v_2^* - \delta\}} a^*(v_2^* - \delta) + \chi_{\{v \geq v_2^{**}\}}(1 - a^*)v_2^{**} - u_2^*, \\
q(v, b_2) &= 0.
\end{aligned}$$

Clearly, for $\varepsilon > 0$ sufficiently small, (a^*, p^*, q^*) is feasible and strictly improves welfare upon (a, p, q), a contradiction.

Suppose $v_2^{**} > v_2^* = v_1^*$.

Let $\varepsilon > 0$ and $\delta > 0$ be such that $a^* \left[F(v_1^* + \varepsilon) - F(v_1^*)\right] = \pi[F(v_2^{**}) - F(v_2^{**} - \delta)]$. For $\varepsilon > 0$ sufficiently small, we have $v_2^{**} - v_1^* - \varepsilon - \delta \geq 0$. Consider a new mechanism (a^*, p^*, q^*) that satisfies

$$
\begin{aligned}
&a^*(v, b_1) = \chi_{\{v \geq v_1^* + \varepsilon\}} a^*, \ p(v, b_1) = \chi_{\{v \geq v_1^*\}} a^*(v_1^* + \varepsilon) - u_1^*, \\
&q(v, b_1) = \frac{1}{c} \left[\chi_{\{v \geq v_1^*\}} a^*(v_2^* - v_1^*) + u_1^* - u_2^*\right], \\
&a(v, b_2) = \chi_{\{v \geq v_2^* + \varepsilon\}} a^* + \chi_{\{v \geq v_2^{**} - \delta\}} \left(1 - a^*\right), \\
&p(v, b_2) = \chi_{\{v \geq v_2^* + \varepsilon\}} a^*(v_2^* + \varepsilon) + \chi_{\{v \geq v_2^{**} - \delta\}} (1 - a^*)(v_2^{**} - \delta) - u_2^*, \\
&q(v, b_2) = 0.
\end{aligned}
$$

Clearly, for $\varepsilon > 0$ sufficiently small, (a^*, p^*, q^*) is feasible and strictly improves welfare upon (a, p, q), a contradiction.

Suppose $v_2^{**} = v_2^* = v_1^*$ and $a^* < 1$.

Let $\varepsilon > 0$ and $\delta > 0$ be such that $[(1 - \pi)a^* + \pi] \left[F(v_1^* + \varepsilon) - F(v_1^*)\right] = (1 - \pi)\delta[1 - F(v_1^* + \varepsilon)]$. For $\varepsilon > 0$ sufficiently small, we have $\delta \leq 1 - a^*$. Consider a new mechanism (a^*, p^*, q^*) that satisfies

$$
\begin{aligned}
&a^*(v, b_1) = \chi_{\{v \geq v_1^* + \varepsilon\}} (a^* + \delta), \ p(v, b_1) = \chi_{\{v \geq v_1^*\}} (a^* + \delta)(v_1^* + \varepsilon) - u_1^*, \\
&q(v, b_1) = \frac{1}{c} \left(u_1^* - u_2^*\right), \\
&a(v, b_2) = \chi_{\{v \geq v_2^* + \varepsilon\}}, \ p(v, b_2) = \chi_{\{v \geq v_2^* + \varepsilon\}} - u_2^*, \\
&q(v, b_2) = 0.
\end{aligned}
$$

Clearly, for $\varepsilon > 0$ sufficiently small, (a^*, p^*, q^*) is feasible and strictly improves welfare upon (a, p, q), a contradiction.

Suppose $v_2^{**} = v_2^* = v_1^*$ and $a^* = 1$.

In this case, the first-best allocation rule is achieved. Hence, it must be the case that the total verification cost is strictly positive, i.e., $u_1^* > u_2^* \geq 0$. Let $u_2^* - u_1^* \geq \varepsilon > 0$. Consider a new mechanism (a^*, p^*, q^*) that satisfies

$$a^*(v, b_1) = \chi_{\{v \geq v_1^*\}},\ p(v, b_1) = \chi_{\{v \geq v_1^*\}} v_1^* - u_1^* + \varepsilon,$$
$$q(v, b_1) = \frac{1}{c}\left(u_1^* - u_2^* - \varepsilon\right),$$
$$a(v, b_2) = \chi_{\{v \geq v_2^*\}},\ p(v, b_2) = \chi_{\{v \geq v_2^*\}} - u_2^*,$$
$$q(v, b_2) = 0.$$

Clearly, for $\varepsilon > 0$ sufficiently small, (a^*, p^*, q^*) is feasible and strictly improves welfare upon (a, p, q), a contradiction.

Lastly, I show that (BB) holds with equality. Suppose not. Then we can increase u_1^* and u_2^* by the same amount. The resulting new mechanism is feasible and gives the same welfare. In particular, (BC) holds with strict inequality in the new mechanism. Then we can repeat the above argument and construct another feasible mechanism which strictly improves welfare upon (a, p, q), a contradiction.

■

By Theorem 9, $v_1^*, v_2^*, v_2^{**}, a^2, u_1^*, u_2^*, \beta, \eta, \lambda, \mu^1, \mu^2, \alpha^3, \xi_1$ and ξ_2 satisfy the following first-order conditions:

$$(1-\pi)\left[(\beta - v_1^* - \lambda\varphi(v_1^*))f(v_1^*) + (1+\lambda)\rho[1-F(v_1^*)] + (1+\lambda)\rho(v_2^* - v_1^*)f(v_1^*)\right]$$
$$-\eta - \mu^2 = 0, \tag{B.27}$$
$$\pi(\beta - v_2^* - \lambda\varphi(v_2^*))f(v_2^*) - (1-\pi)(1+\lambda)\rho[1-F(v_1^*)] + \mu^2 = 0, \tag{B.28}$$
$$(1-a^*)(\beta - v_2^{**} - \lambda\varphi(v_2^{**}))f(v_2^{**}) = 0, \tag{B.29}$$
$$\pi\int_{v_2^*}^{v_2^{**}} [v + \lambda\varphi(v) - \beta]\, f(v)\mathrm{d}v$$
$$+ (1-\pi)\left[\int_{v_1^*}^{\overline{v}} [v + \lambda\varphi(v) - \beta]\, f(v)\mathrm{d}v - (1+\lambda)\rho(v_2^* - v_1^*)[1-F(v_1^*)]\right]$$
$$-\eta v_1^* + \mu^2(v_2^* - v_1^*) - \alpha^3 = 0, \tag{B.30}$$
$$\eta + \mu^1 + \mu^2 - (1-\pi)(\lambda + \rho + \lambda\rho) + \xi_1 = 0, \tag{B.31}$$
$$-\mu^1 - \mu^2 - \pi\lambda + (1-\pi)(1+\lambda)\rho + \xi_2 = 0. \tag{B.32}$$

Furthermore, (S) and (BB) become:

$$(1-\pi)a^*[1-F(v_1^*)] + \pi a^*[F(v_2^{**}) - F(v_2^*)] + \pi[1-F(v_2^{**})] = S, \tag{B.33}$$
$$-(1-\pi)u_1^* + (1-\pi)a^* v_1^*[1-F(v_1^*)] - \pi u_2^* + \pi a^* v_2^*[1-F(v_2^*)] + \pi(1-a^*)v_2^{**}[1-F(v_2^{**})]$$
$$-(1-\pi)\rho(u_1^* - u_2^*) - (1-\pi)\rho a^*(v_2^* - v_1^*)[1-F(v_1^*)] = 0. \tag{B.34}$$

Proof of Proposition 4.

1. Suppose, on the contrary, that $u_1^* > u_2^* \geq 0$. In this case, $\xi_1 = \mu^1 = \mu^2 = 0$. (B.31) implies that $\eta = (1-\pi)(\lambda + \rho + \lambda\rho)$. (B.32) implies $\xi_2 = \pi\lambda - (1-\pi)(1+\lambda)\rho$. Since $\xi_2 \geq 0$, we have $\lambda[\pi - \rho(1-\pi)] \geq \rho(1-\pi)$ which implies that $\rho < \pi/(1-\pi)$, a contradiction.

2. Since $S < 1$, we have $u_1^* = u_2^*$ by the first result of Proposition 4. It suffices to show that $v_1^* = v_2^*$. Suppose, on the contrary, that $v_2^* > v_1^*$. In this case, $\mu^2 = 0$. Combining (B.31) and (B.32) yields $\eta - \lambda + \xi_1 + \xi_2 = 0$. Since $\xi_1, \xi_2 \geq 0$, we have $\eta \leq \lambda$. Taking the difference of

(B.27) divided by $(1-\pi)f(v_1^*)$ and (B.28) divided by $\pi f(v_2^*)$ gives

$$[1+(1+\lambda)\rho]\left(v_2^*-v_1^*\right)+\lambda\left[\varphi(v_2^*)-\varphi(v_1^*)\right]+(1+\lambda)\rho\frac{1-F(v_1^*)}{f(v_1^*)}$$
$$+(1+\lambda)\frac{\rho(1-\pi)}{\pi}\frac{1-F(v_1^*)}{f(v_2^*)}-\frac{\eta}{(1-\pi)f(v_1^*)}=0. \tag{B.35}$$

Since $v_2^*>v_1^*$, $f(v_2^*)\le f(v_1^*)$ and $\eta\le\lambda$, we have

$$0\ge[1+(1+\lambda)\rho]\left(v_2^*-v_1^*\right)+\lambda\left[\varphi(v_2^*)-\varphi(v_1^*)\right]+(1+\lambda)\frac{\rho}{\pi}\frac{1-F(v_1^*)}{f(v_1^*)}-\frac{\lambda}{(1-\pi)f(v_1^*)}$$
$$>\frac{\rho}{\pi}\frac{1-F(v_1^*)}{f(v_1^*)}+\lambda\left[\frac{\rho}{\pi}\frac{1-F(v_1^*)}{f(v_1^*)}-\frac{1}{(1-\pi)f(v_1^*)}\right]\ge 0,$$

where the last inequality holds since $1-F(v_1^*)\ge S$ and $\rho\ge\pi/[S(1-\pi)]$. A contradiction. Hence, $v_1^*=v_2^*$.

■

Proof of Proposition 5. Suppose, on the contrary, that $u_1^*>u_2^*\ge 0$. In this case, $\xi_1=\mu^1=\mu^2=0$. (B.31) implies that $\eta=(1-\pi)(\lambda+\rho+\lambda\rho)$. Taking the difference of (B.27) divided by $(1-\pi)f(v_1^*)$ and (B.28) divided by $\pi f(v_2^*)$ gives

$$[1+(1+\lambda)\rho]\left(v_2^*-v_1^*\right)+\lambda\left[\varphi(v_2^*)-\varphi(v_1^*)\right]+(1+\lambda)\rho\frac{1-F(v_1^*)}{f(v_1^*)}$$
$$+(1+\lambda)\frac{\rho(1-\pi)}{\pi}\frac{1-F(v_1^*)}{f(v_2^*)}-\frac{\eta}{(1-\pi)f(v_1^*)}=0. \tag{B.35}$$

Suppose $S\le(1-\pi)\left[1-F(b_1)\right]$. Since (BC) holds with equality and $u_1^*\ge 0$, we have $a^*\ge b_1/v_1^*$. By (S), we have

$$(1-\pi)\frac{b_1}{v_1^*}\left[1-F(v_1^*)\right]\le S.$$

Since $S\le(1-\pi)\left[1-F(b_1)\right]$, there exists a unique Let $\hat{v}(S,b_1,\pi)\in[b_1,\overline{v}]$ such that the above inequality holds with equality when $v_1^*=\hat{v}(S,b_1,\pi)$, where $\hat{v}$ is strictly decreasing in S and π and strictly increasing in b_1. Then $v_1^* \geq \hat{v}(S, b_1, \pi)$. Hence, $v_2^* - v_1^* \leq \varphi(v_2^*) - \varphi(v_1^*) \leq \overline{v} - \varphi\left(\hat{v}(S, b_1, \pi)\right)$.

Since $v_2^* \geq v_1^*$, $f(v_2^*) \leq f(v_1^*)$ and $\eta = (1-\pi)(\lambda + \rho + \lambda\rho)$, we have

$$\begin{aligned}
0 \leq & [1 + (1+\lambda)\rho]\left(v_2^* - v_1^*\right) + \lambda\left[\varphi(v_2^*) - \varphi(v_1^*)\right] + (1+\lambda)\frac{\rho}{\pi}\frac{1 - F(v_1^*)}{f(v_2^*)} - \frac{\lambda + \rho + \lambda\rho}{f(v_1^*)} \\
< & (1+\lambda)(1+\rho)\left[\overline{v} - \varphi\left(\hat{v}(S, b_1, \pi)\right)\right] + (1+\lambda)\frac{\rho}{\pi}\frac{1 - F(v_1^*)}{f(v_2^*)} - \frac{(1+\lambda)\rho}{f(v_1^*)} \\
\leq & (1+\lambda)\left\{(1+\rho)\left[\overline{v} - \varphi\left(\hat{v}(S, b_1, \pi)\right)\right] + \frac{\rho}{\pi}\frac{1 - F\left(\hat{v}(S, b_1, \pi)\right)}{f(\overline{v})} - \frac{\rho}{f\left(\hat{v}(S, b_1, \pi)\right)}\right\}.
\end{aligned}$$

Note that the term in the braces is strictly increasing in S and converges to $-\rho/f(\overline{v}) < 0$ as S goes to zero. Hence, there exists $\hat{S}$ such that $u_1^* = u_2^*$ if $S < \hat{S}$. ■

Proof of Proposition 6. Suppose, on the contrary, that $u_1^* > u_2^* \geq 0$. In this case, $\xi_1 = \mu^1 = \mu^2 = 0$. (B.31) implies that $\eta = (1-\pi)(\lambda + \rho + \lambda\rho)$. Taking the difference of (B.27) divided by $(1-\pi)f(v_1^*)$ and (B.28) divided by $\pi f(v_2^*)$ gives

$$\begin{aligned}
& [1 + (1+\lambda)\rho]\left(v_2^* - v_1^*\right) + \lambda\left[\varphi(v_2^*) - \varphi(v_1^*)\right] + (1+\lambda)\rho\frac{1 - F(v_1^*)}{f(v_1^*)} \\
& + (1+\lambda)\frac{\rho(1-\pi)}{\pi}\frac{1 - F(v_1^*)}{f(v_2^*)} - \frac{\eta}{(1-\pi)f(v_1^*)} = 0.
\end{aligned} \tag{B.35}$$

Suppose $S \geq (1-\pi)\left[1 - F(b_1)\right]$. Then by (S),

$$\pi[1 - F(v_2^*)] \geq (1-\pi)\left[1 - F(b_1)\right] - S. \tag{B.36}$$

Since $S \geq (1-\pi)\left[1 - F(b_1)\right]$, there exists a unique $\hat{v}(S, b_1, \pi) \in [b_1, \overline{v}]$ such that (B.36) holds with equality, where $\hat{v}$ is strictly decreasing in b_1, S and π. Then $v_2^* \leq \hat{v}(S, b_1, \pi)$. Hence, $v_2^* - v_1^* \leq \varphi(v_2^*) - \varphi(v_1^*) \leq \varphi\left(\hat{v}(S, b_1, \pi)\right) - b_1$.

Since $v_2^* \geq v_1^*$, $f(v_2^*) \leq f(v_1^*)$ and $\eta = (1-\pi)(\lambda+\rho+\lambda\rho)$, we have

$$\begin{aligned}
0 \leq & [1+(1+\lambda)\rho]\left(v_2^* - v_1^*\right) + \lambda\left[\varphi(v_2^*) - \varphi(v_1^*)\right] + (1+\lambda)\frac{\rho}{\pi}\frac{1-F(v_1^*)}{f(v_2^*)} - \frac{\lambda+\rho+\lambda\rho}{f(v_1^*)} \\
< & (1+\lambda)(1+\rho)\left[\varphi\left(\hat{v}(S,b_1,\pi)-b_1\right)\right] + (1+\lambda)\frac{\rho}{\pi}\frac{1-F(v_1^*)}{f(v_2^*)} - \frac{(1+\lambda)\rho}{f(v_1^*)} \\
\leq & (1+\lambda)\left\{(1+\rho)\left[\varphi\left(\hat{v}(S,b_1,\pi)\right)-b_1\right] + \frac{\rho}{\pi}\frac{1-F\left(b_1\right)}{f\left(\hat{v}(S,b_1,\pi)\right)} - \frac{\rho}{f\left(b_1\right)}\right\}.
\end{aligned}$$

Note that the term in the braces is strictly decreasing in b_1 and converges to $-\rho/f(\overline{v}) < 0$ as b_1 goes to $\overline{v}$. Hence, there exists $\hat{b}_1$ such that $u_1^* = u_2^*$ if $b_1 > \hat{b}_1$. ■

B.5. Extensions and discussions

B.5.1. Per-unit price constraint

Proof of Theorem 10. The proof of Theorem 7 can easily modified to prove Theorem 10. It suffices to show that (a^*, p^*) satisfies (PC) (instead of (BC)):

$$\begin{aligned}
p^*(\overline{v}, b) =& \overline{v}a(\overline{v}, b) - \int_{\underline{v}}^{\overline{v}} a^*(v,b)\mathrm{d}v - u(\underline{v}, b) \\
\leq& \overline{v}a(\overline{v}, b) - \int_{\underline{v}}^{\overline{v}} a(v,b)\mathrm{d}v - u(\underline{v}, b) \\
\leq& a(\overline{v}, b)b \\
=& a^*(\overline{v}, b)b,
\end{aligned}$$

where the third line holds by the same argument used in the proof of Theorem 7 and the last line holds since $a^*(\overline{v}, b) = a(\overline{v}, b)$ by construction. Hence, there exists v_1^* and v_2^* such that the optimal allocation rule satisfies $a(v, b_1) = \chi_{\{v \geq v_1^*\}} \min\left\{\frac{u^*}{v_1^* - b_1}, 1\right\}$ and $a(v, b_2) = \chi_{\{v \geq v_2^*\}}$. ■

Lemma 38 *Suppose Assumption 3 holds, and the principal does not inspect agents. In an optimal mechanism of $\mathcal{P}'_{PC}$, it is without loss of generality to assume that $u(\underline{v}, b_1) = u(\underline{v}, b_2)$.*

Proof. The proof of Lemma 2 can easily modified to prove Lemma 38. It suffices to show that

(a^*, p^*) satisfies (PC) (instead of (BC)). Note that $a^*(\overline{v}, b) = a(\overline{v}, b)$ by construction and the rest of the proof follows from a similar argument used in the proof of Theorem 7. ■

Lemma 39 *Suppose Assumptions 2 and 3 hold, and the principal does not inspect agents. In an optimal mechanism of $\mathcal{P}'_{PC}$, the allocation rule satisfies*

$$\int_{\underline{v}}^{v} a(v, b_2) f(v) \mathrm{d}v \geq \int_{\underline{v}}^{v} a(v, b_1) f(v) \mathrm{d}v, \ \forall v. \tag{B.37}$$

Proof. The proof of Lemma 3 applies. ■

Proof of Theorem 11. The proof of Theorem 8 can easily modified to prove Theorem 11. It suffices to show that (a^*, p^*) satisfies (PC) (instead of (BC)). Note that $a^*(\overline{v}, b) = a(\overline{v}, b)$ by construction and the rest of the proof follows from a similar argument used in the proof of Theorem 7. ■

B.5.2. Monetary penalty

Proof of Lemma 9. Consider types $t := (v, b)$ and $\hat{t}$ such that $p(\hat{t}) + \max\{\theta(\hat{t}, n), \theta(\hat{t}, b)\} \leq b$. Then (IC) requires that

$$\begin{aligned}&a(t)v - p(t) - (1 - q(t))\theta(t, n) - q(t)\theta(t, b)\\ \geq &a(\hat{t})v - p(\hat{t}) - (1 - q(\hat{t}))\theta(\hat{t}, n) - q(\hat{t})\theta(\hat{t}, b)\end{aligned}$$

Consider an alternative mechanism $(a^*, p^*, q^*, \theta^*)$ with $a^* = a$ and $q^* = q$. Let $\theta^*(t, n) = \theta^*(t, b) = 0$ for all t and $\theta^*(\hat{t}, b) = c$ for all $\hat{t}$ such that $\hat{b} \neq b$. Let $p^*(t) = p(t) + (1 - q(t))\theta(t, n) + q(t)\theta(t, b)$. Since $p(t) + \max\{\theta(t, n), \theta(t, b)\} \leq b$, we have $p^*(t) \leq b$, i.e., (BC) holds. It is easy to see that the new mechanism also satisfies (IR), (BB) and (S) and does not affect the welfare.

Finally, I show that (IC) holds. Consider types $t := (v, b)$ and $\hat{t}$ such that $p^*(\hat{t}) + c \leq b$. If $\hat{b} = b$, then (BC) in the old mechanism implies that $p(\hat{t}) + \max\{\theta(\hat{t}, n), \theta(\hat{t}, b)\} \leq b$. Hence,

$$\begin{aligned}
&a^*(t)v - p^*(t) \\
=&a(t)v - p(t) - (1 - q(t))\theta(t, n) - q(t)\theta(t, b) \\
\geq&a(\hat{t})v - p(\hat{t}) - (1 - q(\hat{t}))\theta(\hat{t}, n) - q(\hat{t})\theta(\hat{t}, b) \\
=&a^*(\hat{t})v - p^*(\hat{t}) + q(\hat{t})\theta(\hat{t}, \hat{b}) - q(\hat{t})\theta(\hat{t}, b) \\
=&a^*(\hat{t})v - p^*(\hat{t}).
\end{aligned}$$

If $\hat{b} \neq b$, then $b \geq p^*(\hat{t}) + c = p(\hat{t}) + (1 - q(\hat{t}))\theta(\hat{t}, n) + q(\hat{t})\theta(\hat{t}, \hat{b}) + c \geq p(\hat{t}) + \max\{\theta(\hat{t}, n), \theta(\hat{t}, b)\}$. Hence,

$$\begin{aligned}
&a^*(t)v - p^*(t) \\
=&a(t)v - p(t) - (1 - q(t))\theta(t, n) - q(t)\theta(t, b) \\
\geq&a(\hat{t})v - p(\hat{t}) - (1 - q(\hat{t}))\theta(\hat{t}, n) - q(\hat{t})\theta(\hat{t}, b) \\
=&a^*(\hat{t})v - p^*(\hat{t}) + q(\hat{t})\theta(\hat{t}, \hat{b}) - q(\hat{t})\theta(\hat{t}, b) \\
\geq&a^*(\hat{t})v - p^*(\hat{t}) - q^*(\hat{t})\theta^*(\hat{t}, b).
\end{aligned}$$

The last inequality holds since $\theta(\hat{t}, \hat{b}) \geq 0$ and $\theta^*(\hat{t}, b) = c \geq \theta(\hat{t}, b)$. ■

B.5.3. Punishing the innocent or without verification

Lemma 40 *An optimal mechanism of $\mathcal{P}_{PI}$ satisfies (i) $\theta(t, \hat{b}) = 1$ for $\hat{b} \neq b$, (ii) $p(t) < b$ implies that $\theta(t, n) = \theta(t, b) = 0$ and (iii) $(1 - \theta(t, n))\theta(t, b) = 0$ for almost all t.*

Proof. By the standard argument, (IC-v) implies that a is non-decreasing and

$$p(t) + (1 - q(t))\theta(t, n)c + q(t)\theta(t, b)c = a(t)v - \int_0^v a(\nu, b)\mathrm{d}\nu - u(0, b).$$

Consider a new mechanism $(a^*, p^*, q^*, \theta^*)$. Let $a^* = a$. Thus, (S) holds. Let $\theta^*(\hat{t}, b) = 1$ and $q^*(\hat{t}) = q(\hat{t})\theta(\hat{t}, b)$ for $b \neq \hat{b}$. Let $p^*(t) = \min\{p(t) + (1 - q(t))\theta(t, n)c + q(t)\theta(t, b)c, b\}$. Thus, (BC) holds. Since $p^*(t) \geq p(t)$ and $q^*(t) \leq q(t)$, (BB) holds. Let $\theta^*(t, n) = 0$ if $q^*(t) = 1$ and otherwise

$$\theta^*(t,n) = \min\left\{\frac{p(t) + (1 - q(t))\theta(t,n)c + q(t)\theta(t,b)c - p^*(t)}{(1 - q^*(t))c}, 1\right\}.$$

Finally, let $\theta^*(t, b) = 0$ if $q^*(t) = 0$ and otherwise

$$\theta^*(t,b) = \frac{p(t) + (1 - q(t))\theta(t,n)c + q(t)\theta(t,b)c - p^*(t) - (1 - q^*(t))\theta^*(t,n)c}{q^*(t)c} \geq 0.$$

Then, $\theta^*(t, b) > 0$ if and only if $\theta^*(t, n) = 1$. Furthermore, $p^*(t) \geq p(t)$. Hence, $\theta^*(t, b) > 0$ implies that

$$\theta^*(t,b) \leq \frac{(1 - q(t))\theta(t,n) + q(t)\theta(t,b) - 1 + q^*(t)}{q^*(t)} \leq 1.$$

Note also that, by construction, $p(t) < b$ implies that $\theta^*(t, n) = \theta^*(t, b) = 0$. By construction,

$$p(t) + (1 - q(t))\theta(t,n)c + q(t)\theta(t,b)c = p^*(t) + (1 - q^*(t))\theta^*(t,n)c + q^*(t)\theta^*(t,b)c.$$

Hence (IR) holds. Consider a type $t = (v, b)$ and $\hat{t}$ such that $p(\hat{t}) \leq p^*(\hat{t}) \leq b$. Then

$$\begin{aligned}
&a^*(t)v - p^*(t) - (1 - q^*(t))\theta^*(t,n)c - q^*(t)\theta^*(t,b)c \\
=&a(t)v - p(t) - (1 - q(t))\theta(t,n)c - q(t)\theta(t,b)c \\
\geq&a(\hat{t})v - p(\hat{t}) - (1 - q(\hat{t}))\theta(\hat{t},n)c - q(\hat{t})\theta(\hat{t},b)c \\
\geq&a^*(\hat{t})v - p^*(\hat{t}) - (1 - q^*(\hat{t}))\theta^*(\hat{t},n)c - q^*(\hat{t})c.
\end{aligned}$$

If $\hat{b} = b$, the last inequality holds trivially. If $\hat{b} \neq b$, the last inequality holds since $p^*(\hat{t}) + (1 - q^*(\hat{t}))\theta(\hat{t}, n)c \geq p(\hat{t}) + (1 - q(\hat{t}))\theta(\hat{t}, n)c$ and $q^*(\hat{t}) = q(\hat{t})\theta(\hat{t}, b)$. Hence, (IC) holds. Thus, we have verified that $(a^*, p^*, q^*, \theta^*)$ is feasible. Since $p^*(t) \geq p(t)$, we have $(1 - q(t))\theta(t, n)c + q(t)\theta(t, b)c \geq (1 - q^*(t))\theta^*(t, n)c + q^*(t)\theta^*(t, b)c$. Furthermore, $q^*(t) \leq q(t)$. For a positive measure set of t, one of the above two inequalities holds strictly. Hence, $(a^*, p^*, q^*, \theta^*)$ strictly improves welfare. ■

Lemma 41 *An optimal mechanism of $\mathcal{P}_{PI}$ satisfies $\theta(t, b) = 0$ for almost all t.*

Proof. Fix a mechanism (a, p, q, θ). Suppose $\theta(t, b) > 0$ on a positive measure set of t. Consider a new mechanism $(a^*, p^*, q^*, \theta^*)$ with $a^* = a$ and $p^* = p$. If $\theta(t, b) > 0$, let $q^*(t) = q(t)[1 - \theta(t, b)] < q(t)$, $\theta^*(t, b) = 0$ and

$$\theta^*(t, n) = \frac{(1 - q(t))\theta(t, n) + q(t)\theta(t, b)}{1 - q^*(t)} = 1.$$

If $\theta(t, b) = 0$, let $q^*(t) = q(t)$, $\theta^*(t, b) = 0$ and $\theta^*(t, n) = \theta(t, n)$. By construction,

$$(1 - q(t))\theta(t, n)c + q(t)\theta(t, b)c = (1 - q^*(t))\theta^*(t, n)c + q^*(t)\theta^*(t, b)c.$$

Clearly, the new mechanism satisfies (IR), (BC), (BB) and (S) and strictly improves welfare. Consider types $t = (v, b)$ and $\hat{t}$ such that $p(\hat{t}) = p^*(\hat{t}) \leq b$. Then

$$\begin{aligned}
&a^*(t)v - p^*(t) - (1 - q^*(t))\theta^*(t, n)c - q^*(t)\theta^*(t, b)c \\
=&a(t)v - p(t) - (1 - q(t))\theta(t, n)c - q(t)\theta(t, b)c \\
\geq&a(\hat{t})v - p(\hat{t}) - (1 - q(\hat{t}))\theta(\hat{t}, n)c - q(\hat{t})c \\
=&a^*(\hat{t})v - p^*(\hat{t}) - (1 - q^*(\hat{t}))\theta^*(\hat{t}, n)c - q^*(\hat{t})c.
\end{aligned}$$

If $\theta(\hat{t}, \hat{b}) = 0$, the last equality holds trivially. If $\theta(\hat{t}, \hat{b}) > 0$, the last equality holds since $\theta(\hat{t}, n) = \theta^*(\hat{t}, n) = 1$. ■

APPENDIX TO CHAPTER 4

C.1. Omitted proofs in Section 4.3

Before proceeding to the proofs, I first define *symmetric mechanisms* formally. Let $\sigma_{i,j} : W^n \to W^n$ denote the function that interchanges the ith and the jth coordinates, i.e.,

$$\sigma_{i,j}(w_1, \dots, w_n) = (w_1, \dots, w_{i-1}, w_j, w_{i+1}, \dots, w_{j-1}, w_i, w_{j+1}, \dots, w_n),\ \forall (w_1, \dots, w_n).$$

We say that an allocation rule $\boldsymbol{q}$ is *symmetric* if q_1 is such that $q_1(\boldsymbol{w}) = q_1(\sigma_{i,j}(\boldsymbol{w}))$ for all $i, j \neq 1$, $q_i(\boldsymbol{w}) = q_1(\sigma_{1,i}(\boldsymbol{w}))$ and $\sum_i q_i(\boldsymbol{w}) \leq 1$ for all $\boldsymbol{w}$. We say that a mechanism $(\boldsymbol{q}, \boldsymbol{t})$ is *symmetric* if its allocation rule $\boldsymbol{q}$ is symmetric.

Proof of Lemma 11. By construction, $H(w(s,\alpha)|\alpha) = s$ for all $s \in [0,1]$ and $\alpha \in \mathbb{A}$. Taking derivative of both sides of the equation with respect to α yields

$$h(w(s,\alpha)|\alpha)w_\alpha(s,\alpha) + H_\alpha(w(s,\alpha)|\alpha) = 0,$$

or equivalently,

$$-\frac{H_\alpha(w(s,\alpha)|\alpha)}{h(w(s,\alpha)|\alpha)} = w_\alpha(s,\alpha). \tag{C.1}$$

If that the information structures are supermodular ordered, then $-H_\alpha(w|\alpha)/h(w|\alpha)$ is strictly increasing in w. Because $w(s,\alpha)$ is strictly increasing in s, $w_\alpha(s,\alpha)$ is strictly increasing in s. Hence, for all $s, s' \in (0,1)$, $s' > s$ and $\alpha' > \alpha''$ we have

$$\begin{aligned} w(s',\alpha') - w(s',\alpha'') &= \int_{\alpha''}^{\alpha'} w_\alpha(s',\alpha)\mathrm{d}\alpha \\ &> \int_{\alpha''}^{\alpha'} w_\alpha(s,\alpha)\mathrm{d}\alpha \\ &= w(s,\alpha') - w(s,\alpha''). \end{aligned}$$

That is, $w(\cdot,\cdot)$ is strictly supermodular. ∎

Lemma 42 *Let (c, d) be an interval of the real line, and J and Q be two non-decreasing functions. Assume that for some measure h on $\mathbb{R}$ we have*

$$\int_c^d J(w)\mathrm{d}h(w) = 0.$$

Then $\int_c^d J(w)Q(w)\mathrm{d}h(w) \geq 0$.

Proof. This lemma is a corollary of Lemma 1 in Persico (2000). ■

Lemma 43 *Suppose that $Q : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}$ is non-decreasing on $[w(0, \alpha_i), w(1, \alpha_i)]$, then*

$$\mathbb{E}_{w_i}\left[-\frac{H_{\alpha_i}(w_i|\alpha_i)}{h(w_i|\alpha_i)} Q(w_i) \middle| \alpha_i \right] \geq 0, \tag{C.2}$$

where the equality holds if Q is constant.

Proof. Because both Q and $-\frac{H_{\alpha_i}(w_i|\alpha_i)}{h(w_i|\alpha_i)}$ are non-decreasing on $[w(0, \alpha_i), w(1, \alpha_i)]$, it suffices to show that

$$\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} H_{\alpha_i}(w_i|\alpha_i)\mathrm{d}w_i = 0. \tag{C.3}$$

On the one hand, by integration by parts,

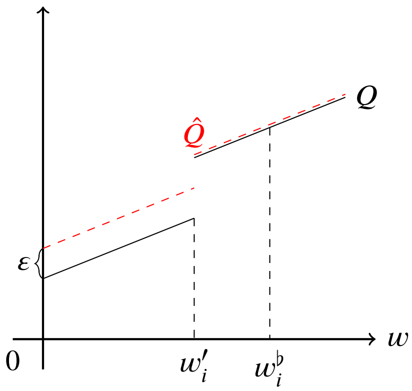
$$\begin{aligned}\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} H(w_i|\alpha_i)\mathrm{d}w_i &= w_i H(w_i|\alpha_i)\Big|_{w(0,\alpha_i)}^{w(1,\alpha_i)} - \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} w_i \mathrm{d}H(w_i|\alpha_i), \\ &= w(1, \alpha_i) - \mu.\end{aligned}$$

Taking derivative with respect to α_i yields

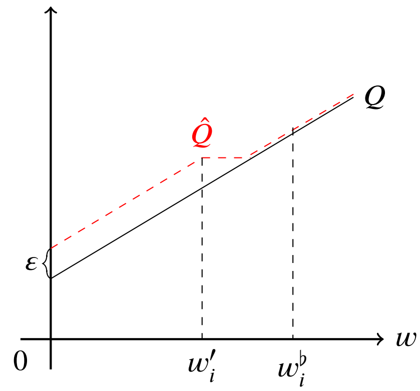
$$\frac{\partial}{\partial \alpha_i} \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} H(w_i|\alpha_i)\mathrm{d}w_i = w_{\alpha_i}(1, \alpha_i). \tag{C.4}$$

On the other hand, by the chain rule, we have

$$\frac{\partial}{\partial \alpha_i} \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} H(w_i|\alpha_i)\mathrm{d}w_i = w_{\alpha_i}(1, \alpha_i) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} H_{\alpha_i}(w_i|\alpha_i)\mathrm{d}w_i. \tag{C.5}$$



(a) Case I



(b) Case II

Figure 14: Proof of Lemma 12

Comparing (C.4) and (C.5) proves (C.3). By Lemma 42, inequality (C.2) holds. If Q is constant, the equality holds by (C.3). ■

Proof of Lemma 12. Define $w^\flat := \sup\left\{w_i \,\middle|\, Y(w_i') > 0,\ \forall w(0,\alpha^*) \le w_i' \le w_i\right\}$. By the continuity of Y, we have $Y(w^\flat) = 0$ and $w^\flat > w(0,\alpha^*)$. The proof is by construction. There are four cases to consider.

Case I: Suppose that there exists $w_i' \in (w(0,\alpha^*), w^\flat)$ such that Q is discontinuous at w_i'.

Let $Q(w_i'^{+})$ denote the right-hand limit of Q at w_i', and $Q(w_i'^{-})$ the corresponding left-hand limit. Let $0 < \varepsilon \le \min\left\{\min_{w(0,\alpha^*)\le w_i \le w_i'} \frac{Y(w_i)}{H(w_i'|\alpha^*)}, Q(w_i'^{+}) - Q(w_i'^{-})\right\}$. Define $\hat{Q}$ as follows. If $w_i \le w(0,\alpha^*)$, then $\hat{Q}(w_i) := Q(w_i)$; and if $w_i > w(0,\alpha^*)$, then

$$\hat{Q}(w_i) := Q(w_i) + \varepsilon \chi_{\{w_i \le w_i'\}},$$

where $\chi_{\{w_i \le w_i'\}}$ is an indicator function. (See Figure 14a for an illustration.) By construction, $\hat{Q}(w) \ge Q(w)$ for all $w \in W$ and the inequality holds strictly on a positive measure set. It is also clear that $\hat{Q}$ satisfies (MON). We now verify that $\hat{Q}$ satisfies (IA′) and (F′). Because $\chi_{\{w_i \le w_i'\}}$

is non-increasing on $[w(0,\alpha^*), w(1,\alpha^*)]$, by Lemma 43, we have

$$
\begin{aligned}
&\mathbb{E}\left[-\frac{H_{\alpha_i}(w_i|\alpha^*)}{h(w_i|\alpha^*)}\hat{Q}(w_i)\middle|\alpha_i=\alpha^*\right],\\
=&\mathbb{E}\left[-\frac{H_{\alpha_i}(w_i|\alpha^*)}{h(w_i|\alpha^*)}Q(w_i)\middle|\alpha_i=\alpha^*\right]+\varepsilon\mathbb{E}\left[-\frac{H_{\alpha_i}(w_i|\alpha^*)}{h(w_i|\alpha^*)}\chi_{\{w_i\le w_i'\}}\middle|\alpha_i=\alpha^*\right],\\
\le& C'(\alpha^*)+0=C'(\alpha^*).
\end{aligned}
$$

Hence, $\hat{Q}$ satisfies (IA$'$). Finally, let

$$
\hat{Y}(w_i) := \int_{w_i}^{\overline{\theta}} \left[H(z|\alpha^*)^{n-1}-\hat{Q}(z)\right] h(z|\alpha^*)\mathrm{d}z.
$$

If $w_i \le w_i'$, then $\hat{Y}(w_i) = Y(w_i)-\varepsilon[H(w_i'|\alpha^*)-H(w_i|\alpha^*)] \ge Y(w_i)-\varepsilon H(w_i'|\alpha^*) \ge 0$. If $w_i > w_i'$, then $\hat{Y}(w_i) = Y(w_i) \ge 0$. Hence, $\hat{Q}$ satisfies (F$'$).

Case II: Suppose that Q is continuous on $[w(0,\alpha^*), w^\flat]$.

We first show that there exists $w_i' \in (w(0,\alpha^*), w^\flat)$ such that $Q(w_i') < Q(w^\flat)$. Suppose, to the contrary, that $Q(w_i) = Q(w^\flat)$ for all $w_i \in (w(0,\alpha^*), w^\flat)$. If $Q(w^\flat) \ge H(w^\flat|\alpha^*)^{n-1}$, then $Y(w(0,\alpha^*)) = \int_{w(0,\alpha^*)}^{w^\flat}[H(z|\alpha^*)^{n-1}-Q(z)]h(z|\alpha^*)\mathrm{d}z < 0$, a contradiction. If $Q(w^\flat) < H(w^\flat|\alpha^*)^{n-1}$, then, by the continuity of Q and H, there exists $\delta > 0$ such that $Q(w_i) < H(w_i|\alpha^*)^{n-1}$ for all $w_i \in [w^\flat, w^\flat+\delta]$. Hence,

$$
0 = Y(w^\flat) = \int_{w^\flat}^{w^\flat+\delta}[H(z|\alpha^*)^{n-1}-Q(z)]h(z|\alpha^*)\mathrm{d}z + Y(w^\flat+\delta) > Y(w^\flat+\delta),
$$

a contradiction. Thus, there exists $w_i' \in (w(0,\alpha^*), w^\flat)$ such that $Q(w_i') < Q(w^\flat)$.

By the continuity of Q, there exists $w_i'' \in (w_i', w^\flat)$ such that $Q(w_i'') = \frac{1}{2}\left(Q(w_i')+Q(w^\flat)\right)$. Let

$0 < \varepsilon \le \min\left\{\min_{w(0,\alpha^*)\le w_i \le w_i''} \frac{Y(w_i)}{H(w_i''|\alpha^*)}, Q(w_i'') - Q(w_i')\right\}$. Let

$$\hat{Q}(w_i) := \begin{cases} \max\{Q(w_i') + \varepsilon, Q(w_i)\} & \text{if } w_i > w_i', \\ Q(w_i) + \varepsilon & \text{if } w(0,\alpha^*) < w_i \le w_i', \\ Q(w_i) & \text{if } w_i \le w(0,\alpha^*). \end{cases}$$

(See Figure 14b for an illustration.) Note that if $w_i \ge w_i''$ then $Q(w_i) \ge Q(w_i'') \ge Q(w_i') + \varepsilon$. Hence, $\hat{Q}(w_i) = Q(w_i)$ for $w_i \ge w_i''$. By construction, $\hat{Q}(w) \ge Q(w)$ for all $w \in W$ and the inequality holds strictly on a positive measure set. Clearly, $\hat{Q}$ satisfies (MON). We now verify that $\hat{Q}$ satisfies (IA$'$) and (F$'$). It is easy to verify that $\hat{Q} - Q$ is non-increasing on $[w(0,\alpha^*), w(1,\alpha^*)]$ and therefore $\hat{Q}$ satisfies (IA$'$) by Lemma 43. Finally, if $w_i \ge w_i''$, then $\hat{Y}(w_i) = Y(w_i)$. If $w_i < w_i''$, then

$$\begin{aligned} \hat{Y}(w_i) = &\int_{w_i}^{w^\flat} \left[H(z|\alpha^*)^{n-1} - \hat{Q}(z)\right] h(z|\alpha^*)\mathrm{d}z, \\ =&Y(w_i) - \int_{w_i}^{w_i''} \left[\hat{Q}(z) - Q(z)\right] h(z|\alpha^*)\mathrm{d}z, \\ \ge&Y(w_i) - \varepsilon \left[H(w_i''|\alpha^*) - H(w_i|\alpha^*)\right], \\ \ge&Y(w_i) - \varepsilon H(w_i''|\alpha^*) \ge 0. \end{aligned}$$

Hence, $\hat{Q}$ satisfies (F$'$).

Case III: Suppose that Q is continuous on $[w(0,\alpha^*), w^\flat)$ and $Q(w^{\flat-}) < H(w^\flat|\alpha^*)^{n-1}$.

Define $R(w_i) := Y(w_i)/(H(w^\flat|\alpha^*) - H(w_i|\alpha^*))$ for $w_i < w^\flat$. Then by L'Hopital's rule,

$$\lim_{w_i \to w^{\flat-}} R(w_i) = H(w^\flat|\alpha^*)^{n-1} - Q(w^{\flat-}) > 0.$$

Let $0 < \varepsilon \le \min\left\{\inf_{w(0,\alpha^*)\le w_i < w^\flat} R(w_i), Q(w^{\flat+}) - Q(w^{\flat-})\right\}$. Define $\hat{Q}$ as follows. If $w_i \le w(0,\alpha^*)$, then $\hat{Q}(w_i) := Q(w_i)$; and if $w_i > w(0,\alpha^*)$, then $\hat{Q}(w_i) := Q(w_i) + \varepsilon\chi_{\{w_i < w^\flat\}}$. By construction, $\hat{Q}(w) \ge Q(w)$ for all $w \in W$ and the inequality holds strictly on a positive measure set. Clearly, $\hat{Q}$ satisfies (MON). We can verify that $\hat{Q}$ satisfies (IA$'$) following the arguments in Case

I. Finally, if $w_i < w^\flat$, then $\hat{Y}(w_i) = Y(w_i) - \varepsilon[H(w^\flat|\alpha^*) - H(w_i|\alpha^*)] \geq Y(w_i) - R(w_i)[H(w^\flat|\alpha^*) - H(w_i|\alpha^*)] = 0$. If $w_i \geq w^\flat$, then $\hat{Y}(w_i) = Y(w_i) \geq 0$. Hence, $\hat{Q}$ satisfies (F′).

Case IV: Suppose that Q is continuous on $[w(0,\alpha^*), w^\flat)$ and $Q(w^{\flat-}) \geq H^{n-1}(w^\flat|\alpha^*)$.

We first show that $Q(w^{\flat-}) = H^{n-1}(w^\flat|\alpha^*)$. Suppose to the contrary that $Q(w^{\flat-}) > H^{n-1}(w^\flat|\alpha^*)$. Then by the continuity of Q and H on $[w(0,\alpha^*), w^\flat)$, there exists $\delta > 0$ such that $Q(w_i) > H^{n-1}(w_i|\alpha^*)$ for all $w_i \in (w^\flat - \delta, w^\flat)$. Then

$$Y(w^\flat - \delta) = \int_{w^\flat-\delta}^{w^\flat} [H(z|\alpha^*)^{n-1} - Q(z)]h(z|\alpha^*)\mathrm{d}z + Y(w^\flat) < 0,$$

a contradiction to that Q is feasible. Hence, $Q(w^{\flat-}) = H^{n-1}(w^\flat|\alpha^*)$. Second, we show that there exists $w_i' \in (w(0,\alpha^*), w^\flat)$ such that $Q(w_i') < Q(w^{\flat-})$. Suppose to the contrary that $Q(w_i) = Q(w^{\flat-})$ for all $w_i \in (w(0,\alpha^*), w^\flat)$, then $Y(w(0,\alpha^*)) = \int_{w(0,\alpha^*)}^{w^\flat}[H^{n-1}(z|\alpha^*) - Q(z)]h(z|\alpha^*)\mathrm{d}z + Y(w^\flat) < 0$, a contradiction. Hence, there exists $w_i' \in (w(0,\alpha^*), w^\flat)$ such that $Q(w_i') < Q(w^{\flat-})$. The rest of the proof follows from that of Case II. ■

C.1.1. Proof of Lemma 14

The proof uses a network-flow approach (see Che et al. (2013b) for detailed discussions of this approach). For simplicity, I prove here the result for the case of finite W. By a similar argument to that of the proof of Theorem 5 in Che et al. (2013b), the result generalizes to the case of continuum W. I abuse notation a bit and let f denote the probability mass function in the case of finite W. In this case, (4.6) becomes:

$$\sum_{\boldsymbol{w}\in A^n}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w}) \leq n\sum_{w\in A} f(w)Q(w) \leq \sum_{\boldsymbol{w}\in A^n}\prod_{i=1}^{n} f(w_i),\ \forall A \subset W. \tag{C.6}$$

The proof is similar to that of Theorem 3 in Che et al. (2013b). Before proceeding to the proof, I first introduce some notations and definitions. Let $D_i := \{(w_i, i)|w_i \in W\}$ and $D := \cup_{i=1}^{n} D_i$, where the latter is known as the *disjoint union* of the individual posterior estimate spaces. To simplify notation, we write typical elements of D as w_i instead of (w_i, i). Given an interim allocation rule Q, define a circulation network as (N, E, k, d) as follows. The node set is $N := D \cup W^n \cup \{\circ\}$ consisting of demand nodes D, supply nodes W^n, and a circulation node $\circ$. Directed edges $E \subset N \times N$ specify the pairs of nodes that can carry flows. There are three different kinds of edges:

- Edges from supply nodes to demand nodes: $(\tilde{\boldsymbol{w}}, w_i) \in E$ if $\tilde{w}_i = w_i$.
- Edges from demand nodes to the circulation node $\circ$: $(w_i, \circ) \in E$ for all $w_i \in D$.
- Edges from the circulation node $\circ$ to supply nodes: $(\circ, \boldsymbol{w}) \in E$ for all $\boldsymbol{w} \in W^n$.

Let $d(v, N')$ and $k(v, N')$ denote a lower and upper bound for the total flow from node v to subset $N' \subset N \setminus \{v\}$. There are three different kinds of flow capacities:

- Flow capacities from supply nodes: For each supply node $\boldsymbol{w} \in W^n$, let

$$d(\boldsymbol{w}, N') = \prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w})$$

  if $N' \supset \{w_1, \dots, w_n\}$ or else $d(\boldsymbol{w}, N') = 0$; and let

$$k(\boldsymbol{w}, N') = \prod_{i=1}^{n} f(w_i)$$

  if $N' \cap \{w_1, \dots, w_n\} \neq \emptyset$ or else $k(\boldsymbol{w}, N') = 0$.

- Flow capacities from demand nodes: For each demand node $w_i \in D$, let

$$k(w_i, N') = d(w_i, N') = f(w_i)Q(w_i)$$

  if $\circ \in N'$ or else $k(w_i, N') = d(w_i, N') = 0$.

- Flow capacities from $\circ$: Let $d(\circ, N') = 0$ and $k(\circ, N') = K$ for some $K > 0$ sufficiently large.

A *feasible circulation flow* on (N, E, k, d) is a function $\zeta : E \to \mathbb{R}_+$ that satisfies the *capacity* *constraints*

$$d(v,N') \le \sum_{v'\in N':(v,v')\in E} \zeta(v,v') \le k(v,N'),\ \forall v \in N, \forall N' \subset N \backslash \{v\},$$

and the *flow conservation law*

$$\sum_{v'\in N':(v,v')\in E} \zeta(v,v') = \sum_{v'\in N':(v',v)\in E} \zeta(v',v),\ \forall v \in N.$$

By Theorem 1 in Che et al. (2013b), an interim allocation $\boldsymbol{Q}$ is implementable by an ex post allocation rule $\boldsymbol{q}$ satisfying $\sum_{i=1}^{n} q_i(\boldsymbol{w}) \ge \rho(\boldsymbol{w})$ for all $\boldsymbol{w}$ if and only if there exists a feasible circulation flow for the network (N, E, k, d) defined above. It is easy to verify that for every $v \in N$, $k(v,\cdot)$ and $d(v,\cdot)$ are *paramodular*:

1. $k(v,\cdot)$ is *submodular*: For any $N', N'' \subset N$, $k(v,N')+k(v,N'') \ge k(v,N'\cup N'')+k(v,N'\cap N'')$.

2. $d(v,\cdot)$ is *supermodular*: For any $N', N'' \subset N$, $d(v,N') + d(v,N'') \le d(v,N'\cup N'') + d(v,N'\cap N'')$.

3. $k(v,\cdot)$ and $d(v,\cdot)$ are *compliant*: For any $N', N'' \subset N$, $k(v,N')-d(v,N'') \ge k(v,N'\backslash N'')-d(v,N''\backslash N')$.

Hence, by Theorem 1 in Hassin (1982), a feasible circulation flow $\zeta : E \to \mathbb{R}_+$ exists if and only if

$$\sum_{v\in N\backslash N'} d(v,N') \le \sum_{v\in N'} k(v,N\backslash N'),\ \forall N' \subset N, \tag{C.7}$$

which requires that the sum of lower bounds on the flows entering N' does not exceed the sum of upper bounds on the flows exiting N'.

Necessity: Suppose that the interim allocation rule $\boldsymbol{Q}$ is the reduced form of an ex post allocation rule $\boldsymbol{q}$ satisfying $\sum_{i=1}^{n} q_i(\boldsymbol{w}) \ge \rho(\boldsymbol{w})$ for all $\boldsymbol{w}$. Then, by Theorem 1 in Che et al. (2013b) and Theorem 1 in Hassin (1982), (C.7) holds. Let $N' = \cup_{i=1}^{n}\{(w_i,i)|w_i \in A\} \subset D$, where $A \subset W$ is a measurable set. The right-hand side of (C.7) becomes

$$\begin{aligned}\sum_{v\in N'} k(v, N\backslash N') &= \sum_{w_i\in N'} k(w_i, \circ)\\ &= \sum_{w_i\in N'} f(w_i)Q(w_i)\\ &= n\sum_{w\in A} f(w)Q(w)\end{aligned}$$

and the left-hand side of (C.7) becomes

$$\begin{aligned}\sum_{v\in N'} d(v, N') &= \sum_{\boldsymbol{w}:\{w_i\}\cap N'\neq\emptyset} d(\boldsymbol{w}, N')\\ &= \sum_{\boldsymbol{w}\in A^n}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w}),\end{aligned}$$

which proves the first inequality in (C.6). Let $N' = N\backslash \cup_{i=1}^{n} \{(w_i, i)|w_i \in A\}$. The right-hand side of (C.7) becomes

$$\sum_{v\in N'} k(v, N\backslash N') = \sum_{\boldsymbol{w}\in A^n}\prod_{i=1}^{n} f(w_i)$$

and the left-hand side of (C.7) becomes

$$\begin{aligned}\sum_{v\in N'} d(v, N') &= \sum_{i=1}^{n}\sum_{w_i\in A} d(w_i, \circ)\\ &= n\sum_{w\in A} f(w)Q(w),\end{aligned}$$

which proves the second inequality in (C.6).

Sufficiency: Because ρ is symmetric, by a similar argument to that in the proof of Theorem 7 in Che et al. (2013b), (C.6) holds if and only if

$$\sum_{\boldsymbol{w}\in\prod_i A_i}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w}) \le \sum_{i=1}^{n}\sum_{w_i\in A_i} f(w_i)Q(w_i) \le \sum_{\boldsymbol{w}\in\cup_i(A_i\times W^{n-1})}\prod_{i=1}^{n} f(w_i),\ \forall\prod_{i=1}^{n} A_i \subset W^n. \quad \text{(C.8)}$$

For completeness, I include a proof of this claim in Lemma 44.

Suppose first that $\circ \notin N'$. Let $A_i = N' \cap D_i$ for all i. In this case,

$$
\begin{aligned}
\sum_{v\in N'} d(v, N') &= d(\circ, N' \cap W^n) + \sum_{\boldsymbol{w}\in W^n\setminus N'} d(\boldsymbol{w}, N' \cap D) \\
&= \sum_{\boldsymbol{w}\in W^n\setminus N'} d(\boldsymbol{w}, N' \cap D) \\
&\leq \sum_{\boldsymbol{w}\in W^n\setminus N'} d(\boldsymbol{w}, N' \cap D) \\
&\leq \sum_{\boldsymbol{w}\in \cup_i (A_i\times W^{n-1})} d(\boldsymbol{w}, N' \cap D) \\
&= \sum_{\boldsymbol{w}\in \prod_{i=1}^n A_i} d(\boldsymbol{w}, N' \cap D) \\
&= \sum_{\boldsymbol{w}\in \prod_{i=1}^n A_i} \prod_{i=1}^n f(w_i)\rho(\boldsymbol{w}) \\
&\leq \sum_{i=1}^n \sum_{w_i\in A_i} f(w_i)Q(w_i) + \sum_{w_i\in D\cap N'} k(w_i, \circ) \\
&= \sum_{v\in N'} k(v, N\setminus N').
\end{aligned}
$$

Suppose next $\circ \in N'$. Then if $W^n \not\subset N'$, we have $\sum_{v\in N'} k(v, N\setminus N') \geq k(\circ, N\setminus N') = K > \sum_{v\in N\setminus N'} d(v, N')$ for K sufficiently large. Otherwise, if $W^n \subset N'$, then let $A_i = D_i\setminus N'$ for all i and

$$
\begin{aligned}
\sum_{v\in N'} k(v, N\setminus N') &= \sum_{\boldsymbol{w}\in \cup_i (A_i\times W^{n-1})} k(\boldsymbol{w}, D\setminus N') \\
&= \sum_{\boldsymbol{w}\in A_i\times W^{n-1}} \prod_{i=1}^n f(w_i) \\
&\geq \sum_{i=1}^n \sum_{w_i\in A_i} f(w_i)Q(w_i) \\
&= \sum_{v\in D\setminus N'} d(v, N').
\end{aligned}
$$

Hence, if (C.6) holds, then (C.7) also holds. The conclusion follows by Theorem 1 in Hassin (1982) and Theorem 1 in Che et al. (2013b).

Lemma 44 (C.6) *holds if and only if* (C.8) *holds.*

Proof. Clearly, if (C.8) holds, then (C.6) holds. Suppose that (C.6) holds. I only prove here that the first inequality in (C.6) holds. Virtually the same argument can be applied to prove the second inequality in (C.8).

Suppose, to the contrary, that there exists $\prod_{i=1}^n A_i \subset W^n$ such that the first inequality in (C.8) is violated. Suppose that $\prod_{i=1}^n A_i$ is minimal in the sense that for all proper subsets $\prod_{i=1}^n A_i' \subsetneq \prod_{i=1}^n A_i$, the first inequality in (C.8) holds. Let $\overline{A} := \cup_i A_i$. I want to show that the first inequality in (C.6) is violated for $\overline{A}$, which is a contradiction.

To show this, I show that starting from $\prod_{i=1}^n A_i$, I can construct a finite sequence of sets $\prod_{i=1}^n A_i = \mathcal{S}^1 \subsetneq \mathcal{S}^2 \subsetneq \cdots \subsetneq \mathcal{S}^M = \overline{A}^n$ such that the first inequality in (C.8) is violated for all $\mathcal{S}^m$. The sequence is constructed inductively:

Step 1. Let $\mathcal{S}^1 := \prod_{i=1}^n A_i$.

Step m. If $\mathcal{S}^{m-1} = \overline{A}^n$, then we are done. Otherwise, there exist $j, k \in \{1, \dots, n\}$ such that $B_j := A_j \backslash \mathcal{S}_k^{m-1} \neq \emptyset$ or $B_k := A_k \backslash \mathcal{S}_j^{m-1} \neq \emptyset$. Let $\mathcal{S}^m := (\mathcal{S}_j^{m-1} \cup B_k) \times (\mathcal{S}_k^{m-1} \cup B_j) \times \prod_{i \neq j,k} \mathcal{S}_i^{m-1}$.

Because there are a finite number of agents, the construction stops after a finite number of steps. Next I show that if the first inequality in (C.8) is violated for $\mathcal{S}^m$, then it is also violated for $\mathcal{S}^{m+1}$. Recall that the first inequality in (C.8) is violated for $\prod_{i=1}^n A_i$:

$$\sum_{i=1}^n \sum_{w_i \in A_i} f(w_i) Q(w_i) < \sum_{\boldsymbol{w} \in \prod_{i=1}^n A_i} \prod_{i=1}^n f(w_i) \rho(\boldsymbol{w}) \tag{C.9}$$

Because $\prod_{i=1}^{n} A_i$ is chosen minimally, we have

$$\sum_{i=1}^{n}\sum_{w_i\in A_i} f(w_i)Q(w_i) - \sum_{w_j\in B_j} f(w_j)Q(w_j) - \sum_{w_k\in B_k} f(w_k)Q(w_k)$$
$$\geq \sum_{\boldsymbol{w}\in(A_j\setminus B_j)\times(A_k\setminus B_k)\times\prod_{i\neq j,k} A_i}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w}).$$

Hence,

$$\sum_{w_j\in B_j} f(w_j)Q(w_j) + \sum_{w_k\in B_k} f(w_k)Q(w_k)$$
$$< \sum_{\boldsymbol{w}\in\prod_{i=1}^{n} A_i}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w}) - \sum_{\boldsymbol{w}\in(A_j\setminus B_j)\times(A_k\setminus B_k)\times\prod_{i\neq j,k} A_i}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w}).$$

For $\mathcal{S}^{m+1} = (\mathcal{S}_j^m \cup B_k)\times(\mathcal{S}_k^m \cup B_j)\times\prod_{i\neq j,k}\mathcal{S}_i^m$, we have

$$\begin{aligned}
&\sum_{i=1}^{n}\sum_{w_i\in\mathcal{S}_i^m} f(w_i)Q(w_i) + \sum_{w_j\in B_k} f(w_j)Q(w_j) + \sum_{w_k\in B_j} f(w_k)Q(w_k)\\
&=\sum_{i=1}^{n}\sum_{w_i\in\mathcal{S}_i^m} f(w_i)Q(w_i) + \sum_{w_j\in B_j} f(w_j)Q(w_j) + \sum_{w_k\in B_k} f(w_k)Q(w_k)\\
&< \sum_{\boldsymbol{w}\in\mathcal{S}^m}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w}) + \sum_{\boldsymbol{w}\in A_j\times A_k\times\prod_{i\neq j,k} A_i}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w}) - \sum_{\boldsymbol{w}\in(A_j\setminus B_j)\times(A_k\setminus B_k)\times\prod_{i\neq j,k} A_i}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w})\\
&= \sum_{\boldsymbol{w}\in\mathcal{S}^m}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w}) + \sum_{\boldsymbol{w}\in A_k\times A_j\times\prod_{i\neq j,k} A_i}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w}) - \sum_{\boldsymbol{w}\in(A_k\setminus B_k)\times(A_j\setminus B_j)\times\prod_{i\neq j,k} A_i}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w})\\
&\leq \sum_{\boldsymbol{w}\in\mathcal{S}^m}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w}) + \sum_{\boldsymbol{w}\in(\mathcal{S}_j^m\cup B_k)\times(\mathcal{S}_k^m\cup B_j)\times\prod_{i\neq j,k}\mathcal{S}_i^m}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w})\\
&\quad - \sum_{\boldsymbol{w}\in\mathcal{S}_j^m\times\mathcal{S}_k^m\times\prod_{i\neq j,k}\mathcal{S}_i^m}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w})\\
&= \sum_{\boldsymbol{w}\in\mathcal{S}^{m+1}}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w})
\end{aligned}$$

where the second and the fourth lines hold because ρ is symmetric, and the fifth line holds because $\sum_{\boldsymbol{w}\in\prod_{i=1}^{n} A_i}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w})$ is supermodular over $\prod_{i=1}^{n} A_i$. ■

C.1.2. Solving ($\mathcal{P}'$-α^*)

Recall that the sub-problem ($\mathcal{P}'$-α^*) is

$$V(\alpha^*) := \max_Q \mathbb{E}\left[wQ(w)|\,\alpha^*\right],$$

subject to

$$Y(w) := \int_w^{\overline{\theta}} [H(z|\alpha^*)^{n-1} - Q(z)]h(z|\alpha^*)\mathrm{d}z \geq 0, \ \forall w \in [\underline{\theta}, \overline{\theta}]. \tag{F$'$}$$

$$Q(w) \text{ is non-decreasing in } w, \tag{MON}$$

$$\mathbb{E}\left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)}Q(w)\bigg|\,\alpha^*\right] \leq C'(\alpha^*). \tag{IA$'$}$$

For brevity, denote $w(0,\alpha^*)$ by $\underline{w}$, $w(1,\alpha^*)$ by $\overline{w}$, $h(w|\alpha^*)$ by $h(w)$, $H(w|\alpha^*)$ by $H(w)$ and $H_\alpha(w|\alpha^*)$ by $H_\alpha(w)$. Let $X(w) := \int_0^w H_\alpha(z)Q(z)\mathrm{d}z$ for all $w \in [\underline{w}, \overline{w}]$. Then this is a control problem with state variables X, Y and Q, and a control variable $a \geq 0$. The evolution of the state variables is governed by

$$X'(w) = H_\alpha(w)Q(w), \tag{C.10}$$

$$Y'(w) = -[H(w)^{n-1} - Q(w)]h(w), \tag{C.11}$$

$$Q'(w) = a(w), \tag{C.12}$$

where the last equality holds if $Q(w)$ is differentiable at w. The non-negativity constraint for a guarantees that Q is non-decreasing. This implies some regularity on Q, but still leaves some problems to apply control theory directly. First, we have to allow for (upward) jumps in the state variable Q. Second, Q is not guaranteed to be piecewise continuous and piecewise continuously differentiable.

These problems can be circumvented by solving the maximization problem under the additional restriction that Q is Lipschitz continuous with global Lipschitz constant K:

$$Q \in \mathcal{L}^K := \left\{ Q : W \to [0,1] | |Q(z) - Q(z')| \leq K|z - z'| \; \forall z, z' \in [0,1] \right\}.$$

We define the maximization problem ($\mathcal{P}^K$-α^*) as ($\mathcal{P}'$-α^*) subject to the additional constraint $Q \in \mathcal{L}^K$.

We say that Q is a *feasible solution* of ($\mathcal{P}'$-α^*) if it satisfies (MON), (F′) and (IA′), and an *optimal solution* of ($\mathcal{P}'$-α^*) if it maximizes $\mathbb{E}\left[wQ(w) | \, \alpha^* \right]$ subject to (MON), (F′) and (IA′). Similarly, we say that $Q \in \mathcal{L}^K$ a *feasible solution* of ($\mathcal{P}^K$-α^*) if it satisfies (MON), (F′) and (IA′), and $Q \in \mathcal{L}^K$ an *optimal solution* of ($\mathcal{P}^K$-α^*) if it maximizes $\mathbb{E}\left[wQ(w) | \, \alpha^* \right]$ subject to (MON), (F′) and (IA′).

Lemma 45 in Appendix C.1.2 shows that an optimal solution of ($\mathcal{P}'$-α^*) exists, and for every $K > 0$, an optimal solution of ($\mathcal{P}^K$-α^*) exists. Lemma 46 in Appendix C.1.2 shows that there exists an optimal solution of ($\mathcal{P}'$-α^*), which is the pointwise limit of the optimal solutions of ($\mathcal{P}^K$-α^*).

The rest of Appendix C.1.2 is organized as follows. Appendix C.1.2 introduces and proves Lemmas 45 and 46. Appendix C.1.2 gives the necessary conditions that an optimal solution of ($\mathcal{P}^K$-α^*) must satisfy. Appendix C.1.2 proves Theorem 15. Appendix C.1.2 proves Lemma 15.

Existence of optimal solutions

Before introducing and proving Lemmas 45 and 46, I first introduce some notations. I abuse notation a bit and let h denote the probability measure on W corresponding to $H(w)$. In what follows, let $L_2(h)$ denote the set of measurable functions whose absolute value raised to the 2nd power has finite integral. For brevity, denote $L_2(h)$ by L_2. Because L_2 is the dual of L_2 under the duality $\langle f, g \rangle = \mathbb{E}_w \left[f(w)g(w) | \alpha = \alpha^* \right]$, topologize L_2 with its weak*, or $\sigma(L_2, L_2)$, topology.

Lemma 45 *The following two statements are true.*

1. *An optimal solution of* ($\mathcal{P}'$-α^*) *exists.*

2. *For every $K > 0$, an optimal solution of $(\mathcal{P}^K\text{-}\alpha^*)$ exists.*

Proof. The proof is based on Mierendorff (2009).

1. Let $\{Q^\nu\}$ be a sequence of feasible solutions of $(\mathcal{P}'\text{-}\alpha^*)$ such that

$$\int_{\underline{w}}^{\overline{w}} zQ^\nu(z)h(z)\mathrm{d}z \to V(\alpha^*).$$

By Helly's selection theorem, there exists a subsequence $\{Q^{\nu_\kappa}\}$ and a non-decreasing function Q such that Q^{ν_κ} converges pointwise to Q. Let $\mathscr{D}$ collect all $Q : W \to [0,1]$ that satisfies (F′), (MON) and (IA′). Consider $\mathscr{D}$ as a subset of L_2. Recall that m is the probability measure on corresponding to $H(z)$. Then $\mathscr{D}$ is $\sigma(L_2, L_2)$ compact by a proof similar to that of Lemma 5.4 in Border (1991) and Lemma 8 in Mierendorff (2011). Therefore, after taking subsequences again, Q^{ν_κ} converges to Q in $\sigma(L_2, L_2)$ topology and $Q \in \mathscr{D}$. Because $z \in L_2$ and $h \in L_2$, the weak convergence of $\{Q^{\nu_\kappa}\}$ implies that

$$\int_{\underline{w}}^{\overline{w}} zQ(z)h(z)\mathrm{d}z = V(\alpha^*).$$

2. Let $\{Q^\nu\}$ be a sequence of feasible solutions of $(\mathcal{P}^K\text{-}\alpha^*)$ such that

$$\int_{\underline{w}}^{\overline{w}} zQ^\nu(z)h(z)\mathrm{d}z \to V^K(\alpha^*).$$

After taking subsequences, we can assume that Q^ν converges to Q pointwise and in $\sigma(L_2, L_2)$ topology, and $Q \in \mathscr{D}$ as in part 1. Because $Q^\nu \in \mathcal{L}^K$, for all $z, z' \in W$,

$$|Q(z) - Q(z')| = \lim_{\nu\to\infty} |Q^\nu(z) - Q^\nu(z')| \le K|z - z'|.$$

Hence, $Q \in \mathcal{L}^K$.

∎

Lemma 46 *Let $\{Q^K\}$ be a sequence of optimal solutions of ($\mathcal{P}^K$-α^*) where $K \to \infty$. Then there exists a feasible solution Q of ($\mathcal{P}'$-α^*) and a subsequence Q^{K_ν} such that Q^{K_ν} converges to Q for almost every $w \in W$. Furthermore, Q is optimal, i.e.,*

$$\int_{\underline{w}}^{\overline{w}} zQ(z)h(z)\mathrm{d}z = V(\alpha^*).$$

Proof. The proof is based on Reid (1968) and Mierendorff (2009). After taking a subsequence, we can assume that Q^K converges pointwise to a feasible solution $\hat{Q}$ of ($\mathcal{P}'$-α^*) (see the proof of Lemma 45). To show the optimality of $\hat{Q}$, let Q be an optimal solution of ($\mathcal{P}'$-α^*). We can extend Q to $\mathbb{R}$ by setting $Q(z) := 0$ for $z < \underline{w}$ and $Q(z) := 1$ for $z > \overline{w}$. Define $Q_d : \mathbb{R} \to [0, 1]$ as

$$Q_d(z) := \frac{1}{d} \int_{z-d}^{z} Q(\zeta)\mathrm{d}\zeta, \ \forall z \in \mathbb{R}.$$

By the Lebesgue differentiation theorem (see, e.g., Theorem 3.21 in Folland (1999)), $Q_d(z) \to Q(z)$ for almost every $z \in W$ as $d \to 0$. Because Q is non-decreasing and $Q(z) \in [0, 1]$ for all z, Q_d is non-decreasing, $Q_d \leq Q$, and $Q_d(z) \in [0, 1]$ for all z. Furthermore, $Q_d \in \mathcal{L}^{\frac{1}{d}}$: For all $z > z'$,

$$\begin{aligned}
0 \leq Q_d(z) - Q_d(z') =& \frac{1}{d}\left(\int_{z-d}^{z} Q(\zeta)\mathrm{d}\zeta - \int_{z'-d}^{z'} Q(\zeta)\mathrm{d}\zeta\right) \\
=& \frac{1}{d}\left(\int_{z'}^{z} Q(\zeta)\mathrm{d}\zeta - \int_{z'-d}^{z-d} Q(\zeta)\mathrm{d}\zeta\right) \\
\leq& \frac{1}{d}\int_{z'}^{z} Q(\zeta)\mathrm{d}\zeta \\
\leq& \frac{1}{d}(z - z').
\end{aligned}$$

Finally, Q_d satisfies (F′) because $Q_d \leq Q$ and Q satisfies (F′).

Define $\tilde{Q}_d := Q_d$ if $-\int_{\underline{w}}^{\overline{w}} H_{\alpha_i}(z)Q_d(z)\mathrm{d}z \leq C'(\alpha^*)$ and otherwise $\tilde{Q}_d := \beta_d Q_d + (1-\beta_d)/n$, where

$$\beta_d = \frac{C'(\alpha^*)}{-\int_{\underline{w}}^{\overline{w}} H_{\alpha_i}(z)Q_d(z)\mathrm{d}z}.$$

Then by Lemma 43, $-\int_{\underline{w}}^{\overline{w}} H_{\alpha_i}(z)\tilde{Q}_d(z)\mathrm{d}z \leq C'(\alpha^*)$. Thus, $\tilde{Q}_d$ is a feasible solution of ($\mathcal{P}^K$-α^*), where $K = 1/d$. Because $\beta_d \to 0$, $\tilde{Q}_d \to Q$ almost everywhere as $d \to 0$. By the dominated convergence theorem,

$$\int_{\underline{w}}^{\overline{w}} z\tilde{Q}_d(z)h(z)\mathrm{d}z \to \int_{\underline{w}}^{\overline{w}} zQ(z)h(z)\mathrm{d}z, \text{ as } d \to 0$$

and

$$\int_{\underline{w}}^{\overline{w}} zQ^K(z)h(z)\mathrm{d}z \to \int_{\underline{w}}^{\overline{w}} z\hat{Q}(z)h(z)\mathrm{d}z, \text{ as } K \to \infty.$$

Let $d = 1/K$. Then for all K, $\tilde{Q}_d$ is a feasible solution of ($\mathcal{P}^K$-α^*) and therefore

$$\int_{\underline{w}}^{\overline{w}} z\tilde{Q}_d(z)h(z)\mathrm{d}z \leq \int_{\underline{w}}^{\overline{w}} zQ^K(z)h(z)\mathrm{d}z.$$

Hence,

$$\int_{\underline{w}}^{\overline{w}} z\hat{Q}(z)h(z)\mathrm{d}z = \int_{\underline{w}}^{\overline{w}} zQ(z)h(z)\mathrm{d}z.$$

This completes the proof. ■

Solving ($\mathcal{P}^K$-α^*)

The problem ($\mathcal{P}^K$-α^*) can be summarized as follows:

$$\max_{X,Y,Q,a} \int_{\underline{w}}^{\overline{w}} zQ(z)h(z)\mathrm{d}z, \qquad (\mathcal{P}^K\text{-}\alpha^*)$$

subject to

$$X'(z) = H_\alpha(z)Q(z), \tag{C.10}$$

$$Y'(z) = -[H(z)^{n-1} - Q(z)]h(z), \tag{C.11}$$

$$Q'(z) = a(z), \tag{C.12}$$

$$X(\underline{w}) = 0,\ \ X(\overline{w}) \geq -C'(\alpha^*), \tag{C.13}$$

$$Y(\underline{w}) = 0,\ \ Y(\overline{w}) = 0, \tag{C.14}$$

$$Q(\underline{w}) \geq 0,\ \ Q(\overline{w}) \leq 1, \tag{C.15}$$

$$0 \leq a(z) \leq K, \tag{C.16}$$

$$Y(z) \geq 0. \tag{C.17}$$

We say that some property holds virtually everywhere if the property is fulfilled at all z except for a countable number of z's. We use the following abbreviation for "virtually everywhere": $v.e.$ By Theorem 6.7.15 in Seierstad and Sydsæter (1987), we have

Lemma 47 *Let (X, Y, Q, a) be an admissible pair that solves $(\mathcal{P}^K\text{-}\alpha^*)$. Then there exist a number λ_0, vector functions $(\lambda_X, \check{\lambda}_Y, \lambda_Q)$ and $(\underline{\eta}_a, \overline{\eta}_a)$, and a non-decreasing function η_Y, all having one-*

sided limits everywhere, such that the following condition holds:

$$\lambda_0 = 0 \text{ or } \lambda_0 = 1, \tag{C.18}$$

$$(\lambda_0, \lambda_X(z), \check{\lambda}_Y(z), \lambda_Q(z), \eta_Y(\overline{w}) - \eta_Y(\underline{w})) \neq 0, \ \forall z, \tag{C.19}$$

$$\lambda_Q(z)a(z) \geq \lambda_Q(z)a, \ \forall a \in (0, K), v.e. \tag{C.20}$$

$$\lambda_Q(z) - \overline{\eta}_a(z) + \underline{\eta}_a(z) = 0, v.e. \tag{C.21}$$

$$\eta_Y \text{ is constant on any interval where } Y > 0. \tag{C.22}$$

$$\lambda_X \text{ and } \lambda_Q \text{ are continuous.} \tag{C.23}$$

$$\lambda'_X(z) = 0, \ v.e. \tag{C.24}$$

$$\lambda'_Q(z) = -\left[\lambda_0 z + \lambda_X(z)\frac{H_\alpha(z)}{h(z)} + \check{\lambda}_Y(z)\right] h(z) + \eta_Y(z)h(z), \ v.e. \tag{C.25}$$

$$\check{\lambda}_Y(z) + \eta_Y(z) \text{ is continuous,} \tag{C.26}$$

$$\check{\lambda}'_Y(z) + \eta'_Y(z) = 0, \ v.e. \tag{C.27}$$

$$\lambda_X(\overline{w}) \geq 0 (= 0 \text{ if } X(\overline{w}) > -C'(\alpha^*)), \tag{C.28}$$

$$\lambda_Q(\overline{w}) \leq 0 (= 0 \text{ if } Q(\overline{w}) < 1), \tag{C.29}$$

$$\lambda_Q(\underline{w}) \leq 0 (= 0 \text{ if } Q(\underline{w}) > 0). \tag{C.30}$$

$$\underline{\eta}_a(z) \geq 0 (= 0 \text{ if } a(z) > 0), \tag{C.31}$$

$$\overline{\eta}_a(z) \geq 0 (= 0 \text{ if } a(z) < K). \tag{C.32}$$

In what follows, I assume that (X, Y, Q, a) is an admissible pair that solves ($\mathcal{P}^K$-α^*) and $(X, Y, Q, a, \lambda_0, \lambda_X, \check{\lambda}_Y, \lambda_Q, \underline{\eta}_a, \overline{\eta}_a, \eta_Y)$ satisfy the conditions in Lemma 47. I begin the analysis by simplifying the conditions in Lemma 47.

Because λ_X is continuous and $\lambda'_X(z) = 0$ virtually everywhere, $\lambda_X(z)$ is constant on $[\underline{w}, \overline{w}]$. I abuse notation a bit and denote this constant by λ_X. Then (C.28) is equivalent to

$$\lambda_X \geq 0 (= 0 \text{ if } X(\overline{w}) > -C'(\alpha^*)).$$

Similarly, because $\check{\lambda}_Y + \eta_Y$ is continuous and $\check{\lambda}'_Y(z) + \eta'_Y(z) = 0$ virtually everywhere, $\check{\lambda}_Y(z) + \eta_Y(z)$ is constant on $[\underline{w}, \overline{w}]$. We can assume without loss of generality that $\check{\lambda}_Y(z) + \eta_Y(z) = 0$ for all $z \in [\underline{w}, \overline{w}]$. Let $\lambda_Y := 2\check{\lambda}_Y$. Then $\eta_Y = -\lambda_Y/2$ and condition (C.22) is equivalent to

$$\lambda_Y(z) \text{ is constant on any interval where } Y(z) > 0, \tag{C.33}$$

and (C.25) is equivalent to

$$\lambda'_Q(z) = -\left[\lambda_0 z + \lambda_X(z)\frac{H_{\alpha_i}(z)}{h(z)} + \lambda_Y(z)\right] h(z), \text{ v.e.}$$

Furthermore, η_Y is non-decreasing if and only if λ_Y is non-increasing. Because λ_Y has one-sided limits everywhere, we can assume without loss of generality that $\lambda_Y(\underline{w}) = \lim_{z\to\underline{w}} \lambda_Y(z)$ and $\lambda_Y(\overline{w}) = \lim_{z\to\overline{w}} \lambda_Y(z)$. Finally, (C.20), (C.21), (C.31) and (C.32) can be simplified to for virtually all $z \in (\underline{w}, \overline{w})$: If $0 < a(z) < K$, then $\lambda_Q(z) = \overline{\eta}_a(z) = \underline{\eta}_a(z) = 0$. If $a(z) = 0$, then $\overline{\eta}_a(z) = 0$ and $-\underline{\eta}_a(z) = \lambda_Q(z) \leq 0$. If $a(z) = K$, then $\underline{\eta}_a(z) = 0$ and $\overline{\eta}_a(z) = \lambda_Q(z) \geq 0$.

Then the conditions in Lemma 47 can be simplified as follows:

Corollary 10 *Let (X, Y, Q, a) be an admissible pair for ($\mathcal{P}^K$-α^*). If (X, Y, Q, a) is optimal, then there exist a constant λ_X, a continuous and piecewise continuously differentiable function λ_Q, and* *a non-increasing function λ_Y such that the following holds:*

$$\lambda_X \geq 0 \ (= 0 \ \text{if} \ X(\bar{w}) > -C'(\alpha^*)). \tag{C.34}$$

$$\lambda_Q'(z) = -\left[z + \lambda_X \frac{H_{\alpha_i}(z)}{h(z)} + \lambda_Y(z)\right] h(z), \ \text{v.e.} \tag{C.35}$$

$$\lambda_Y \text{ is constant on any interval where } Y > 0. \tag{C.36}$$

$$\lambda_Q(\overline{w}) \leq 0 (= 0 \ \text{if} \ Q(\overline{w}) < 1). \tag{C.37}$$

$$\lambda_Q(\underline{w}) \leq 0 (= 0 \ \text{if} \ Q(\underline{w}) > 0). \tag{C.38}$$

$$a(z) = \begin{cases} = 0 & \text{if} \ \lambda_Q(z) \leq 0, \\ \in [0, K] & \text{if} \ \lambda_Q(z) = 0, \quad \text{v.e.} \\ = K & \text{if} \ \lambda_Q(z) \geq 0. \end{cases} \tag{C.39}$$

Proof. We prove Corollary 10 by proving the following two lemmas.

Lemma 48 $\lambda_Q(\underline{w}) = \lambda_Q(\overline{w}) = 0.$

Proof. By the transversality condition (C.30), $\lambda_Q(\underline{w}) \leq 0$ and equality holds if $Q(\underline{w}) > 0$. Suppose, to the contrary, that $\lambda_Q(\underline{w}) < 0$. Then $Q(\underline{w}) = 0$. By continuity there exists $\delta > 0$ such that $\lambda_Q(z) < 0$ for all $z \in (\underline{w}, \underline{w} + \delta)$. Hence, by (C.20), $a(z) = 0$ for all $z \in (\underline{w}, \underline{w} + \delta)$. This implies that $Q(z) = 0$ for all $z \in (\underline{w}, \underline{w} + \delta)$. Let $z \in (\underline{w}, \underline{w} + \delta)$, then

$$0 = Y(\underline{w}) = \int_{\underline{w}}^{z} H(\zeta)^{n-1} h(z) \mathrm{d}z + Y(z) > Y(z),$$

a contradiction. Hence, $\lambda_Q(\underline{w}) = 0$. A similar argument proves that $\lambda_Q(\overline{w}) = 0$. ■

Lemma 49 *(Non-triviality)* $\lambda_0 = 1$.

Proof. Suppose, to the contrary, that $\lambda_0 = 0$. Then

$$\lambda_Q'(z) = -\left[\lambda_X \frac{H_\alpha(z)}{h(z)} + \lambda_Y(z)\right] h(z), \ \text{v.e.}$$

Hence,

$$\lambda_Q(\overline{w}) = \lambda_Q(\underline{w}) - \int_{\underline{w}}^{\overline{w}} \left[\lambda_X \frac{H_\alpha(z)}{h(z)} + \lambda_Y(z) \right] h(z) \mathrm{d}z,$$
$$= \lambda_Q(\underline{w}) - \int_{\underline{w}}^{\overline{w}} \lambda_Y(z) h(z) \mathrm{d}z.$$

Because $\lambda_Q(\underline{w}) = \lambda_Q(\overline{w}) = 0$, we have

$$\int_{\underline{w}}^{\overline{w}} \lambda_Y(z) h(z) \mathrm{d}z = 0.$$

Because λ_Y is non-decreasing, it must be that $\lambda_Y(\underline{w}) \geq 0$ and $\lambda_Y(\overline{w}) \leq 0$. Suppose that $\lambda_X > 0$. Then because $H_{\alpha_i}(z)/h(z)$ is strictly decreasing and $\lambda_Y(z)$ is non-increasing, $\lambda_Q(H^{-1}(\cdot))$ is strictly convex. Hence, $\lambda_Q(z) < 0$ for all $z \in (\underline{w}, \overline{w})$, and therefore $a(z) = 0$ for all $z \in (\underline{w}, \overline{w})$. That is, Q is constant. However, if Q is constant, then $X(\overline{w}) = 0 > -C'(\alpha^*)$, a contradiction to that $\lambda_X > 0$. Hence, $\lambda_X = 0$. Then

$$\lambda_Q(z) = -\int_{\underline{w}}^{z} \lambda_Y(\zeta) h(\zeta) \mathrm{d}\zeta.$$

Suppose that $\lambda_Y(\underline{w}) = \lambda_Y(\overline{w}) = 0$, then $\lambda_Y(z) = 0$ for all $z \in (\underline{w}, \overline{w})$. Hence, $\lambda_Q(z) = 0$ for all $z \in (\underline{w}, \overline{w})$ and $\eta_Y(\overline{w}) - \eta_Y(\underline{w}) = -\lambda_Y(\overline{w})/2 +$
$lambda_Y(\underline{w})/2 = 0$. Then

$$(\lambda_0, \lambda_X(z), \lambda_Y(z), \lambda_Q(z), \eta_Y(\overline{w}) - \eta_Y(\underline{w})) = 0, \ \forall z,$$

which is a contradiction to (C.19). Hence, $\lambda_Y(\underline{w}) > 0$ and $\lambda_Y(\overline{w}) < 0$. Thus, $\lambda_Q(z) < 0$ for all $z \in (\underline{w}, \overline{w})$ and therefore Q is constant. Hence, $Y(z) > 0$ for all $z \in [\underline{w}, \overline{w}]$. This, by (C.33), implies that λ_Y is constant on $(\underline{w}, \overline{w})$, which is a contradiction to the fact that that $\lambda_Y(\underline{w}) > 0$ and $\lambda_Y(\overline{w}) < 0$. Hence, $\lambda_0 = 1$. ■

This completes the proof of Corollary 10. ■

Before proceeding, I first introduce some notations and proves two technical lemmas (Lemmas 50 and 52) that will be useful for later proof. For the ease of notation, I suppress the dependence of φ^{λ_X}, J^{λ_X}, $\overline{\varphi}^{\lambda_X}$ and $\overline{J}^{\lambda_X}$ on α^*. For each $w \in W$, define

$$m_Y(w) := -\int_{\underline{w}}^{w} \lambda_Y(z)h(z)\mathrm{d}z.$$

It follows from (C.35) that for any $\underline{z}, \overline{z} \in W$ and $\underline{z} < \overline{z}$, we have

$$\begin{aligned}\lambda_Q(\overline{z}) =& \lambda_Q(\underline{z}) - \int_{\underline{z}}^{\overline{z}} \left[z + \lambda_X \frac{H_{\alpha_i}(z)}{h(z)} + \lambda_Y(z)\right] h(z)\mathrm{d}z,\\ =& \lambda_Q(\underline{z}) - \int_{\underline{z}}^{\overline{z}} \left[\varphi^{\lambda_X}(H(z)) + \lambda_Y(z)\right] h(z)\mathrm{d}z.\end{aligned}$$

If $\underline{z} = \underline{w}$, then

$$\lambda_Q(\overline{z}) = \lambda_Q(\underline{w}) - J^{\lambda_X}(H(\overline{z})) + m_Y(\overline{z}). \tag{C.40}$$

Hence, for virtually all $z \in (\underline{w}, \overline{w})$,

$$a(z) = \begin{cases} = 0 & \text{if } \lambda_Q(\underline{w}) + m_Y(z) \leq J^{\lambda_X}(H(z)),\\ \in [0, K] & \text{if } \lambda_Q(\underline{w}) + m_Y(z) = J^{\lambda_X}(H(z)),\\ = K & \text{if } \lambda_Q(\underline{w}) + m_Y(z) \geq J^{\lambda_X}(H(z)).\end{cases}$$

Lemma 50 *For all $t \in [0, 1]$,*

$$\lambda_Q(\underline{w}) + m_Y(H^{-1}(t)) \geq \overline{J}^{\lambda_X}(t).$$

Proof. The proof of Lemma 50 uses the following lemma.

Lemma 51 (Reid) *Suppose that $\lambda_Q(\underline{w}) + m_Y(H^{-1}(t)) = J^{\lambda_X}(t)$ for $t \in \{\underline{t}, \bar{t}\}$. Let $a, b \in \mathbb{R}$ and $l(t) = a + bt$. If $J^{\lambda_X}(t) \geq l(t)$ for all $t \in [\underline{t}, \bar{t}]$, then*

$$\lambda_Q(\underline{w}) + m_Y(H^{-1}(t)) \geq l(t),\ \forall t \in [\underline{t}, \bar{t}].$$

Proof. Suppose, to the contrary, that $\lambda_Q(\underline{w}) + m_Y(H^{-1}(t)) < l(t)$ for some $t \in (\underline{t}, \bar{t})$. Then by continuity there exist $\varepsilon > 0$ and $t_1, t_2 \in (\underline{t}, \bar{t})$ such that $\underline{t} < t_1 < t_2 < \bar{t}$, $\lambda_Q(\underline{w}) + m_Y(H^{-1}(\tau)) < l(\tau) - \varepsilon$ for $\tau \in (t_1, t_2)$, and

$$\lambda_Q(\underline{w}) + m_Y(H^{-1}(t_1)) = l(t_1) - \varepsilon,$$
$$\lambda_Q(\underline{w}) + m_Y(H^{-1}(t_2)) = l(t_2) - \varepsilon.$$

On the one hand, this implies that $\lambda_Y((H^{-1}(\cdot)) = -m'_Y(H^{-1}(\cdot))$ cannot be constant on (t_1, t_2).

On the other hand, $\lambda_Q(\underline{w}) + m_Y(H^{-1}(\tau)) < l(\tau) - \varepsilon < J^{\lambda_X}(\tau)$ for $\tau \in (t_1, t_2)$. Hence, $a(H^{-1}(\tau)) = 0$ for $\tau \in (t_1, t_2)$, which implies that $Y(H^{-1}(\tau)) > 0$ on the interval (t_1, t_2). To see this, note that $Y'(z) = Q(z) - H^{n-1}(z)$ is strictly decreasing if Q is constant. Hence, Y is strictly concave on $(H^{-1}(t_1), H^{-1}(t_2))$. For any $\tau \in (t_1, t_2)$ there exists $\lambda \in (0, 1)$ such that $H^{-1}(\tau) = \lambda H^{-1}(t_1) + (1 - \lambda)H^{-1}(t_2)$. By strict concavity, $Y(H^{-1}(\tau)) > \lambda Y(H^{-1}(t_1)) + (1 - \lambda)Y(H^{-1}(t_2)) \geq 0$. By (C.36), $Y(H^{-1}(\cdot)) > 0$ on (t_1, t_2) implies that $\lambda_Y(H^{-1}(\cdot))$ is constant on (t_1, t_2), a contradiction. ■

By (C.40) and Lemma 48, $\lambda_Q(\underline{w}) + m_Y(H^{-1}(0)) = 0 = J^{\lambda_X}(0)$ and $\lambda_Q(\underline{w}) + m_Y(H^{-1}(1)) = J^{\lambda_X}$. Then, by Lemma 51, Lemma 50 holds. ■

Lemma 52 *If $K > \overline{K} := \max_{z \in W}(n-1)H(z)^{n-2}h(z)$, then*

$$\lambda_X \leq \bar{\lambda}_X := \left[\min_{z \in W} \frac{\partial}{\partial z}\left[-\frac{H_\alpha(z)}{h(z)}\right]\right]^{-1}.$$

Proof. The proof of Lemma 52 uses Lemmas 53, 54 and 55.

Lemma 53 (interior solution) *Suppose that $a(z) \in (0, K)$ for $z \in (\underline{z}, \overline{z})$, then $\lambda_Y(z) = -\varphi^{\lambda_X}(H(z))$ for virtually every $z \in (\underline{z}, \overline{z})$.*

Proof. If $a(z) \in (0, K)$ for $z \in (\underline{z}, \overline{z})$, then $\lambda_Q(\underline{w}) + m_Y(z) = J^{\lambda_X}(H(z))$ for virtually every

$z \in (\underline{z}, \overline{z})$. Differentiating this equality with respect to z yields for virtually every $z \in (\underline{z}, \overline{z})$:

$$-\lambda_Y(z)h(z) = \varphi^{\lambda_X}(H(z))h(z),$$

Because $h > 0$, $-\lambda_Y(z) = \varphi^{\lambda_X}(H(z))$ for virtually every $z \in (\underline{z}, \overline{z})$. ∎

Lemma 54 (constant Q) *Suppose that $a(z) = 0$ on $(\underline{z}, \overline{z})$ with $\underline{z} < \overline{z}$ and let $(\underline{z}, \overline{z})$ be chosen maximally. Then*

$$\lambda_Q(z) = 0,$$
$$\lambda_Q(\underline{w}) + m_Y(z) = J^{\lambda_X}(H(z)),$$

for $z = \underline{z}$ if $\underline{z} > \underline{w}$, and $z = \overline{z}$ if $\overline{z} < \overline{w}$. Furthermore,

$$\varphi^{\lambda_X}(H(\underline{z})) + \lambda_Y(\underline{z}^-) \geq 0, \text{ if } \underline{z} > \underline{w},$$
$$\varphi^{\lambda_X}(H(\overline{z})) + \lambda_Y(\overline{z}^+) \leq 0, \text{ if } \overline{z} < \overline{w}.$$

Proof. Because $a(z) = 0$ on $(\underline{z}, \overline{z})$, then

$$\lambda_Q(\underline{w}) + m_Y(z) \geq J^{\lambda_X}(H(z)), \text{ v.e. } z \in (\underline{z}, \overline{z}).$$

Suppose that $\underline{z} > \underline{w}$ and let $S_- := \{z < \underline{z} | a(z) > 0\}$. Since $(\underline{z}, \overline{z})$ is chosen maximally, $Q(z) < Q(\underline{z})$ for all $z < \underline{z}$. Furthermore, because Q is absolutely continuous, $S_- \cap [\underline{z} - \delta, \underline{z}]$ has positive measure for every $\delta > 0$. Hence, there exists a sequence $\{z_k\} \in S_-$ converging to $\underline{z}$ with $\lambda_Q(\underline{w}) + m_Y(z_k) \geq J^{\lambda_X}(H(z_k))$ for all k. By continuity, if $\underline{z} > \underline{w}$, then $\lambda_Q(\underline{w}) + m_Y(\underline{z}) = J^{\lambda_X}(H(\underline{z}))$, and therefore $\lambda_Q(\underline{z}) = \lambda_Q(\underline{w}) + m_Y(\underline{z}) - J^{\lambda_X}(H(\underline{z}))$. A similar argument proves that $\lambda_Q(\overline{z}) = 0$ and $\lambda_Q(\overline{w}) + m_Y(\overline{z}) = J^{\lambda_X}(H(\overline{z}))$ if $\overline{z} < \overline{w}$.

If $\underline{z} > \underline{w}$, then for virtually all $z \in S_-$,

$$
\begin{aligned}
0 &= \lambda_Q(\underline{z}) \\
&= \lambda_Q(z) - \int_z^{\underline{z}} \left[\varphi^{\lambda_X}(H(\zeta)) + \lambda_Y(\zeta)\right] h(\zeta)\mathrm{d}\zeta \\
&\geq - \int_z^{\underline{z}} \left[\varphi^{\lambda_X}(H(\zeta)) + \lambda_Y(\zeta)\right] h(\zeta)\mathrm{d}\zeta.
\end{aligned}
$$

Thus, there exists a sequence $\{z_k\} \in S_-$ converging to $\underline{z}$ such that

$$\int_{z_k}^{\underline{z}} \left[\varphi^{\lambda_X}(H(\zeta)) + \lambda_Y(\zeta)\right] h(\zeta)\mathrm{d}\zeta \geq 0, \ \forall k.$$

Hence,

$$\varphi^{\lambda_X}(H(\underline{z})) + \lambda_Y(\underline{z}^-) \geq 0, \text{ if } \underline{z} > \underline{w}.$$

A similar argument proves that

$$\varphi^{\lambda_X}(H(\overline{z})) + \lambda_Y(\overline{z}^+) \leq 0, \text{ if } \overline{z} < \overline{w}.$$

■

Lemma 55 (a(z)=K) *Suppose that $a(z) = K$ on $(\underline{z}, \overline{z})$ with $\underline{z} < \overline{z}$ and let $(\underline{z}, \overline{z})$ be chosen maximally. Then*

$$
\begin{aligned}
\lambda_Q(z) &= 0, \\
\lambda_Q(\underline{w}) + m_Y(z) &= J^{\lambda_X}(H(z)),
\end{aligned}
$$

for $z = \underline{z}$ if $\underline{z} > \underline{w}$, and $z = \overline{z}$ if $\overline{z} < \overline{w}$. Furthermore,

$$
\begin{aligned}
\varphi^{\lambda_X}(H(\underline{z})) + \lambda_Y(\underline{z}^-) &\leq 0, \ \textit{if } \underline{z} > \underline{w}, \\
\varphi^{\lambda_X}(H(\overline{z}))^{\lambda_X} + \lambda_Y(\overline{z}^+) &\geq 0, \ \textit{if } \overline{z} < \overline{w},
\end{aligned}
$$

and

$$\varphi^{\lambda_X}(H(\underline{z})) + \lambda_Y(\underline{z}^+) \leq 0, \text{ if } \underline{z} = \underline{w},$$
$$\varphi^{\lambda_X}(H(\overline{z})) + \lambda_Y(\overline{z}^-) \geq 0, \text{ if } \overline{z} = \overline{w}.$$

Proof. The proofs for the case $\underline{z} > \underline{w}$ and the case $\overline{z} < \overline{w}$ are very similar to that of Lemma 54 and neglected here. We now show that the third inequality holds for $\underline{z} = \underline{w}$. Note that by the transervasality condition, $\lambda_Q(\underline{w}) \leq 0$. For virtually all $z \in (\underline{w}, \overline{z})$,

$$\begin{aligned} 0 &\leq \lambda_Q(z) \\ &= \lambda_Q(\underline{w}) - \int_{\underline{w}}^{z} \left[\varphi^{\lambda_X}(H(\zeta)) + \lambda_Y(\zeta)\right] h(\zeta)\mathrm{d}\zeta \\ &\leq - \int_{\underline{w}}^{z} \left[\varphi^{\lambda_X}(H(\zeta)) + \lambda_Y(\zeta)\right] h(\zeta)\mathrm{d}\zeta, \end{aligned}$$

That is,

$$\int_{\underline{w}}^{z} \left[\varphi(H(\zeta)) + \lambda_Y(\zeta)\right] h(\zeta)\mathrm{d}\zeta \leq 0, \text{ v.e. } z \in (\underline{w}, \overline{z}).$$

Let z goes to $\underline{w}$, and this proves the third inequality for $\underline{z} = \underline{w}$. A similar argument proves the fourth inequality for $\overline{z} = \overline{w}$. ∎

Suppose, to the contrary, that $\lambda_X > \overline{\lambda}_X$. Then $\varphi^{\lambda_X}(H(z))$ is strictly decreasing. Suppose that there exists an interval $(\underline{z}, \overline{z})$ such that $a(z) \in (0, K)$ for $z \in (\underline{z}, \overline{z})$. Then, by Lemma 53, $\lambda_Y(z) = -\varphi^{\lambda_X}(H(z))$ for virtually every $z \in (\underline{z}, \overline{z})$. Thus, λ_Y is strictly increasing on $(\underline{z}, \overline{z})$, which is a contradiction to the fact that λ_Y is non-increasing. Because a is piecewise continuous by assumption, $a(z) \in \{0, K\}$ for almost every $z \in W$.

Suppose that there exists an interval $(\underline{z}, \overline{z})$ such that $a(z) = K$ on $(\underline{z}, \overline{z})$ and let $(\underline{z}, \overline{z})$ be chosen maximally. Then $Y'(z) = Q(z) - H(z)^{n-1}$ which is strictly increasing because $K > \max_{z \in W}(n-1)H(z)^{n-2}h(z)$, and therefore $Y(z)$ is strictly convex on $[\underline{z}, \overline{z}]$. This implies that $Y(z) > 0$ on $[\underline{z}, \overline{z}]$ except at most one point. Suppose that $Y(z) > 0$ for all $z \in (\underline{z}, \overline{z})$. Then λ_Y is constant on $(\underline{z}, \overline{z})$, and $\lambda_Q(H^{-1}(t)) = \lambda_Q(\underline{z}) - \int_{H(\underline{z})}^{t} \left[\varphi^{\lambda_X}(\tau) + \lambda_Y(H^{-1}(\tau))\right] \mathrm{d}\tau$ is strictly convex on $(H(\underline{z}), H(\overline{z}))$. By Lemma 55 and the transversality condition, $\lambda_Q(\underline{z}) \leq 0$ and $\lambda_Q(\overline{z}) \leq 0$. Then the strict convexity of $\lambda_Q(H^{-1}(t))$ implies that $\lambda_Q(z) < 0$ for all $z \in (\underline{z}, \overline{z})$. However, $a(z) = K$ on $(\underline{z}, \overline{z})$ implies that $\lambda_Q(z) \geq 0$ for virtually every $z \in (\underline{z}, \overline{z})$, a contradiction. Hence, there exists a unique $z_0 \in (\underline{z}, \overline{z})$ such that $Y(z_0) = 0$, and therefore $Y(\underline{z}) > 0$ and $Y(\overline{z}) > 0$. Because $Y(\underline{w}) = Y(\overline{w}) = 0$ we have $\underline{w} < \underline{z} < \overline{z} < \overline{w}$. Note that this also implies that λ_Y is constant on a neighborhood of and therefore continuous at $z \in \{\underline{z}, \overline{z}\}$. By Lemma 55, we have

$$\varphi^{\lambda_X}(H(\underline{z})) + \lambda_Y(\underline{z}) \leq 0,$$
$$\varphi^{\lambda_X}(H(\overline{z})) + \lambda_Y(\overline{z}) \geq 0.$$

Hence,

$$\lambda_Y(\underline{z}) \leq -\varphi^{\lambda_X}(H(\underline{z})) < -\varphi^{\lambda_X}(H(\overline{z})) \leq \lambda_Y(\overline{z}),$$

where the second inequality holds because φ^{λ_X} is strictly decreasing and H is strictly increasing. However, this is a contradiction to that λ_Y is non-increasing. Hence, $a(z) = 0$ for almost all $z \in W$.

Because Q is absolutely continuous, this implies that Q is constant on W. However, by Lemma 43, $X(\overline{w}) = 0 > -C'(\alpha^*)$ when Q is constant on W, which implies that $\lambda_X = 0$, a contradiction to the supposition that $\lambda_X > \bar{\lambda}_X > 0$. Hence, $\lambda_X \leq \bar{\lambda}_X$. ■

Proof of Theorem 15

Let $\{Q^\nu\}$ be a sequence of optimal solutions of ($\mathcal{P}^K$-α^*) where $K = K_\nu > \overline{K}$ for each ν, $\overline{K}$ is defined is Lemma 52, and $K_\nu \to \infty$ as $\nu \to \infty$. After taking a subsequence, we can assume that $\{Q^\nu\}$ converges pointwise. Let Q^∞ denote the almost everywhere limit of this sequence. Denote the corresponding joint variables associated with Q^ν by λ^ν. By Lemma 52, $\{\lambda_X^\nu\}$ is bounded for ν sufficiently large. After taking a subsequence, we can assume that $\{\lambda_X^\nu\}$ converges, and let $\lambda_X^\infty := \lim_{\nu\to\infty} \lambda_X^\nu$. By Lemma 46, Q^∞ is an optimal solution.

For brevity, let φ^ν (or φ^∞) denote $\varphi^{\lambda_X^\nu}$ (or $\varphi^{\lambda_X^\infty}$), J^ν (or J^∞) denote $J^{\lambda_X^\nu}$ (or $J^{\lambda_X^\infty}$), $\overline{\varphi}^\nu$ (or $\overline{\varphi}^\infty$) denote

$\overline{\varphi}^{\lambda_X^\nu}$ (or $\overline{\varphi}^{\lambda_X^\infty}$), $\overline{J}^\nu$ (or $\overline{J}^\infty$) denote $\overline{J}^{\lambda_X^\nu}$ (or $\overline{J}^{\lambda_X^\infty}$).

Because Q^ν satisfies (IA′) with equality for all ν and Q^∞ is the pointwise limit of $\{Q^\nu\}$, Q^∞ satisfies (IA′) with equality. By a similar argument, $Y^\infty(\underline{w}) = 0$. Lemmas 57 and 58 below show that Q^∞ satisfies the two pooling properties when $\lambda_X = \lambda_X^\infty$. Finally, Lemma 59 below proves that that $\lambda_X^\infty = \lambda_X^*$, where $\lambda_X^* > 0$ is such that inequality (4.15) holds. By the arguments in Section 4.3.2, this completes the proof Theorem 15.

Before introducing and proving Lemmas 57, 58 and 59, I first prove the following technical lemma, which is used in the proofs of Lemmas 57 and 58.

Lemma 56 *The following four statements are true.*

1. *The sequence $\{\varphi^\nu\}$ is uniformly convergent with limit φ^∞.*
2. *The sequence $\{J^\nu\}$ is uniformly convergent with limit J^∞.*
3. *The sequence $\{\varphi^{\nu\prime}\}$ is uniformly convergent with limit $\varphi^{\infty\prime}$.*
4. *The sequence $\{\overline{J}^\nu\}$ is uniformly convergent with limit $\overline{J}^\infty$.*

Proof. Let $\gamma_1 := \max_{z\in W} |H_{\alpha_i}(z)/h(z)| > 0$,

$$\gamma_2 := \max_{z\in W} \left| \frac{\partial}{\partial z} \left[-\frac{H_{\alpha_i}(z)}{h(z)} \right] \right| > 0,$$

and $\gamma_3 := \max_{z\in W} 1/h(z) > 0$. Here γ_1, γ_2 and γ_3 are well define because H and h are twice continuously differentiable and W is compact.

1.
$$|\varphi^\nu(t) - \varphi^\infty(t)| = \left|\lambda_X^\nu - \lambda_X^\infty\right| \left| \frac{H_{\alpha_i}(H^{-1}(t))}{h(H^{-1}(t))} \right| \le \gamma_1 \left|\lambda_X^\nu - \lambda_X^\infty\right| \to 0,$$

as $\nu \to \infty$. Hence, the sequence $\{\varphi^\nu\}$ is uniformly convergent with limit φ^∞.

2.

$$\begin{aligned}|J^{\nu}(t) - J^{\infty}(t)| &= \int_0^t |\varphi^{\nu}(\tau) - \varphi^{\infty}(\tau)| \, \mathrm{d}\tau \\ &\le t\gamma_1 \left|\lambda_X^{\nu} - \lambda_X^{\infty}\right| \\ &\le \gamma_1 \left|\lambda_X^{\nu} - \lambda_X^{\infty}\right| \to 0,\end{aligned}$$

as $\nu \to 0$. Hence, the sequence $\{J^{\nu}\}$ is uniformly convergent with limit J^{∞}.

3.

$$\left|\varphi^{\nu\prime}(t) - \varphi^{\infty\prime}(t)\right| = \left|\lambda_X^{\nu} - \lambda_X^{\infty}\right| \left| \frac{\partial}{\partial z} \left[\frac{H_{\alpha_i}(H^{-1}(t))}{h(H^{-1}(t))} \right] \frac{1}{h(H^{-1}(t))} \right| \le \gamma_2\gamma_3 \left|\lambda_X^{\nu} - \lambda_X^{\infty}\right| \to 0,$$

as $\nu \to \infty$. Hence, the sequence $\{\varphi^{\nu\prime}\}$ is uniformly convergent with limit $\varphi^{\infty\prime}$.

4. Because the sequence $\{J^{\nu}\}$ is uniformly convergent with limit J^{∞}, for any $\varepsilon > 0$ there exists $\overline{\nu} > 0$ such that for all $\nu > \overline{\nu}$, $|J^{\infty}(t) - J^{\nu}(t)| \le \varepsilon$ for all $t \in [0,1]$. Fix $t \in [0,1]$. Let $t_1, t_2, \beta \in [0,1]$ be such that $\beta t_1 + (1-\beta)t_2 = t$. Then for any $\nu > \overline{\nu}$

$$\begin{aligned}\overline{J}^{\infty}(t) \le& \beta J^{\infty}(t_1) + (1-\beta)J^{\infty}(t_2) \\ \le& \beta J^{\nu}(t_1) + (1-\beta)J^{\nu}(t_2) + \beta|J^{\infty}(t_1) - J^{\nu}(t_1)| + (1-\beta)|J^{\infty}(t_2) - J^{\nu}(t_2)| \\ \le& \beta J^{\nu}(t_1) + (1-\beta)J^{\nu}(t_2) + \varepsilon\end{aligned}$$

Hence,

$$\overline{J}^{\infty}(t) \le \min\left\{\beta J^{\nu}(t_1) + (1-\beta)J^{\nu}(t_2) | \beta t_1 + (1-\beta)t_2 = t\right\} + \varepsilon = \overline{J}^{\nu}(t) + \varepsilon.$$

Similarly, we can show that $\overline{J}^{\nu}(t) \le \overline{J}^{\infty}(t) + \varepsilon$. Hence, $|\overline{J}^{\infty}(t) - \overline{J}^{\nu}(t)| \le \varepsilon$. Because this holds for any $t \in [0,1]$, we have that the sequence $\{\overline{J}^{\nu}\}$ is uniformly convergent with limit $\overline{J}^{\infty}$.

■

Lemma 57 *Suppose that $J^\infty(H(z)) > \overline{J}^\infty(H(z))$ for $z \in (\underline{z}, \overline{z})$ with $\underline{z} < \overline{z}$ and let $(\underline{z}, \overline{z})$ be chosen maximally. Then Q^∞ is constant on $(\underline{z}, \overline{z})$.*

Proof. For each $0 < \delta < (\overline{z} - \underline{z})/2$, let $\varepsilon(\delta) := \min_{z \in [\underline{z}+\delta, \overline{z}-\delta]}\{J^\infty(H(z)) - \overline{J}^\infty(H(z))\}$. Then $\varepsilon(\delta)$ is non-increasing in δ and converges to zero as δ converges to zero. Fix $\delta_0 > 0$. Let $\varepsilon_0 := \frac{1}{4}\varepsilon(\delta_0) > 0$. There exist $0 < \delta_1 < \delta_2 < \delta_0$ such that $\varepsilon(\delta_1) = \varepsilon_0$ and $\varepsilon(\delta_2) = 2\varepsilon_0$. I claim that there exists $\overline{v}$ such that for all $v > \overline{v}$,

$$J^v(H(z)) - \bar{J}^v(H(z)) \geq \frac{7\varepsilon_0}{2} \text{ if } z \in [\underline{z} + \delta_0, \overline{z} - \delta_0], \tag{C.41}$$

$$J^v(H(z)) - \bar{J}^v(H(z)) \geq \frac{\varepsilon_0}{2} \text{ if } z \in [\underline{z} + \delta_1, \overline{z} - \delta_1], \tag{C.42}$$

$$J^v(H(z)) - \bar{J}^v(H(z)) \leq \frac{5\varepsilon_0}{2} \text{ if } J^\infty(H(z)) - \overline{J}^\infty(H(z)) \leq 2\varepsilon_0. \tag{C.43}$$

We begin by prove (C.42). Because the sequence $\{J^v\}$ is uniformly convergent with limit J^∞, there exists $\overline{v}$ such that for all $v > \overline{v}$, $|J^v(t) - J^\infty(t)| < \varepsilon_0/8$ for all $t \in [0, 1]$. Let $t \in [H(\underline{z} + \delta_1), H(\overline{z} - \delta_1)]$. Then, by construction, $J^\infty(t) - \overline{J}^\infty(t) > \varepsilon_0$. Hence, there exists $\beta, t_1, t_2 \in [0, 1]$ such that $\beta t_1 + (1 - \beta)t_2 = t$ and $\beta J^\infty(t_1) + (1 - \beta)J^\infty(t_2) < J^\infty(t) - 3\varepsilon_0/4$. Then

$$\begin{aligned}
\bar{J}^v(t) \leq& \beta J^v(t_1) + (1 - \beta)J^v(t_2) \\
\leq& \beta J^\infty(t_1) + \beta|J^v(t_1) - J^\infty(t_1)| + (1 - \beta)J^\infty(t_2) + (1 - \beta)|J^v(t_2) - J^\infty(t_2)| \\
\leq& \beta J^\infty(t_1) + (1 - \beta)J^\infty(t_2) + \frac{\varepsilon_0}{8} \\
\leq& J^\infty(t) - \frac{3\varepsilon_0}{4} + \frac{\varepsilon_0}{8} \\
\leq& J^v(t) + |J^v(t) - J^\infty(t)| - \frac{5\varepsilon_0}{8} \\
\leq& J^v(t) + \frac{\varepsilon}{8} - \frac{5\varepsilon_0}{8} \\
=& J^v(t) - \frac{\varepsilon_0}{2}.
\end{aligned}$$

Hence, $J^v(H(z)) - \bar{J}^v(H(z)) \geq \varepsilon/2$ for all $z \in [\underline{z} + \delta_1, \overline{z} - \delta_1]$. A similar argument proves (C.41).

To show (C.43), let z be such that $J^\infty(H(z)) - \overline{J}^\infty(H(z)) \le 2\varepsilon_0$. For any $\beta, t_1, t_2 \in [0,1]$ such that $\beta t_1 + (1-\beta)t_2 = t$ and $t = H(z)$, we have

$$\begin{aligned}
&\beta J^\nu(t_1) + (1-\beta)J^\nu(t_2) \\
\ge&\beta J^\infty(t_1) - \beta|J^\nu(t_1) - J^\infty(t_1)| + (1-\beta)J^\infty(t_2) - (1-\beta)|J^\nu(t_2) - J^\infty(t_2)| \\
\ge&\beta J^\infty(t_1) + (1-\beta)J^\infty(t_2) - \frac{\varepsilon_0}{8} \\
\ge&\overline{J}^\infty(t) - \frac{\varepsilon_0}{8} \\
=&J^\infty(t) - [J^\infty(t) - \overline{J}^\infty(t)] - \frac{\varepsilon_0}{8} \\
\ge&J^\infty(t) - 2\varepsilon_0 - \frac{\varepsilon_0}{8}, \\
\ge&J^\nu(t) - |J^\nu(t) - J^\infty(t)| - 2\varepsilon_0 - \frac{\varepsilon_0}{8}, \\
\ge&J^\nu(t) - \frac{\varepsilon_0}{8} - 2\varepsilon_0 - \frac{\varepsilon_0}{8}, \\
\ge&J^\nu(t) - \frac{5\varepsilon_0}{2}.
\end{aligned}$$

Hence,

$$\overline{J}^\nu(t) := \min\{\beta J^\nu(t_1) + (1-\beta)J^\nu(t_2) | \beta, t_1, t_2 \in [0,1] \text{ and } \beta t_1 + (1-\beta)t_2 = t\} \ge J^\nu(t) - \frac{5\varepsilon_0}{2}.$$

Because $\varepsilon(\delta_1) = \varepsilon_0$ and $\varepsilon(\delta_2) = 2\varepsilon_0$, by continuity

$$\begin{aligned}
h_\delta := \min\Big\{ & h(\{z \in [\underline{z}+\delta_1, \underline{z}+\delta_2] | J^\infty(H(z)) - \overline{J}^\infty(H(z)) \le 2\varepsilon_0\}), \\
& h(\{z \in [\overline{z}-\delta_2, \overline{z}-\delta_1] | J^\infty(H(z)) - \overline{J}^\infty(H(z)) \le 2\varepsilon_0\})\Big\} > 0.
\end{aligned}$$

Fix $\nu > \overline{\nu}$ such that $K_\nu > 1/h_\delta$. Suppose that there exists $(b_1, b_2) \subset (\underline{z}+\delta_0, \overline{z}-\delta_0)$ such that $a^\nu(z) > 0$. Then $\lambda_Q(b_1), \lambda_Q(b_2) \ge 0$. Because λ_Y^ν is non-increasing, we have $\lambda_Y^\nu(b_2) \le \lambda_Y^\nu(b_1)$. Note that $\overline{J}^\nu$ is linear and therefore $\overline{\varphi}^\nu$ is constant on $(\underline{z}, \overline{z})$. Hence, we have either $-\lambda_Y^\nu(b_2) \ge \overline{\varphi}^\nu(H(z))$ for all $z \in (\underline{z}, \overline{z})$, or $-\lambda_Y^\nu(b_1) \le \overline{\varphi}^\nu(H(z))$ for all $z \in (\underline{z}, \overline{z})$. Assume without loss of generality that $-\lambda_Y^\nu(b_2) \ge \overline{\varphi}^\nu(H(z))$ for all $z \in (\underline{z}, \overline{z})$. For any $z \in [\overline{z}-\delta_2, \overline{z}-\delta_1]$ with $J^\infty(H(z)) - \overline{J}^\infty(H(z)) \le 2\varepsilon_0$, we have

$$
\begin{aligned}
\lambda_Q^\nu(z) &= \lambda_Q^\nu(b_2) - \int_{b_2}^{z} \left[\lambda_Y^\nu(\zeta) + \varphi^\nu(H(\zeta))\right] h(\zeta)\mathrm{d}\zeta \\
&\geq \int_{b_2}^{z} \overline{\varphi}^\nu(H(\zeta))h(\zeta)\mathrm{d}\zeta - \int_{b_2}^{z} \varphi^\nu(H(\zeta))h(\zeta)\mathrm{d}\zeta \\
&= \overline{J}^\nu(H(z)) - \overline{J}^\nu(H(b_2)) - J^\nu(H(z)) + J^\nu(H(b_2)) \\
&= J^\nu(H(b_2)) - \overline{J}^\nu(H(b_2)) - \left[J^\nu(H(z)) - \overline{J}^\nu(H(z))\right] \\
&\geq \frac{7\varepsilon_0}{2} - \frac{5\varepsilon_0}{2} = \varepsilon_0 > 0,
\end{aligned}
$$

where the first inequality holds because $\lambda_Q^\nu(b_2) \geq 0$ and $-\lambda_Y^\nu(\zeta) \geq -\lambda_Y^\nu(b_2) \geq \overline{\varphi}^\nu(H(\zeta))$ for all $\zeta \geq b_2$. That is, $a^\nu(z) = K_\nu$ for almost every $z \in [\overline{z}-\delta_2, \overline{z}-\delta_1]$ with $J^\infty(H(z)) - \overline{J}^\infty(H(z)) \leq 2\varepsilon_0$. However, this is a contradiction to that $K_\nu > 1/h_\delta$ because $0 \leq Q \leq 1$.

Hence, $a^\nu(z) = 0$ for almost every $z \in [\underline{z} - \delta_0, \overline{z} + \delta_0]$ for ν sufficiently large. Let ν goes to infinity and we have Q^∞ is constant on $[\underline{z} - \delta_0, \overline{z} + \delta_0]$. Because this is true for any $\delta_0 > 0$, we have that Q^∞ is constant on $(\underline{z}, \overline{z})$. ■

Lemma 58 *Suppose that $Y^\infty(z) > 0$ for all $z \in (\underline{z}, \overline{z})$ with $\underline{z} < \overline{z}$ and let $(\underline{z}, \overline{z})$ be chosen maximally. Then $\overline{\varphi}^\infty$ is constant on $(H(\overline{z}), H(\underline{z}))$.*

Proof. Suppose, to the contrary, that $\overline{\varphi}^\infty(H(\overline{z})^-) > \overline{\varphi}^\infty(H(\underline{z}))$. Because $\overline{\varphi}^\infty$ is non-decreasing and right-continuous, there exists $\overline{\delta} > 0$ such that $\overline{\varphi}^\infty(H(\overline{z} - \delta)) > \overline{\varphi}^\infty(H(\underline{z} + \delta))$ for all $\delta \in (0, \overline{\delta})$. Fix $\delta \in (0, \min\{\overline{\delta}/2, (\overline{z} - \underline{z})/4\})$. Because $\overline{J}^\infty$ is convex and $\overline{\varphi}^\infty$ is non-constant on $(\underline{z} + \delta, \overline{z} - \delta)$, we have

$$
\overline{J}^\infty(H(z)) < \overline{J}^\infty(H(\underline{z}+\delta)) + [H(z) - H(\underline{z}+\delta)]\frac{\overline{J}^\infty(H(\overline{z} - \delta)) - \overline{J}^\infty(H(\underline{z} + \delta))}{H(\overline{z} - \delta) - H(\underline{z} + \delta)}, \ \forall z \in (\underline{z}+\delta, \overline{z}-\delta).
$$

Let

$$\varepsilon_1 := \min_{z\in[\underline{z}+2\delta,\overline{z}-2\delta]} \left\{ \overline{J}^{\infty}(H(\underline{z}+\delta)) + [H(z) - H(\underline{z}+\delta)] \frac{\overline{J}^{\infty}(H(\overline{z}-\delta)) - \overline{J}^{\infty}(H(\underline{z}+\delta))}{H(\overline{z}-\delta) - H(\underline{z}+\delta)} - \overline{J}^{\infty}(H(z)) \right\} > 0,$$

$$\varepsilon_2 := \min_{z\in[\underline{z}+\delta,\overline{z}-\delta]} Y^{\infty}(z) > 0,$$

and

$$M_1 := 2 \max_{z\in[\underline{z},\overline{z}]} |\varphi^{\infty}(H(z))| > 0.$$

Because the sequence $\{\overline{J}^{\nu}\}$ is uniformly convergent with limit $\overline{J}^{\infty}$, the sequence $\{Y^{\nu}\}$ is uniformly convergent with limit Y^{∞}, and the sequence $\{\varphi^{\nu}\}$ is uniformly convergent with limit φ^{∞}, there exists $\overline{\nu}$ such that for $\nu > \overline{\nu}$, $|Y^{\nu}(z) - Y^{\infty}(z)| < \varepsilon_2/2$ for all $z \in W$, $|\overline{J}^{\infty}(t) - \overline{J}^{\nu}(t)| \leq \varepsilon_1/8$ for all $t \in [0,1]$, and $|\varphi^{\nu}(t) - \varphi^{\infty}| \leq M/2$ for all $t \in [0,1]$. Then for all $\nu > \overline{\nu}$ and $z \in [\underline{z}+2\delta, \overline{z}-2\delta]$ we have

$$\begin{aligned}
&\overline{J}^{\nu}(H(\underline{z}+\delta)) + [H(z) - H(\underline{z}+\delta)] \frac{\overline{J}^{\nu}(H(\overline{z}-\delta)) - \overline{J}^{\nu}(H(\underline{z}+\delta))}{H(\overline{z}-\delta) - H(\underline{z}+\delta)} - \overline{J}^{\nu}(H(z)) \\
\geq &\overline{J}^{\infty}(H(\underline{z}+\delta)) - \left|\overline{J}^{\infty}(H(\underline{z}+\delta)) - \overline{J}^{\nu}(H(\underline{z}+\delta))\right| - \overline{J}^{\infty}(H(z)) - \left|\overline{J}^{\infty}(H(z)) - \overline{J}^{\nu}(H(z))\right| \\
&+ \frac{H(z) - H(\underline{z}+\delta)}{H(\overline{z}-\delta) - H(\underline{z}+\delta)} \left[\overline{J}^{\infty}(H(\overline{z}-\delta)) - \left|\overline{J}^{\infty}(H(\overline{z}-\delta)) - \overline{J}^{\nu}(H(\overline{z}-\delta))\right| \right. \\
&\left. - \overline{J}^{\infty}(H(\underline{z}+\delta)) - \left|\overline{J}^{\infty}(H(\underline{z}+\delta)) - \overline{J}^{\nu}(H(\underline{z}+\delta))\right|\right] \\
\geq &\overline{J}^{\infty}(H(\underline{z}+\delta)) + [H(z) - H(\underline{z}+\delta)] \frac{\overline{J}^{\infty}(H(\overline{z}-\delta)) - \overline{J}^{\infty}(H(\underline{z}+\delta))}{H(\overline{z}-\delta) - H(\underline{z}+\delta)} - \overline{J}^{\infty}(H(z)) - \frac{\varepsilon_1}{2} \\
\geq &\frac{\varepsilon_1}{2}.
\end{aligned} \tag{C.44}$$

For all $\nu > \overline{\nu}$ and $z \in [\underline{z}+\delta, \overline{z}-\delta]$ we have

$$Y^{\nu}(z) \geq Y^{\infty}(z) - |Y^{\nu}(z) - Y^{\infty}(z)| \geq \frac{\varepsilon_2}{2} > 0.$$

For all $\nu > \overline{\nu}$ and $z \in [\underline{z} + \delta, \overline{z} - \delta]$, we have

$$\varphi^{\nu}(H(z)) \leq \varphi^{\infty}(H(z)) + |\varphi^{\nu}(H(z)) - \varphi^{\infty}(H(z))| \leq \frac{M_1}{2} + \frac{M_1}{2} = M_1.$$

Finally, let $M_3 := \min_{z\in[\underline{z},\overline{z}]} h(z) > 0$ and

$$M_2 := \left| \frac{\overline{J}^{\nu}(H(\overline{z} - \delta)) - \overline{J}^{\nu}(H(\underline{z} + \delta))}{H(\overline{z} - \delta) - H(\underline{z} + \delta)} \right| > 0.$$

Fix $\nu > \overline{\nu}$ such that $K_\nu > 1/\min\{z_0 + \varepsilon_1/8M_1M_3, z_0 + \varepsilon_1/8M_2M_3, \overline{z} - \delta\}$. Because $Y^\nu(z) > 0$ on $[\underline{z}+\delta, \overline{z}-\delta]$, λ_Y is constant and therefore $m_Y^\nu(H^{-1}(\cdot))$ is affine on $[H(\underline{z}+\delta), H(\overline{z}-\delta)]$. By Lemmas 48 and 50, we have $m_Y^\nu(z) \geq \overline{J}^{\nu}(H(z))$ for all $z \in W$. In particular, $m_Y^\nu(\underline{z} + \delta) \geq \overline{J}^{\nu}(H(\underline{z} + \delta))$ and $m_Y^\nu(\overline{z} - \delta) \geq \overline{J}^{\nu}(H(\overline{z} - \delta))$. Hence, for all $z \in [\underline{z} + \delta, \overline{z} - \delta]$,

$$m_Y^\nu(z) \geq \overline{J}^{\nu}(H(\underline{z} + \delta)) + [H(z) - H(\underline{z} + \delta)]\frac{\overline{J}^{\nu}(H(\overline{z} - \delta)) - \overline{J}^{\nu}(\underline{z} + \delta)}{H(\overline{z} - \delta) - H(\underline{z} + \delta)}.$$

Suppose that $J^\nu(H(z_0)) = \overline{J}^{\nu}(H(z_0))$ for some $z_0 \in [\underline{z} + 2\delta, \overline{z} - 2\delta]$. Then for all $z \in \left(z_0, \overline{z} - \delta\right)$ such that $H(z) - H(z_0) \leq \min\{\varepsilon_1/8M_1, \varepsilon_2/8M_2\}$ we have

$$\begin{aligned}
&J^\nu(H(z)) \\
=&J^\nu(H(z_0)) + \int_{H(z_0)}^{H(z)} \varphi^\nu(\tau)\mathrm{d}\tau \\
\leq&\overline{J}^{\nu}(H(z_0)) + M_1(H(z) - H(z_0)) \\
\leq&\overline{J}^{\nu}(H(\underline{z} + \delta)) + [H(z) - H(\underline{z} + \delta) - H(z) + H(z_0)]\frac{\overline{J}^{\nu}(H(\overline{z} - \delta)) - \overline{J}^{\nu}(H(\underline{z} + \delta))}{H(\overline{z} - \delta) - H(\underline{z} + \delta)} - \frac{\varepsilon_1}{2} + \frac{\varepsilon_1}{8} \\
\leq&m_Y^\nu(z) + [H(z) - H(z_0)]M_2 - \frac{\varepsilon_1}{2} + \frac{\varepsilon_1}{8} \\
\leq&m_Y^\nu(z) + \frac{\varepsilon_1}{8} - \frac{\varepsilon_1}{2} + \frac{\varepsilon_1}{8} = m_Y^\nu(z) - \frac{\varepsilon_1}{4},
\end{aligned}$$

where the second inequality holds by (C.44). That is, $J^\nu(H(z)) < m_Y^\nu(z)$, and therefore $a^\nu(z) = K_\nu$.

Because $H(z) - H(z_0) \leq M_3(z - z_0)$, we have $a^\nu(z) = K_\nu$ for all $z \in (z_0, \min\{z_0 + \varepsilon_1/8M_1M_3, z_0 + \varepsilon_1/8M_2M_3, \overline{z} - \delta\})$, a contradiction to that $K_\nu > 1/\min\{z_0 + \varepsilon_1/8M_1M_3, z_0 + \varepsilon_1/8M_2M_3, \overline{z} - \delta\}$ because $0 \leq Q \leq 1$.

Hence, $J^\nu(H(z)) > \overline{J}^\nu(H(z))$ for all $z \in [\underline{z} + 2\delta, \overline{z} - 2\delta]$. This implies that $\overline{\varphi}^\nu$ is constant on $[H(\underline{z}-2\delta), H(\overline{z}+2\delta)]$. Clearly, $\{\overline{\varphi}^\nu\}$ converges uniformly on $[H(\underline{z}-2\delta), H(\overline{z}+2\delta)]$ and $\lim_{\nu\to\infty} \overline{\varphi}^\nu$ is constant on $[H(\underline{z}-2\delta), H(\overline{z}+2\delta)]$. Because, on $[H(\underline{z}-2\delta), H(\overline{z}+2\delta)]$, $\{\overline{J}^\nu\}$ converges uniformly to $\overline{J}^\infty$, each $\overline{J}^\nu$ is differentiable with derivative $\overline{\varphi}^\nu$, and $\{\overline{\varphi}^\nu\}$ converges uniformly, we have $\overline{J}^\infty$ is differentiable on $[H(\underline{z} - 2\delta), H(\overline{z} + 2\delta)]$ and its derivative $\overline{\varphi}^\infty(t) = \lim_{\nu\to\infty} \overline{\varphi}^\nu(t)$ for all $t \in [H(\underline{z}-2\delta), H(\overline{z}+2\delta)]$. Thus, $\overline{\varphi}^\infty$ is constant on $[H(\underline{z}-2\delta), H(\overline{z}+2\delta)]$ and $2\delta < \overline{\delta}$, a contradiction to the supposition. Hence, $\overline{\varphi}^\infty$ is constant on $(H(\underline{z}), H(\overline{z}))$. ■

Corollary 11 *Suppose that $\overline{\varphi}^\infty(H(z))$ is constant on $(\underline{z}, \overline{z})$ with $\underline{z} < \overline{z}$ and let $(\underline{z}, \overline{z})$ be chosen maximally. Then $Y^\infty(\underline{z}) = Y^\infty(\overline{z}) = 0$, i.e.,*

$$\int_{\underline{z}}^{\overline{z}} \left[H(\zeta)^{n-1} - Q^\infty(\zeta)\right] h(\zeta) \mathrm{d}\zeta = 0.$$

Proof. This is an immediate corollary of Lemma 58. Suppose, to the contrary, that $Y^\infty(\overline{z}) > 0$. Then by Lemma 58, $\overline{\varphi}^\infty(H(\cdot))$ is constant on a neighborhood of $\overline{z}$, a contradiction to the fact that $(\underline{z}, \overline{z})$ is chosen maximally. Hence, $Y^\infty(\overline{z}) = 0$. Similarly, $Y^\infty(\underline{z}) = 0$. ■

Lemma 59 *$\lambda_X^\infty = \lambda_X^*$, where $\lambda_X^* > 0$ is such that inequality* (4.15) *holds .*

Proof. For any $\lambda_X > 0$, recall that $Q^+(\cdot, \lambda_X)$ and $Q^-(\cdot, \lambda_X)$ are defined as follows: If

$$J^{\lambda_X}(H(w|\alpha^*), \alpha^*) > \overline{J}^{\lambda_X}(H(w|\alpha^*), \alpha^*),\ \forall w \in (\underline{w}, \overline{w}),$$

and let $(\underline{w}, \overline{w})$ be chosen maximally, then let

$$Q^+(w, \lambda_X) := \frac{\frac{1}{n}[H(\overline{w}|\alpha^*)^n - H(\underline{w}|\alpha^*)^n]}{H(\overline{w}|\alpha^*) - H(\underline{w}|\alpha^*)},\ \forall w \in (\overline{w}, \underline{w}).$$

Otherwise, let $Q^+(w, \lambda^X) := H(w|\alpha^*)^{n-1}$. If $\overline{\varphi}^{\lambda_X}(H(\cdot|\alpha^*), \alpha^*)$ is constant on $(\underline{w}, \overline{w})$ with $\underline{w} < \overline{w}$ and let $(\underline{w}, \overline{w})$ be chosen maximally, then let

$$Q^-(w, \lambda_X) := \frac{\frac{1}{n}[H(\overline{w}|\alpha^*)^n - H(\underline{w}|\alpha^*)^n]}{H(\overline{w}|\alpha^*) - H(\underline{w}|\alpha^*)}, \ \forall w \in (\overline{w}, \underline{w}).$$

Otherwise, let $Q^-(z, \lambda_X, \alpha^*) := H(z|\alpha^*)^{n-1}$. For the ease of notation, denote $Q^+(w, \lambda_X^\infty)$ (or $Q^-(w, \lambda_x^\infty)$) by $Q^+(w)$ (or $Q^-(w)$). Note that all Q^+, Q^- and Q^∞ are implementable and non-decreasing, allocate the object with probability one, and satisfy the two pooling properties. Hence, by the arguments in Section 4.3.2, for $Q \in \{Q^+, Q^-, Q^\infty\}$,

$$\begin{aligned} &\int_{\underline{w}}^{\overline{w}} \left[z + \lambda_X^\infty \frac{H_{\alpha_i}(z)}{h(z)} \right] Q(z)h(z)\mathrm{d}z + \lambda_X^\infty C'(\alpha^*) \\ &= \int_{\underline{w}}^{\overline{w}} \overline{\varphi}^\infty(H(z))H(z)^{n-1}\mathrm{d}H(z) + \lambda_X^\infty C'(\alpha^*). \end{aligned} \tag{C.45}$$

Next we show that $Y^+ \leq Y^\infty \leq Y^-$. Let $S^+ := \{z \in W | Y^+(z) > 0\}$, $S^- := \{z \in W | Y^-(z) > 0\}$ and $S := \{z \in W | Y^\infty(z) > 0\}$. By construction,

$$S^+ = \cup\{(\underline{z}, \overline{z}) | J^\infty(H(z)) > \overline{J}^\infty(H(z)) \ \forall z \in (\underline{z}, \overline{z})\},$$

and

$$S^- = \cup\{(\underline{z}, \overline{z}) | \overline{\varphi}^\infty(H(\cdot)) \text{ is constant on } (\underline{z}, \overline{z})\}.$$

It follows from Lemma 57 and Lemma 58 that $S^+ \subset S \subset S^-$. If $z \notin S^-$, then $Y^+(z) = Y^\infty(z) = Y^-(z) = 0$. If $z \in S^- \backslash S$, then $Y^+(z) = Y^\infty(z) = 0 < Y^-(z)$. Consider $z \in S \subset S^-$, then there exists an interval $(\underline{z}, \overline{z})$ with $\underline{z} < z < \overline{z}$ such that $\overline{\varphi}^\infty(H(\cdot))$ is constant on $(\underline{z}, \overline{z})$. Let $(\underline{z}, \overline{z})$ be chosen maximally, then by construction $Y^-(\overline{z}) = Y^-(\underline{z}) = 0$. By Corollary 11, $Y^\infty(\overline{z}) = Y^\infty(\underline{z}) = 0$. For any $z \in (\underline{z}, \overline{z})$,

$$Y^-(z) - Y^\infty(z) = \int_z^{\overline{z}} \left[Q^\infty(\zeta) - Q^-(\zeta)\right] \mathrm{d}H(\zeta).$$

Then for any $t \in (H(\underline{z}), H(\overline{z}))$,

$$\left[Y^-(H^{-1}(t)) - Y^\infty(H^{-1}(t))\right]' = Q^-(H^{-1}(t)) - Q^\infty(H^{-1}(t)),$$

which is non-increasing on $(H(\underline{z}), H(\overline{z}))$ because Q^∞ is non-decreasing and Q^- is constant on $(\underline{z}, \overline{z})$ by construction. Hence, $Y^-(H^{-1}(t)) - Y^\infty(H^{-1}(t))$ is concave on $(H(\underline{z}), H(\overline{z}))$. Because $Y^-(\underline{z}) - Y^\infty(\underline{z}) = 0$ and $Y^-(\overline{z}) - Y^\infty(\overline{z}) = 0$, we have $Y^-(z) - Y^\infty(z) \geq 0$ for all $z \in (\underline{z}, \overline{z})$. Thus, $Y^-(z) - Y^\infty(z) \geq 0$ for all $z \in S$. If $z \in S \backslash S^+$, then $Y^-(z) \geq Y^\infty(z) \geq 0 = Y^+(z)$. Finally, consider $z \in S^+ \subset S$, it suffices to show that $Y^+(z) \leq Y^\infty(z)$. By construction, there exists an interval $(\underline{z}, \overline{z})$ with $\underline{z} < z < \overline{z}$ such that $\overline{J}^\infty(H(z)) < J^\infty(H(z))$ for all $z \in (\underline{z}, \overline{z})$. Let $(\underline{z}, \overline{z})$ be chosen maximally, then by construction $Y^+(\overline{z}) = Y^+(\underline{z}) = 0$. For any $z \in (\underline{z}, \overline{z})$

$$Y^+(z) - Y^\infty(z) = \int_z^{\overline{z}} \left[Q^\infty(\zeta) - Q^-(\zeta)\right] \mathrm{d}H(\zeta) - Y^\infty(\overline{z}).$$

Then for any $t \in (H(\underline{z}), H(\overline{z}))$,

$$\left[Y^+(H^{-1}(t)) - Y^\infty(H^{-1}(t))\right]' = Q^+(H^{-1}(t)) - Q^\infty(H^{-1}(t)),$$

which is constant on $(H(\underline{z}), H(\overline{z}))$ because Q^∞ is constant on $(\underline{z}, \overline{z})$ by Lemma 57 and Q^- is constant on $(\underline{z}, \overline{z})$ by construction. Hence, $Y^+(H^{-1}(t)) - Y^\infty(H^{-1}(t))$ is affine on $(H(\underline{z}), H(\overline{z}))$. Because $Y^+(\underline{z}) = 0 \leq Y^\infty(\underline{z})$ and $Y^+(\overline{z}) = 0 \leq Y^\infty(\overline{z})$, we have $Y^+(z) - Y^\infty(z) \leq 0$ for all $z \in (\underline{z}, \overline{z})$. Thus, $Y^+(z) - Y^\infty(z) \leq 0$ for all $z \in S^+$.

Furthermore, for any implementable allocation rule Q, we have

$$\begin{aligned}
&\int_{\underline{w}}^{\overline{w}} zQ(z)\mathrm{d}H(z) \\
&= \int_{\underline{w}}^{\overline{w}} zY'(z)\mathrm{d}z + \int_{\underline{w}}^{\overline{w}} zH(z)^{n-1}\mathrm{d}H(z) \\
&= \int_{\underline{w}}^{\overline{w}} zH(z)^{n-1}\mathrm{d}H(z) - \int_{\underline{w}}^{\overline{w}} Y(z)\mathrm{d}z. \qquad \text{(C.46)}
\end{aligned}$$

Hence,

$$\int_{\underline{w}}^{\overline{w}} zQ^{+}(z)\mathrm{d}H(z) \geq \int_{\underline{w}}^{\overline{w}} zQ^{\infty}(z)\mathrm{d}H(z) \geq \int_{\underline{w}}^{\overline{w}} zQ^{-}(z)\mathrm{d}H(z).$$

Because $\lambda_X^{\infty} > 0$, combining this and (C.45) yields

$$\int_{\underline{w}}^{\overline{w}} H_{\alpha}(z)Q^{+}(z)\mathrm{d}z \leq -C'(\alpha^{*}) \leq \int_{\underline{w}}^{\overline{w}} H_{\alpha}(z)Q^{-}(z)\mathrm{d}z. \tag{4.15}$$

By Lemma 15, there exists an unique $\lambda_X > 0$ such that (4.15) holds. Hence, $\lambda_X^{\infty} = \lambda_X^{*}$. ■

The arguments in Lemma 59 also proves the following corollary:

Corollary 12 *For any non-decreasing implementable Q that allocates the object with probability one and satisfies the two pooling properties, the following inequality holds:*

$$\int_{\underline{w}}^{\overline{w}} zQ^{+}(z,\lambda_X)\mathrm{d}H(z) \geq \int_{\underline{w}}^{\overline{w}} zQ(z)\mathrm{d}H(z) \geq \int_{\underline{w}}^{\overline{w}} zQ^{-}(z,\lambda_X)\mathrm{d}H(z).$$

Proof of Lemma 15

I break the proof into several lemmas. For each λ_X, recall that $Q^{+}(\cdot,\lambda_X)$ is the "steepest" allocation rule associated with λ'_X, and $Q^{-}(\cdot,\lambda_X)$ is the "least steep" allocation rule associated with λ_X. By the argument in the proof of Corollary 8,

$$\int_{\underline{w}}^{\overline{w}} -\frac{H_{\alpha}(z)}{h(z)}Q^{-}(z,\lambda_X)h(z)\mathrm{d}z \leq \int_{\underline{w}}^{\overline{w}} -\frac{H_{\alpha}(z)}{h(z)}Q^{+}(z,\lambda_X)\mathrm{d}z.$$

Let $\lambda'_X > \lambda_X$. First, I show in Lemma 60 that $Q^{-}(\cdot,\lambda_X)$ is steeper than $Q^{+}(\cdot,\lambda'_X)$. Second, I show in Lemma 62 that this implies that the marginal benefit of information acquisition given $Q^{-}(\cdot,\lambda_X)$ is higher than that given $Q^{+}(\cdot,\lambda'_X)$:

$$\int_{\underline{w}}^{\overline{w}} -\frac{H_{\alpha}(z)}{h(z)}Q^{+}(z,\lambda'_X)h(z)\mathrm{d}z < \int_{\underline{w}}^{\overline{w}} -\frac{H_{\alpha}(z)}{h(z)}Q^{-}(z,\lambda_X)h(z)\mathrm{d}z.$$

Furthermore, if λ_X is sufficiently large, then $Q^+(\cdot, \lambda_X)$ is constant and $\int_{\underline{w}}^{\overline{w}} -\frac{H_\alpha(z)}{h(z)} Q^+(z, \lambda_X)h(z)\mathrm{d}z = 0 < C'(\alpha^*)$; if $\lambda_X = 0$, then $Q^-(\cdot, \lambda_X) = H^{n-1}(\cdot)$ and $\int_{\underline{w}}^{\overline{w}} -\frac{H_\alpha(z)}{h(z)} Q^-(z, \lambda_X)h(z)\mathrm{d}z > C'(\alpha^*)$ because (4.7) is violated. Hence, there exists a unique $\lambda_X > 0$ such that inequality (4.15) holds.

Lemma 60 *Let $\lambda'_X > \lambda_X$. Suppose that $\overline{\varphi}^{\lambda_X}$ is constant on $(\underline{t}, \bar{t})$ with $\underline{t} < \bar{t}$ and let $(\underline{t}, \bar{t})$ be chosen maximally. Then there exists $\delta > 0$ such that $J^{\lambda'_X}(t) > \overline{J}^{\lambda'_X}(t)$ for all $t \in (\underline{t} - \delta, \bar{t} + \delta)$.*

The proof of Lemma 60 relies on the following technical lemma.

Lemma 61 *Let $t \in (0, 1)$. If $J^{\lambda_X}(t) = \overline{J}^{\lambda_X}(t)$, then $\overline{J}^{\lambda_X}$ is continuously differentiable at t with derivative $\overline{\varphi}^{\lambda_X}(t) = \varphi^{\lambda_X}(t)$ and $\varphi^{\lambda_X\prime}(t) \geq 0$. Furthermore, $J^{\lambda_X}(t) = \overline{J}^{\lambda_X}(t)$ if and only if*

$$J^{\lambda_X}(\tau) \geq (\tau - t)\varphi^{\lambda_X}(t) + J^{\lambda_X}(t), \ \forall \tau \in [0, 1]. \tag{C.47}$$

Proof. For ease of notation, I suppress the dependence of $J, \overline{J}, \varphi$ and $\overline{\varphi}$ on λ_X. Let $t \in (0, 1)$. Suppose that $J(t) = \overline{J}(t)$. Suppose, to the contrary, that $\overline{J}$ is not continuously differentiable at t, then $\overline{\varphi}(t^-) < \overline{\varphi}(t^+)$. Then either $\varphi(t) > \overline{\varphi}(t^-)$ or $\varphi(t) < \overline{\varphi}(t^+)$. Assume without loss of generality that $\varphi(t) < \overline{\varphi}(t^+)$. Because φ is continuous and $\overline{\varphi}$ is non-decreasing, there exists $\delta > 0$ such that $\varphi(\tau) < \overline{\varphi}(t^+) \leq \overline{\varphi}(\tau)$ for all $\tau \in (t, t + \delta)$. Then

$$J(t + \delta) = J(t) + \int_t^{t+\delta} \varphi(\tau)\mathrm{d}\tau < \overline{J}(t) + \int_t^{t+\delta} \overline{\varphi}(\tau)\mathrm{d}\tau = \overline{J}(t + \delta),$$

a contradiction. Hence, $\overline{J}$ is continuously differentiable at t. It follows from a similar argument that $\overline{\varphi}(t) = \varphi(t)$ with $\varphi'(t) \geq 0$. Furthermore, for all $\tau \in [0, 1]$,

$$J(\tau) \geq \overline{J}(\tau) \geq (\tau - t)\varphi(t) + \overline{J}(t) = (\tau - t)\varphi(t) + J(t),$$

where the second inequality holds because $\overline{J}$ is convex.

Suppose that (C.47) holds. Then $\tau :\to (\tau - t)\varphi(t) + J(t)$ is a convex function below J. Because $\overline{J}$

is the greatest convex function below J, we have

$$\overline{J}(\tau) \geq (\tau - t)\varphi(t) + J(t) \ \forall \tau \in [0,1].$$

If $\tau = t$, then $\overline{J}(t) \geq J(t)$. Hence, $\overline{J}(t) = J(t)$. ■

Proof of Lemma 60. First, I claim that $\overline{J}^{\lambda_X}(\underline{t}) = J^{\lambda_X}(\underline{t})$ and $\overline{J}^{\lambda_X}(\bar{t}) = J^{\lambda_X}(\bar{t})$. To see that $\overline{J}^{\lambda_X}(\underline{t}) = J^{\lambda_X}(\underline{t})$, suppose to the contrary that $\overline{J}^{\lambda_X}(\underline{t}) > J^{\lambda_X}(\underline{t})$. Then $\overline{\varphi}^{\lambda_X}(t)$ is constant in a neighborhood of $\underline{t}$. A contradiction to that $(\underline{t}, \bar{t})$ is chosen maximally. A similar argument proves that $\overline{J}^{\lambda_X}(\bar{t}) = J^{\lambda_X}(\bar{t})$.

Consider $t \in (\underline{t}, \bar{t})$. Let $\beta \in (0,1)$ be such that $\beta\underline{t} + (1-\beta)\bar{t} = t$. It suffices to show that

$$J^{\lambda'_X}(t) > \beta J^{\lambda'_X}(\underline{t}) + (1-\beta)J^{\lambda'_X}(\bar{t}).$$

Because $\overline{\varphi}^{\lambda_X}$ is constant on $(\underline{t}, \bar{t})$, $\overline{J}^{\lambda_X}(\underline{t}) = J^{\lambda_X}(\underline{t})$ and $\overline{J}^{\lambda_X}(\bar{t}) = J^{\lambda_X}(\bar{t})$, we have

$$\begin{aligned} J^{\lambda_X}(t) &\geq \overline{J}^{\lambda_X}(t) \\ &= \beta\overline{J}^{\lambda_X}(\underline{t}) + (1-\beta)\overline{J}^{\lambda_X}(\bar{t}) \\ &= \beta J^{\lambda_X}(\underline{t}) + (1-\beta)J^{\lambda_X}(\bar{t}). \end{aligned}$$

Hence,

$$\begin{aligned} 0 \leq& J^{\lambda_X}(t) - \beta J^{\lambda_X}(\underline{t}) - (1-\beta)J^{\lambda_X}(\bar{t}) \\ =& \int_0^t H^{-1}(\zeta)\mathrm{d}\zeta - \beta\int_0^{\underline{t}} H^{-1}(\zeta)\mathrm{d}\zeta - (1-\beta)\int_0^{\bar{t}} H^{-1}(\zeta)\mathrm{d}\zeta \\ &+ \lambda_X\left[\int_0^t \frac{H_\alpha(H^{-1}(\zeta))}{h(H^{-1}(\zeta))}\mathrm{d}\zeta - \beta\int_0^{\underline{t}} \frac{H_\alpha(H^{-1}(\zeta))}{h(H^{-1}(\zeta))}\mathrm{d}\zeta - (1-\beta)\int_0^{\bar{t}} \frac{H_\alpha(H^{-1}(\zeta))}{h(H^{-1}(\zeta))}\mathrm{d}\zeta\right]. \end{aligned}$$

Because $H^{-1}(\cdot)$ is strictly increasing, $\int_0^t H^{-1}(\zeta)\mathrm{d}\tau$ is strictly convex in t and therefore

$$\int_0^t H^{-1}(\zeta)\mathrm{d}\zeta - \beta\int_0^{\underline{t}} H^{-1}(\zeta)\mathrm{d}\zeta - (1-\beta)\int_0^{\bar{t}} H^{-1}(\zeta)\mathrm{d}\zeta < 0.$$

Hence,

$$\int_0^t \frac{H_\alpha(H^{-1}(\zeta))}{h(H^{-1}(\zeta))}\mathrm{d}\zeta - \beta\int_0^{\underline{t}} \frac{H_\alpha(H^{-1}(\zeta))}{h(H^{-1}(\zeta))}\mathrm{d}\zeta - (1-\beta)\int_0^{\bar{t}} \frac{H_\alpha(H^{-1}(\zeta))}{h(H^{-1}(\zeta))}\mathrm{d}\zeta > 0.$$

Then

$$\begin{aligned}
& J^{\lambda'_X}(t) - \beta J^{\lambda'_X}(\underline{t}) - (1-\beta)J^{\lambda'_X}(\bar{t}) \\
=& \int_0^t H^{-1}(\zeta)\mathrm{d}\zeta - \beta\int_0^{\underline{t}} H^{-1}(\zeta)\mathrm{d}\zeta - (1-\beta)\int_0^{\bar{t}} H^{-1}(\zeta)\mathrm{d}\zeta \\
& + \lambda'_X\left[\int_0^t \frac{H_\alpha(H^{-1}(\zeta))}{h(H^{-1}(\zeta))}\mathrm{d}\zeta - \beta\int_0^{\underline{t}} \frac{H_\alpha(H^{-1}(\zeta))}{h(H^{-1}(\zeta))}\mathrm{d}\zeta - (1-\beta)\int_0^{\bar{t}} \frac{H_\alpha(H^{-1}(\zeta))}{h(H^{-1}(\zeta))}\mathrm{d}\zeta\right] \\
>& \int_0^t H^{-1}(\zeta)\mathrm{d}\zeta - \beta\int_0^{\underline{t}} H^{-1}(\zeta)\mathrm{d}\zeta - (1-\beta)\int_0^{\bar{t}} H^{-1}(\zeta)\mathrm{d}\zeta \\
& + \lambda_X\left[\int_0^t \frac{H_\alpha(H^{-1}(\zeta))}{h(H^{-1}(\zeta))}\mathrm{d}\zeta - \beta\int_0^{\underline{t}} \frac{H_\alpha(H^{-1}(\zeta))}{h(H^{-1}(\zeta))}\mathrm{d}\zeta - (1-\beta)\int_0^{\bar{t}} \frac{H_\alpha(H^{-1}(\zeta))}{h(H^{-1}(\zeta))}\mathrm{d}\zeta\right] \\
\geq& J^{\lambda_X}(t) - \beta J^{\lambda_X}(\underline{t}) - (1-\beta)J^{\lambda_X}(\bar{t}) \\
\geq& 0.
\end{aligned}$$

Consider $\underline{t}$. Since $\overline{\varphi}^{\lambda_X}$ is constant on $(\underline{t}, \bar{t})$ with $J^{\lambda_X}(\underline{t}) = \overline{J}^{\lambda_X}(\underline{t})$ and $J^{\lambda_X}(\bar{t}) = \overline{J}^{\lambda_X}(\bar{t})$, by Lemma 61, we have

$$\begin{aligned}
\overline{J}^{\lambda_X}(\bar{t}) =& J^{\lambda_X}(\bar{t}) \\
\geq& (\bar{t} - \underline{t})\varphi^{\lambda_X}(\underline{t}) + J^{\lambda_X}(\underline{t}) \\
=& (\bar{t} - \underline{t})\varphi^{\lambda_X}(\underline{t}) + \overline{J}^{\lambda_X}(\underline{t}) = \overline{J}^{\lambda_X}(\bar{t}).
\end{aligned}$$

Hence, $J^{\lambda_X}(\bar{t}) = (\bar{t} - \underline{t})\varphi^{\lambda_X}(\underline{t}) + J^{\lambda_X}(\underline{t})$. Thus,

$$\int_{\underline{t}}^{\bar{t}} H^{-1}(\tau)\mathrm{d}\tau - (\bar{t} - \underline{t})H^{-1}(\underline{t}) = \lambda_X\left[(\bar{t} - \underline{t})\frac{H_\alpha(H^{-1}(\underline{t}))}{h(H^{-1}(\underline{t}))} - \int_{\underline{t}}^{\bar{t}} \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))}\right].$$

Since $H^{-1}(t)$ is strictly increasing, the left-hand side of the above equality is strictly positive. Hence for $\lambda'_X > \lambda_X > 0$ we have

$$\int_{\underline{t}}^{\bar{t}} H^{-1}(\tau)\mathrm{d}\tau - (\bar{t}-\underline{t})H^{-1}(\underline{t}) < \lambda'_X \left[(\bar{t}-\underline{t})\frac{H_\alpha(H^{-1}(\underline{t}))}{h(H^{-1}(\underline{t}))} - \int_{\underline{t}}^{\bar{t}} \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))}\right],$$

i.e.,

$$J^{\lambda'_X}(\bar{t}) < (\bar{t}-\underline{t})\varphi^{\lambda'_X}(\underline{t}) + J^{\lambda'_X}(\underline{t}).$$

By Lemma 61, $J^{\lambda'_X}(\underline{t}) > \overline{J}^{\lambda'_X}(\underline{t})$.

A similar argument proves that $J^{\lambda'_X}(\bar{t}) > \overline{J}^{\lambda'_X}(\bar{t})$. By continuity, there exists $\delta > 0$ such that $J^{\lambda'_X}(t) > \overline{J}^{\lambda'_X}(t)$ for all $t \in (\underline{t}-\delta, \bar{t}+\delta)$. ■

Lemma 62 *Let $[\underline{z}, \overline{z}] \subset [\underline{w}, \overline{w}]$ with $\underline{z} < \overline{z}$, and $z^0 \in (\underline{z}, \overline{z})$. Suppose that $Q : [\underline{\theta}, \overline{\theta}] \to [0,1]$ and $\check{Q} : [\underline{\theta}, \overline{\theta}] \to [0,1]$ satisfying that*

$$\int_{\underline{z}}^{\overline{z}} Q(z)h(z)\mathrm{d}z = \int_{\underline{z}}^{\overline{z}} \check{Q}(z)h(z)\mathrm{d}z, \tag{C.48}$$

and

$$Q(z) \geq \check{Q}(z) \text{ if } z > z^0, \text{ and } Q(z) \leq \check{Q}(z) \text{ if } z < z^0. \tag{C.49}$$

Then

$$\int_{\underline{z}}^{\overline{z}} -\frac{H_{\alpha_i}(z)}{h(z)}[Q(z)-\check{Q}(z)]h(z)\mathrm{d}z \geq 0, \tag{C.50}$$

where the inequality holds strictly if the set $\{z \in [\underline{z}, \overline{z}] | Q(z) \neq \check{Q}(z)\}$ has a positive measure.

Proof. Because $-\frac{H_\alpha(w)}{h(w)}$ is strictly increasing in w, and Q and $\check{Q}$ satisfy (C.49), we have

$$\int_{\underline{z}}^{\overline{z}} \left[-\frac{H_{\alpha_i}(z)}{h(z)} + \frac{H_{\alpha_i}(z^0)}{h(z^0)}\right][Q(z)-\check{Q}(z)]h(z)\mathrm{d}z \geq 0,$$

where the inequality holds strictly if the set $\{z \in [\underline{z}, \overline{z}] | Q(z) \neq \check{Q}(z)\}$ has a positive measure. This implies inequality (C.50) by (C.48). ■

C.1.3. Sufficient conditions for the first-order approach

In this section I provide sufficient conditions for the first-order approach to be valid. Let $\pi(\alpha_i)$ denote an agent i's payoff from choosing α_i given mechanism $(\boldsymbol{q}, \boldsymbol{t})$ and $\alpha_j = \alpha^*$ for all $j \neq i$. Then

$$\pi(\alpha_i) := U(w(0, \alpha_i)) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} \left[1 - H(w_i|\alpha_i)\right] Q(w_i)\mathrm{d}w_i - C(\alpha_i),$$

where Q is defined by (4.2) for $\alpha_j = \alpha^*$ for all $j \neq i$. Then

$$\begin{aligned}
\pi'(\alpha_i) =& U'(w(0, \alpha_i))w_{\alpha_i}(0, \alpha_i) + \left[1 - H(w(1, \alpha_i)|\alpha_i)\right] Q(w(1, \alpha_i))w_{\alpha}(1, \alpha_i) \\
& - \left[1 - H(w(0, \alpha_i)|\alpha_i)\right] Q(w(0, \alpha_i))w_{\alpha_i}(0, \alpha_i) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -H_{\alpha_i}(w_i|\alpha_i)Q(w_i)\mathrm{d}w_i - C'(\alpha_i) \\
=& \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -H_{\alpha_i}(w_i|\alpha_i)Q(w_i)\mathrm{d}w_i - C'(\alpha_i),
\end{aligned}$$

where the second line holds because $H(w(1, \alpha_i)|\alpha_i) = 1$, $H(w(0, \alpha_i)|\alpha_i) = 0$, and $U'(w(0, \alpha_i)) = Q(w(0, \alpha_i))$ by the envelope condition. A sufficient condition for the first-order approach to be valid is that $\pi'(\alpha_i)$ is strictly decreasing for all non-decreasing implementable allocation rule Q. If the support of the conditional expectation $[w(0, \alpha_i), w(1, \alpha_i)]$ is invariant, then $\pi'(\alpha_i)$ is strictly decreasing if $-H_{\alpha_i}(w_i|\alpha_i)$ has the *single-crossing property* in $(\alpha_i; w_i)$ and $C'(\alpha_i)$ is strictly decreasing. Suppose that C is twice continuously differentiable, then

$$\begin{aligned}
\pi''(\alpha_i) =& \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -\frac{\partial^2 H(w_i|\alpha_i)}{\partial \alpha_i^2} Q(w_i)\mathrm{d}w_i - H_{\alpha_i}(w(1, \alpha_i)|\alpha_i)w_{\alpha_i}(1, \alpha_i)Q(w(1, \alpha_i)) \\
& + H_{\alpha_i}(w(0, \alpha_i)|\alpha_i)w_{\alpha_i}(0, \alpha_i)Q(w(0, \alpha_i)) - C''(\alpha_i), \\
\leq & \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -\frac{\partial^2 H(w_i|\alpha_i)}{\partial \alpha_i^2} Q(w_i)\mathrm{d}w_i - H_{\alpha_i}(w(1, \alpha_i)|\alpha_i)w_{\alpha_i}(1, \alpha_i)Q(w(1, \alpha_i)) - C''(\alpha_i).
\end{aligned}$$

The inequality holds because $H_{\alpha_i}(w(0, \alpha_i)|\alpha_i) \geq 0$ and $w_{\alpha_i}(0, \alpha_i) \leq 0$ when the information structures are supermodular ordered. The following proposition from Shi (2012) gives sufficient conditions for $\pi''(\alpha_i) < 0$ for the two leading examples.

Proposition 15 (Shi (2012)) *The following conditions are sufficient for first order approach:*

- *In the linear experiments, if* $\alpha_i C''(\alpha_i) \geq f\left(\overline{\theta}\right)(\overline{\theta}-\mu)^2$ *for all* α_i*, then* $\pi''(\alpha_i) < 0$ *when either* $F(\theta)$ *is convex, or* $F(\theta) = \theta^b$ *(*$b > 0$*) with support* $[0, 1]$.

- *In the normal experiments,* $\pi''(\alpha_i) < 0$ *if* $\sqrt{\beta^3/\left[\alpha_i^3(\alpha_i+\beta)^5\right]} < 2\sqrt{2\pi}C''(\alpha_i)$ *for all* α_i.

I conclude this section by proving the following lemma which is used in the proof of Proposition 10.

Lemma 63 *If* $\pi''(\alpha_i) < 0$ *for all* α_i *and all non-decreasing implementable allocation rule* Q*, then*

$$\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -H_{\alpha_i}(w_i|\alpha_i)H(w_i|\alpha_i)^{n-1}\mathrm{d}w_i - C'(\alpha_i) \text{ is strictly decreasing in } \alpha_i. \tag{4.16}$$

Proof. In particular, $\pi''(\alpha_i) < 0$ if $Q(w_i) = H(w_i|\alpha_i)^{n-1}$ for all w_i. Then

$$\begin{aligned}
&\frac{\partial}{\partial\alpha_i}\left[\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -H_{\alpha_i}(w_i|\alpha_i)H(w_i|\alpha_i)^{n-1}\mathrm{d}w_i - C'(\alpha_i)\right]\\
=&\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -\frac{\partial^2 H(w_i|\alpha_i)}{\partial\alpha_i^2}Q(w_i)\mathrm{d}w_i - H_{\alpha_i}(w(1,\alpha_i)|\alpha_i)w_{\alpha_i}(1,\alpha_i)Q(w(1,\alpha_i))\\
&+ H_{\alpha_i}(w(0,\alpha_i)|\alpha_i)w_{\alpha_i}(0,\alpha_i)Q(w(0,\alpha_i)) - C''(\alpha_i) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -(n-1)H_{\alpha_i}(w_i|\alpha_i)^2 H(w_i|\alpha_i)^{n-2}\mathrm{d}w_i,\\
=&\pi''(\alpha_i) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -(n-1)H_{\alpha_i}(w_i|\alpha_i)^2 H(w_i|\alpha_i)^{n-2}\mathrm{d}w_i\\
<&0,
\end{aligned}$$

where $Q(w_i) = H(w_i|\alpha_i)^{n-1}$ for all w_i. ■

C.1.4. Efficient asymmetric mechanisms

Proof of Theorem 16

As in the symmetric case, I prove Theorem 16 by proving the following two lemmas. Define

$$Y(\boldsymbol{w}) := 1 - \prod_{i=1}^{n} H(w_i|\alpha_i^*) - \sum_{i=1}^{n}\int_{w_i}^{w(1,\alpha_i^*)} Q_i(z_i)\mathrm{d}H(z_i|\alpha_i^*), \forall \boldsymbol{w} \in \prod_{i=1}^{n}[w(0,\alpha_i^*), w(1,\alpha_i^*)].$$

Recall that $1 - \prod_{i=1}^{n} H(w_i|\alpha_i^*)$ is the probability with which there exists an agent i whose type is above w_i; and $\sum_{i=1}^{n}\int_{w_i}^{w(1,\alpha_i^*)} Q_i(z_i)\mathrm{d}H(z_i|\alpha_i^*)$ is the probability with which an agent whose type is above w_i receives the object. Let $\underline{\boldsymbol{w}} := (w(0,\alpha_1^*),\dots,w(0,\alpha_n^*))$. Then $Y(\underline{\boldsymbol{w}})$ is the difference between 1 and the probability with which some agent receives the object. Clearly, (4.18) is violated if and only if $Y(\underline{\boldsymbol{w}}) > 0$.

Lemma 64 *Suppose that the information structures are supermodular ordered, and $\alpha_i = \alpha_i^*$ for all i. Let $\boldsymbol{Q}$ be any interim allocation rule satisfying eqrefeq:AFprime* (MON), (AIA′) *and* $Y(\underline{\boldsymbol{w}}) > 0$. *Then, for any i, there exists $\hat{\boldsymbol{Q}}$ satisfying* (F′), (MON) *and* (AIA′) *such that $\hat{Q}_j = Q_j$ for $j \neq i$ and*

$$\hat{Q}_i(w_i) \geq Q_i(w_i),\ \forall w_i \in [w(0,\alpha_i^*), w(1,\alpha_i^*)], \tag{C.51}$$

and strict inequality holds for a set of w_i with positive measure.

Proof. Fix i. Define $Y_i(w_i) := \inf_{w_{-i}} Y(\boldsymbol{w})$ for all $w_i \in [w(0,\alpha_i^*), w(1,\alpha_i^*)]$. By Theorem 3 in Milgrom and Segal (2002), Y_i is differentiable and $Y_i'(w_i) = -h(w_i|\alpha_i^*)\prod_{j\neq i} H(w_j^*|\alpha_j^*) + Q_i(w_i)h(w_i|\alpha_i^*)$ where w_{-i}^* is such that $Y(w_i, w_{-i}^*) = Y_i(w_i)$ for all $w_i \in (w(0,\alpha_i^*), w(1,\alpha_i^*))$. Note that

$$Y(w(0,\alpha_i^*), w_{-i}) = 1 - \int_{w(0,\alpha_i^*)}^{w(1,\alpha_i^*)} Q_i(z_i)\mathrm{d}H(z_i|\alpha_i^*) - \sum_{j\neq i}^{n}\int_{w_j}^{w(1,\alpha_j^*)} Q_j(z_j)\mathrm{d}H(z_j|\alpha_j^*),$$

which is strictly increasing in w_j for all $j \neq i$. Hence, $Y_i(w(0,\alpha_i^*)) = Y(\underline{w}) > 0$. Define $w^\flat := \sup\left\{w_i \,\middle|\, Y_i(w_i') > 0\ \forall w(0,\alpha_i^*) \leq w_i' \leq w_i\right\}$. By the continuity of Y_i, we have $Y_i(w^\flat) = 0$ and $w^\flat > w(0,\alpha_i^*)$. There are four cases to consider.

Case I: Suppose that there exists $w_i' \in (w(0,\alpha_i^*), w^\flat)$ such that Q_i is discontinuous at w_i'. Let $Q_i(w_i'^{+})$ denote the right-hand limit of Q_i at w_i', and $Q_i(w_i'^{-})$ the corresponding left-hand limit. Let $0 < \varepsilon \leq \min\left\{\min_{w(0,\alpha_i^*)\leq w_i \leq w_i'} \frac{Y_i(w_i)}{H(w_i'|\alpha_i^*)}, Q_i(w_i'^{+}) - Q_i(w_i'^{-})\right\}$. Define $\boldsymbol{Q}$ as follows. If $w_i \leq$

$w(0,\alpha^*)$, then $\hat{Q}_i(w_i) := Q_i(w_i)$; and if $w_i > w(0,\alpha^*)$, then

$$\hat{Q}_i(w_i) := Q_i(w_i) + \varepsilon \chi_{\{w_i \le w_i'\}},$$

where $\chi_{\{w_i \le w_i'\}}$ is the indicator function. Let $\hat{Q}_j := Q_j$ for all $j \neq i$. By construction, $\hat{Q}_i(w) \geq Q_i(w)$ for all $w_i \in W_i$ and the inequality holds strictly on a positive measure set. By a similar argument to that in the proof of Lemma 12, $\hat{Q}_i$ satisfies (MON) and (AIA′). We now verify that $\hat{\boldsymbol{Q}}$ satisfies (AF′). If $w_i \le w_i'$, then $\hat{Y}(w_i, w_{-i}) = Y(w_i, w_{-i}) - \varepsilon[H(w_i'|\alpha^*) - H(w_i|\alpha^*)] \geq Y(w_i, w_{-i}) - \varepsilon H(w_i'|\alpha_{-i}^*) \geq 0$ for all w_{-i}. If $w_i > w_i'$, then $\hat{Y}(w_i, w_{-i}) = Y(w_i, w_{-i}) \geq 0$ for all w_{-i}. That is, $\hat{\boldsymbol{Q}}$ satisfies (AF′).

Case II: Suppose that Q_i is continuous on $[w(0,\alpha_i^*), w^\flat]$. We first show that there exists $w_i' \in (w(0,\alpha_i^*), w^\flat)$ such that $Q_i(w_i') < Q_i(w^\flat)$. Suppose, to the contrary, that $Q_i(w_i) = Q_i(w^\flat)$ for all $w_i \in (w(0,\alpha_i^*), w^\flat)$. Let w_{-i}^* be such that $Y(w^\flat, w_{-i}^*) = Y_i(w^\flat) = 0$. If $Q_i(w^\flat) \geq \prod_{j\neq i} H(w_j^*|\alpha_j^*)$, then $Y(w(0,\alpha_i^*), w_{-i}^*) = Y(w^\flat, w_{-i}^*) + \int_{w(0,\alpha_i^*)}^{w^\flat} \left[\prod_{j\neq i} H(w_j^*|\alpha_j^*) - Q_i(z)\right] h(z|\alpha_i^*)\mathrm{d}z < 0$, a contradiction. Hence, $Q_i(w^\flat) < \prod_{j\neq i} H(w_j^*|\alpha_j^*)$. Then, by the continuity of Q_i and H, there exists $\delta > 0$ such that $Q_i(w_i) < \prod_{j\neq i} H(w_j^*|\alpha_j^*)$ for all $w_i \in [w^\flat, w^\flat + \delta]$. Moreover,

$$0 = Y(w^\flat, w_{-i}^*) = \int_{w^\flat}^{w^\flat+\delta} \left[\prod_{j\neq i} H(w_j^*|\alpha_j^*) - Q(z)\right] h(z|\alpha_i^*)\mathrm{d}z + Y(w^\flat+\delta, w_{-i}^*) > Y(w^\flat+\delta, w_{-i}^*),$$

a contradiction. Hence, there exists $w_i' \in (w(0,\alpha_i^*), w^\flat)$ such that $Q_i(w_i') < Q_i(w^\flat)$.

By the continuity of Q_i, there exists $w_i'' \in (w_i', w^\flat)$ such that $Q_i(w_i'') = \frac{1}{2}\left(Q_i(w_i') + Q_i(w^\flat)\right)$. Let $0 < \varepsilon \le \min\left\{\min_{w(0,\alpha_i^*)\le w_i \le w_i''} \frac{Y_i(w_i)}{H(w_i''|\alpha_i^*)}, Q_i(w_i'') - Q_i(w_i')\right\}$. Let $\hat{Q}_j := Q_j$ for $j \neq i$ and

$$\hat{Q}_i(w_i) := \begin{cases} \max\{Q_i(w_i') + \varepsilon, Q_i(w_i)\} & \text{if } w_i > w_i', \\ Q_i(w_i) + \varepsilon & \text{if } w(0,\alpha_i^*) < w_i \le w_i', \\ Q_i(w_i) & \text{if } w_i \le w(0,\alpha_i^*). \end{cases}$$

Note that if $w_i \geq w_i''$ then $Q_i(w_i) \geq Q_i(w_i'') \geq Q_i(w_i') + \varepsilon$. Thus, $\hat{Q}_i(w_i) = Q_i(w_i)$ for $w_i \geq w_i''$. By construction, $\hat{Q}_i(w_i) \geq Q(w_i)$ for all $w_i \in W_i$ and the inequality holds strictly on a positive measure set. By a similar argument to that in the proof of Lemma 12, $\hat{Q}_i$ satisfies (MON) and (AIA′). We now verify that $\hat{\boldsymbol{Q}}$ satisfies (AF′). If $w_i \geq w_i''$, then $\hat{Y}(w_i, w_{-i}) = Y(w_i, w_{-i}) \geq 0$ for all w_{-i}. If $w_i < w_i''$, then for all w_{-i},

$$\begin{aligned}\hat{Y}(w_i, w_{-i}) =& Y(w_i, w_{-i}) - \int_{w_i}^{w_i''} \left[\hat{Q}_i(z) - Q_i(z)\right] h_i(z|\alpha_i^*)\mathrm{d}z,\\ \geq& Y(w_i, w_{-i}) - \varepsilon \left[H(w_i''|\alpha_i^*) - H(w_i|\alpha_i^*)\right],\\ \geq& Y_i(w_i) - \varepsilon H(w_i''|\alpha^*) \geq 0.\end{aligned}$$

Hence, $\hat{\boldsymbol{Q}}$ satisfies (AF′).

Case III: Let w_{-i}^* be such that $Y(w^\flat, w_{-i}^*) = Y_i(w^\flat) = 0$. Suppose that Q_i is continuous on $[w(0, \alpha_i^*), w^\flat)$ and $Q_i(w^{\flat-}) < \prod_{j\neq i} H(w_j^*|\alpha_j^*)$. Define $R(w_i) := Y_i(w_i)/(H(w^\flat|\alpha_i^*) - H(w_i|\alpha_i^*))$ for $w_i < w^\flat$. Then, by Theorem 3 in Milgrom and Segal (2002) and L'Hopital's rule,

$$\lim_{w_i \to w^{\flat-}} R(w_i) = \prod_{j\neq i} H_j(w_j^*|\alpha_j^*) - Q_i(w^{\flat-}) > 0.$$

Let $0 < \varepsilon \leq \min\left\{\inf_{w(0,\alpha_i^*)\leq w_i < w^\flat} R(w_i), Q_i(w^{\flat+}) - Q_i(w^{\flat-})\right\}$. Let $\hat{Q}_j := Q_j$ for all $j \neq i$. If $w_i \leq w(0, \alpha_i^*)$, then $\hat{Q}_i(w_i) \equiv Q_i(w_i)$; and if $w_i > w(0, \alpha_i^*)$, then $\hat{Q}_i(w_i) \equiv Q_i(w_i) + \varepsilon\chi_{\{w_i < w^\flat\}}$.

By construction, $\hat{Q}_i(w_i) \geq Q_i(w_i)$ for all $w_i \in W_i$ and the inequality holds strictly on a positive measure set. One can verify that $\hat{Q}_i$ satisfies (MON) and (AIA′) by an argument similar to that in the proof of Lemma 12. Finally, if $w_i < w^\flat$, then $\hat{Y}(w_i, w_{-i}) = Y(w_i, w_{-i}) - \varepsilon[H(w^\flat|\alpha_i^*) - H(w_i|\alpha_i^*)] \geq Y_i(w_i) - R(w_i)[H(w^\flat|\alpha^*) - H(w_i|\alpha^*)] = 0$ for all w_{-i}. If $w_i \geq w^\flat$, then $\hat{Y}(w_i, w_{-i}) = Y(w_i, w_{-i}) \geq 0$ for all w_{-i}. Hence, $\hat{\boldsymbol{Q}}$ satisfies (AF′).

Case IV: Let w_{-i}^* be such that $Y(w^\flat, w_{-i}^*) = Y_i(w^\flat) = 0$. Suppose that Q_i is continuous on $[w(0, \alpha_i^*), w^\flat)$ and $Q_i(w^{\flat-}) \geq \prod_{j\neq i} H(w_j^*|\alpha_j^*)$. We first show that $Q_i(w^{\flat-}) = \prod_{j\neq i} H(w_j^*|\alpha_j^*)$. Suppose, to the contrary, that $Q_i(w^{\flat-}) > \prod_{j\neq i} H(w_j^*|\alpha_j^*)$. Then, by the continuity of Q_i and H on

$[w(0,\alpha^*), w^\flat)$, there exists $\delta > 0$ such that $Q_i(w_i) > \prod_{j\neq i} H(w_j^*|\alpha_j^*)$ for all $w_i \in (w^\flat - \delta, w^\flat)$. Then

$$Y(w^\flat - \delta, w_{-i}) = \int_{w^\flat-\delta}^{w^\flat} \left[\prod_{j\neq i} H(w_j^*|\alpha_j^*) - Q_i(z)\right] h(z|\alpha_i^*)\mathrm{d}z < 0,$$

a contradiction. Hence, $Q_i(w^{\flat-}) = \prod_{j\neq i} H(w_j^*|\alpha_j^*)$.

Next, we show that there exists $w_i' \in (w(0,\alpha_i^*), w^\flat)$ such that $Q_i(w_i') < Q_i(w^{\flat-})$. Suppose, to the contrary, that $Q_i(w_i) = Q_i(w^{\flat-})$ for all $w_i \in (w(0,\alpha^*), w^\flat)$. Then

$$Y(w(0,\alpha^*), w_{-i}^*) = \int_{w(0,\alpha^*)}^{w^\flat} \left[\prod_{j\neq i} H(w_j^*|\alpha_j^*) - Q_i(z)\right] h(z|\alpha_i^*)\mathrm{d}z < 0,$$

a contradiction. Hence, there exists $w_i' \in (w(0,\alpha_i^*), w^\flat)$ such that $Q_i(w_i') < Q_i(w^{\flat-})$. The rest of the proof follows that of Case II. ■

Lemma 65 *Suppose that the information structures are supermodular ordered, and $\alpha_i = \alpha_i^*$ for all i. Let $\boldsymbol{Q}$ and $\hat{\boldsymbol{Q}}$ be two implementable allocation rules satisfying* (4.5). *Let $\boldsymbol{q}$ be an ex-post allocation rule that implements $\boldsymbol{Q}$. Then there exists an ex-post allocation rule $\hat{\boldsymbol{q}}$ that implements $\hat{\boldsymbol{Q}}$ and satisfies*

$$\mathbb{E}_{\boldsymbol{w}}\left[\sum_i (w_i + \gamma \sum_{j\neq i} w_j)\hat{q}_i(\boldsymbol{w}) \middle| \alpha_i = \alpha_i^* \ \forall i\right] > \mathbb{E}_{\boldsymbol{w}}\left[\sum_i (w_i + \gamma \sum_{j\neq i} w_j) q_i(\boldsymbol{w}) \middle| \alpha_i = \alpha_i^* \ \forall i\right].$$

The proof of Lemma 65 relies on the following technical lemma. For each i, let h_i denote the probability measure on $[w(0,\alpha_i^*), w(1,\alpha_i^*)]$ corresponding to $H(w_i|\alpha_i^*)$, then

Lemma 66 *Let $\boldsymbol{Q} : \prod_i [w(0,\alpha_i^*), w(1,\alpha_i^*)] \to [0,1]^n$ be an interim allocation rule and $\rho : \prod_i [w(0,\alpha_i^*), w(1,\alpha_i^*)] \to [0,1]$ be a measurable function. Then there exists an ex post allocation rule $\boldsymbol{q}$ that implements $\boldsymbol{Q}$ and satisfies*

$$\sum_i q_i(\boldsymbol{w}) \geq \rho(\boldsymbol{w}) \text{ for almost all } \boldsymbol{w} \in \prod_i [w(0,\alpha_i^*), w(1,\alpha_i^*)]$$

if and only if for each measurable set $\boldsymbol{A} = (A_1, \ldots, A_n)$ *where* $A_i \subset [w(0, \alpha_1^*), w(0, \alpha_i^*)]$ *for all i, the following inequality holds:*

$$\int_{\boldsymbol{A}} \rho(\boldsymbol{w}) \mathrm{d}h_1(w_1) \ldots \mathrm{d}h_n(w_n) \leq \sum_i \int_{A_i} Q(w_i) \mathrm{d}h_i(w_i) \leq \int_{\boldsymbol{A}} \rho(\boldsymbol{w}) \mathrm{d}h_1(w_1) \ldots \mathrm{d}h_n(w_n). \tag{C.52}$$

The proof of Lemma 14 can be readily extended to prove Lemma 66 and is neglected here. With Lemma 66 in hand, the proof of Lemma 13 can be readily extended to prove Lemma 65 and is also neglected here. Theorem 16 follows immediately given Lemmas 64 and 65.

Other omitted proofs

Lemma 67 *Suppose that the second-order condition of the agents' optimization problem is satisfied, and the information structures are uniformly supermodular ordered. Let* $\boldsymbol{\alpha}^*$ *be a socially optimal information choice. Suppose, in addition, that* (4.19) *holds. Then* (AIA′) *holds with equality for all* i*. Furthermore,*

$$\pi^s(\boldsymbol{\alpha}^*) = (1-\gamma)\left[\mu + \sum_{i=1}^n b(\alpha_i^*) C'(\alpha_i^*)\right] - \sum_i C(\alpha_i^*). \tag{4.20}$$

Proof. Let $\boldsymbol{\alpha}^*$ be a socially optimal information choice. Suppose, to the contrary, that $\boldsymbol{\alpha}^*$ is such that (4.19) holds with strictly inequality for some i. Then

$$\begin{aligned} V(\boldsymbol{\alpha}^*) &= \sum_{i=1}^n \int_{w(0,\alpha_i^*)}^{w(1,\alpha_i^*)} w \prod_{j \neq i} H(w|\alpha_j^*) h(w|\alpha_i^*) \mathrm{d}w \\ &= \int_{\underline{\theta}}^{\overline{\theta}} w \mathrm{d} \prod_i H(w|\alpha_i^*) \\ &= \overline{\theta} - \int_{\underline{\theta}}^{\overline{\theta}} \prod_i H(w|\alpha_i^*) \mathrm{d}w, \end{aligned}$$

where the last line holds by integration by parts. Hence,

$$\begin{aligned}\frac{\partial}{\partial \alpha_i}\pi^s(\boldsymbol{\alpha}^*) &= (1-\gamma)\frac{\partial}{\partial \alpha_i}V(\boldsymbol{\alpha}^*) - C'(\alpha_i^*) \\ &= (1-\gamma)\int_{\underline{\theta}}^{\overline{\theta}} -H_{\alpha_i}(w|\alpha_i^*)\prod_{j\neq i} H(w|\alpha_j^*)\mathrm{d}w - C'(\alpha_i^*),\end{aligned}$$

which is strictly decreasing in α_i if the second-order condition of the agents' optimization problem is satisfied. Because $\boldsymbol{\alpha}^*$ is such that (4.19) holds with strictly inequality for i,

$$\frac{\partial}{\partial \alpha_i}\pi^s(\boldsymbol{\alpha}^*) < -\gamma C'(\alpha_i^*) \leq 0.$$

Hence, $\boldsymbol{\alpha}^*$ is not optimal, a contradiction. Hence, if $\boldsymbol{\alpha}^*$ is chosen optimally, then (AIA′) holds with equality for all i.

Let $\boldsymbol{\alpha}^*$ be such that (4.19) holds. Because the information structures are uniformly supermodular ordered, and (AIA′) holds with equality, we have

$$\int_{w(0,\alpha_i^*)}^{w(1,\alpha_i^*)} \frac{w-\mu}{b(\alpha_i)}\prod_{j\neq i} H(w|\alpha_j^*)h(w|\alpha_i^*)\mathrm{d}w = C'(\alpha_i^*).$$

Hence,

$$\begin{aligned}V(\boldsymbol{\alpha}^*) &= \sum_{i=1}^{n}\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} w\prod_{j\neq i} H(w|\alpha_j^*)h(w|\alpha_i^*)\mathrm{d}w \\ &= \sum_{i=1}^{n} b(\alpha_i^*)C'(\alpha_i^*) + \mu\sum_{i=1}^{n}\int_{w(0,\alpha^*)}^{w(1,\alpha^*)}\prod_{j\neq i} H(w|\alpha_j^*)h(w|\alpha_i^*)\mathrm{d}w \\ &= \sum_{i=1}^{n} b(\alpha_i^*)C'(\alpha_i^*) + \mu\int_{\underline{\theta}}^{\overline{\theta}} w\mathrm{d}\prod_{i} H(w|\alpha_i) \\ &= \sum_{i=1}^{n} b(\alpha_i^*)C'(\alpha_i^*) + \mu.\end{aligned}$$

Hence,

$$\pi^s(\boldsymbol{\alpha}^*) = (1-\gamma)\left[\mu + \sum_{i=1}^{n} b(\alpha_i^*)C'(\alpha_i^*)\right] - \sum_{i} C(\alpha_i^*). \tag{4.20}$$

This completes the proof. ■

BIBLIOGRAPHY

N. I. Al-Najjar. Aggregation and the law of large numbers in large economies. *Games and Economic Behavior*, 47(1):1–35, 2004.

E. Ben-Porath, E. Dekel, and B. L. Lipman. Optimal allocation with costly verification. *American Economic Review*, 104(12):3779–3813, 2014.

J.-P. Benoit and V. Krishna. Multiple-object auctions with budget constrained bidders. *Review of Economic Studies*, pages 155–179, 2001.

D. Bergemann and J. Välimäki. Information acquisition and efficient mechanism design. *Econometrica*, 70(3):1007–1033, 2002.

D. Bergemann, X. Shi, and J. Välimäki. Information acquisition in interdependent value auctions. *Journal of the European Economic Association*, 7(1):61–89, 2009.

B. Biais, D. Martimort, and J.-C. Rochet. Competing mechanisms in a common value environment. *Econometrica*, 68(4):799–837, 2000.

K. C. Border. Implementation of reduced form auctions: A geometric approach. *Econometrica*, pages 1175–1187, 1991.

K. C. Border and J. Sobel. Samurai accountant: A theory of auditing and plunder. *The Review of Economic Studies*, 54(4):525–540, 1987. doi: 10.2307/2297481.

S. Brusco and G. Lopomo. Simultaneous ascending bid auctions with privately known budget constraints. *Journal of Industrial Economics*, 56(1):113–142, 2008.

Y.-K. Che and I. Gale. Standard auctions with financially constrained bidders. *The Review of Economic Studies*, 65(1):1–21, 1998.

Y.-K. Che and I. Gale. The optimal mechanism for selling to a budget-constrained buyer. *Journal of Economic Theory*, 92(2):198–233, 2000.

Y.-K. Che and I. Gale. Revenue comparisons for auctions when bidders have arbitrary types. *Theoretical Economics*, 1(1):95–118, 2006.

Y.-K. Che, I. Gale, and J. Kim. Assigning resources to budget-constrained agents. *The Review of Economic Studies*, 80:73–107, 2013a.

Y.-K. Che, J. Kim, and K. Mierendorff. Generalized reduced-form auctions: A network-flow approach. *Econometrica*, 81(6):2487–2520, 2013b.

D. Coady, M. Grosh, and J. Hoddinott. Targeting outcomes redux. *The World Bank Research Observer*, 19(1):61–85, 2004.

W. Cook, W. Cunningham, W. Pulleyblank, and A. Schrijver. *Combinatorial Optimization*. Wiley Series in Discrete Mathematics and Optimization. Wiley, 2011. ISBN 9781118031391.

J. Crémer, Y. Spiegel, and C. Z. Zheng. Auctions with costly information acquisition. *Economic Theory*, 38(1):41–72, 2009.

J. Currie and F. Gahvari. Transfers in cash and in-kind: theory meets the data. *Journal of Economic Literature*, 46(2):333–383, 2008.

G. Folland. *Real analysis: modern techniques and their applications*. Pure and applied mathematics. Wiley, 1999. ISBN 9780471317166.

A. Frank. *Connections in Combinatorial Optimization*. Oxford Lecture Series in Mathematics and Its Applications. OUP Oxford, 2011. ISBN 9780199205271.

D. Gale and M. Hellwig. Incentive-compatible debt contracts: The one-period problem. *The Review of Economic Studies*, 52(4):647–663, 1985. doi: 10.2307/2297737.

J.-J. Ganuza and J. S. Penalva. Signal orderings based on dispersion and the supply of private information in auctions. *Econometrica*, 78(3):1007–1030, 2010. ISSN 1468-0262. doi: 10.3982/ECTA6640.

J. Glazer and A. Rubinstein. On optimal rules of persuasion. *Econometrica*, 72(6):1715–1736, 2004. ISSN 1468-0262. doi: 10.1111/j.1468-0262.2004.00551.x.

P. Haile, K. Hendricks, and R. Porter. Recent us offshore oil and gas lease bidding: A progress report. *International Journal of Industrial Organization*, 28(4):390–396, 2010.

M. Harris and R. M. Townsend. Resource allocation under asymmetric information. *Econometrica: Journal of the Econometric Society*, pages 33–64, 1981.

J. Harsanyi. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, 63, 1955.

R. Hassin. Minimum cost flow with set-constraints. *Networks*, 12(1):1–21, 1982.

J. P. Johnson and D. P. Myatt. On the simple economics of advertising, marketing, and product design. *The American Economic Review*, pages 756–784, 2006.

K. L. Judd. The law of large numbers with a continuum of iid random variables. *Journal of Economic theory*, 35(1):19–25, 1985.

J.-J. Laffont and J. Robert. Optimal auction with financially constrained buyers. *Economics Letters*, 52(2):181–186, 1996.

E. L. Lehmann. Comparing location experiments. *The Annals of Statistics*, 16(2):pp. 521–533, 1988. ISSN 00905364.

D. Levin and J. L. Smith. Equilibrium in auctions with entry. *The American Economic Review*, pages 585–599, 1994.

T. R. Lewis and D. E. Sappington. Supplying information to facilitate price discrimination. *International Economic Review*, pages 309–327, 1994.

Y. Li. Mechanism design with costly verification and limited punishments. 2016.

A. Malakhov and R. V. Vohra. Optimal auctions for asymmetrically budget constrained bidders. *Review of Economic Design*, 12(4):245–257, 2008.

E. Maskin and J. Riley. Optimal auctions with risk averse buyers. *Econometrica: Journal of the Econometric Society*, pages 1473–1518, 1984a.

E. S. Maskin. Auctions, development, and privatization: Efficient auctions with liquidity-constrained buyers. *European Economic Review*, 44(4):667–681, 2000.

E. S. Maskin and J. G. Riley. Optimal auctions with risk averse buyers. *Econometrica*, 52(6):1473–1518, 1984b.

S. Matthews. Information acquisition in discriminatory auctions. *Bayesian models in economic theory, ed. by M. Boyer, and R. Kihlstrom*, 49:1477–1500, 1984a.

S. A. Matthews. On the implementability of reduced form auctions. *Econometrica*, 52(6):1519–1522, 1984b.

S. A. Matthews. On the implementability of reduced form auctions. *Econometrica: Journal of the Econometric Society*, pages 1519–1522, 1984c.

K. Mierendorff. Optimal dynamic mechanism design with deadlines. *University of Bonn*, 2009.

K. Mierendorff. Asymmetric reduced form auctions. *Economics Letters*, 110(1):41–44, 2011.

K. Mierendorff. Optimal dynamic mechanism design with deadlines. *Journal of Economic Theory*, 161:190–222, 2016.

P. Milgrom and I. Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601, 2002. ISSN 1468-0262. doi: 10.1111/1468-0262.00296.

D. Mookherjee and I. Png. Optimal auditing, insurance, and redistribution. *The Quarterly Journal of Economics*, 104(2):399–415, 1989. doi: 10.2307/2937855.

R. B. Myerson. Incentive compatibility and the bargaining problem. *Econometrica: journal of the Econometric Society*, pages 61–73, 1979.

R. B. Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981.

T. Mylovanov and A. Zapechelnyuk. Mechanism design with ex-post verification and limited punishments. 2014.

M. M. Pai and R. Vohra. Optimal auctions with financially constrained buyers. *Journal of Economic Theory*, 150:383–425, 2014a.

M. M. Pai and R. V. Vohra. Optimal auctions with financially constrained buyers. *Journal of Economic Theory*, 150:383–425, 2014b. ISSN 00220531. doi: 10.1016/j.jet.2013.09.015.

R. Patel and C. Urgan. Costly inspection and money burning in internal capital markets. 2017.

N. Persico. Information acquisition in auctions. *Econometrica*, 68(1):135–148, 2000.

D. Quint and K. Hendricks. Indicative bidding in auctions with costly entry. Technical report, Working paper, 2013.

W. T. Reid. A simple optimal control problem involving approximation by monotone functions. *Journal of Optimization Theory and Applications*, 2(6):365–377, 1968.

M. Richter. Mechanism design with budget constraints and a continuum of agents. 2015.

J.-C. Rochet and L. A. Stole. The economics of multidimensional screening. *Econometric Society Monographs*, 35:150–197, 2003.

J. Rong, N. Sun, and D. Wang. A new evaluation criterion for allocation mechanisms with application to vehicle license allocations in china. 2015.

A. Seierstad and K. Sydsæter. *Optimal Control Theory with Economic Applications*. Advanced textbooks in economics. North-Holland, 1987. ISBN 9780444879233.

X. Shi. Optimal auctions with information acquisition. *Games and Economic Behavior*, 74(2): 666–686, 2012.

M. Stegeman. Participation costs and efficient auctions. *Journal of Economic Theory*, 71(1):228–259, 1996.

D. Szalay. Contracts with endogenous information. *Games and Economic Behavior*, 65(2):586–625, 2009.

J. Toikka. Ironing without control. *Journal of Economic Theory*, 146(6):2510–2526, 2011.

R. M. Townsend. Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory*, 21(2):265 – 293, 1979. ISSN 0022-0531. doi: http://dx.doi.org/10.1016/0022-0531(79)90031-0.

R. M. Townsend. Information constrained insurance: the revelation principle extended. *Journal of Monetary Economics*, 21(2):411–450, 1988.

W. Vickrey. Measuring marginal utility by reactions to risk. *Econometrica*, 13(4):319–333, 1945.

R. V. Vohra. *Mechanism Design: a linear programming approach*, volume 47. Cambridge University Press, 2011.

R. V. Vohra. Optimization and mechanism design. *Mathematical programming*, pages 1–21, 2012.

L. Ye. Optimal auctions with endogenous entry. *Contributions in Theoretical Economics*, 4(1), 2004.