NORTHWESTERN UNIVERSITY

Approximate Optimality of Simple Mechanisms

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Computer Science

By

Yingkai Li

EVANSTON, ILLINOIS

June 2022

ABSTRACT

Approximate Optimality of Simple Mechanisms

Yingkai Li

We consider general utility models and information structures of the agents and illustrate when economic conclusions for designing simple mechanisms in classical settings extends for general environments. We show that whether economic conclusions can be generalized depends on the details of the generalizations. For example, in single-item auction, competition and non-anonymity are not crucial factors for revenue maximization when agents have linear utilities [Yan, 2011, Alaei et al., 2018], and these conclusions extend for broad classes of non-linear utilities. In comparison, the economic conclusions we derived for exogenous information settings often fail when the information is endogenous. For example, in multi-dimensional information acquisition problems, scoring the agent separately is without loss when the signals are exogenous, but suffers a great loss when the signals are endogenous. In selling information problems, price discrimination and commitment to revealing partial information are crucial for revenue maximization if the agent has an exogenous signal about the unknown state [Bergemann et al., 2021]. However, pricing for full information is approximately optimal when the signal is endogenous.

Acknowledgements

I want to thank my advisor Jason Hartline for his guidance during my PhD study. It has been a great pleasure to work with him for these years, and I have learned a lot from him, including but not limited to how to develop a good taste for researches, how to solve hard math questions, or how to communicate with other researchers to convey the novelty of our results. These great experience will continue to be helpful for me when I pursue my career in academics in the future.

I also want to express my gratitude to Harry Pei, who essentially acted as my second advisor from the economic department. Without his help, I wouldn't be able to discover many interesting topics that economists are excited about, or learn how to think as an economist when conducting economic researches. He really helped me enter the greater domain of economics, which eventually lead me to accomplish great work beyond the classical computer science perspective.

I am also grateful to Brendan Lucier, Nicole Immorlica, Alex Slivkins and Moshe Babaioff for their hospitality when I visited Microsoft Research for internship during the summer of 2020 and 2021. This is the first time I had an internship at an industrial research lab, and it is really exciting for me to see how theory is conducted in a more applied perspective and connected to real-world applications.

Moreover, I want to thank all my dissertation committee members, Jason Hartline, Harry Pei, Brendan Lucier, Yannai Gonczarowski and Annie Liang. My thesis has benefited a lot from the discussions with them, and their valuable comments and suggestions.

I'm also very fortunate to have worked with many other brilliant researchers in this field: Bo Li, Yaonan Jin, Pinyan Lu, Yuan Zhou, Yiding Feng, Jing Chen, Xiaotie Deng, Edmund Lou, Liren Shan, Yifan Wu, Aleck Johnsen, Yining Wang, Jason Gaitonde, Bar Light, Haoran Ye, Kefan Dong, Qin Zhang, Xiaowei Wu, Inbal Talgam-Cohen, Tal Alon, Paul Dütting, Hedyeh Beyhaghi, Matt Weinberg, Haifeng Xu, and Boli Xu. I would not be able to write many nice papers without your help.

It has been a wonderful time for me, both academically and personally, to spend my PhD at Northwestern University. It is pleasure to be a part of this amazing community, which includes many great faculties: Samir Khuller, Ben Golub, Konstantin Makarychev, Aravindan Vijayaraghavan and Yingni Guo; and many of my colleges and friends: Modibo Camara, Aidao Chen, Xue Chen, Saba Ahmadi, Huck Bennett, Abhratanu Dutta, Sanchit Kalhan, Jinshuo Dong, Matthew vonAllmen, Shirley Zhang, Michalis Mamakos, Sheng Long, Aravind Reddy, Leif Rasmussen, Sheng Yang, Chenhao Zhang, Charles Cui, Anant Shah, Jingyuan Wang, Anran Li and Xiaoyun Qiu.

Finally and mostly importantly, I would like to thank my parents, Tianwang Li and Lianying Wang, my wife Wenyang Li, and all other family members. I could not accomplish any of this without your love and support.

Table of Contents

List of Figures

CHAPTER 1

Introduction

The *method of approximation* quantifies the extent to which a theory can be generalized from ideal models and enables the separation of details from salient features of the model. Given an (possibly complicated, detail-dependent) optimal mechanism for an objective like revenue, there may exist other (simple, detail-free) mechanisms that *approximate* (i.e., attain a "large" fraction of) the optimal revenues. In this case we may say that the theory behind the simple mechanisms generalizes from the ideal model. Otherwise, it does not.[1]

There are extensive studies of simple mechanisms with approximation guarantees for the classical mechanism design problems with specific assumptions on the agents' utility models. For example, Yan [2011] shows that when agents have linear utilities, sequential posted pricings, which arrange the agents in an order and offer while-supplies-last posted prices, guarantee an $e/(e-1)$-approximation, i.e., the best order and prices achieves at least 63.2% of the optimal auction revenue. This implies that simultaneity and competition are not necessary drivers for revenue maximization for linear utility agents. Another example is information acquisition, in order to elicit high-dimensional information, it is often without loss to elicit marginal information on each dimensional separately [Lambert, 2011]. In the model of selling information, Bergemann et al. [2021] show that when the information of the agents is exogenous, pricing for full information cannot guarantee any

[1]See the survey of Hartline [2012] for detailed discussion of the method of approximation in economics.

non-trivial fraction of the optimal revenue. This implies that price discrimination is a crucial factor for revenue maximization given exogenous information.

We study the problem that to what extend the economic conclusions from the classical mechanism design literature generalize for agents with general utility models and general information structures. In particular, in this thesis, we focus on two specific generalizations: from linear utilities to non-linear utilities, and from exogenous information to endogenous information. We show that many economic lessons we obtained for linear utilities generalize for broad class of non-linear utilities, while the economic lessons we obtained for exogenous information fail for endogenous information. Therefore, there is no unified solutions for generalizing the economic conclusions to complex models, and more investigation is required for understanding the performance of simple mechanisms in various general environments. In the next section, we will summarize the main results in this thesis.

1.1. Main Contributions

1.1.1. Non-linear Utilities

For classical auction design for agents with linear utilities, Bulow and Roberts [1989] show that the marginal revenue maximization mechanism is revenue optimal, drawing a close connection between the classical microeconomics and the auction theory. Yan [2011] shows that sequential posted pricings guarantee an $e/(e-1)$-approximation, which suggests the relative irrelevance of simultaneity and competition for revenue maximization. Jin et al. [2019] further show that when the agents' (non-identical) value distributions satisfy a concavity property, a.k.a., "regular distributions", posting an anonymous price guarantees a 2.62-approximation, which implies that discrimination across different agents is not essential either.

We generalize these approximation results from linear agents to non-linear agents.[2] From this generalization, not only do we observe that the main drivers of good mechanisms are similar for non-linear agents, but also that non-linearity itself is not a main concern that necessitates specialized mechanism designs (beyond the approach of our generalization).

Bulow and Roberts [1989], as later interpreted by Alaei et al. [2013], show that to design optimal mechanisms for linear agents, it is without loss to restrict attention to *pricing-based mechanisms*, i.e., mechanisms where the menu offered to each agent is equivalent to a distribution over posted prices. The multi-agent mechanism design problem can be decomposed as single-agent mechanism design problems through the reduced-form approach of Border [1991]. From Bulow and Roberts [1989], the solution to these single-agent problems for linear agents are (possibly randomized) price postings and the optimal mechanism can be interpreted as marginal revenue maximization. Thus, every mechanism for linear agents is equivalent to a pricing-based mechanism.

Pricing-based mechanisms can be generalized to non-linear agents by considering *per-unit* prices, i.e., given per-unit price p, an agent can purchase any lottery with winning probability $q \in [0, 1]$ and pay price $p \cdot q$ in expectation. For non-linear agents (e.g., agents with budget constraints), not all mechanisms can be interpreted as pricing-based mechanisms and, in fact, pricing-based mechanisms are not generally optimal. Nonetheless, we show that these mechanisms are approximately optimal for large families of non-linear

[2]In this thesis, we write "agents with linear utilities" as "linear agents" for short, and "agents with non-linear utilities" as "non-linear agents".

agents. For these families we say that the non-linear agents *resemble* linear agents. We introduce a reduction framework. Given a pricing-based mechanism that guarantees a β-approximation (i.e., achieves at least $1/\beta$ fraction of the optimal objective) for linear agents and given non-linear agents that are ζ-*resemblant*[3] of linear agents and satisfy the von Neumann-Morgenstern expected utility representation [Morgenstern and von Neumann, 1953], the reduction framework transforms the aforementioned pricing-based mechanism for linear agents into an analogous pricing-based mechanism for the non-linear agents. The non-linear agent mechanism guarantees a $\beta\zeta$-approximation bound.

The reduction framework can be combined with approximation results for linear agents to show that simple mechanisms such as marginal revenue maximization, sequential posted pricing, and anonymous pricing are approximately optimal for non-linear agents that resemble linear agents, and the economic lessons (e.g., non-cruciality of simultaneity, competition, discrimination) derived from those mechanisms for linear agents can be lifted to non-linear agents. As an example, agents with independent private budget and regular valuation distribution are 3-resemblant of linear agents, which implies that the approximation of sequential posted pricing for such non-linear agents is $3e/(e-1)$.

This thesis characterizes broad families of non-linear agents that are ζ-resemblant for small constant factors ζ (e.g., agents with independent private budget and regular valuation distribution) and families that are not (e.g., agents whose budget and value are correlated). For non-linear agents that are ζ-resemblant, pricing-based mechanisms are approximately optimal wherever they are approximately optimal for linear agents; thus, non-linearity of utility can be viewed as a detail that can be omitted from the

[3]We measure the resemblance of agents in terms of the (topological) closeness of their revenue curves, as defined in Bulow and Roberts [1989]. We provide the details in Chapter 2.

model without significantly altering the main take-aways. On the other hand, with utility models that are not ζ-resemblant for modest ζ, non-linearity is a crucial feature that needs specific study for identifying forms of mechanisms lead to good economic outcomes.

Our reduction framework can be applied more broadly for non-linear agents beyond the expected utility theory with the restriction to *posted pricing mechanisms*[4] (e.g., sequential posted pricing, anonymous pricing). For instance, when agents have stochastic outside options – which can be viewed as a special form of non-linear utility that does not satisfy expected utility theory – are 2-resemblant under a concavity assumption. Thus, for such agents, sequential posted pricing is approximately optimal and the economic lessons from previous discussions generalize.

1.1.2. Information Acquisition

We consider the problem of an uniformed principal acquiring information from a strategic agent. In particular, we formalize the problem as optimizing scoring rules for reporting the expectation of a incentivizing the agent to exert effort on acquiring additional information.

Proper scoring rules incentivize a forecaster to reveal her true belief about an unknown and probabilistic state. The principal publishes a scoring rule that maps the reported belief and the realized state to a reward for the forecaster. The forecaster reports her belief about the state. The state is realized and the principal rewards the forecaster according to the scoring rule. A scoring rule is proper if the forecaster's optimal strategy, under any belief she may possess, is to report that belief. Proper scoring rules are also designed for directly eliciting a statistic of the distribution such as its expectation.

[4]Posted pricing mechanisms are pricing-based mechanisms where prices posted to each agent do not depend on actions of other agents.

Not all proper scoring rules work well in any a given scenario. This thesis considers a mathematical program for optimization of scoring rules where (a) the objective captures the incentive for the forecaster to exert effort and (b) the boundedness constraints prevent the principal from scaling the scores arbitrarily. For (a), we focus on a simple binary model of effort where the forecaster does or does not exert effort and with this effort the forecaster obtains a refined posterior distribution from the prior distribution on the unknown state (e.g., by obtaining a signal that is correlated with the state). We adopt the objective that takes the perspective of the forecaster at the point of the decision with knowledge of both the prior and the distributions of posteriors that is obtained by exerting effort. We want a scoring rule that maximizes the difference in expected scores for the posterior distribution and prior distribution. For (b), we impose the ex post constraint that the score is in a bounded range, i.e., without loss, between zero and one. Notice that this program would be meaningless without a constraint on the scores - otherwise the score could be scaled arbitrarily - and it would be meaningless without considering the difference in scores between posterior and prior - otherwise any bounded scoring rule scaled towards zero plus a constant close to the upper bound would be near optimal.

We solve for the optimal scoring rule for reporting the expectation in single-dimensional space. As we expect for single-dimensional mechanism design problems for an agent with linear utility [Myerson, 1981b], the optimal scoring rule is a step function (which induces a V-shaped scoring rule with its lower tip at the expectation of the prior belief). To implement this V-shaped scoring rule, it is sufficient for the designer to know the prior mean instead of the details on the distribution over posteriors. We also demonstrate a first result for prior-independent analysis of scoring rules. Among scoring rules for reporting the expectation, the quadratic scoring rule is within a constant factor of optimal.

For multi-dimensional forecasting, without concern of acquiring additional information, a simple choice of the principal is to elicit information separately across different dimensions, and the aggregated information would still be accurate.

When the agent can acquire a costly signal, we show that the gap between acquiring information separately and optimally can be linear in the size of the dimensions. For symmetric distributions, we give an analytical characterization of the optimal scoring rule as inducing a V-shaped utility function. For multi-dimensional forecasting without a symmetry assumption, we identify a V-shaped scoring rule that gives an 8-approximation. This scoring rule can be interpreted as scoring the dimension for which the agent's posterior in the optimal single-dimensional scoring rule gives the highest utility. Equivalently, it can be implemented by letting the agent select which dimension to score and only scoring that dimension (after exerting effort to learn the posterior mean of all dimensions). Moreover, while optimal mechanisms generally depend on the distribution over posteriors, our approximation bounds are proved for simple mechanisms (V-shaped scoring rules) that depend only on the prior mean, and do not require detailed knowledge of the distribution over posteriors.

1.1.3. Selling Information

We consider the problem of maximizing the revenue of the data broker, where the agent can endogenously acquire additional information. Specifically, there is an unknown state and both the data broker and the agent have a common prior over the set of possible states.

The data broker can offer a menu of information structures for revealing the states with associated prices to the agent. Then the agent picks the expected utility maximization entry from the menu, and pays the corresponding price to the data broker. The agent has a private valuation for information and can acquire additional costly information upon receiving the signal from the data broker. The literature has acknowledged the possibility for the agents to conduct their own experiments to be privately informed of the states [e.g., Bergemann, Bonatti, and Smolin, 2018]. The distinct feature in our model is that the decision for acquiring additional information is endogenous. Specifically, after receiving the signal from the data broker, the agent can subsequently acquire additional information with costs. For example, the agent is a decision maker who chooses an action to maximize her expected utility based on her posterior belief over the states. The agent will first acquire information from the data broker, and based on her posterior, she can potentially conduct more experiments to refine her belief before taking the action. Another example captured in our model is where the agent is a firm that sells products to consumers, and the information the data broker provides is a market segmentation of the consumers. The firm has a private and convex cost for producing different level of qualities for the product, and the firm can conduct his own experiments (e.g., sending surveys to potential consumers) with additional costs to further segment the market after receiving the information from the data broker.[5] In addition, in our model, we allow the firm to repeat the market research until it is not beneficial to do so, i.e., when the cost of information exceeds the marginal benefits of information. This captures the situation

[5]Yang [2020] studies a similar model, where the firm cannot conduct his own market research to refine his knowledge. Moreover, the cost function of the firm is linear in Yang [2020], which leads to a qualitatively different result compared to our model. See Section 4.3 for a detailed discussion.

that the firm can decide the date for announcing the product to the market, and before the announcement, the firm sends out surveys to potential consumers each day to learn the segmentation of the markets. At the end of each day, the firm receives an informative signal through the survey, and decides whether to continue the survey in next day, or stop the survey and announce the product with corresponding market prices to the public.

When the private information of the agent is exogenous, [e.g., Bergemann, Bonatti, and Smolin, 2018] show that in the revenue optimal mechanism, the menu complexity can be linear in the number of action choices of the agent in the worst case, and posting a deterministic price for revealing full information cannot guarantee a constant approximation to the optimal. This suggests that third degree price discrimination is crucial for revenue maximization in exogenous information setting.

In this model of allowing the agent to acquire information endogenously, we impose a linearity assumption on the agent's private preference over different experiments, i.e., the value of the agent for any posterior distribution is simply the product of her private type and the value of the posterior distribution. In the examples we provided in previous paragraphs, both the decision maker who chooses an optimal action to maximize her payoff based on the posterior belief and the firm that sells products to consumers to maximize the revenue satisfy the linear valuation assumption. Essentially, this condition assumes that the private type of the agent represents her value for additional information, and there is a linear structure on the preference. It excludes the situation where the private type of the agent represents an exogenous private signal correlated with the states. We show that with linear valuations, when the agent can acquire additional costly information, there exists a threshold type θ^* such that (1) for any type $\theta \geq \theta^*$, the optimal mechanism reveals full information to the agent; and (2) for any type $\theta < \theta^*$, the optimal mechanism may reveal partial information and the individual rational constraint always binds. The first statement is the standard no distortion at the top observation in the optimal mechanisms. The second statement suggests that the optimal mechanism may discriminate lower types of the agent by offering the options of revealing partial information to the agent with lower prices. Moreover, the allocations and the prices for those lower types are set such that the agent is exactly indifferent between participation and choosing the outside option (by conducting her own experiments with additional costs). Our characterizations suggest that the optimal mechanisms for selling information may be complex and contain a continuum of menu entries when the information is endogenous. However, posted pricing for revealing full information achieves at least half of the optimal revenue in the worst case. This suggests that price discrimination is not crucial for approximating the optimal revenue in endogenous information setting, which leads to a sharp contrast to the exogenous information setting.

1.2. Related Work

Non-linear Utilities. Frameworks for reducing approximation for non-linear agents to approximation for linear agents has also been studied in Alaei et al. [2013]. This reduction framework converts the marginal revenue mechanism for linear agents to a mechanisms for non-linear agents and general objectives. Their reduction framework is also applicable to other DSIC, IIR, deterministic mechanisms for linear agents. Unlike our framework which uses single-agent price-posting mechanisms (induced from price-posting payoff curves) as a building-block, Alaei et al. [2013] convert mechanisms for linear agents into mechanisms for non-linear agents with single-agent ex ante optimal mechanisms (induced from optimal payoff curves) as components. From the mechanism designer's perspective, identifying ex ante optimal mechanisms for a single non-linear agents can be much harder than identifying ex ante optimal price-posting mechanisms (e.g., private budget utility, risk averse utility). Furthermore, due to this difference, the implementation of the reduction framework together with its outcome mechanisms in Alaei et al. [2013] is more complex than ours. In general, the framework in Alaei et al. [2013] converts DSIC mechanisms for linear agents into Bayesian incentive compatible mechanisms for non-linear agents.

Mechanism design for non-linear agents is well studied in the literature. In this work, as applications of our general framework, we focus on three specific non-linear models, agents with budget constraints, agents with risk averse attitudes, and agents with endogenous valuation.

Laffont and Robert [1996] and Maskin [2000] study the revenue-maximization and welfare-maximization problems for symmetric agents with *public* budgets in single-item environments. Boulatov and Severinov [2018] generalize their results to agents with i.i.d. values but asymmetric public budgets. Che and Gale [2000] consider the single agent problem with *private* budget and valuation distribution that satisfies declining marginal revenues, and characterize the optimal mechanism by a differential equation. Devanur and Weinberg [2017a] consider the single agent problem with private budget and an arbitrary valuation distribution, characterize the optimal mechanism by a linear program, and use an algorithmic approach to construct the solution. Pai and Vohra [2014] generalize the characterization of the optimal mechanism to symmetric agents with uniformly distributed private budgets. Richter [2019] shows that a price-posting mechanism is optimal for selling a divisible good to a continuum of agents with private budgets if their valuations are regular with decreasing density. For more general settings, no closed-form characterizations are known. However, the optimal mechanism can be solved by a polynomial-time solvable linear program over interim allocation rules [cf. Alaei et al., 2012, Che et al., 2013].

Most results for agents with risk-averse utilities consider the comparative performance of the first- and second-price auctions, cf., Holt Jr [1980], Che and Gale [2006]. Matthews [1983] and Maskin and Riley [1984], however, characterize the optimal mechanisms for symmetric agents for constant absolute risk aversion and more general risk-averse models. Baisa [2017] shows that the optimal mechanism for risk averse agents departs from the linear agents, since the optimal mechanism does not allocate to the highest bidder, and can better screen the agents through allocating the item to a group of agents with lotteries. Gershkov et al. [2021b] show that if the seller can make positive transfer to the agents, the optimal mechanism features the property that under equilibrium, all agents face no uncertainty in the realized utility.

The model for agents with endogenous valuation has been studied extensively in Tan [1992], King et al. [1992], Gershkov et al. [2021a], Akbarpour et al. [2021] where agents can make costly investment before the auction. This is a generalization of the model for agents with entry costs [Celik and Yilankaya, 2009]. This main focus of the literature is to characterize the optimal mechanisms in restricted settings. For example, Gershkov et al. [2021a] characterize the revenue optimal symmetric mechanism for symmetric buyers.[6] The reduction framework in this thesis implies that sequentially offering a price to each

[6]Gershkov et al. [2021a] also showed that even for symmetric buyers, symmetric mechanism may not be revenue optimal among all possible mechanisms.

agent is a constant approximation for both welfare and revenue maximization when there are multiple asymmetric buyers. Akbarpour et al. [2021] consider approximating the optimal welfare when it is computationally intractable to find the optimal allocation. They show that any algorithm that excludes bossy negative externalities can be converted to a mechanism that guarantees the same approximation ratio to the optimal welfare. They restrict attention to full information equilibrium, while our analysis applies to settings with private valuations.

It is well known that simple mechanisms generate robust performance guarantees for both welfare maximization [Roughgarden et al., 2017] and revenue maximization [Carroll, 2017, Bei et al., 2019]. Moreover, simple mechanisms are approximately optimal under natural assumptions of type distributions. For single item auction and linear agents, Jin et al. [2019] show that the tight ratio between anonymous pricing and the optimal mechanism is 2.62 under regularity assumption, and Yan [2011] shows that the tight approximation ratio is ${}^{e}/_{(e-1)}$ for sequential posted pricing. The approximate optimality of sequential posted pricing can be generalized to multi-item settings when agents have unit-demand valuations [Chawla et al., 2010, Cai et al., 2016]. For non-linear agents, given matroid environments, Chawla et al. [2011] show that a simple lottery mechanism is a constant approximation to the optimal pointwise individually rational mechanism for agents with monotone-hazard-rate valuations and private budgets. In contrast, our approximation results are with respect to the optimal mechanism under interim individually rationality which can be arbitrarily larger than the benchmark from Chawla et al. [2011]. For multiple items, Cheng et al. [2018] shows that selling items separately or as a bundle is approximately optimal for a single agent with additive valuation. Our analyses uses one of their lemmas.

Scoring Rules. Characterizations of scoring rules for eliciting the mean and for eliciting a finite-state distribution play a prominent role in our analysis. Previous works show, in various contexts, that scoring rules are proper if and only if their induced utility functions are convex. McCarthy [1956] characterized proper scoring rules for eliciting the full distribution on a finite set of states. Osband and Reichelstein [1985] characterized continuously differentiable scoring rules that elicit multiple statistics of a probability distribution. Lambert [2011] characterized the statistics that admit proper scoring rules and characterized the uniformly-Lipschitz-continuous scoring rules for the mean of a single-dimensional state. Abernethy and Frongillo [2012] characterized the proper scoring rules for the marginal means of multi-dimensional random states in the interior of the report space. We augment this characterization by showing that the induced utility function converges to a limit on the boundary of the report space. This augmentation enables us to write the mathematical program that optimizes over the whole report space.

Most of the prior work looking at incentives of eliciting information considers a fundamentally different model from ours. This prior work typically focuses on the incentives of the forecaster to exert effort to obtain a signal (a.k.a., a data point), but then assumes that this data point is reported directly (and cannot itself be misreported). In this space, Cai, Daskalakis, and Papadimitriou [2015] considers the learning problem where the principal aims to acquire data to train a classifier to minimize squared error less the cost of eliciting the data points from individual agents. The mechanism for soliciting the data from the agents trades off cost (in incentivizing effort) for accuracy of each individual point. Chen, Immorlica, Lucier, Syrgkanis, and Ziani [2018] and Chen and Zheng [2019] consider the estimation of the mean of a population data. Their objective is to minimize the variance of the resulting estimator subject to a budget constraint on the cost of procuring the data (from incentivizing effort).

A few papers have considered incentivizing effort under a proper scoring rule for a single-dimensional state. Osband [1989] considers incentivizing the forecaster to reduce variance under constraints that result in the optimal scoring rule being quadratic. Zermeno [2011] considers a slightly different model and derives that the optimal scoring rule has V-shaped utility; our work begins with such a result for our model. Neyman, Noarov, and Weinberg [2021] consider a forecaster with access to costly samples of a Bernoulli distribution and characterizes optimal scoring rules in the limit as the sample cost approaches zero. Our main contrasting result is the approximate optimality of the V-shaped scoring rule for binary effort and forecasts over multi-dimensional state spaces.

There are several papers on optimizing scoring rules following the model proposed in this thesis. Hartline et al. [2021a] extend the framework to the setting where the agent's effort is multi-dimensional (e.g., corresponding to independent tasks) and the agent can independently exert effort in each dimension. The main result of this extension is that the intuition that linking incentives across different dimensions is beneficial generalizes. The authors propose a generalization of the V-shaped scoring rule that is approximately optimal, which requires the agent to predict k states correctly instead of one (where k is a constant depending on the primitives). Hartline et al. [2021b] extend the framework to the setting where the agent's effort is continuous (but single-dimensional) and the cost of the agent's effort is private to the agent. In this case the principal benefits from offering several scoring rules (and agents with different costs choose different ones), each offered scoring rule is V-shaped. The model also allows for the principal to have negative utility for payments to the agent. Chen and Yu [2021] consider our objective of maximizing the incentives of binary effort in a max-min design framework. For example, they show that the quadratic scoring rule is max-min optimal over a large family of distributional settings. Kong [2021] generalizes the framework from single-agent scoring rules to multi-agent peer prediction, i.e., without ground truth. In peer prediction, the designer needs to cross reference the reports of different agents to verify the informativeness of the report.

Scoring rules are also widely studied in the literature on peer prediction where ground truth is unknown and agent reports must be compared to each other. Frongillo and Witkowski [2017] considers the optimization goal of incentive for effort in single-task peer prediction. The differences in this model result in incomparable results.

Selling Information. There is a large literature on selling information to agents with uncertainty over the states. Those papers can be classified into two categories according to the agents' private types. The first category is when the agents' private types represent their willingness to pay for different signal structures [e.g., Yang, 2020, Smolin, 2020, Liu, Shen, and Xu, 2021]. In this case, the private types of the agents are assumed to be independent from the realization of the state, and hence the private types do not affect the belief updating process for receiving the signals. The second category is when the agents' private types represent their private signals that are informative about the states [e.g., Admati and Pfleiderer, 1986, 1990, Bergemann, Bonatti, and Smolin, 2018, Bergemann, Cai, Velegkas, and Zhao, 2021]. In this case, agents have heterogeneous prior beliefs for the states, and they will update their posteriors accordingly upon receiving the signals from the seller. Note that in both lines of work, the private types of the agents are given exogenously, and hence it is a pure adverse selection model. Our model assumes that the private types of the agents represents their preferences for different signal structures, which are assumed to be independent from the realization of the states. The distinct feature is that we allow agents to endogenously acquire costly signals that are informative about the unknown states. Thus the main focus of this thesis is the interaction between adverse selection and moral hazard, and we will provide characterizations of the optimal mechanisms in this setting.

Our problem is also relevant to the literature of mechanism design with endogenous information. Crémer and Khalil [1992] considers the model of endogenous information in a contract design model. The main distinction from their model and ours is the timeline of the agent. In their paper, the agent gather information before signing the contract, while in our model, the agent can observe additional information after the interaction with the data broker. This difference in timeline also distinguishes our model from the literature on auction with endogenous entry [Menezes and Monteiro, 2000] and auction with buyer optimal learning [Shi, 2012, Mensch, 2021], where those papers assume that the agents make the information acquisition decision before interacting with the seller, and the mechanism offered by the seller distorts the agents' incentives to learn their valuation.

1.3. Organization

1.4. Bibliographic Notes

The content in this thesis is based the following research papers: "Optimal Auctions vs. Anonymous Pricing: Beyond Linear Utility" by Feng, Hartline, and Li [2019], "Simple Mechanisms for Non-linear Agents" by Feng, Hartline, and Li [2020], "Optimization of Scoring Rules" by Hartline, Li, Shan, and Wu [2020] and "Selling Data to an Agent with Endogenous Information" by Li [2021].

CHAPTER 2

Auctions for Agents with Non-linear Utilities

2.1. Preliminaries

In this chapter, we consider general payoff maximization in single-item auction for non-linear agents. For example, welfare maximization, revenue maximization and their convex combinations are special cases of payoff maximization.

Agent Models. There is a set of agents N where $|N| = n$. An agent's *utility model* is defined as (Θ, Φ, u) where Θ, Φ, and u are the type space, distribution and utility function. The outcome for an agent is the distribution over the pair (x, p), where allocation $x \in \{0, 1\}$ and payment $p \in \mathbb{R}_+$. The utility function of each player u is a mapping from her private type and the outcome to her utility for the outcome. There are several specific utility models we are interested in this thesis.

- **Linear utility:** For each agent $i \in N$, her private type is her value v_i of the good. Given allocation x and payment p, her utility is $v_i \cdot x - p$. In the following sections, we will drop the subscripts when we discuss the single agent problems.
- **Private-budget utility:** Each agent $i \in N$ has a private value v_i and private budget constraint w_i. We refer to the pair (v_i, w_i) as the private type of the agent. The valuation v_i for each agent i is sampled from the valuation distribution F_i and her budget w_i is sampled from the budget distribution G_i. We assume that F_i and G_i are independent distributions. We also use F_i and G_i to denote the

cumulative probability function for the valuation and budget of agent i. Given allocation x and payment p, her utility is $v_i x - p$ if the payment does not exceed her budget, i.e., $p \leq w_i$. Otherwise, her utility is $-\infty$.

Note that when the support of budget distribution G is a singleton $\{w\}$, it is equivalent to assume that the agent has a (deterministic) public budget w. We name the utility model of such agents as *public-budget utility*.

- **Risk-averse utility:** For each agent $i \in N$, her private type is her value $v_i \in [0, \bar{v}_i]$ of the good. Given allocation x and payment p, the utility function u is a concave function mapping from the wealth $v_i \cdot x - p$ of the agent to her utility. In the later discussion on revenue maximization, we restrict attention to a very specific form of risk aversion studied in Fu et al. [2013], which is both computationally and analytically tractable: utility functions that are linear up to a given capacity C and then flat. Given allocation x and payment p, an agent has utility $\min\{v_i \cdot x - p, C\}$. We refer to this utility function as *capacitated utility*. The capacity C is encoded in the utility function and is not necessarily identical across agents.
- **Endogenous valuation:** Each agent $i \in N$ can make costly investments before the auction by taking action $a_i \in \mathbb{R}$. For agent i with private type θ_i, the cost for action a_i is $C_i(a_i)$ and the value for the item is $v_i(a_i, \theta_i) = a_i + \theta_i$. Given allocation x and payment p, agent i taking action a_i has utility $x \cdot v_i(a_i, \theta_i) - p - C_i(a_i)$. This is the model presented in Gershkov et al. [2021a].[1] Note that in this

[1]Gershkov et al. [2021a] characterized the single-agent revenue optimal mechanism for slightly more general classes of valuation functions. To simplify the presentation, in this thesis, we only illustrate the proof for this special form of valuation function, and the same technique can be easily extended to broader settings.

endogenous utility model, the agent can be equivalently modeled as one with convex preference over allocations, which does not satisfy the expected utility characterization.

Mechanisms. In this thesis, we consider the sealed-bid mechanisms: in a mechanism $\{(x_i, p_i)\}_{i\in N}$, agents simultaneously submit sealed bids $\{b_i\}_{i\in N}$ from their type spaces to the mechanism, and each agent i gets allocation $x_i(\{b_i\}_{i\in N})$ with payment $p_i(\{b_i\}_{i\in N})$. The outcome of mechanisms is a distribution of the allocation payment pair (x_i, p_i) for each agent i where the allocation is a probability $x_i \in [0, 1]$ and the price is $p_i \in \mathbb{R}_+$. An allocation is feasible if $\sum_i x_i \leq 1$.[2]

We consider mechanisms that satisfy *Bayesian incentive compatible* (BIC), i.e., no agent can gain strictly higher expected utility than reporting her private type truthfully if all other agents are reporting their private types truthfully, and *interim individual rational* (IIR), i.e., the expected utility is non-negative for all agents and all private types if all agents are reporting their private types truthfully mechanisms. For later discussion, we also define *dominant strategy incentive compatible* (DSIC) for a mechanism if no agent can gain strictly higher expected utility than reporting her private type truthfully, regardless of other agents' report.

Payoff Curves. The payoff function of the seller is a mapping from the lotteries of each agent, to a real value. We assume that the payoff function satisfies expected utility theory,[3] i.e., the payoff for a distribution over lotteries is the corresponding expected

[2]Our results can be extended to more general feasibility constraints. We will not provide full details of this extension in this thesis. See Feng et al. [2020] for more discussions.

[3]In contrast, we do not restrict the agents to satisfy the expected utility theory.

payoff.[4] Moreover, the payoff function is additive separable across different agents. In this paragraph, we define the *payoff curves*, and introduce the *revenue curves* and *welfare curves* as special cases of the payoff curves. Specifically, the revenue contribution from agent i is her expected payment p_i, and the welfare contribution from from agent i is her expected value for realized allocation x_i.[5] In addition, we define the *optimal payoff curves* and *price-posting payoff curves* as follows.

Definition 2.1.1. *Given ex ante constraint q, the* optimal payoff curve $R(q)$ *is a mapping from quantile q to the optimal ex ante payoff for the single agent problem, i.e., the optimal payoff of the mechanism which in expectation sells the item with probability q.*

Fact 2.1.1. *The optimal payoff curve is concave.*

Fact 2.1.1 holds because the space of mechanisms is closed under convex combination. We also study mechanisms based on simple per-unit posted posting.

Definition 2.1.2. *Posting* per-unit price p *is offering a menu* $\{(x, x \cdot p) : x \in [0, 1]\}$ *to the agent. A budgeted agent with value v and budget w given* per-unit price p *will purchase the lottery $x = \min\{1, {}^{w}/_{p}\}$ if $v \geq p$, and purchase the lottery $x = 0$ otherwise.*

Definition 2.1.3. *The* market clearing price p^q *for the ex ante constraint q is the per-unit price such that the item is sold with probability q.*

[4]For example, the seller may care about the ex ante welfare of the agents, i.e., the sum of the ex ante utility of the agents when each agent is assigned with a lottery.

[5]Note that there are alternative definitions for welfare of non-linear agents. For example, when agents are risk averse, an alternative definition for welfare contribution from agent i is the sum of her payment p_i and her utility $u_i(x_i, p_i)$. Whether non-linear agents resemble linear agents under this alternative welfare definition is left as an open question.

Definition 2.1.4. *Given ex ante constraint q, the* price-posting payoff curve $P(q)$ *is a mapping from quantile q to the optimal price-posting payoff for the single agent problem, i.e., the optimal payoff of the price posting mechanism which sells the item with probability q in expectation over the type distribution and the probabilities of the selected lottery.*

Price-posting payoff curves are not generally concave, we can iron it to get the concave hull of the price-posting payoff curves.

Definition 2.1.5. *The ironed price-posting payoff curve $\bar{P}$ is the concave hull of the price-posting payoff curve P.*

Next we review the relation between the optimal revenue curves and the concave hull of the price-posting revenue curves for linear agents.

Lemma 2.1.2 (Bulow and Roberts, 1989)**.** *The optimal revenue curve R of a linear agent is equal to her ironed price-posting revenue curve $\bar{P}$.*

A similar result holds for the welfare curve. Note that the price-posting welfare curve is always concave for linear agents.

Lemma 2.1.3. *The optimal welfare curve R of a linear agent is equal to her price-posting welfare curve P, both are concave and $R = P = \bar{P}$.*

In general, for agents with budgets, the optimal payoff (e.g., revenue or welfare) curves and the concave hull of the price-posting payoff curves are not equivalent, and the ex ante optimal mechanism is more complicated and extracts strictly higher payoff than the optimal price posting mechanism and randomizations over price posting mechanisms.

Ex Ante Relaxation. Next we provide the benchmark, the ex ante relaxation. For auctions with downward-closed feasibility constraints, any sequence of ex ante quantiles $\{q_i\}_{i\in N}$ is ex ante feasible if there exists a randomized, ex post feasible allocation such that the probability agent i receives an item, i.e., marginal allocation probability for agent i, equals q_i. We denote the set of ex ante feasible quantiles by EAF. Note that $\{q_i\}_{i\in N} \in \text{EAF}$ if and only if $\sum_i q_i \leq 1$. The optimal ex ante payoff given a specific collection of payoff curves $\{R_i\}_{i\in N}$ is

$$\text{EAR}(\{R_i\}_{i\in N}) = \max_{\{q_i\}_{i\in N}\in\text{EAF}} \sum\nolimits_{i\in N} R_i(q_i).$$

Pricing-based Mechanisms and Posted Pricing Mechanisms. In Bayesian mechanism design, the taxation principle suggests that it is without loss to focus on menu mechanisms: Fixing any agent, the mechanism offers a menu of outcomes (i.e., her allocation and payment) to the agent, where the menu depends on other agents' bids. Among all such menu mechanisms, there are two subclasses of mechanisms closely related to price posting which allow simple implementations – *pricing-based mechanisms* and *posted pricing mechanisms*. The subclass of *pricing-based mechanisms* consider mechanisms where the menu (offered by the mechanism) is equivalent to posting a per-unit price. Furthermore, a pricing-based mechanism is called a *posted pricing mechanism* if the menu (a.k.a., per-unit price) offered to each agent is invariant of other agents' bids.

2.2. Reduction Framework

2.2.1. Reduction Framework for Sequential Posted Pricing

In this section, we introduce the definition of ζ-resemblance to quantify the single-agent approximation by price-posting in non-linear utility models. As a warm up, we introduce a reduction framework which extends approximation results of posted pricing mechanisms for linear agents to non-linear agents that satisfy the definition. In next section, we discuss a more general reduction framework for pricing-based mechanisms.

As we discussed in Section 2.1, the taxation principle suggests that it is without loss t focus on menu mechanisms in Bayesian mechanism design. For non-linear agents, the menu offered in the Bayesian optimal mechanism are complicated even in single-agent environments. For example, to maximize the revenue from a single agent with private budget, the menu size of the optimal mechanism is exponential to the size of the support of the budget distribution [Devanur and Weinberg, 2017a]. In contrast, for linear agents, there exist posted pricing mechanisms that is optimal (resp. approximately optimal) in the single-agent (resp. multi-agent) environments [Myerson, 1981a, Riley and Zeckhauser, 1983, Yan, 2011, Alaei et al., 2018]. Here we introduce a reduction framework that extends the approximation bounds of posted pricing mechanisms for linear agents to non-linear agents.

To simplify the presentation, we focus on the reduction framework on a canonical class of posted pricing mechanisms – *sequential posted pricing mechanisms* (see Definition 2.2.1 for a formal definition). A generalization of the framework to other posted pricing mechanisms is straightforward and we include more discussions in Appendix A.1.

Note that given the ex ante probability q, the payoff of posting the market clearing price is uniquely determined by the price-posting payoff curve and quantile q. Thus, for simplicity, we define the sequential posted pricing in quantile space.[6]

Definition 2.2.1. *A* sequential posted pricing mechanism *is parameterized by* $(\{o_i\}_{i\in N}, \{q_i\}_{i\in N})$ *where* $\{o_i\}_{i\in N}$ *denotes an order of the agents and* $\{q_i\}_{i\in N}$ *denotes the quantile corresponding to the per-unit prices to be offered to agents if the item is not sold to previous agents.*[7]

According to the definition, the payoff of the sequential posted pricing mechanism with parameters $(\{o_i\}_{i\in N}, \{q_i\}_{i\in N})$ is uniquely determined by the price-posting payoff curves $\{P_i\}_{i\in N}$ of the agents. Specifically,

$$\mathrm{SPP}(\{P_i\}_{i\in N}, \{o_i\}_{i\in N}, \{q_i\}_{i\in N}) = \sum_{i\in N}\left(\prod_{j:o(j)<o(i)}(1-q_j)\right)P_i(q_i)\ .$$

and the optimal payoff among the class of sequential posted pricing mechanisms is

$$\mathrm{SPP}(\{P_i\}_{i\in N}) = \max_{\{o_i\}_{i\in N}, \{q_i\}_{i\in N}} \mathrm{SPP}(\{P_i\}_{i\in N}, \{o_i\}_{i\in N}, \{q_i\}_{i\in N}).$$

As we mentioned above, Yan [2011] shows the following approximation guarantee for sequential posted pricing.

[6]The reason for defining posted pricings in quantile space is that the mapping from quantiles to prices is not generally pinned down by the payoff curve (specifically, for the welfare objective) for non-linear agents. As the actual prices to be posted are not important in our reduction framework, it is convenient to remain in quantile space. Any sequential posted pricing mechanism defined in quantile space can be converted to a sequential posted pricing mechanism in price space [e.g., Chawla et al., 2010]. Thus, in this thesis, without loss of generality, we will focus on the sequential posted pricing mechanisms in quantile space.

[7]In the sequential posted pricing mechanism, each agent may only get a lottery for winning the item. We assume that the lottery is realized immediately after each agent's purchase decision. The per-unit prices are offered to each agent if and only if the item is not sold to previous agents given the realization.

Theorem 2.2.1 (Yan, 2011)**.** *For linear agents with the price-posting payoff curves* $\{P_i\}_{i \in N}$*, there exists a sequential posted pricing mechanism* $(\{o_i\}_{i \in N}, \{q_i\}_{i \in N})$ *that is an* $^e/_{(e-1)}$*-approximation to the ex ante relaxation, i.e.,* $\mathrm{SPP}(\{P_i\}_{i \in N}, \{o_i\}_{i \in N}, \{q_i\}_{i \in N}) \geq (1 - {}^1/_e) \cdot \mathrm{EAR}(\{\bar{P}_i\}_{i \in N})$*.*

To quantify the extent to which a non-linear agent resemble a linear agent, we start with the following observation. For a linear agent, the ironed price-posting payoff curve equals the optimal payoff curve. However, for a non-linear agent, the Bayesian optimal mechanisms are not posted pricing mechanisms in general. In other words, for a non-linear agent, the ironed price-posting payoff curve is not generally equivalent to the optimal payoff curve. Hence, we introduce ζ*-resemblance* of an agent to measure her ironed price-posting payoff curve resemble her optimal payoff curve.

Definition 2.2.2 (ζ-resemblance)**.** *An agent's ironed price-posting payoff curve* $\bar{P}$ *is* ζ-resemblant *to her optimal payoff curve* R*, if for all* $q \in [0, 1]$*, there exists* $q \leq q^\dagger$ *such that* $\bar{P}(q) \geq {}^1/_\zeta \cdot R(q^\dagger)$*. Such an agent is* ζ-resemblant*.*

Smaller ζ-resemblance guarantee implies that such non-linear agents resemble linear agents better, since the approximation guarantee for sequential posted pricing mechanisms for linear agents can be lifted to those non-linear agents with an additional factor ζ (Theorem 2.2.2). Note that the ζ-resemblant property is equivalent to show the approximation of posted pricing mechanisms for a continuum of i.i.d. (non-linear) agents with unit-demand and limit supply. In Sections 2.3 to 2.5, we give small constant bound on this resemblant property under several canonical non-linear utility models for both welfare maximization and revenue maximization.

To extend the approximation of sequential posted pricing mechanisms for linear agents to non-linear agents, we need to reduce a non-linear agent to her linear agent analog as follows.

Definition 2.2.3. *Fix any set of (non-linear) agents with price-posting payoff curves* $\{P_i\}_{i\in N}$. *The* linear agents analog *is a set of linear agents whose price-posting payoff curves are* $\{P_i\}_{i\in N}$ *and the optimal payoff curves are* $\{\bar{P}_i\}_{i\in N}$.

Note that the linear agent analog is well-defined for both welfare maximization and revenue maximization.[8] Based on the definition of ζ-resemblance and the linear agent analog, we present a reduction framework that converts sequential posted pricing mechanisms for linear agents to non-linear agents, and approximately preserves its payoff approximation guarantee.

Theorem 2.2.2. *Fix any set of (non-linear) agents with price-posting payoff curves* $\{P_i\}_{i\in N}$ *that are* ζ*-resemblant to their optimal payoff curves* $\{R_i\}_{i\in N}$. *If there exists a sequential posted pricing mechanism* $(\{o_i\}_{i\in N}, \{q_i\}_{i\in N})$ *that is a* γ*-approximation to the ex ante relaxation for linear agents analog with price-posting payoff curves* $\{P_i\}_{i\in N}$, *i.e.,* $\mathrm{SPP}(\{P_i\}_{i\in N}, \{o_i\}_{i\in N}, \{q_i\}_{i\in N}) \geq {}^1/_\gamma \cdot \mathrm{EAR}(\{\bar{P}_i\}_{i\in N})$, *then this mechanism is also a* $\gamma\zeta$*-approximation to the ex ante relaxation for non-linear agents, i.e.,* $\mathrm{SPP}(\{P_i\}_{i\in N}, \{o_i\}_{i\in N}, \{q_i\}_{i\in N}) \geq {}^1/_{\gamma\zeta} \cdot \mathrm{EAR}(\{R_i\}_{i\in N})$.

[8] The price-posting revenue (resp. welfare) curve $P(q)$ of a linear agent uniquely pins down her valuation distribution as $v(q) = \frac{P(q)}{q}$ (resp. $v(q) = P'(q)$). For general payoff function, given the price-posting payoff curves $\{P_i\}_{i\in N}$ of the non-linear agents, there may not exist distributions for linear agents such that their price-posting payoff curves coincide with $\{P_i\}_{i\in N}$. However, both the payoffs for sequential posted pricing mechanisms and the ex ante relaxation are well defined given the payoff curves, and Theorem 2.2.1 holds for payoff curves that does not correspond to any distributions of the agents. Hence, we can refer to the linear agents analog even without the existence of the underlying distributions.

Proof. Let $\{q_i^\dagger\}_{i\in N}$ be the profile of optimal ex ante quantiles for optimal payoff curves $\{R_i\}_{i\in N}$. Since the ironed price-posting payoff curves $\{\bar{P}_i\}_{i\in N}$ are ζ-resemblant to the optimal payoff curves $\{R_i\}_{i\in N}$, there exists a sequence of quantiles $\{q_i^\ddagger\}_{i\in N}$ such that for any agent i, $q_i^\ddagger \le q_i^\dagger$ and $\bar{P}(q_i^\ddagger) \ge {}^1\!/\!_\zeta \cdot R(q_i^\dagger)$. Note that since $\sum_i q_i^\ddagger \le \sum_i q_i^\dagger \le 1$, $\{q_i^\ddagger\}_{i\in N}$ is also feasible for ex ante relaxation. Therefore,

$$\mathrm{EAR}(\{R_i\}_{i\in N}) = \sum_{i\in N} R_i(q_i^\dagger) \le \zeta \cdot \sum_{i\in N} \bar{P}_i(q_i^\ddagger) \le \zeta \cdot \mathrm{EAR}(\{\bar{P}_i\}_{i\in N}).$$

Since the expected payoff of the sequential posted pricing mechanism $(\{o_i\}_{i\in N}, \{q_i\}_{i\in N})$ only depends on the price posting payoff curves, not on the agents' utility models, we have

$$\mathrm{SPP}(\{P_i\}_{i\in N}, \{o_i\}_{i\in N}, \{q_i\}_{i\in N}) \ge {}^1\!/\!_\gamma \cdot \mathrm{EAR}(\{\bar{P}_i\}_{i\in N}) \ge {}^1\!/\!_{\gamma\zeta} \cdot \mathrm{EAR}(\{R_i\}_{i\in N}),$$

and Theorem 2.2.2 holds. □

The reduction framework (Theorem 2.2.2) seems to be an immediate consequence from the definition of sequential posted pricing and definition of ζ-resemblance. In the later sections, We will discuss its extensions to other (probably more general) classes of mechanisms by adopting the same method. Specifically, in Appendix A.1, we show that how a similar reduction framework hold for other formats of posted pricing mechanisms – oblivious posted pricing where mechanisms cannot control the order of agents, and anonymous pricing where mechanisms need to post an identical price to all agents. In Section 2.2.2, we show that when the agents satisfy the expected utility representation, any deterministic, dominant strategy incentive compatible mechanism can be converted to approximately preserve the approximation ratio for non-linear agents.

As an application of the reduction framework in Theorem 2.2.2, consider (non-linear) agents with private budget utility. Optimal mechanism for agents with private budget utility have been studied in the literature (e.g. Che and Gale, 2000, Devanur and Weinberg, 2017a for single-agent, Pai and Vohra, 2014 for i.i.d. agents and Alaei et al., 2012 for non-i.i.d. agents). The characterization of these optimal mechanisms are complicated even for simple distributions (e.g., value and budget drawn i.i.d. from $[0, 1]$ uniformly). However, with the reduction framework (Theorem 2.2.2 for posted pricing mechanism and Theorem 2.2.3 for pricing-based mechanism), due to the resemblance between price-posting payoff curve and optimal payoff curve, we can extend the simple mechanism (i.e., sequential/oblivious posted pricing mechanism and marginal payoff mechanism) from linear agents to private-budgeted agents with better approximation guarantees.

2.2.2. Reduction Framework for Pricing-based Mechanisms

Following the discussion in Section 2.2.1, in this section we introduce the reduction framework for pricing-based mechanisms. For this reduction framework, we focus on agents satisfying the von Neumann-Morgenstern expected utility representation.

Recall that by taxation principle, it is without loss to consider menu mechanisms. The class of pricing-based mechanisms is ones whose menu offered to each agent is posting a per-unit price. For linear agents, every mechanism (e.g., the Bayesian optimal mechanism) can be implemented as a pricing-based mechanism. Here, our reduction framework extends the approximation bounds of deterministic, dominant strategy incentive compatible (DSIC), interim individual rational (IIR), pricing-based mechanisms for linear agents to non-linear agents whose utility models satisfy the expected utility representation.

Due to the technical reason, we make the following assumption on agents' utility models. Note that this assumption is satisfied for most common utility models, e.g., linear utility, budget utility, risk averse utility.

Assumption 1. *The item is* the ordinary good*, i.e., when offered a per-unit price for the item to the agent, her demand is weakly decreasing in price.*

Based on the definition of ζ-resemblance and linear agent analog, we present the meta-theorem (Theorem 2.2.3): a reduction framework that converts every deterministic, DSIC, IIR, pricing-based mechanism for linear agents to a DSIC, IIR, pricing-based mechanism for non-linear agents, and approximately preserves its payoff approximation guarantee.

Theorem 2.2.3 (Reduction Framework)**.** *Fix any set $\mathcal{A}$ of (non-linear) agents with price-posting payoff curves $\{P_i\}_{i\in N}$ and optimal payoff curves $\{R_i\}_{i\in N}$. For any deterministic, DSIC, IIR, pricing-based mechanism $\mathcal{M}_L$ for linear agents, there is a pricing-based mechanism $\mathcal{M}$ for non-linear agents $\mathcal{A}$ that is DSIC, IIR, and satisfies*

i. Identical payoff: mechanism $\mathcal{M}$ for non-linear agents $\mathcal{A}$ has the same payoff as mechanism $\mathcal{M}_L$ for the linear agents analog $\mathcal{A}_L$. Denote the payoff of mechanism $\mathcal{M}$ as $\mathcal{M}(\{P_i\}_{i\in N})$.

ii. Identical feasibility: mechanism $\mathcal{M}$ for non-linear agents $\mathcal{A}$ has the same distribution over outcomes as mechanism $\mathcal{M}_L$ for the linear agents analog $\mathcal{A}_L$.

Denote by γ the approximation of mechanism $\mathcal{M}_L$ for the linear agents analog $\mathcal{A}_L$ to the ex ante relaxation of $\mathcal{A}_L$, i.e., $\mathcal{M}_L(\{P_i\}_{i\in N}) \geq {}^{1}\!/\!{}_{\gamma} \cdot \mathrm{EAR}(\{\bar{P}_i\}_{i\in N})$. If each non-linear agent in $\mathcal{A}$ is ζ-resemblant, then mechanism $\mathcal{M}$ for non-linear agents $\mathcal{A}$ is $\gamma\,\zeta$-approximation to the ex ante relaxation of $\mathcal{A}$, i.e., $\mathcal{M}(\{P_i\}_{i\in N}) \geq {}^{1}\!/\!{}_{\gamma\,\zeta} \cdot \mathrm{EAR}(\{R_i\}_{i\in N})$.

In Section 2.2.2.1, we present the implementation of the reduction framework. In Section 2.2.2.2, we show how it achieves the claimed properties in Theorem 2.2.3. Finally, in Section 2.2.2.3, we discuss the consequence of the reduction framework for the marginal payoff mechanism (i.e., the Bayesian optimal mechanism) for linear agents.

2.2.2.1. Implementation in Theorem 2.2.3. Algorithm 1 describes the implementation of Theorem 2.2.3.[9] This implementation includes two notations $\hat{q}_i^{\mathcal{M}_L}\left(\{q_j\}_{j\in N\setminus\{i\}}\right)$ and $x^{\hat{q}}(\theta)$ which we define below.

For any deterministic DSIC, IIR mechanism $\mathcal{M}_L$ for linear agents, it can be represented by a mapping from the quantiles of other agents to a threshold quantile for each agent. The agent wins when her quantile is below the threshold and loses when her quantile is above the threshold. We denote the function that maps the profile of other agent quantiles $\{q_j\}_{j\in N\setminus\{i\}}$ to a quantile threshold for agent i as $\hat{q}_i^{\mathcal{M}_L}\left(\{q_j\}_{j\in N\setminus\{i\}}\right)$.

For any non-linear agent model (Θ, F, u), the single-agent pricing problem identifies the per-unit (market clearing) price $p^{\hat{q}}$ to offer the agent for any ex ante allocation constraint $\hat{q}$. Denote the allocation probability selected by an agent with type θ when offered per-unit price $p^{\hat{q}}$ as $\hat{x}^{\hat{q}}(\theta)$. For every type θ, define function $H_\theta(q) = \hat{x}^{q}(\theta)$. Note that under the ordinary good assumption (Assumption 1) $H_\theta(q)$ is weakly increasing in q for all type

[9]The construction is a simplification of a construction in Alaei et al. [2013].

θ under (Assumption 1), and thus can be viewed as the cumulative density function of a distribution. See Lemma 2.2.4.

Algorithm 1: Reduction Framework for Pricing-based Mechanism

Input: Non-linear agents $\{(\Theta_i, F_i, u_i)\}_{i \in N}$; and deterministic, DSIC, IIR mechanism $\mathcal{M}_L$ for linear agents

1 For each agent i with private type θ_i, map the type to a random quantile q_i according to the distribution H_{i,θ_i} with cdf $H_{i,\theta_i}(q) = \hat{x}_i{}^q(t_i)$.
`/*` $H_i(q)$ `is well-defined. See Lemma 2.2.4 */`

2 For each agent i, calculate quantile threshold as $\hat{q}_i = \hat{q}_i^{\mathcal{M}_L}\left(\{q_j\}_{j \in N \setminus \{i\}}\right)$.
`/*` $\hat{q}_i^{\mathcal{M}_L}(\cdot)$ `is well-defined since` $\mathcal{M}_L$ `is deterministic and DSIC. */`

3 For each agent i, set payment $p_i = p^{\hat{q}_i}\, x_i^{\hat{q}_i}(\theta_i)$, and allocation $x_i = 1$ if $q_i < \hat{q}_i$ and $x_i = 0$ otherwise.

Lemma 2.2.4. *For an ordinary good (Assumption 1), the allocation probability $x^q(\theta)$ is weakly increasing in q for all type θ.*

Proof. For an ordinary good by definition, the agent's expected allocation probability is weakly decreasing in the price. Thus, the per-unit price in each q ex ante mechanism (with respect to the price-posting payoff curve P) is weakly decreasing in q. Now consider the q ex ante mechanism with respect to the ironed price-posting payoff curve $\bar{P}$ for all quantile q. The per-unit price is monotone (by the previous argument) on quantiles that are not in ironed intervals. Within an ironed interval, the mechanism is a mix over two end-points of non-ironed intervals which linearly interpolates between the end-points and is thus monotone. □

2.2.2.2. Proof of Theorem 2.2.3. We first show the implementation (Algorithm 1) is DSIC, IIR and satisfies both identical payoff and identical feasibility properties.

Lemma 2.2.5. *Given a deterministic, DSIC, IIR mechanism $\mathcal{M}_L$ for linear agents, the mechanism $\mathcal{M}$ from the implementation (Algorithm 1) is DSIC, IIR, and satisfies identical payoff and identical feasibility properties in Theorem 2.2.3.*

Proof. Since mechanism $\mathcal{M}_L$ is deterministic and DSIC, Algorithm 1 is well-defined. Since for each agent i, her type θ_i is drawn from F_i and q_i is drawn from H_i condition on θ_i, the (unconditional) distribution of q_i is uniform on $[0, 1]$. Thus, from each agent i's perspective, the other agents' quantiles are distributed independently and uniformly on $[0, 1]$. This agent faces a distribution over ex ante posted pricing that is identical to the distribution of quantile thresholds in the mechanism $\mathcal{M}_L$. Thus, DSIC and the identical payoff property is satisfied. Since $\mathcal{M}_L$ is IIR, $\mathcal{M}$ is also IIR. Finally, note that the distribution of q_i is uniform on $[0, 1]$, identical feasibility property is satisfied by construction. □

We now show that the implementation extends the approximation guarantee of mechanism $\mathcal{M}_L$ for linear agents. Note that this is immediately implied by the identical payoff property and the following lemma.

Lemma 2.2.6. *For agents with ironed price-posting payoff curves $\{\bar{P}_i\}_{i\in N}$ and the optimal payoff curves $\{R_i\}_{i\in N}$, if each agent is ζ-resemblant, the ex ante relaxation on the ironed price-posting payoff curve is a ζ-approximation to the ex ante relaxation on the optimal payoff curves, i.e.,* $\mathrm{EAR}(\{\bar{P}_i\}_{i\in N}, \mathcal{X}) \geq {}^{1}/_{\zeta} \cdot \mathrm{EAR}(\{R_i\}_{i\in N})$.

Proof. Let $\{q_i^\dagger\}_{i\in N} \in \mathrm{EAF}(\mathcal{X})$ be the profile of optimal ex ante quantiles for optimal payoff curves $\{R_i\}_{i\in N}$. Since the ironed price-posting payoff curves $\{\bar{P}_i\}_{i\in N}$ are ζ-resemblant to the optimal payoff curves $\{R_i\}_{i\in N}$, there exists a sequence of quantiles

$\{q_i\}_{i\in N}$ such that for any agent i, $q_i \leq q_i^\dagger$ and $\bar{P}(q_i) \geq {}^1\!/\!\zeta \cdot R(q_i^\dagger)$. Note that $\{q_i\}_{i\in N}$ is also feasible. Therefore,

$$\text{EAR}(\{R_i\}_{i\in N}) = \sum_{i\in N} R_i(q_i^\dagger) \leq \zeta \cdot \sum_{i\in N} \bar{P}_i(q_i) \leq \zeta \cdot \text{EAR}(\{\bar{P}_i\}_{i\in N}). \qquad \square$$

2.2.2.3. Application on Marginal Payoff Mechanism. In Bulow and Roberts [1989], authors introduce the marginal revenue mechanism and show its revenue-optimality for linear agents. The marginal revenue mechanism can be easily extended to other payoff objectives and we denote its extensions as the marginal payoff mechanisms. The ex ante relaxation gives an upper bound on the Bayesian optimal mechanism. For linear agents, the gap between the ex ante relaxation and the Bayesian optimal mechanisms (i.e., marginal payoff mechanisms) is precisely determined by the optimal payoff curves.

Definition 2.2.4. *The* ex ante gap *for the optimal payoff curves* $\{R_i\}_{i\in N}$ *is the ratio between the ex ante relaxation* $\text{EAR}(\{R_i\}_{i\in N})$ *and the payoff of the Bayesian optimal mechanism for linear agents* $\text{OPT}(\{R_i\}_{i\in N})$.

In single-item environments, the ex ante gap γ is at most ${}^1\!/\!(1 - {}^1\!/\!\sqrt{2\pi})$ [Yan, 2011]. By our framework Theorem 2.2.3 on the marginal payoff mechanisms, we obtain the marginal payoff mechanism for non-linear agents, and its approximation guarantee.

Definition 2.2.5. *The* marginal payoff mechanism*, denoted by* MPM *(defined in Algorithm 1) corresponds to the linear agent marginal revenue mechanism. Denote the payoff of* MPM *for agents with price-posting payoff curves* $\{P_i\}_{i\in N}$ *as* $\text{MPM}(\{P_i\}_{i\in N})$.

Proposition 2.2.7. *Given agents with the ironed price-posting payoff curves* $\{\bar{P}_i\}_{i\in N}$ *and the optimal payoff curves* $\{R_i\}_{i\in N}$*, if each agent is* ζ*-resemblant, the worst case ratio* *between the the marginal payoff mechanism with respect to price-posting payoff curves and the ex ante relaxation on the optimal payoff curves is* $\zeta\gamma$, *i.e.,* $\mathrm{MPM}(\{P_i\}_{i\in N}) \geq {}^1/_{\zeta\gamma} \cdot \mathrm{EAR}(\{R_i\}_{i\in N})$, *where* γ *is the ex ante gap with curves* $\{\bar{P}_i\}_{i\in N}$.

2.3. Budgeted Agent

2.3.1. Welfare Maximization

For agents with budget constraints, the ex ante optimal mechanism might be complicated and hard to characterize. However, as we show below, without any assumption on the valuation distribution or the budget distribution except the independence, posting the market clearing price guarantees a 2-approximation in welfare.

Theorem 2.3.1. *An agent with private budget has the price-posting welfare curve* P *that is* 2*-resemblant to her optimal welfare curve* R *if the budget is drawn independently from the valuation.*

The proof of Theorem 2.3.1 generalizes the price decomposition technique from Abrams [2006] and extends it for welfare analysis.

Fix an arbitrary ex ante constraint q, denote EX as the q ex ante welfare-optimal mechanism, and **Payoff**[EX] as its welfare. We want to decompose EX into two mechanisms $\mathrm{EX}^{\dagger}$ and $\mathrm{EX}^{\ddagger}$ according to the market clearing price p^q and bound the welfare from those two mechanisms separately. The decomposed mechanism may violate the incentive constraint for budgets, and we refer to this setting as the random-public-budget utility model. Note that the market clearing price is the same in both the private budget model and the random-public-budget utility model. Intuitively, mechanism $\mathrm{EX}^{\dagger}$ contains per-unit prices at most the market clearing price, while mechanism $\mathrm{EX}^{\ddagger}$ contains per-unit prices at least the market clearing price. Both mechanisms $\mathrm{EX}^{\dagger}$ and $\mathrm{EX}^{\ddagger}$ satisfy the ex ante constraint q, and the sum of their welfare upper bounds the original ex ante mechanism EX, i.e., $\mathbf{Payoff}[\mathrm{EX}] \leq \mathbf{Payoff}\big[\mathrm{EX}^{\dagger}\big] + \mathbf{Payoff}\big[\mathrm{EX}^{\ddagger}\big]$.

To construct $\mathrm{EX}^{\dagger}$ and $\mathrm{EX}^{\ddagger}$ that satisfy the properties above, we first introduce a characterization of all incentive compatible mechanisms for a single agent with private-budget utility, and her behavior in the mechanisms.

Definition 2.3.1. *An* allocation-payment function $\tau : [0, 1] \to \mathbb{R}_+$ *is a mapping from the allocation x to the payment p.*

Lemma 2.3.2. *For a single agent with private-budget utility, in any incentive compatible mechanism, for all types with any fixed budget, the mechanism provides a convex and non-decreasing allocation-payment function, and subject to this allocation-payment function, each type will purchase as much as she wants until the budget constraint binds, or the unit-demand constraint binds, or the value binds (i.e., her marginal utility becomes zero).*

Proof. Myerson [1981a] show that any mechanisms (x, p) for a single linear agent is incentive compatible (the agent does not prefer to misreport her value) if and only if a) $x(v)$ is non-decreasing; b) $p(v) = vx(v) - \int_0^v x(t)dt$. Thus, given any non-decreasing allocation x, the payment p is uniquely pined down by the incentive constraints.

Comparing with the linear utility, the incentive compatibility in the private-budget utility guarantees that the agent does not prefer to misreport either her value or budget. If we relax the incentive constraints such that she is only allowed to misreport her value, Myerson result already shows that for any fixed budget level w, the allocation $x(v, w)$ is non-decreasing in v and the payment $p(v,w) = vx(v,w) - \int_0^v x(t,w)dt$ is uniquely pined down. We define the allocation-payment function $\tau_w(\hat{x}) = \max\{p(v,w) + v \cdot (\hat{x} - x(v,w)) : x(v,w) \leq \hat{x}\}$ if $\hat{x} \leq x(\bar{v},w)$; and ∞ otherwise. Given the characterization of allocation and payment above, this allocation-payment function is well-defined, non-decreasing and convex. □

Remark 2.3.2. *Unlike Myerson's result which give a sufficient and necessary condition for incentive compatible mechanisms for linear agents, Lemma 2.3.2 only characterizes a necessary condition for private-budget utility.*[10] *This condition is already enough for our arguments in Section 2.3.1.*

Now we give the construction of $\text{EX}^\dagger$ and $\text{EX}^\ddagger$ by constructing their allocation-payment functions. The decomposition is illustrated in Figure 2.1. For agent with budget w, let τ_w be the allocation-payment function in mechanism EX, and x^*_w be the utility maximization allocation for a linear agent with value equal to the market clearing price p^q, i.e., $x^*_w = \text{argmax}\{x : \tau'_w(x) \leq p^q\}$. For agents with budget w, we define the allocation-payment functions $\tau^\dagger_w$ and $\tau^\ddagger_w$ for $\text{EX}^\dagger$ and $\text{EX}^\ddagger$ respectively below,

$$\tau^\dagger_w(x) = \begin{cases} \tau_w(x) & \text{if } x \leq x^*_w, \\ \infty & \text{otherwise;} \end{cases} \qquad \tau^\ddagger_w(x) = \begin{cases} \tau_w(x^*_w + x) - \tau_w(x^*_w) & \text{if } x \leq 1 - x^*_w, \\ \infty & \text{otherwise.} \end{cases}$$

By construction, for each type of the agent, the allocation from EX is upper bounded by the sum of the allocation from $\text{EX}^\dagger$ and $\text{EX}^\ddagger$, which implies that the welfare from EX is

[10] This characterization is only necessary because it relaxes the incentive constraints for misreporting the private budget.

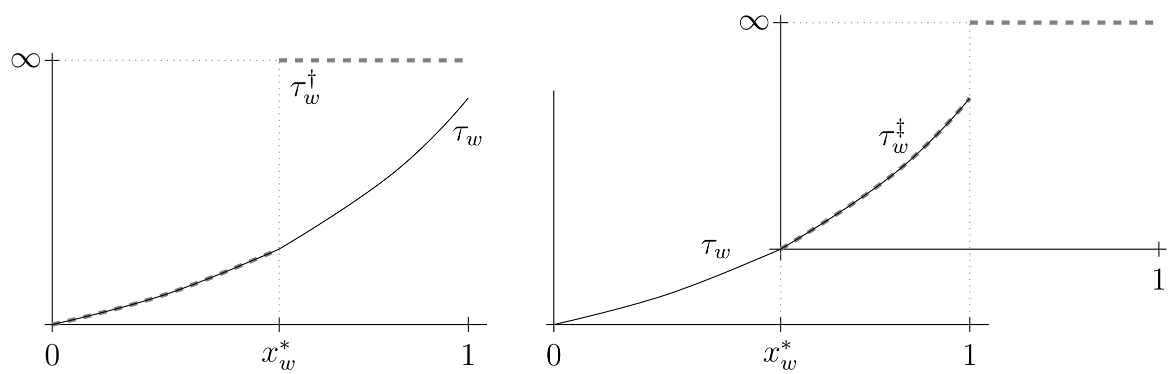

Figure 2.1. Depicted are allocation-payment function decomposition. The black lines in both figures are the allocation-payment function τ_w in ex ante optimal mechanism EX; the gray dashed lines are the allocation-payment function $\tau_w^\dagger$ and $\tau_w^\ddagger$ in $\text{EX}^\dagger$ and $\text{EX}^\ddagger$, respectively.

upper bounded by the sum of the welfare from $\text{EX}^\dagger$ and $\text{EX}^\ddagger$, and the requirements for the decomposition are satisfied.

As sketched above, we separately bound the welfare in $\text{EX}^\dagger$ and $\text{EX}^\ddagger$ by the welfare from posting the market clearing price.

Lemma 2.3.3. *For a single agent with random-public-budget utility, independently distributed value and budget, and any ex ante constraint q, the welfare from posting the market clearing price p^q is at least the welfare from* $\text{EX}^\dagger$*, i.e.,* $P(q) \geq \textbf{Payoff}\big[\text{EX}^\dagger\big]$.

Proof. Consider agent with type (v, w) and agent with type (v', w), where both value v and v' are higher than the market clearing price p^q. Notice that the allocations for these two types are the same in $\text{EX}^\dagger$ and in market clearing, since the per-unit price in both mechanisms is at most p^q which makes the mechanisms unable to distinguish these two types.

Let $x^\dagger$ be the allocation rule in $\mathrm{EX}^\dagger$ and let x^q be the allocation rule in posting the market clearing price p^q. For any value $v \geq p^q$, the expected allocation for types with value v is lower in $\mathrm{EX}^\dagger$ than in market clearing, i.e., $\mathbf{E}_w\left[x^\dagger(v,w)\right] \leq \mathbf{E}_w[x^q(v,w)]$. Otherwise suppose the types with value v^* has strictly higher allocation in $\mathrm{EX}^\dagger$ for some value $v^* \geq p^q$, i.e, $\mathbf{E}_w\left[x^\dagger(v^*,w)\right] > \mathbf{E}_w[x^q(v^*,w)]$. By the fact stated in previous paragraph, we have that for any budget w and any value $v, v^* \geq p^q$, $x^q(v,w) = x^q(v^*,w)$, $x^\dagger(v,w) = x^\dagger(v^*,w)$, and the expected allocation in $\mathrm{EX}^\dagger$ is

$$\mathbf{E}_{v,w}\left[x^\dagger(v,w)\right] \geq \Pr[v \geq p^q] \cdot \mathbf{E}_{v,w}\left[x^\dagger(v,w)|v \geq p^q\right] = \Pr[v \geq p^q] \cdot \mathbf{E}_w\left[x^\dagger(v^*,w)\right]$$

$$> \Pr[v \geq p^q] \cdot \mathbf{E}_w[x^q(v^*,w)] = \Pr[v \geq p^q] \cdot \mathbf{E}_{v,w}[x^q(v,w)|v \geq p^q] = q,$$

where the qualities hold due to the independence between the value and the budget. Note that this implies that $\mathrm{EX}^\dagger$ violates the ex ante constraint q, a contradiction. Further, for any type with value $v \geq p^q$, $\mathbf{E}_w\left[x^\dagger(v,w)\right] \leq \mathbf{E}_w[x^q(v,w)]$ implies that the allocation in market clearing "first order stochastic dominantes" the allocation in $\mathrm{EX}^\dagger$, i.e., for any threshold $v^\dagger$, the expected allocation from all types with value $v \geq v^\dagger$ in market clearing is at least the expected allocation from those types in $\mathrm{EX}^\dagger$. Taking expectation over the valuation and the budget, the expected welfare from market clearing is at least the welfare from $\mathrm{EX}^\dagger$, i.e., $P(q) \geq \mathbf{Payoff}\left[\mathrm{EX}^\dagger\right]$. □

Lemma 2.3.4. *For a single agent with random-public-budget utility, independently distributed value and budget, and any ex ante constraint q; the welfare from market clearing is at least the welfare from* $\mathrm{EX}^\ddagger$*, i.e.,* $P(q) \geq \mathbf{Payoff}\left[\mathrm{EX}^\ddagger\right]$.

Proof. In both $\mathrm{EX}^{\ddagger}$ and market clearing, types with value lower than p^q will purchase nothing, so we only consider the types with value at least p^q in this proof. Consider any type (v, w) where $v \geq p^q$, its allocation in market clearing is at least its allocation in $\mathrm{EX}^{\ddagger}$, because the per-unit price in $\mathrm{EX}^{\ddagger}$ is higher. Thus, the welfare from market clearing is at least the welfare from $\mathrm{EX}^{\ddagger}$, i.e., $P(q) \geq \mathbf{Payoff}\big[\mathrm{EX}^{\ddagger}\big]$. □

Proof of Theorem 2.3.1. Combining Lemma 2.3.3 and 2.3.4, for any quantile q, we have

$$R(q) = \mathbf{Payoff}[\mathrm{EX}] \leq \mathbf{Payoff}\big[\mathrm{EX}^{\dagger}\big] + \mathbf{Payoff}\big[\mathrm{EX}^{\ddagger}\big] \leq 2P(q) \leq \max_{q' \leq q} 2\bar{P}(q'). \qquad \square$$

2.3.2. Revenue Maximization

In this section we analyze the resemblance of revenue curves for an agent with budget. We show that approximate resemblance is satisfied under weaker assumptions on the valuation distribution or the budget distribution. For simplicity, in this section, we use the notation $\mathbf{Payoff}_w[\cdot]$ to denote the revenue given any mechanism if the budget of the agent is w, and $\mathbf{Payoff}[\cdot]$ to denote the revenue by taking expectation over the budget w.

2.3.2.1. Public Budget. In this section, we consider the simpler setting where agents have public budgets, i.e., the budget distribution is a point mass. For an agent with a public budget, we show that the ironed price-posting revenue curve is 1-resemblant to her optimal revenue curve if her valuation distribution is regular (Theorem 2.3.5) and for an agent with general valuation distribution, the ironed price-posting revenue curve is 2-resemblant to her optimal revenue curve (Theorem 2.3.7).

Theorem 2.3.5. *An agent with public budget and regular valuation distribution has the ironed price-posting revenue curve $\bar{P}$ that equals (i.e. 1-resemblant) her optimal revenue curve R.*

To prove Theorem 2.3.5, it is sufficient to show for any quantile $\hat{q} \in [0, 1]$, the $\hat{q}$ ex ante optimal mechanism is a price-posting mechanism, i.e., $R(\hat{q}) = P(\hat{q})$. To show this, we write the ex ante optimal mechanism as an optimization program, and apply Lagrangian relaxation on the budget constraint. This leads to a new optimization program similar to an agent with linear utility but with a Lagrangian objective function. Following the technique that price-posting revenue curve indicates the ex ante optimal mechanism for a linear agent, we consider the *Lagrangian price-posting revenue curve* which characterizes the ex ante optimal mechanism for the Lagrangian objective function. See further discussion about this technique in Alaei et al. [2013] and Feng and Hartline [2018]. The detailed proof of Theorem 2.3.5 is deferred to Appendix A.2.

For an agent with a general valuation distribution, resemblance follows from a characterization of the ex ante optimal mechanism from Alaei et al. [2013].

Lemma 2.3.6 (Alaei et al., 2013)**.** *For a single agent with public budget, the $q \in [0, 1]$ ex ante optimal mechanism has a menu with size at most two.*

Theorem 2.3.7. *An agent with public budget has the ironed price-posting revenue curve $\bar{P}$ that is 2-resemblant to her optimal revenue curve R.*

Proof. By Lemma 2.3.6, the allocation rule x_q of the ex ante revenue maximization mechanism for the single agent with public budget has a menu of size at most two. We decompose its allocation into x_L and x_H as illustrated in Figure 2.2. Note that

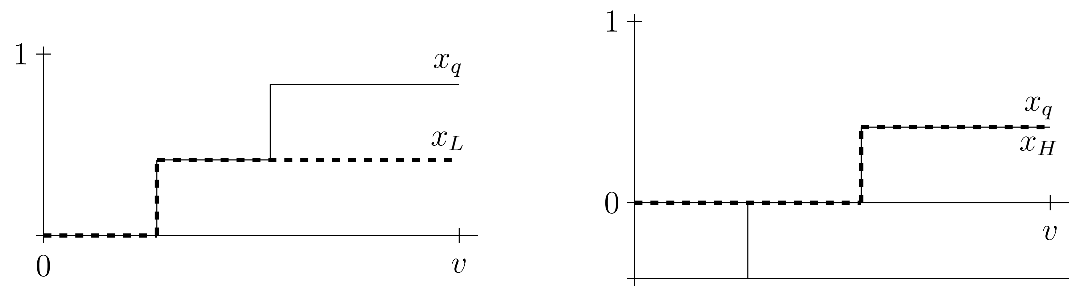

Figure 2.2. The thin solid line is the allocation rule for the optimal ex ante mechanism. The thick dashed line on the left side is the allocation of the decomposed mechanism with lower price, while the thick dashed line on the right side is the allocation of the decomposed mechanism with higher price.

both allocation x_L and x_H are (randomized) price-posting allocation rules, and neither allocation violates the allocation constraint q. Thus,

$$R(q) = \mathbf{Payoff}[x_q] = \mathbf{Payoff}[x_L] + \mathbf{Payoff}[x_H] \leq 2 \max_{q^\dagger \leq q} \bar{P}(q^\dagger). \qquad \square$$

2.3.2.2. Private Budget. In this section, we study the resemblance of the ironed price-posting revenue curve and the optimal revenue curve for agents with private budget. For linear agents, those two curves are equivalent for any valuation distribution. However, for an agent with private budget, the gap between them can be unbounded. Specifically, when the budget distribution is correlated with the valuation distribution, posting prices is not a constant approximation to the optimal revenue for a single agent even with strong regularity assumption on the marginal valuation distribution and budget distribution.

Example 2.3.3 (necessity of the independence between the value and budget distributions)**.** *Fix a large constant h. Consider a single agent with value v drawn from $[1, h]$ with density function $\frac{h}{h-1}\frac{1}{v^2}$, and budget $w = 2h - v$, i.e., her value and budget are fully correlated. A mechanism which charges the agent $v - 2\epsilon$ with probability $1 - \frac{\epsilon}{h}$, or w*

with probability $\frac{\epsilon}{h}$ *for sufficient small positive* ϵ *is incentive compatible and has revenue* $O(\ln h)$*. However, the revenue of the anonymous pricing is* $O(1)$*.*

Therefore, in this section, we focus on the case when the budget distribution is independent with the valuation distribution for each agent. Note that even with the independence assumption, without any further assumption on the valuation or the budget distribution, posting prices is not approximately optimal even for a single agent, see the following example as an illustration. Therefore, we consider mild assumption on either the valuation distribution or the budget distribution and show the corresponding resemblant property.

Example 2.3.4. *Consider the budget distribution is the discrete equal revenue distribution, i.e.,* $g(i) = {}^{1}\!/_{\varpi \cdot i^2}$*, where* $\varpi = {}^{\pi^2}\!/_{6}$*. Let the quantile function of the valuation distribution be* $q(i) = {}^{1}\!/_{\ln i}$*. The optimal price posting revenue is a constant. Next consider the pricing function* $\tau(x) = \frac{1}{1-x}$*. From this pricing function, the value* v_i *corresponding to payment* i *is* $v_i = i^2$*. Note that the revenue from this payment function is infinity, i.e.,*

$$\begin{aligned}
\textbf{Payoff}[\tau] &\geq \lim_{m\to\infty} \sum_{i=1}^{m} (i \cdot q(v_i) \cdot g(i)) \\
&= \frac{1}{2\varpi} \lim_{m\to\infty} \sum_{i=1}^{m} \frac{1}{i \cdot \ln i} \\
&= \frac{1}{2\varpi} \lim_{m\to\infty} \ln \ln m \to \infty.
\end{aligned}$$

Therefore, the gap between price posting and the optimal mechanism is infinite.

First we show that regularity on the valuation distribution is sufficient to guarantee the resemblance between the ironed price-posting revenue curves and the optimal revenue curve, without further assumption on the budget distribution.

Theorem 2.3.8. *A single agent with private-budget utility and regular valuation distribution has an ironed price-posting revenue curve $\bar{P}$ that is* 3*-resemblant to her optimal revenue curve R, if her value and budget are independently distributed.*

Fix an arbitrary ex ante constraint q, denote EX as the q ex ante revenue-optimal mechanism, and **Payoff**[EX] as its revenue. We decompose EX into two mechanisms $\mathrm{EX}^{\dagger}$ and $\mathrm{EX}^{\ddagger}$ according to the market clearing price p^q. Intuitively, the per-unit prices in $\mathrm{EX}^{\dagger}$ for all types are at most the market clearing price and the per-unit prices in $\mathrm{EX}^{\ddagger}$ for all types are larger than the market clearing price. The details of the decomposition is specified in Section 2.3.1, and we will bound the revenue from those two mechanisms separately.

Lemma 2.3.9. *For a single agent with random-public-budget utility, independently distributed value and budget, and any ex ante constraint q; the revenue of* $\mathrm{EX}^{\dagger}$ *is at most the revenue from posting the market clearing price, i.e., $P(q) \geq \mathbf{Payoff}\left[\mathrm{EX}^{\dagger}\right]$.*

Proof. The ex ante allocation of $\mathrm{EX}^{\dagger}$ is at most the ex ante allocation of EX, i.e., q. Combining with the fact that the per-unit prices in $\mathrm{EX}^{\dagger}$ for all types are weakly lower than the market clearing price, its revenue is at most the revenue of posting the market clearing price. □

For the revenue bound of $\mathrm{EX}^{\ddagger}$, we consider two different cases: (1) the market clearing price is larger than the monopoly reserve; and (2) the market clearing price is smaller than the monopoly reserve.

Lemma 2.3.10. *For a single private-budget agent with independently distributed value and budget and regular value distribution, if the market clearing price $p^q = {}^{P(q)}/_{q}$ is larger* *than the monopoly reserve, i.e.,* $p^q = P(q)/q \geq m^*$*, the revenue of posting the market clearing price is at least the revenue of* $\mathrm{EX}^{\ddagger}$*, i.e.,* $P(q) \geq \textbf{Payoff}\big[\mathrm{EX}^{\ddagger}\big]$*.*

Proof. In both $\mathrm{EX}^{\ddagger}$ and the mechanism that posts the market clearing price, the types with value lower than the market clearing price p^q will purchase nothing, so we only consider the types with value at least p^q in this proof. Each budget level is considered separately.

For types with budget $w \leq p^q$, by posting the market clearing price p^q, those types always pay their budgets w, which is at least the revenue from those types in $\mathrm{EX}^{\ddagger}$.

For types with budget $w > p^q$, by posting the market clearing price p^q, those types always pay p^q. Since the budget constraints do not bind for these types, it is helpful to consider the price-posting revenue curve without budget, which we denote by P^L. The regularity of the valuation distribution guarantees that P^L is concave. The concavity of P^L implies that higher prices above m^* extracts lower revenue than p^q. Since the per-unit prices in $\mathrm{EX}^{\ddagger}$ for all types are at least p^q, the concavity of P^L guarantees that the expected revenue of posting p^q for types with budget larger than p is at least the expected revenue for those types in $\mathrm{EX}^{\ddagger}$. Combining these bounds above, we have $P(q) \geq \textbf{Payoff}\big[\mathrm{EX}^{\ddagger}\big]$. □

Lemma 2.3.11. *For a single private-budget agent with independently distributed value and budget and regular value distribution, if the market clearing price* $p^q = {}^{P(q)}/_{q}$ *is smaller than the monopoly reserve, there exists* $q^{\dagger} \leq q$ *such that the market clearing revenue from* $q^{\dagger}$ *is a 2-approximation to the revenue from* $\mathrm{EX}^{\ddagger}$*, i.e.,* $2P(q^{\dagger}) \geq \textbf{Payoff}\big[\mathrm{EX}^{\ddagger}\big]$*.*

Proof. Note that any price that is at least p^q is feasible for the ex ante constraint q. We consider posting a random price $\boldsymbol{p} = \max\{p^q, \boldsymbol{p}_0\}$ with $\boldsymbol{p}_0$ drawn identically to

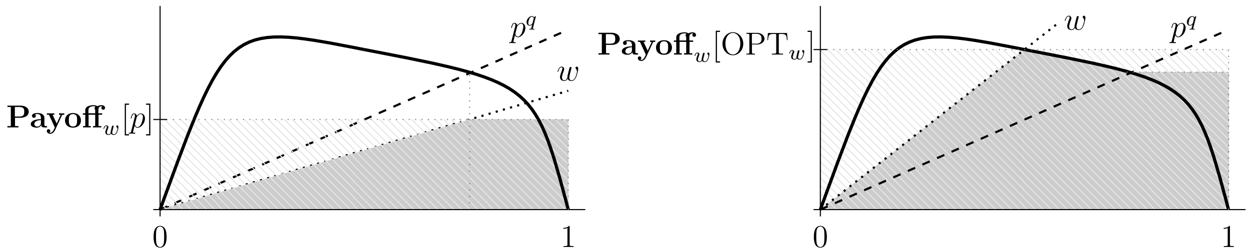

Figure 2.3. In the geometric proof of Lemma 2.3.11, the upper bound on the expected revenue of $\mathrm{EX}^{\ddagger}$ ($\mathbf{Payoff}_w[p]$ and $\mathbf{Payoff}_w[\mathrm{OPT}_w]$ on the left and right, respectively) is the area of the light gray striped rectangle and the revenue from posting random price $\boldsymbol{p}$ is the area of the dark gray region. By geometry, the latter is at least half of the former. The black curve is the price-posting revenue curve with no budget constraint P^L. The figure on the left depicts the small-budget case (i.e., $w < p^q$), and the figure on the right depicts the large-budget case (i.e., $w \geq p^q$).

the agents value distribution. Fixing the budget of the agent w, consider the following geometric argument [cf. Dhangwatnotai et al., 2015]. For both sides of Figure 2.3, the area of the light gray stripped rectangle upper bounds the revenue of $\mathrm{EX}^{\ddagger}$ and the area of the dark gray region is the expected revenue from posting random price $\boldsymbol{p}$. Consequently, concavity of the price-posting revenue curve with no budget constraint P^L (by regularity of the value distribution) implies that a triangle with half the area of the light gray rectangle is contained within the dark gray region and, thus, the random price is a 2-approximation. As the random price does not depend on the budget w, the same bound holds when w is random. Of course, the optimal deterministic price that is at least p^q is only better than the random price and the lemma is shown. The remainder of this proof verifies that the geometry of the regions described above is correct.

The left side of Figure 2.3 depicts the fixed budgets w that are at most p^q. The area of the light gray striped rectangle upper bounds the revenue of $\mathrm{EX}^{\ddagger}$ as follows. Let $\mathbf{Payoff}_w[p]$ be the expected revenue from posting price p to types with budget w. Under both $\mathrm{EX}^{\ddagger}$ and the market clearing price p^q, types with value below the market clearing price pay zero. For the remaining types, in $\mathrm{EX}^{\ddagger}$ they pay at most their budget and in market clearing they pay exactly their budget. Thus, $\mathbf{Payoff}_w\left[\mathrm{EX}^{\ddagger}\right] \leq \mathbf{Payoff}_w[p^q] = w\,(1 - F(p^q))$ where, recall, $1 - F(p^q)$ is the probability the agent's value is at least the market clearing price p^q. Of course, $w\,(1 - F(p^q))$ is the height and area (its width is 1) of the light gray striped region on the left side of Figure 2.3.

The right side of Figure 2.3 depicts the fixed budgets w that are at least p^q. The area of the light gray striped rectangle upper bounds the revenue of $\mathrm{EX}^{\ddagger}$ as follows. Let OPT_w be the optimal mechanism to types with budget w without ex ante constraint and $\mathbf{Payoff}_w[\mathrm{OPT}_w]$ be its expected revenue from these types. Clearly, $\mathbf{Payoff}_w\left[\mathrm{EX}^{\ddagger}\right] \leq \mathbf{Payoff}_w[\mathrm{OPT}_w]$ as the latter optimizes with relaxed constraints of the former. Laffont and Robert [1996] show that OPT_w posts the minimum between budget w and the monopoly reserve m^* when the agent has public budget and regular valuation. As the budget does not bind for this price, its revenue is given by the price-posting revenue curve with no budget constraint, i.e., $\mathbf{Payoff}_w[\mathrm{OPT}_w] = P^L(1 - F(\min\{w, m^*\}))$. Of course, this revenue is the height and area (its width is 1) of the light gray striped region on the right side of Figure 2.3.

Next, we will show that the revenue of posting the random price $\boldsymbol{p}$ is the grey shaded areas illustrated in Figure 2.3 (in both cases). A random price from the value distribution, i.e., $\boldsymbol{p}_0$, corresponds to a uniform random quantile constraint, i.e., drawing uniformly from the horizontal axis. Since we truncate the lower end of the price distribution at the market clearing price p^q, the revenue from quantiles greater than q equals the revenue from the market clearing price. For any fixed w, when $\boldsymbol{p} \in [p^q, w]$, the budget does not bind and the revenue of posting price $\boldsymbol{p}$ is $P^L(\boldsymbol{q})$ where P^L is the price-posting revenue curve without budget; and when $\boldsymbol{p} > w$, the revenue of posting price $\boldsymbol{p}$ is $w\boldsymbol{q}$. Thus, the revenue from a random price is given by the integral of the area under the curve defined by $\boldsymbol{q}w$ when $\boldsymbol{p} \geq w$, by $P^L(\boldsymbol{q})$ when $\boldsymbol{p} \in [w, p^q]$ and this interval exists, and by $\min(w, p^q)$ when $\boldsymbol{p} = p^q$, i.e., when $\boldsymbol{p}_0 \leq p^q$. This area is the dark gray region. □

Proof of Theorem 2.3.8. Fix any ex ante constraint q. If the market clearing price $p^q = {}^{P(q)}/_{q}$ is at least the monopoly reserve, Lemma 2.3.9 and Lemma 2.3.10 imply that $\textbf{Payoff}\big[\mathrm{EX}^{\dagger}\big] \leq P(q)$, and $\textbf{Payoff}\big[\mathrm{EX}^{\ddagger}\big] \leq P(q)$, thus, $P(q)$ is a 2-approximation to $\textbf{Payoff}\big[\mathrm{EX}^{\dagger}\big] + \textbf{Payoff}\big[\mathrm{EX}^{\dagger}\big] = \textbf{Payoff}[\mathrm{EX}]$, i.e., $R(q)$. If the market clearing price p^q is smaller than the monopoly reserve, let $q^{\dagger} = \operatorname{argmax}_{q' \leq q} P(q')$, Lemma 2.3.9 and Lemma 2.3.11 imply that $\textbf{Payoff}\big[\mathrm{EX}^{\dagger}\big] \leq P(q) \leq P(q^{\dagger})$, and $\textbf{Payoff}\big[\mathrm{EX}^{\ddagger}\big] \leq 2P(q^{\dagger})$, thus, $P(q^{\dagger})$ is a 3-approximation to $R(q)$. Thus, the agent is 3-resemblant for ex ante optimization. □

We also consider the assumption that the budget exceeds its expectation with constant probability at least ${}^{1}/_{\kappa}$. This assumption on budget distribution is also studied in Cheng et al. [2018]. Notice that a common distribution assumption, monotone hazard rate, is a special case of it with $\kappa = e$ [cf. Barlow and Marshall, 1965].

Theorem 2.3.12. *A single agent with private-budget utility has an ironed price-posting revenue curve $\bar{P}$ that is $(1 + 3\kappa - {}^{1}/_{\kappa})$-resemblant to her optimal revenue curve R, if her value and budget are independently distributed, and the probability the budget exceeds its expectation is ${}^{1}/_{\kappa}$.*

The proof of Theorem 2.3.12 also uses the similar decomposition technique as in Theorem 2.3.1 and 2.3.8.

Let w^* denote the expected budget of the agent. For any ex ante constraint q, denote EX as the q ex ante revenue optimal mechanism.

Our analysis here is similar to the analysis for welfare, i.e., the price decomposition technique. Consider the decomposition of EX into three mechanisms $\text{EX}^{\dagger}$, $\text{EX}^{\S}$ and $\text{EX}^{\ddagger}$ such that mechanism $\text{EX}^{\dagger}$ contains per-unit prices at most the market clearing price, mechanism $\text{EX}^{\ddagger}$ contains per-unit prices at least the expected budget, while mechanism $\text{EX}^{\S}$ contains per-unit prices between the market clearing price and the expected budget. All mechanisms satisfy the ex ante constraint q, and the sum of their welfare is upper bounded by the welfare of the original ex ante mechanism EX, i.e., $\textbf{Payoff}[\text{EX}] \leq \textbf{Payoff}\big[\text{EX}^{\dagger}\big] + \textbf{Payoff}\big[\text{EX}^{\S}\big] + \textbf{Payoff}\big[\text{EX}^{\ddagger}\big]$. Note that in the special case where the market clearing price is larger than the expected budget, i.e., $p^q > w^*$, $\text{EX}^{\S}$ does not exist and mechanism EX is decomposed into $\text{EX}^{\dagger}$ and $\text{EX}^{\ddagger}$.

We construct the allocation-payment functions $\tau_w^{\dagger}$, $\tau_w^{\ddagger}$ and $\tau_w^{\S}$ for $\text{EX}^{\dagger}$, $\text{EX}^{\ddagger}$, and $\text{EX}^{\S}$ respectively. For each budget w, let τ_w be the allocation-payment function for types with budget w in mechanism EX, and x_w^* be the utility maximization allocation for the agent with value and budget equal to the market clearing price p^q, i.e., $x_w^* = \text{argmax}\{x : \tau_w'(x) \leq p^q\}$. Let $x_w^{\sharp}$ be the utility maximization allocation for the agent with value and budget equal to the expected budget w^*, i.e., $x_w^{\sharp} = \text{argmax}\{x : \tau_w'(x) \leq w^*\}$. Then the allocation-payment functions $\tau_w^\dagger$, $\tau_w^\ddagger$ and $\tau_w^\S$ are defined respectively as follows,

$$\tau_w^\dagger(x) = \begin{cases} \tau_w(x) & \text{if } x \le x_w^*, \\ \infty & \text{otherwise;} \end{cases} \qquad \tau_w^\S(x) = \begin{cases} \tau_w(x_w^* + x) - \tau_w(x_w^*) & \text{if } x \le x_w^\sharp - x_w^*, \\ \infty & \text{otherwise;} \end{cases}$$

$$\tau_w^\ddagger(x) = \begin{cases} \tau_w(x_w^\sharp + x) - \tau_w(x_w^\sharp) & \text{if } x \le 1 - x_w^\sharp, \\ \infty & \text{otherwise.} \end{cases}$$

The revenue contribution from $\mathrm{EX}^\dagger$ is bounded in Lemma 2.3.9. Next we illustrate how to bound the revenue from $\mathrm{EX}^\ddagger$ and $\mathrm{EX}^\S$ respectively using the revenue from price-posting.

Lemma 2.3.13. *For a single agent with private-budget utility, independently distributed value and budget, for any quantile q, there exists $q^\dagger \in [0, q]$ such that $(1+\kappa-1/\kappa)\cdot P(q^\dagger) \ge$ $\mathbf{Payoff}\big[\mathrm{EX}^\ddagger\big]$.*

Proof. Let w^* be the expected budget and let $\bar{p} = \max\{w^*, p^q\}$. Let $\bar{q}$ be the quantile corresponding to value $\bar{p}$ and let $q^\dagger = \mathrm{argmax}_{q' \le q} P(q')$. Thus $P(\bar{q}) \le P(q^\dagger)$. Moreover, by the construction of the decomposition, the per-unit price in $\mathrm{EX}^\ddagger$ is larger than $\bar{p}$. Similar to the proof of Lemma 2.3.10, we only consider the types with value at least $\bar{p}$.

Let $\mathbf{Payoff}_w\big[\tau_w^\ddagger\big]$ be the expected revenue of providing the allocation-payment function $\tau_w^\ddagger$ in $\mathrm{EX}^\ddagger$ to the types with budget w; and let $\mathbf{Payoff}_w[p]$ be the expected revenue of posting price p to the types with budget w. The following three facts allow comparison of $\mathbf{Payoff}\big[\mathrm{EX}^\ddagger\big]$ to $P(q^\dagger)$:

(a) Posting the price $\bar{p}$ makes the budget constraints bind for the types with budget at most w^*, so $\mathbf{Payoff}_w\left[\tau_w^\ddagger\right] \leq \mathbf{Payoff}_w[\bar{p}]$ for all $w \leq w^*$.

(b) $\mathbf{Payoff}_w\left[\tau_w^\ddagger\right] \leq \frac{w}{w^*}\mathbf{Payoff}_{w^*}\left[\tau_w^\ddagger\right]$ for all $w \geq w^*$. This is because if the type (v, w^*) pays her budget w^* (i.e., the budget constraint binds), her payment is a (w/w^*)-approximation to the payment from the type (v, w), since the type (v, w) pays at most w. Moreover, if the type (v, w^*) pays less than her budget w^* (i.e., the unit-demand constraint binds, or the value binds), her allocation is equal to the allocation from the type (v, w) for $w \geq w^*$. Hence, their payments are the same.

(c) Since the revenue of posting price $\bar{p}$ to an agent with budget w^* is at most the revenue to an agent with budget $w > w^*$; with the assumption that budgets exceed the expectation w^* with probability at least $1/\kappa$, it implies that

$$\mathbf{Payoff}_{w^*}[\bar{p}] \cdot \frac{1}{\kappa} \leq \mathbf{E}\left[\mathbf{Payoff}_w[\bar{p}] \;\middle|\; w \geq w^*\right] \cdot \mathbf{Pr}[w \geq w^*] \leq P(\bar{q}).$$

We upper bound the revenue of $\text{EX}^\ddagger$ as follows,

$$\begin{aligned}
\mathbf{Payoff}\left[\text{EX}^\ddagger\right] &= \int_{\underline{w}}^{w^*} \mathbf{Payoff}_w\left[\tau_w^\ddagger\right] dG(w) + \int_{w^*}^{\bar{w}} \mathbf{Payoff}_w\left[\tau_w^\ddagger\right] dG(w) \\
&\leq \int_{\underline{w}}^{w^*} \mathbf{Payoff}_w[\bar{p}]\, dG(w) + \int_{w^*}^{\bar{w}} \frac{w}{w^*}\mathbf{Payoff}_{w^*}\left[\tau_w^\ddagger\right] dG(w) \\
&\leq (1 - \frac{1}{\kappa})P(\bar{q}) + \frac{\int_{w^*}^{\bar{w}} w dG(w)}{w^*}\mathbf{Payoff}_{w^*}[\bar{p}] \\
&\leq (1 - \frac{1}{\kappa})P(\bar{q}) + \mathbf{Payoff}_{w^*}[\bar{p}] \leq (1 + \kappa - \frac{1}{\kappa})P(q^\dagger)
\end{aligned}$$

where the first inequality is due to facts (a) and (b); in the second inequality, the first term is due to $\mathbf{Pr}[w \leq w^*] \leq 1 - {}^1/_\kappa$, the revenue $\mathbf{Payoff}_w[\bar{p}]$ is monotone increasing in w, and by definition $\int_{\underline{w}}^{\bar{w}} \mathbf{Payoff}_w[\bar{p}]\, dG(w) = P(\bar{q})$, and the second term is due to fact (a); and the last inequality is due to $P(\bar{q}) \leq P(q^\dagger)$ and fact (c). □

Lemma 2.3.14. *For a single agent with private-budget utility, independently distributed value and budget, when $p^q \leq w^*$, there exists $q^\dagger \leq q$ such that the price-posting revenue from $q^\dagger$ is a $(2\kappa - 1)$-approximation to the revenue from* $\mathrm{EX}^\S$*, i.e.,* $(2\kappa - 1)P(q^\dagger) \geq \mathbf{Payoff}\left[\mathrm{EX}^\S\right]$.

Proof. Let $q^\dagger = \operatorname{argmax}_{q' \leq q} P(q')$. Suppose the support of the budget distribution is from $[\underline{w}, \bar{w}]$. Let $\tilde{p}$ be the price larger than the market clearing price p^q and smaller than the expected budget w^* that maximizes revenue without the budget constraint. Consider the following calculation with justification below.

$$
\begin{aligned}
\mathbf{Payoff}\left[\mathrm{EX}^\S\right] &= \int_{\underline{w}}^{w^*} \mathbf{Payoff}_w\left[\tau_w^\S\right] dG(w) + \int_{w^*}^{\bar{w}} \mathbf{Payoff}_w\left[\tau_w^\S\right] dG(w) \\
&\overset{(a)}{\leq} \int_{\underline{w}}^{w^*} \mathbf{Payoff}_{w^*}\left[\tau_w^\S\right] dG(w) + \int_{w^*}^{\bar{w}} \frac{w}{w^*} \mathbf{Payoff}_{w^*}\left[\tau_w^\S\right] dG(w) \\
&\overset{(b)}{\leq} \int_{\underline{w}}^{w^*} \mathbf{Payoff}_{w^*}[\tilde{p}]\, dG(w) + \int_{w^*}^{\bar{w}} \frac{w}{w^*} \mathbf{Payoff}_{w^*}[\tilde{p}]\, dG(w) \\
&\overset{(c)}{\leq} (2 - \frac{1}{\kappa})\, \mathbf{Payoff}_{w^*}[\tilde{p}] \\
&\overset{(d)}{\leq} (2\kappa - 1)\, \mathbf{Payoff}[\tilde{p}] \overset{(e)}{\leq} (2\kappa - 1)\, P(q^\dagger).
\end{aligned}
$$

Inequality (a) holds because given the allocation payment function $\tau_w^\S$, the revenue only increases if we increase the budget to w^*, i.e., $\mathbf{Payoff}_w\left[\tau_w^\S\right] \leq \mathbf{Payoff}_{w^*}\left[\tau_w^\S\right]$ for any

$w \leq w^*$. Moreover, for any $w > w^*$, given the allocation payment function $\tau_w^§$, the revenue is either the same for budget w and w^*, or the budget binds for agent with expected budget w^*. Since the revenue from agent with budget w is at most w, we know that $\mathbf{Payoff}_w\left[\tau_w^§\right] \leq {}^w/_{w^*} \cdot \mathbf{Payoff}_{w^*}\left[\tau_w^§\right]$. Note that for allocation payment rule $\tau_w^§$, per-unit prices are larger than the market clearing price p^q and smaller than the expected budget w^*, and budget does not bind for agents with budget w^*. Therefore, by definition, the optimal per-unit price in this range is $\tilde{p}$, $\mathbf{Payoff}_{w^*}\left[\tau_w^§\right] \leq \mathbf{Payoff}_{w^*}[\tilde{p}]$ and inequality (b) holds. Inequality (c) holds because $\int_{\underline{w}}^{w^*} dG(w) \leq 1 - {}^1/_\kappa$ by the assumption that the probability the budget exceeds its expectation is at least κ, and $\int_{w^*}^{\bar{w}} \frac{w}{w^*} dG(w) \leq 1$. Inequality (d) holds because $\mathbf{Payoff}_{w^*}[\tilde{p}] \leq \kappa \cdot \mathbf{Payoff}[\tilde{p}]$ for any randomized prices $\tilde{p}$ according to Cheng et al. [2018]. Inequality (e) holds by the definition of the price-posting revenue curve P and quantile $q^\dagger$, the fact that price $\tilde{p}$ is larger than the market clearing price p^q. □

Proof of Theorem 2.3.12. Let $q^\dagger = \operatorname{argmax}_{q' \leq q} P(q')$. Combining Lemma 2.3.9, 2.3.13 and 2.3.14, we have

$$\mathbf{Payoff}[\mathrm{EX}] \leq \mathbf{Payoff}\left[\mathrm{EX}^\dagger\right] + \mathbf{Payoff}\left[\mathrm{EX}^\ddagger\right] + \mathbf{Payoff}\left[\mathrm{EX}^§\right] \leq \left(1 + 3\kappa - {}^1/_\kappa\right) P(q^\dagger). \quad \square$$

2.4. Risk Averse Agent

Note that the preference of a risk averse agent coincide with a linear agent when the allocation is deterministic, and the welfare optimal mechanism for the single-agent problem with linear utility is deterministic. Thus it is easy to verify that posting price is optimal for welfare maximization under any ex ante constraint and the price-posting welfare curve is 1-resemblant to the optimal welfare curve. Formally, we have the following theorem, with proof omitted.

Theorem 2.4.1. *An agent with risk-averse utility has the price-posting welfare curve P that equals (i.e. 1-resemblant) her optimal welfare curve R.*

Next we consider the revenue maximization problem when agents are risk averse. Specifically, we consider the risk aversion model in Fu et al. [2013], where each agent's utility function has a capacity constraint. Moreover, following Fu et al. [2013], in this section, we consider the mechanisms that are pointwise individual rational, i.e., losers have no payment, and winners pay at most their reported values. Formally, $x = 0$ implies $p = 0$. In Example 2.4.3 at the end of this section, we show that price-posting mechanism is not a constant approximation to the optimal mechanism when we allow the winners to be charged more than their reported values, even when the capacity is as large as the support of the value.

We introduce a definition and two lemmas, which are adapted from Fu et al. [2013]. Let $(\cdot)^+ \triangleq \max\{\cdot, 0\}$.

Definition 2.4.1 (Fu et al., 2013)**.** *A mechanism is a two priced mechanism if, when it serves an agent with quantile q and capacity C, the payment is either $V(q)$ or $V(q) - C$. The probability that agent is charged with payment $V(q)$ is denoted by $x^v(q)$, and the probability that agent is charged with payment $V(q) - C$ is denoted by $x^C(q)$.*

Lemma 2.4.2 (Fu et al., 2013)**.** *The ex ante optimal mechanism for agents with capacitated utility is two priced.*

Lemma 2.4.3 (Fu et al., 2013)**.** *For any agent with capacity C and price-posting revenue curve P, for two priced mechanism with allocation rule $x(q) = x^v(q) + x^C(q)$, the revenue from that agent is upper bounded as*

$$\mathbf{Payoff}[x] \leq \mathbf{E}\big[(P'(q))^+ \cdot x(q)\big] + \mathbf{E}\big[(P'(q))^+ \cdot x^C(q)\big] + \mathbf{E}\big[(V(q) - C)^+ \cdot x^C(q)\big] .$$

Theorem 2.4.4. *A single agent with capacitated risk averse utility, maximum value $\bar{v}$, and capacity $C \leq \bar{v}$, has a price-posting revenue curve P that is $(2 + \ln {}^{\bar{v}}/_{C})$-resemblant to her optimal revenue curve R.*

Proof. For any quantile $\hat{q}$, let x be the optimal allocation that satisfies ex ante allocation constraint $\hat{q}$. By Lemma 2.4.3,

$$R(\hat{q}) = \mathbf{Payoff}[x] \leq \mathbf{E}\big[(P'(q))^+ \cdot x(q)\big] + \mathbf{E}\big[(P'(q))^+ \cdot x^C(q)\big] + \mathbf{E}\big[(V(q) - C)^+ \cdot x^C(q)\big] .$$

Let m^* be the monopoly reserve, and let $q^\dagger = \min\{Q(m^*, P), \hat{q}\}$. By definition, $q^\dagger \leq \hat{q}$. Since the price-posting revenue curve is concave, posting price $V(q^\dagger)$ maximizes expected marginal revenue under ex ante constraint $\hat{q}$. Therefore,

$$\mathbf{E}\big[(P'(q))^+ \cdot x(q)\big] \leq P(q^\dagger) \quad \text{and} \quad \mathbf{E}\big[(P'(q))^+ \cdot x^C(q)\big] \leq P(q^\dagger).$$

When $q^\dagger = Q(m^*, P)$, for any quantile q, $P(q) \leq P(q^\dagger)$. When $q^\dagger = \hat{q} < Q(m^*, P)$, the allocation $x^C(q)$ with ex ante constraint $\hat{q}$ that maximizes $\mathbf{E}\big[(V(q) - C)^+ \cdot x^C(q)\big]$ satisfies that $x^C(q) = 1$ for $q \leq q^\dagger$, and $x^C(q) = 0$ for $q > q^\dagger$. Since the price-posting revenue curve is concave, in this case, $P(q) \le P(q^\dagger)$ when $q \le q^\dagger$. Therefore,

$$\mathbf{E}\big[(V(q) - C)^+ \cdot x^C(q)\big] = \mathbf{E}\left[\left(\frac{P(q)}{q} - C\right)^+ \cdot x^C(q)\right] \le \mathbf{E}\left[\left(\min\left\{\bar{v}, \frac{P(q^\dagger)}{q}\right\} - C\right)^+\right]$$

$$= \int_{\frac{P(q^\dagger)}{\bar{v}}}^{\min\{1, \frac{P(q^\dagger)}{C}\}} \left(\frac{P(q^\dagger)}{q} - C\right) \mathrm{d}q + \int_{\frac{P(q^\dagger)}{\bar{v}}}^{1} (\bar{v} - C)\,\mathrm{d}q \le P(q^\dagger) \ln \frac{\bar{v}}{C}.$$

Combining the above inequalities, we have $R(q) \le P(q^\dagger)(2 + \ln \frac{\bar{v}}{C})$. □

In Theorem 2.4.4, the dependence on $\ln \bar{v}/C$ is necessary even when there is a single agent.

Example 2.4.2 (necessity of the dependence on $\bar{v}/C$). *Fix a constant $\bar{v}$. Consider a single agent with equal revenue distribution. That is, her value v is drawn from $[1, \bar{v}]$ with a density function $1/v^2$ for $v \in [1, \bar{v})$, and a mass point of probability $1/\bar{v}$ on value $\bar{v}$. The revenue for posting any price is* 1*. Suppose the agent has capacity constraint $C \ge 1$, Consider the mechanism that always allocates the item to the agent, and charges her* 0 *if her value v is less than C, and charges her $v - C$ if her value is at least C. The revenue for this mechanism is* $\ln \bar{v}/C$.

Example 2.4.3 (necessity of the restriction to pointwise individually rational mechanisms). *Fix a constant $\bar{v}$. Consider a single agent with equal revenue distribution as in Example 2.4.2. The revenue for posting any price is* 1*. Suppose the agent has capacity constraint $C = \bar{v}$ and consider the mechanism that always allocates the item to the agent, and charges her $v - \bar{v}$ with probability $\frac{1}{2}$, $\bar{v}$ with probability $\frac{1}{2}$. This mechanism is incentive compatible and individually rational. The revenue for this mechanism is half of* *the welfare, which cannot be approximated within a constant fraction by any price-posting mechanism.*

2.5. Endogenous Valuation

When agents can make investment decisions before the auction, we assume that the investment costs are subtracted from the social welfare, i.e., the welfare contribution from agent i when she chooses investment decision a_i and receives allocation x_i is $v_i(a_i, t_i) \cdot x_i - C_i(a_i)$. Note that for agents with endogenous valuation, to apply Theorem 2.2.2 it is also important to specify the timeline for agents to exert costly efforts as it affects the equilibrium payoff of any given mechanism. In this thesis, we assume that the agent can delay the investment decision until she sends a message to the seller. In the case of sequential posted pricing mechanisms, for each agent i, the agent makes the investment decisions after she sees the realized price offered by the seller. Note that the price is infinite if the item is sold to previous agents and agent i will not make any investment given this price. Under this timeline of the model, we can show that agents with endogenous valuation are 1-resemblant for welfare maximization.

Lemma 2.5.1 (Fan and Lorentz, 1954, Gershkov et al., 2021a)**.** *For any function $L : \mathbb{R}^2 \to \mathbb{R}$ such that $L(x, q)$ is supermodular in (x, q) and convex in x, for any pair of allocations $x \prec \hat{x}$,*[11] *we have*

$$\int_0^1 L(x(q), q) \, \mathrm{d}q \leq \int_0^1 L(\hat{x}(q), q) \, \mathrm{d}q.$$

[11] $x \prec \hat{x}$ means that for any $\hat{q} \in [0, 1]$, $\int_0^{\hat{q}} x(q) \, \mathrm{d}q \leq \int_0^{\hat{q}} \hat{x}(q) \, \mathrm{d}q$ and $\int_0^1 x(q) \, \mathrm{d}q = \int_0^1 \hat{x}(q) \, \mathrm{d}q$.

Theorem 2.5.2. *An agent with endogenous valuation has the price-posting welfare curve P that equals (i.e. 1-resemblant) her optimal welfare curve R.*

Proof. Let $L(x, q)$ be the welfare of the agent with type corresponding to quantile q when she makes optimal investment decision given allocation x. By Gershkov et al. [2021a], the function $L(x, q)$ is supermodular in (x, q) and convex in x. For any quantile constraint $\hat{q}$, let $\hat{x}$ be the allocation such that $\hat{x}(q) = 1$ for any $q \leq \hat{q}$ and $\hat{x}(q) = 0$ otherwise. Any mechanism with allocation x that sells the item with probability $\hat{q}$ satisfies $x \prec \hat{x}$. By Lemma 2.5.1, the optimal mechanism that is $\hat{q}$ feasible has allocation rule $\hat{x}$, which is posting a deterministic price to the agent. Thus this agent has price-posting welfare curve P that equals (i.e. 1-resemblant) her optimal welfare curve R. □

For revenue maximization, we show that posted pricing is optimal for the single agent problem given any ex ante constraint if the type distribution satisfies the regularity condition.

Theorem 2.5.3. *An agent with endogenous valuation and regular type distribution has the ironed price-posting revenue curve $\bar{P}$ that equals (i.e. 1-resemblant) her optimal revenue curve R.*

Proof. Let $L(x, q)$ be the virtual value of the agent given allocation x and type with quantile q. By Gershkov et al. [2021a], the function $L(x, q)$ is supermodular in (x, q) and convex in x if the type distribution is regular. Similar to Theorem 2.5.2, for any quantile $\hat{q}$, the optimal mechanism for maximizing the expected virtual value that sells the item with probability at most $\hat{q}$ is posted pricing. Since the expected revenue equals the expected virtual value, this agent has price-posting revenue curve $\bar{P}$ that equals (i.e. 1-resemblant) her optimal revenue curve R. $\square$

CHAPTER 3

Optimization of Scoring Rules for Incentivizing Effort

3.1. Preliminaries

In this section, we present a formal program for the optimization of proper scoring rules for multi-dimensional random states. Section 3.1.1 describes the basic setting for scoring rules and provides an informal description of the optimization problem for scoring rules that elicit the marginal means of the distribution. In Section 3.2.1, we discuss the characterization of proper scoring rules for eliciting the mean with a weak regularity condition. Section 3.2 gives the formal program for optimizing scoring rules for the mean.

A reason for our focus on scoring rules for eliciting the mean is that, even for continuous state spaces, the communication requirements of eliciting the mean are reasonable. The discussion on eliciting full distribution can be found in Hartline et al. [2020, 2021a].

3.1.1. The Scoring Rule Optimization Problem

This section considers the problem of optimizing scoring rules. A scoring rule maps an agent's reported belief about a random state and the realized state to a payoff for the agent. Our model allows the agent to refine her prior belief by exerting a binary effort. Our objective is to maximize the agent's perceived benefit from exerting effort, i.e., the expected difference in score from reporting the prior and posterior distributions.

There is a prior distribution $D \in \Delta(\Omega)$ over the true state $\omega \in \Omega$ where $\Omega \subseteq \mathbb{R}^n$ is any n dimensional space. The distribution D is public information for both the agent and the principal, and in addition, the agent may privately observe a signal about the true state, which induces a posterior G. We denote the probability the agent will obtain the posterior G by $f(G)$. We focus on scoring rules that elicit the mean of the posterior, i.e., the scoring rule asks the agent to report the marginal means of her posterior, and scores the agent based on her report and the realized state. Let μ_G be the mean of posterior G and μ_D be the mean of the prior distribution D. Let $R \subseteq \mathbb{R}^n$ be the report space including all possible posterior means μ_G and let $r \in R$ be the report of the agent. A simple property of means, the report space is the convex hull of the state space. Two constraints on the scoring rules are the boundedness constraint and the proper constraint[1].

Definition 3.1.1. *A scoring rule $S(r, \theta)$ is proper[2] for eliciting mean if for any distribution G and report $r \in R$, we have*

$$\mathbf{E}_{\omega \sim G}[S(\mu_G, \theta)] \geq \mathbf{E}_{\omega \sim G}[S(r, \theta)] \,.$$

Definition 3.1.2. *A scoring rule $S(r, \theta)$ is bounded by B in space $R \times \Omega$ if $S(r, \theta) \in [0, B]$ for any report $r \in R$ and state $\omega \in \Omega$.*

---

[1]These two constraints are natural and standard in the scoring rule literature. For eliciting the mean, the restriction on proper scoring rules is not without loss in the optimization program (3.1).

[2]Our notion of proper scoring rule is weakly proper rather than strictly proper. Most of the literature on scoring rules does not have an objective and to obtain non-trivial results requires scoring rules to be strictly proper. When optimizing scoring rules there is no meaningful difference between strictly proper and proper as the strictness can be arbitrarily small and therefore provide insignificant additional benefit. Note that any weakly proper scoring rule can also be made strictly proper by taking an arbitrarily small convex combination with a strictly proper scoring rule.

The goal for the principal is to design a bounded proper scoring rule that maximizes the difference in expected score between agents who exert effort and those who do not. Next, we will informally define the optimization program.

Informal program. The problem of maximizing the difference in expected score given the maximum score of B, the state space Ω, the report space which is the convex hull of the state space, i.e., $R = \mathrm{conv}(\Omega)$, and the distribution over posteriors f can be written as the following optimization program:

$$\begin{aligned} (3.1) \qquad \max_{S} \quad & \mathbf{E}_{G\sim f,\omega\sim G}[S(\mu_G,\omega) - S(\mu_D,\omega)] \\ \text{s.t.} \quad & S \text{ is a proper scoring rule for eliciting the mean,} \\ & S \text{ is bounded by } B \text{ in space } R \times \Omega. \end{aligned}$$

The above program aims to optimize the incentive for the agent to exert effort. Consider the situation where the agent has a private stochastic cost for obtaining a signal of the true state. If the agent chooses to pay the cost, she sees the realized signal, forms a posterior about the true state, and optimizes according to the posterior. The agent will only choose to pay the cost if her expected gain from obtaining the signal, i.e., the objective value in Program (3.1), is higher than her cost. By designing the optimal scoring rule for Program (3.1), we also maximize the probability that the agent chooses to pay the cost. We will not formally model such costs in this thesis.

3.2. Canonical Scoring Rules

There is a canonical approach for constructing proper scoring rules. In this section we specify Program (3.1) to canonical proper scoring rules. In the next section we show that this specification is without loss for the program. The following definition and proposition are straightforward from first-order conditions and can be found, e.g., in Abernethy and Frongillo [2012].

Definition 3.2.1. *A canonical scoring rule for the mean S is defined by convex utility function $u : R \to \mathbb{R}$ on report space R, subgradient $\xi : R \to \mathbb{R}^n$ of u, and function $\kappa : \Omega \to \mathbb{R}$ on state space Ω as*

$$S(r, \omega) = u(r) + \xi(r) \cdot (\omega - r) + \kappa(\omega). \tag{3.2}$$

Proposition 3.2.1. *Canonical scoring rules are proper.*

Proof. Canonical scoring rules have the following simple interpretation. By making a report r, the agent selects the supporting hyperplane of u at r on which to evaluate the state. This supporting hyperplane has gradient $\xi(r)$ and contains point $(r, u(r))$. The agent's utility is equal to the value of the realized state ω on this hyperplane (plus constant $\kappa(\omega)$ which is independent of the agent's report). With utility given by a random point on a hyperplane, the expected utility is equal to its mean on the hyperplane. When the agent's true posterior belief is that the state has mean r, the agent's expected utility is $u(r)$ (plus a constant equal to the expected value of $\kappa(\cdot)$ under the agent's posterior belief; summarized below as Lemma 3.2.2). Misreporting r' with belief r gives a utility equal to the value of r on the supporting hyperplane with gradient $\xi(r')$ at r'. By convexity of u, a report of r gives the higher utility of $u(r)$. □

The following two lemmas allow the objective and the boundedness constraint of Program (3.1) to be simplified. The first lemma justifies referring to u as the agent's utility function and its proof was observed in the proof of Proposition 3.2.1.

Lemma 3.2.2. *For any canonical scoring rule for the mean S (defined by u, ξ, and κ), the expected utility from belief G and truthfully report of μ_G is*

$$\mathbf{E}_{\omega\sim G}[S(\mu_G,\omega)] = u(\mu_G) + \mathbf{E}_{\omega\sim G}[\kappa(\omega)]\,. \tag{3.3}$$

Lemma 3.2.3. *Fixing utility function u and subgradients ξ and setting the state-function κ to minimize the score bound B, the canonical scoring rule S (defined by u, ξ and κ) satisfies*

$$u(\omega) - u(r) - \xi(r)\cdot(\omega - r) \leq B \tag{3.4}$$

for any report $r \in R$ and state $\omega \in \Omega$.

Proof. Similar to the proof of Proposition 3.2.1, canonical scoring rules (Definition 3.2.1) can be interpreted via supporting hyperplanes of the utility function. The first term on the left-hand side of (3.4) upper bounds the utility that an agent can obtain at state ω, specifically, it is the utility from reporting state ω. The remainder of the left-hand side subtracts the utility that the agent obtains from report r in state ω, i.e., it evaluates, at state ω, the supporting hyperplane of u at report r. Thus, the boundedness constraint requires the difference between the utility function and the value of any supporting hyperplane of the utility function to be bounded at all states $\omega \in \Omega$. Figure 3.1(a) illustrates this bound.

The subgradient in $\{\xi(r) : r \in R\}$ that maximizes the right-hand side of the inequality identifies the range of ex post score of the agent for this scoring rule. To enforce that the score is within $[0, B]$, we select $\kappa(\omega)$ equal to the negative of the lower endpoint of this range so that the score is 0 for the report with the worst score at state ω.

Of course, since the score bound is B, this inequality is tight for some $r \in R$ and $\omega \in \Omega$. □

We now derive the simplified program for canonical scoring rules. The following notation is sufficient to describe this simplified program and is adopted throughout this chapter. For proper scoring rules for eliciting the mean, the posterior mean and report are denoted by r in report space R. The distribution over posterior beliefs induces a distribution over posterior means, slightly abusing notation, we denote both distributions by f. Specifically, $f(r) = \int_{G:\mu_G=r} f(G)\,\mathrm{d}G$, i.e., the density at posterior mean r is equal to the cumulative density of posteriors G with mean $\mu_G = r$. The prior mean of the distribution μ_D is equal to the mean of the posterior means, denoted μ_f, i.e., $\mu_D = \mathbf{E}_{\omega\sim D}[\omega] = \mathbf{E}_{r\sim f}[r] = \mu_f$.

By Lemma 3.2.2, the objective function in Program (3.1) for canonical scoring rules can be simplified as

$$\begin{aligned}&\mathbf{E}_{G\sim f,\omega\sim G}[S(\mu_G, \omega) - S(\mu_D, \omega)]\\&= \int_{\Delta(\Omega)} [u(\mu_G) - u(\mu_D)]\, f(G)\,\mathrm{d}G = \int_R [u(r) - u(\mu_f)]\, f(r)\,\mathrm{d}r.\end{aligned}$$

Note that the simplified objective function does not depend on subgradient ξ or state function κ, the latter of which is cancelled in the score difference. Thus, the value of the objective function is uniquely determined by the utility function u and the distribution over posterior means f. We denote the performance of utility function u given the distribution over posteriors f by

$$\mathrm{Obj}(u, f) = \int_R u(r)\, f(r)\, \mathrm{d}r - u(\mu_f). \tag{3.5}$$

Combining Lemma 3.2.3 with the simplified objective function (3.5), and shifting the utility function by a constant such that $u(\mu_f) = 0$, we get the following optimization program for optimizing over canonical scoring rules. In the next section we show that the restriction to canonical scoring rules is without loss.

$$\begin{aligned}
\text{(3.6)} \quad \mathrm{OPT}(f, B, \Omega) = \max_u \quad & \int_R u(r) f(r)\, \mathrm{d}r \\
\text{s.t.} \quad & u \text{ is a continuous and convex function, and } u(\mu_f) = 0, \\
& \xi(r) \in \nabla u(r), \quad \forall r \in R, \\
& u(\omega) - u(r) - \xi(r) \cdot (\omega - r) \leq B, \quad \forall r \in R, \omega \in \Omega, \\
& R = \mathrm{conv}(\Omega).
\end{aligned}$$

Note that for any distribution f and state space Ω, the optimal objective $\mathrm{OPT}(f, B, \Omega)$ is a linear function of the maximum score B. In most of the chapter, we normalize $B = 1$ and mainly consider the state space $\Omega = [0, 1]^n$. To simplify the notation, we let $\mathrm{OPT}(f) = \mathrm{OPT}(f, 1, [0, 1]^n)$. We will write $\mathrm{OPT}(f, B, \Omega)$ explicitly in Section 3.4 when we discuss general state spaces with bound $B \neq 1$.

3.2.1. Sufficiency of Canonical Scoring Rules

This section provides a partial converse to Proposition 3.2.1 and shows that the restriction to canonical scoring rules is without loss, i.e., Program (3.1) and Program (3.6) are equivalent. The converse will require a weak technical restriction on the set of scoring rules considered.[3] With this restriction, Abernethy and Frongillo [2012] provide a converse to Proposition 3.2.1 for reports in the relative interior of the report space. We generalize their observation to the boundary of the report space when the scoring rule is bounded. The detailed discussion is deferred in Appendix B.1. Formally, we have the following result establishing that Program (3.1) and Program (3.6) are equivalent.

Definition 3.2.2 (Abernethy and Frongillo, 2012). *A scoring rule S is μ-differentiable if all directional derivatives of $\mathbf{E}_{\omega \sim G}[S(\mu_G, \omega)]$ exists for all posteriors G with mean μ_G in the relative interior of R.*

Theorem 3.2.4. *For optimization of the incentive for exerting a binary effort via a bounded and μ-differentiable scoring rule for the mean, it is without loss to consider canonical scoring rules, i.e., Program (3.1) and Program (3.6) are equivalent.*

3.3. Single-dimensional Scoring Rules

In this section, we focus on the special case of single-dimensional state spaces. We characterize the optimal single dimensional scoring rules for eliciting the mean and show that the optimal scoring rules are simple and only depend on the prior mean of the distribution. We compare the quadratic scoring rule to the optimal scoring rule and show

[3] The literature on scoring rules for eliciting the mean, to the best of our knowledge, obtains converses to Proposition 3.2.1 only with restrictions. For example, Lambert [2011] assumes the scoring rules are continuously differentiable in the agent's report. The restriction we employ is weaker than differentiability.

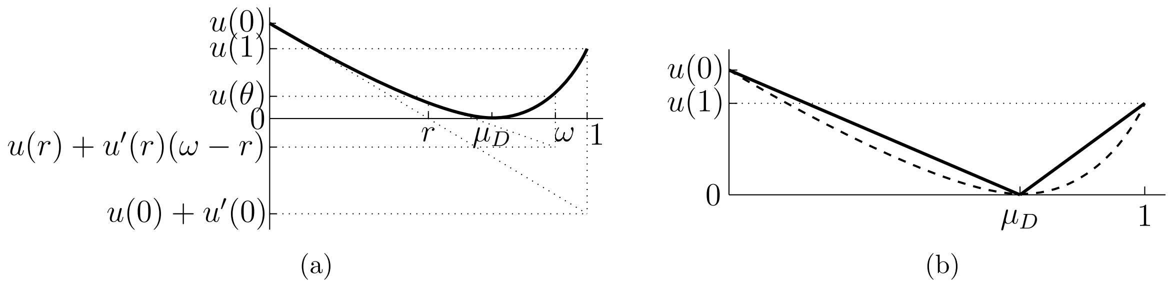

Figure 3.1. The figure on the left hand side illustrates the bounded constraint for proper scoring rule for single dimensional states. The figure on the right hand side characterizes the optimal scoring rule (solid line) for single dimensional states. In this figure, for any convex function u (dotted line) that induces a bounded scoring rule, there exists another convex function $\tilde{u}$ (solid line) which also induces a bounded scoring rule and weakly improves the objective.

that the quadratic scoring rule, though it can be far from optimal for specific distributions over posteriors, it is approximately optimal in the prior-independent setting.

In this section we normalize the state space Ω so that its convex hull, i.e., the report space R, is $[0, 1]$ and the boundedness constraint is given by $B = 1$.

3.3.1. Characterization of Optimal Scoring Rules

In this part, we characterize the optimal proper scoring rules for a single dimensional state. First note that for single dimensional scoring rules, the boundedness constraint of Program (3.6) can be further simplified.

Lemma 3.3.1. *For state space Ω with convex hull $[0, 1]$ and any utility function u, there exists a μ-differentiable proper scoring rule induced by function u which is bounded by $B = 1$ if and only if there exists a set of subgradients $\xi(r) \in \nabla u(r)$ such that*

$$u(1) - u(0) - \xi(0) \leq 1 \text{ and } u(0) - u(1) + \xi(1) \leq 1.$$

Proof. By Lemma B.1.3, it is sufficient to consider only convex function u such that there exists a set of subgradients $\xi(r)$ satisfying constraints that for any $r, \omega \in [0, 1]$

$$u(\omega) - u(r) - \xi(r) \cdot (\omega - r) \leq 1.$$

By convexity of utility u and the monotonicity of subgradients ξ on report space $R = [0, 1]$, it is straightforward to observe that the left-hand side of the boundedness constraint is maximized at $\omega \in \{0, 1\}$ with $r = 1 - \omega$ (see Figure 3.1a). □

With Lemma 3.3.1, Program (3.6) can be written as

$$\begin{aligned} \max_{u} \quad & \int_0^1 u(r) f(r) \,\mathrm{d}r \\ \text{s.t.} \quad & u(r) \text{ is convex and } u(\mu_D) = 0, \\ & \xi(r) \in \nabla u(r), \forall r \in [0, 1], \\ & u(1) - u(0) - \xi(0) \leq 1, \\ & u(0) - u(1) + \xi(1) \leq 1. \end{aligned} \tag{3.7}$$

The main result of this section is the following characterization of the optimal solutions to Program (3.7).

Definition 3.3.1. *A function u is* V-shaped *at μ if there exists parameters a and b such that $u(r) = a\,(r - \mu)$ for $r \leq \mu$ and $u(r) = b\,(r - \mu)$ for $r \geq \mu$.*

Utility functions that are V-shaped at prior mean μ_D are induced by scoring rules with the following simple form. If the agent reports the prior mean her score is zero. For reports above the prior mean, the score is equal to $b\,(\omega - \mu_D)$; and for reports below the prior mean, the score is equal to $a(\omega - \mu_D)$. I.e., as discussed in Section 3.2, the agent's report picks out the supporting hyperplane of the utility function on which to evaluate the state. Note that the implementation of the V-shaped scoring rule only needs the knowledge of the prior mean μ_D, and does not need the distribution over posteriors. We show the following theorem on the optimal solutions of Program (3.7).

Theorem 3.3.2. *For any distribution f over the posterior means with expectation μ_D and state space Ω with convex hull $[0, 1]$, the optimal solutions of Program (3.7) are V-shaped at μ_D with parameters $b = a + {}^1/_{\max\{\mu_D, 1-\mu_D\}}$ and objective value* $\mathrm{OPT}(f) = \mathbf{E}_{r\sim f}[\max(r - \mu_D, 0)]/\max(\mu_D, 1 - \mu_D)$.[4]

Proof. Consider any feasible solution $u(r)$ of Program (3.7). We construct a V-shaped utility function $\tilde{u}(r)$ as

$$\tilde{u}(r) = \begin{cases} -\frac{u(0)}{\mu_D}(r - \mu_D) & \text{for } r \leq \mu_D, \\ \frac{u(1)}{1-\mu_D}(r - \mu_D) & \text{for } r \geq \mu_D. \end{cases}$$

The construction of $\tilde{u}$ is illustrated in Figure 3.1b. It is easy to see that $\tilde{u}$ is convex, $\tilde{u}(\mu_D) = 0$ and $\tilde{u}(r) \geq u(r)$ for any $r \in [0, 1]$. Therefore, the objective value for function $\tilde{u}$ is higher than objective value for function u. Moreover, we have $\tilde{u}(0) = u(0)$, $\tilde{u}(1) = u(1)$, $\tilde{u}'(0) \geq \xi(0)$ and $\tilde{u}'(1) \leq \xi(1)$, which implies $\tilde{u}$ is also a feasible solution to Program (3.7). Thus, an optimal solution is V-shaped.

Next we focus on finding the optimal V-shaped function $\tilde{u}$ for Program (3.7). Let $a = -{}^{u(0)}/_{\mu_D} = \tilde{u}'(0)$ and $b = {}^{u(1)}/_{(1-\mu_D)} = \tilde{u}'(1)$. Since function $\tilde{u}$ satisfies the constraints

[4] By slightly perturbing the utility function u, the V-shaped scoring rule can be transformed into a strictly proper scoring rule with an arbitrarily close objective value.

in Program (3.7), we get

$$b(1-\mu_D) = \tilde{u}(1) \leq 1 + \tilde{u}(0) + \tilde{u}'(0) = 1 - a\cdot\mu_D + a,$$

$$b(1-\mu_D) = \tilde{u}(1) \geq \tilde{u}'(1) + \tilde{u}(0) - 1 = b - a\cdot\mu_D - 1,$$

which implies $b \leq a + {}^1\!/_{(1-\mu_D)}$ and $b \leq a + {}^1\!/_{\mu_D}$. If $b < a + {}^1\!/_{\max\{\mu_D, 1-\mu_D\}}$, then we can either increase b or decrease a to get a better feasible V-shaped utility function. Suppose we fix parameter a, the objective value is pointwise maximized for any report r when $b = a + {}^1\!/_{\max\{\mu_D, 1-\mu_D\}}$.

Next we fix the optimal choice for parameter b. Note that the objective value given any parameter a is

$$\int_0^1 u(r)f(r)\,\mathrm{d}r = \int_0^{\mu_D} a(r-\mu_D)f(r)\,\mathrm{d}r + \int_{\mu_D}^1 \left(a + \frac{1}{\max(\mu_D, 1-\mu_D)}\right)(r-\mu_D)f(r)\,\mathrm{d}r$$

$$= \frac{1}{\max(\mu_D, 1-\mu_D)} \int_{\mu_D}^1 (r-\mu_D)f(r)\,\mathrm{d}r, \tag{3.8}$$

which invariant of parameter a. Therefore, any V-shaped utility function with parameters satisfying $b = a + {}^1\!/_{\max\{\mu_D, 1-\mu_D\}}$ is optimal and obtains objective value given by equation (3.8). □

As mentioned above, we see from Theorem 3.3.2 that the set of utility functions that optimizes Program (3.7) only depends on the prior mean μ_D and not the general shape of the distribution over posterior means f.

An important special case for our subsequent analyses is when the mean of the posteriors is in the center of the report space, i.e., $\mu_D = {}^1\!/_2$ for report space $[0, 1]$. In this case, an optimal utility function u is V-shaped at ${}^1\!/_2$ with $u(0) = u(1) = {}^1\!/_2$. In fact, the symmetric case where f is the uniform distribution on the extremal poster means $\{0, 1\}$ obtains the highest objective value for Program (3.7) with $\text{OPT}(f) = 1/2$. These two observations are fomalized in the following two corollaries.

Corollary 3.3.3. *For any distribution f over the posterior means with expectation $\mu_D = 1/2$, one of the optimal solution of Program (3.7) is symmetric and V-shaped at $1/2$ with $u(0) = u(1) = 1/2$.*

Corollary 3.3.4. *The objective value of any utility function u that is feasible for Program (3.7) on distribution f of posterior means is at most $1/2$, i.e.,* $\text{Obj}(u, f) \leq 1/2$.

Proof. In the characterization of the optimal performance of Theorem 3.3.2, i.e.,

$$\text{OPT}(f) = \mathbf{E}_{r \sim f}[\max(r - \mu_D, 0)] / \max(\mu_D, 1 - \mu_D),$$

it is easy to see that the numerator is maximized and the denominator is minimized in when the distribution of posterior means f is uniform on the extreme points $\{0, 1\}$. For this distribution, the numerator is $1/4$ and the denominator is $1/2$. Thus, $\text{OPT}(f) = 1/2$. □

3.4. Multi-dimensional Scoring Rules

In this section, we focus on the case when the state space is multi-dimensional. We characterize the optimal scoring rule for symmetric distributions over posterior means, and propose a simple scoring rule that is approximately optimal for asymmetric distributions. Then we show that the standard approach in both theory and practice of scoring the agents separately in each dimension is not a good approximation to the optimal multi-dimensional scoring rule.

3.4.1. Optimal Scoring Rules for Symmetric Distributions

This section characterizes the optimal multi-dimensional scoring rule when the distribution over posteriors is symmetric about its center. Program (3.6) is optimized by a symmetric V-shaped utility function. This characterization affords a simple interpretation for rectangular report and state spaces, specifically, the optimal scoring rule can be calculated by taking the maximum score over optimal single-dimensional scoring rules for each dimension, i.e., it is a max-over-separate scoring rule. As these single-dimensional scoring rules depend only on the prior mean, so does the optimal multi-dimensional scoring rule. We first give the characterization and then give the interpretation.

Definition 3.4.1. *A n-dimensional distribution f is* center symmetric *if there exists a center in the report space, i.e., $C \in R$ such that for any $r \in R$, $f(r) = f(2C - r)$.*

Note that for any center symmetric distribution f over posterior means, the mean of the prior coincides with the center of the space, i.e., $\mu_D = C$. The following definition generalizes symmetric V-shaped functions to multi-dimensional state and report spaces.

Definition 3.4.2. *A function u is* symmetric V-shaped *in report and state space $R = \Omega$ with non-empty interior and center C if utility is zero at the center, i.e., $u(C) = 0$, utility is $1/2$ on the boundary, i.e., $u(r) = 1/2$ for $r \in \partial R$, and all other points linearly interpolate between the center and the boundary, i.e., $u(\alpha \cdot r + (1 - \alpha) \cdot C) = \frac{\alpha}{2}$ for any $\alpha \in [0, 1]$ and $r \in \partial R$.*

V-shaped utility functions on convex and center symmetric spaces are bounded and convex, i.e., they are feasible solutions to Program (3.6).

Lemma 3.4.1. *For any convex and center symmetric report and state space $R = \Omega$ with non-empty interior, the center symmetric utility function is convex and bounded for $B = 1$.*

Proof. The following geometry of the utility function is easy verify. First, convexity of report space R implies convexity of u. Second, consider the $n+1$ dimensional space $R \times [-1/2, 1/2]$, where the $n+1$st dimension represents the utility u. The utility function defines a truncated convex cone with vertex equal to $(\mu_D, 0)$ and base at height $1/2$ with cross section R. Consider the point reflection, henceforth, the reflected cone, of this convex cone around its vertex $(\mu_D, 0)$. By basic properties of cones and their point reflections, this reflected cone has the same supporting hyperplanes as the original cone. By the symmetry assumption of R around μ_D, the reflected cone is equal to the mirror reflection of the original cone with respect to the $u = 0$ plane. Consequently, the base of the reflected cone at $u = -1/2$ has cross section equal to R.

We now argue that the utility function satisfies the boundeness constraint, restated for convenience (with report $r \in R$ and state $\omega \in \Omega$):

$$u(\omega) - u(r) - \nabla u(r) \cdot (\omega - r) \leq 1.$$

By definition of the V-shaped utility, we know that the first term is at most $1/2$. The second and third terms, together, can be viewed as subtracting the evaluation, at state ω, of the supporting hyperplane of u at r. The highest point in the reflected cone for any $\omega \in R$ is $-u(\omega)$ and this point lower bounds the value of ω in any of the reflected cones supporting hyperplanes (which are the same as the original cones supporting hyperplanes).

By definition, the reflected cone satisfies $-u(\omega) \geq -1/2$ for $\omega \in R$. We conclude, as desired, that the difference between the first term and the second and third terms is at most 1. □

The following theorem is proved by following a standard approach in multi-dimensional mechanism design, e.g., Armstrong [1996] and Haghpanah and Hartline [2015]. The problem is relaxed onto single-dimensional paths, solved optimally on paths, and it is proven that the solution on paths combine to be a feasible solution on the whole space. Note that in relaxing the problem onto paths, constraints on pairs of reports that are not on the same path are ignored. Similar to the single dimensional V-shaped scoring rule, the implementation of multi-dimensional V-shaped scoring rule only requires the knowledge of the prior mean μ_D.

Theorem 3.4.2. *For any center symmetric distribution f over posterior means in convex report and state space $R = \Omega$, the optimal solution for Program (3.6) is symmetric V-shaped.*

Proof. Consider relaxing the optimization problem on the general space solve it independently on lines through the center. Specifically, consider the conditional distribution of f on the line segment through the center μ_D and the boundary points r and $2\mu_D - r$ on ∂R. Center symmetry implys symmetry on this line segment. By Corollary 3.3.3, the solution to this single-dimensional problem is symmetric V-shaped, i.e., with $u(r) = u(2\mu_D - r) = 1/2$ and $u(\mu_D) = 1/2$.

The solutions on all lines through the center μ_D coincide at μ_D with $u(\mu_D) = 0$. They can be combined, and the resulting utility function u is a symmetric V-shaped function (Definition 3.4.2). Lemma 3.4.1 implies that u is convex and bounded and, thus feasible for the original program. Since it optimizes a relaxation of the original program, it is also optimal for the original program. □

In the remainder of this section we give an interpretation of scoring rules that correspond to V-shaped utility functions on rectangular report and state spaces. On such spaces, these optimal scoring rules can be implemented as the maximum over separate scoring rules (for each dimension). Intuitively, the max-over-separate scoring rule rewards the agent only on the dimension the the agent will receive highest expected payment according to his posterior belief.

The definition of max-over-separate scoring rule is formally introduced in Definition 3.4.3, and it is easy to verify that a max-over-separate scoring rule is proper and bounded if is based on single dimensional scoring rules that are proper and bounded.

Definition 3.4.3. *A scoring rule S is* max-over-separate *if there exists single dimensional scoring rules $(\hat{S}_1, \ldots, \hat{S}_n)$ such that*

(1) For any dimension i, $\hat{S}_i(r_i, \omega_i) = \hat{u}_i(r_i) + \hat{\xi}_i(r_i) \cdot (\omega_i - r_i) + \hat{\kappa}_i(\omega_i)$ where $\hat{\xi}_i(r_i)$ is a subgradient of convex function $\hat{u}_i(r_i)$ and $\hat{\kappa}_i(\omega_i) = \beta_i$ is a constant.

(2) the score is $S(r, \omega) = \hat{S}_i(r_i, \omega_i)$ where $i = \operatorname{argmax}_j \hat{S}_j(r_j, r_j)$.

The incentives of max-over-separate is ensured by the equality of $\hat{S}_j(r_j, r_j)$ (from condition 2) and $\mathbf{E}_{\omega_j \sim G_j}[S_j(r_j, \omega_j)]$ for any marginal posterior distribution G_j on dimension j with mean r_j. Specifically, since the function $\hat{\kappa}_j$ is a constant function of the state, all posteriors G_j with the same mean induce the same expected score.

We conclude the section by showing that, for rectangular report and state spaces, symmetric V-shaped utility functions, which are shown to be optimal by Theorem 3.4.2, can be implemented by max-over-separate scoring rules.

Lemma 3.4.3. *Symmetric V-shaped function u in n-dimensional rectangle report and state space $R = \Omega = \bigtimes_{i=1}^{n}[a_i, b_i]$ with function $\kappa(\omega) = 1/2$ can be implemented as* max-over-separate *scoring rule with single dimensional bounded proper scoring rules $\{\hat{S}_i\}_{i=1}^{n}$ where*

$$\hat{S}_i(r_i, \omega_i) = \begin{cases} -\frac{1}{b_i - a_i}(\omega_i - \mu_{D_i}) + \frac{1}{2} & \textit{for } r_i \leq \mu_{D_i}, \\ \frac{1}{b_i - a_i}(\omega_i - \mu_{D_i}) + \frac{1}{2} & \textit{for } r_i \geq \mu_{D_i}, \end{cases}$$

where $\mu_{D_i} = {}^{(a_i + b_i)}/_2$ is the i^{th} coordinate of the prior mean μ_D.

Proof. First, it is easy to verify that the single dimensional scoring rules $\hat{S}_i$ are proper and bounded in $[0, 1]$. For each dimension i, the utility function for each single dimensional scoring rule $\hat{S}_i$ is V-shaped with

$$\hat{u}_i(r_i) = \begin{cases} -\frac{1}{b_i - a_i}(r_i - \mu_{D_i}) & r_i \leq \mu_{D_i} \\ \frac{1}{b_i - a_i}(r_i - \mu_{D_i}) & r_i \geq \mu_{D_i} \end{cases}, \text{ and} \qquad \hat{\kappa}_i(\omega_i) = 1/2.$$

By Definition 3.4.3, the max-over-separate scoring rule S is $S(r, \omega) = \hat{S}_i(r_i, \omega_i)$ where $i \in \operatorname{argmax}_j \hat{u}_j(r_j)$, and hence the utility function for max-over-separate scoring rule S can be computed as $u(r) = \max_{i \in [n]} \hat{u}_i(r_i)$, which coincides with the symmetric V-shaped function u. □

Corollary 3.4.4. *For any center symmetric distribution f over posterior means in rectangular report and state space $R = \Omega$, a* max-over-separate *scoring rule is optimal.*

Finally, these max-over-separate scoring rules have an indirect choose-and-report implementation where the agent reports the dimension to be scored on and the mean for that dimension. This indirect implementation has a practical advantage that when the communication between the principal and the agent is costly since in n-dimensional spaces, it requires only reporting two rather than n numbers. Note that choose-and-report and max-over-separate are essentially the same scoring rule, with different implementations.

Definition 3.4.4. *A scoring rule S is* choose-and-report *if there exists single dimensional scoring rules $(\hat{S}_1, \ldots, \hat{S}_n)$ such that the agent reports dimension i and mean value r_i, and receives score $S((i, r_i), \omega) = \hat{S}_i(r_i, \omega_i)$.*

An agent's optimal strategy in the choose-and-report scoring rule for proper single-dimensional scoring rules $(\hat{S}_1, \ldots, \hat{S}_n)$ is to choose the dimension i with the highest expected score according to the posterior distribution, i.e., $i = \text{argmax}_j \, \mathbf{E}_{\omega_j \sim G_j}\left[\hat{S}_j(\mu_{G_j}, \omega_j)\right]$, and to report the mean of the posterior for that dimension, i.e., μ_{G_i}. As described above, the advantage of such an indirect scoring rule is that it only requires the agent to report two values to the principal. Lemma 3.4.5 illustrate a nice properties of choose-and-report scoring rules.

Lemma 3.4.5. *The* choose-and-report *scoring rule S defined by proper and bounded single-dimensional scoring rules $(\hat{S}_1, \ldots, \hat{S}_n)$ is itself proper and bounded.*

Proof. Given posterior distribution G, let i be the dimension that maximizes the agent's expected utility under separate scoring rules $\hat{S}_1, \ldots, \hat{S}_n$, i.e.,

$$i = \text{argmax}_j \, \mathbf{E}_{\omega_j \sim G_j}\left[\hat{S}_j(\mu_{G_j}, \omega_j)\right],$$

and let $r_i = \mu_{G_i}$ be the mean of the posterior on dimension i. For report $r = (i, r_i)$ and any other report $r' = (i', r'_i)$, we have

$$\mathbf{E}_{\omega \sim G}[S(r, \omega)] = \mathbf{E}_{\omega_i \sim G_i}\left[\hat{S}_i(r_i, \omega_i)\right] \geq \mathbf{E}_{\omega_{i'} \sim G_{i'}}\left[\hat{S}_{i'}(\mu_{G_{i'}}, \omega_{i'})\right]$$
$$\geq \mathbf{E}_{\omega_{i'} \sim G_{i'}}\left[\hat{S}_{i'}(r'_{i'}, \omega_{i'})\right] = \mathbf{E}_{\omega \sim G}[S(r', \omega)].$$

The first and last equality hold by the definition of choose-and-report proper scoring rules, and the first inequality holds by the definition of dimension i. The second inequality holds since each single dimensional scoring rule is proper. Thus the choose-and-report scoring rule S is proper. Moreover, if each single dimensional proper scoring rule $\hat{S}_i$ is bounded, it is easy to verify that the choose-and-report scoring rule S is also bounded. □

3.4.2. Approximately Optimal Scoring Rules for General Distributions

When the distribution is not symmetric, max-over-separate scoring rules may not be optimal for Program (3.6). However, we show that the optimal max-over-separate scoring rule is approximately optimal for any asymmetric and possibly correlated distribution over a high dimensional rectangular space.

To show this, we symmetrize the distribution over posteriors, and construct a V-shaped scoring rule on the symmetrized distribution. This V-shaped scoring rule can

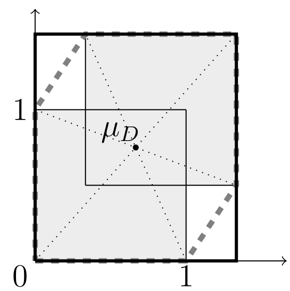

Figure 3.2. This figure depicts a two-dimensional state space. The state space $\Omega = [0, 1]^2$ and its point reflection around the prior mean μ_D are shaded in gray. The extended report and state space are depicted by the region within the thick black rectangle.

be implemented as a max-over-separate scoring rule on the original problem, which only requires the knowledge of prior mean.

Theorem 3.4.6. *For any distribution f over posterior means in n-dimensional rectangular report and state space $R = \Omega = \bigtimes_{i=1}^{n} [a_i, b_i]$, the utility function u of optimal* max-over-separate *scoring rule for Program (3.6) achieves at least $1/8$ of the optimal objective value, i.e.* $\mathrm{Obj}(u, f) \geq 1/8 \cdot \mathrm{OPT}(f, B, \Omega)$.

In the following discussion, we assume without loss of generality that $\mu_{D_i} \geq (a_i + b_i)/2$ for every dimension i. The proof of this theorem introduces the following constructs:

- The *extended report and state space* are $\widetilde{R} = \widetilde{\Omega} = \bigtimes_{i=1}^{n} [a_i, 2\mu_{D_i} - a_i]$. These are rectangular and contain the original report and state spaces $R = \Omega$. See Figure 3.2.
- The *symmetric extended distribution of f* on the extended report space is $\tilde{f}(r) = \frac{1}{2}(f(r) + f(2\mu_D - r))$. Note in this definition that the original distribution f satisfies $f(r) = 0$ for any $r \in \widetilde{R} \setminus R$.

Theorem 3.4.6 now follows by combining the following five lemmas, with proofs provided in Appendix B.2.

Lemma 3.4.7. *Evaluated on any distribution over posterior means f, the optimal* max-over-separate *scoring rule for the distribution f and the state space Ω is at least as good as the optimal scoring rule for the extended distribution $\tilde{f}$ and the extended state space $\widetilde{\Omega}$.*

Lemma 3.4.8. *The symmetric optimizer $\tilde{u}$ for the symmetric extended distribution $\tilde{f}$ and extended state space $\widetilde{\Omega}$ attains the same objective value on the original distribution f, i.e.,* $\mathrm{Obj}(\tilde{u}, f) = \mathrm{OPT}(\tilde{f}, B, \widetilde{\Omega})$.

Lemma 3.4.9. *On extended state space $\widetilde{\Omega}$, the optimal value of Program (3.6) for the symmetric extended distribution $\tilde{f}$ is at least half that for the original distribution f, i.e.,* $\mathrm{OPT}(\tilde{f}, B, \widetilde{\Omega}) \geq \frac{1}{2}\,\mathrm{OPT}(f, B, \widetilde{\Omega})$.

Lemma 3.4.10. *For any distribution over posterior means f, the optimal value of Program (3.6) on the extended state space $\widetilde{\Omega}$ is at least a quarter of that of the original state space Ω, i.e.,* $\mathrm{OPT}(f, B, \widetilde{\Omega}) \geq \frac{1}{4}\,\mathrm{OPT}(f, B, \Omega)$ *or equivalently* $\mathrm{OPT}(f, 4B, \widetilde{\Omega}) \geq \mathrm{OPT}(f, B, \Omega)$.

3.4.3. Inapproximation by Separate Scoring Rules

One way to design the scoring rule for an n-dimensional space is to average independent scoring rules for the marginal distributions of each dimension. In this section we show that the worst-case multiplicative approximation of scoring each dimension separately and scoring optimally is $\Theta(n)$. Moreover, the upperbound $O(n)$ holds for general correlated report distributions, while the lowerbound $\Omega(n)$ holds for independent distributions.

Definition 3.4.5. *A scoring rule S is a* separate scoring rule *if there exists single dimensional scoring rules $(S_1, \ldots, S_n)$ such that $S(r, \omega) = \sum_i S_i(r_i, \omega_i)$.*

Theorem 3.4.11. *In n-dimensional rectangular report and state spaces, the worst-case approximation factor of scoring each dimension separately is $\Theta(n)$.*

Proof. We first argue the upper bound that scoring separately in rectangular report and state spaces guarantees an $O(n)$ approximation. By Theorem 3.4.6, there exists proper and bounded single-dimensional proper scoring rules $(S_1, \ldots, S_n)$ such that the induced max-over-separate S is an 8-approximation to the optimal scoring rule. Let $\hat{S}$ be the separate scoring rule induced by single-dimensional proper scoring rules $(\frac{1}{n}S_1, \ldots, \frac{1}{n}S_n)$. It is easy to verify that scoring rule $\hat{S}$ is bounded, with objective value at least $\frac{1}{n}$ fraction of that for scoring rule S. Thus, separate scoring rule $\hat{S}$ is an $O(n)$ approximation to the optimal scoring rule.

We now give an example of a symmetric distribution over posteriors over the space $R = \Omega = [0, 1]^n$ such that the approximation is $\Omega(n)$. Consider the i.i.d. distribution over posterior means f with marginal distribution f_i dimension i defined by

$$r_i = \begin{cases} 1 & \text{w.p. } {}^1\!/\!{}_{2n}, \\ {}^1\!/\!{}_2 & \text{w.p. } 1 - {}^1\!/\!{}_n, \\ 0 & \text{w.p. } {}^1\!/\!{}_{2n}. \end{cases}$$

The prior mean for each dimension is $^1/_2$ and by Corollary 3.3.3, the optimal scoring rule for each dimension i has V-shaped utility function $\hat{u}_i$ with $\hat{u}_i(0) = \hat{u}_i(1) = {}^1/_2$ and $\hat{u}_i({}^1/_2) = 0$. Thus, the expected objective value for the optimal scoring rule of dimension i

is ${}^{1}\!/\!_{2}\,\mathbf{Pr}_{r_i \sim f_i}[r_i \in \{0, 1\}] = {}^{1}\!/\!_{2n}$. Any average of optimal separate scoring rules, thus, has objective value ${}^{1}\!/\!_{2n}$.

Now consider the max-over-separate scoring rule which has a (multi-dimensional) symmetric V-shaped utility function u and is optimal (see Definition 3.4.2 and Theorem 3.4.2). The objective value is $\mathbf{E}_{r \sim f}[u(r)]$. Importantly $u(r) = 0$ if $r = ({}^{1}\!/\!_{2}, \ldots, {}^{1}\!/\!_{2})$ and, otherwise, $u(r) = {}^{1}\!/\!_{2}$. Thus,

$$\begin{aligned} \mathrm{OPT}(f) &= {}^{1}\!/\!_{2}\,\mathbf{Pr}_{r \sim f}[r \neq ({}^{1}\!/\!_{2}, \ldots, {}^{1}\!/\!_{2})] \\ &= {}^{1}\!/\!_{2}\,(1 - (1 - {}^{1}\!/\!_{n})^{n}) \geq {}^{1}\!/\!_{2}\,(1 - {}^{1}\!/\!_{e}). \end{aligned}$$

Thus, the approximation ratio of optimal separate scoring to optimal scoring is at least ${}^{e\,n}\!/\!_{e\,-\,1}$ (and this bound is tight in the limit of n). □

CHAPTER 4

Selling Data to an Agent with Endogenous Information

4.1. Preliminaries

There is a single agent making decisions facing uncertainly over the state space Ω. Let D be the prior distribution over the states. The agent has a private type $\theta \in \Theta$, and type θ is drawn from a commonly known distribution F with density f. The expected utility of the agent given posterior belief $G \in \Delta(\Omega)$ is $V(G, \theta)$ when her type is θ. We assume that V is convex in G for any type θ.[1] There is a data broker who tries to sell information to the agent to maximize his profit by committing to an information structure that signals the state. Note that an information structure (we will also call this as an experiment) is a mapping $\sigma : \Omega \to \Delta(S)$, where S is the signal space.[2] Let Σ be the set of all possible experiments.

Upon receiving a signal $s \in S$, the agent can conduct her own experiment to further refine her posterior belief on the state with additional costs. Let $\hat{\Sigma} \subseteq \Sigma$ be the set of possible experiments that can be conducted by the agent. The cost of experiment σ given posterior belief G of the agent is denoted by $C^A(\sigma, G) \geq 0$. Let σ^F be the experiment that reveals full information, i.e., $\sigma^F(\omega) = \omega$ for any $\omega \in \Omega$, and let σ^N be the null experiment

[1]The convexity ensures that for any type θ, the agent has higher value for Blackwell more informative experiments.
[2]According to Kamenica and Gentzkow [2011], it is without loss of generality to assume that S is the space of all posterior beliefs, i.e., $S = \Delta(\Omega)$.

that reveals no information with zero cost. In this chapter, we assume that $\sigma^N \in \hat{\Sigma}$,[3] and $C^A(\sigma, G) > 0$ for any $\sigma \neq \sigma^N$.

A mechanism of the data broker is a menu of experiments and associated prices $\{(x_i, p_i)\}$, where x_i is a distribution over experiments. The timeline of the model is illustrated as follows.

(1) The data broker commits to a mechanism $\mathcal{M} = \{(x_i, p_i)\}$.

(2) The agent chooses entry $(x, p) \in \mathcal{M}$, and pays price p to the data broker. The experiment σ is realized according to distribution x and then announced publicly to the agent.

(3) State $\omega \in \Omega$ is realized according to prior D and the data broker sends signal $s \sim \sigma(\omega)$ to the agent.

(4) Upon receiving the signal s, the agent forms posterior belief G, chooses an experiment $\hat{\sigma} \in \hat{\Sigma}$, and pays cost $C^A(\hat{\sigma}, G)$.

(5) The agent receives a signal $s \sim \hat{\sigma}(\omega)$, forms refined posterior belief $\hat{G}$, and receives expected reward $V(\hat{G}, \theta)$.

By the revelation principle, it is without loss to consider the revelation mechanism, that is, the data broker commits to a mapping from types to a distribution over information structures $x : \Theta \to \Delta(\Sigma)$ and the expected payment rule $p : \Theta \to \mathbb{R}$.[4] By slightly overloading the notation, denote

$$V(G, \hat{\Sigma}, \theta) \triangleq \max_{\hat{\sigma} \in \hat{\Sigma}} \mathbf{E}_{\hat{G} \sim \hat{\sigma}|G}\Big[V(\hat{G}, \theta)\Big] - C^A(\hat{\sigma}, G)$$

[3]Intuitively, this assumes that the agent can always choose not to make any additional experiment and pays no extra cost.
[4]The agent observes the realized information structure $\sigma \sim x(\theta)$.

as the maximum utility of the agent given posterior belief G, the set of possible experiments $\hat{\Sigma}$, and private type θ. Here the notation $\hat{\sigma}|G$ represents the distribution over posteriors induced by experiment $\hat{\sigma}$ given the prior belief G. To simplify the notation, we will use $\mathbf{E}_{G\sim x(\theta)|D}[\cdot]$ to represent $\mathbf{E}_{\sigma\sim x(\theta)}\big[\mathbf{E}_{G\sim\sigma|D}[\cdot]\big]$.

Definition 4.1.1. *The mechanism $\mathcal{M} = (x, p)$ is* incentive compatible *if for any type $\theta, \theta' \in \Theta$, we have*

$$\mathbf{E}_{G\sim x(\theta)|D}\Big[V(G, \hat{\Sigma}, \theta)\Big] - p(\theta) \geq \mathbf{E}_{G\sim x(\theta')|D}\Big[V(G, \hat{\Sigma}, \theta)\Big] - p(\theta'),$$

and the mechanism (x, p) is individual rational *if for any type $\theta \in \Theta$, we have*

$$\mathbf{E}_{G\sim x(\theta)|D}\Big[V(G, \hat{\Sigma}, \theta)\Big] - p(\theta) \geq V(D, \hat{\Sigma}, \theta).$$

In this chapter, without loss of generality, we focus on mechanisms (x, p) that are incentive compatible and individual rational. The goal of the data broker is to maximize the expected revenue $\text{Rev}(\mathcal{M}) \triangleq \mathbf{E}_{\theta\sim F}[p(\theta)]$.

For any experiment $\hat{\sigma} \in \hat{\Sigma}$ and any mapping $\kappa : S \to \hat{\Sigma}$, let $\kappa \circ \hat{\sigma}$ represent the experiment such that the agent first conducts experiment $\hat{\sigma}$, and conditional on receiving the signal $s \in S$, the agent continues with experiment $\kappa(s)$ to further refine the posterior belief. For any belief G, let $\hat{G}_{\hat{\sigma},s,G}$ be the posterior belief of the agent when she conducts experiment $\hat{\sigma}$ and receives the signal s. Throughout this chapter, we make the following assumption on the set of possible experiments and the cost function.

Assumption 2. *For any experiment $\hat{\sigma} \in \hat{\Sigma}$ and for any mapping $\kappa : S \to \hat{\Sigma}$, we have $\kappa \circ \hat{\sigma} \in \hat{\Sigma}$. Moreover, for any belief G, we have*

$$C^A(\kappa \circ \hat{\sigma}, G) \leq C^A(\hat{\sigma}, G) + \int_\Omega \int_S C^A(\kappa(s), \hat{G}_{\hat{\sigma},s,G}) \, \mathrm{d}\hat{\sigma}(s|\omega) \, \mathrm{d}G(\omega).$$

Intuitively, Assumption 2 assumes that the set of possible experiments is closed under sequential learning, and the cost function exhibits preference for one-shot learning.[5] This captures the scenario where the agent can repeatedly conduct feasible experiments based on her current posterior belief. Next we illustrate several examples that satisfies the above assumptions.

- $\hat{\Sigma}$ is a singleton. In this case, the unique experiment $\sigma^N \in \hat{\Sigma}$ is null experiment with zero cost.
- $\hat{\Sigma}$ is the set of all possible experiments, i.e., $\hat{\Sigma} = \Sigma$. The cost function C^A is the reduction in information cost, i.e., $C^A(\hat{\sigma}, G) = H(G) - \mathbf{E}_{\hat{G} \sim \hat{\sigma}|G}\left[H(\hat{G})\right]$ where H is any concave function. Possible choices of the information cost function H includes the entropy function [e.g., Sims, 2003] or more generally the uniformly posterior separable cost functions [e.g., Bloedel and Zhong, 2020]. It is easy to verify that uniformly posterior separable cost functions satisfy Assumption 2.
- $\hat{\Sigma}$ is the set of experiments generated by σ^N and $\hat{\sigma}'$ through sequential learning, where σ^N is the one that reveals no additional information with zero cost, and $\hat{\sigma}'$ is an informative experiment that signals the state with fixed cost, i.e., there

[5]Bloedel and Zhong [2020] provide a characterization for the cost function to be indifference for one-shot learning with additional regularity assumptions. In this thesis, we only need to assume weak preference for one-shot learning, and the additional regularity assumptions in Bloedel and Zhong [2020] are not essential.

exists constant $c > 0$ such that $C^A(\hat{\sigma}', G) = c$ for all posterior G. In this case, the agent can choose experiment $\hat{\sigma}'$ as long as it is beneficial for her given her current belief G, and in total the agent pays the cost c multiplies the number of times the experiment $\hat{\sigma}'$ is conducted.[6]

Note that although the general results in this chapter do not require any additional assumption on the valuation function, we will consider the following class of valuation functions in Section 4.2.1 to obtain more structure results on the optimal mechanism. Essentially, we will focus on the setting that the private type of the agent represents her value for acquiring additional information. In particular, the valuation function of the agent is linear.

Definition 4.1.2. *The valuation $V(G, \theta)$ is* linear *if there exists a function $v(G)$ such that $V(G, \theta) = v(G) \cdot \theta$ for any posterior G and any type θ.*

Next we introduce two canonical settings that satisfy the linear valuation assumption.

Example 4.1.3. *Consider the model of a decision maker trying to make a prediction over the states Ω. In this chapter, the agent is the decision maker who chooses an action from the action space A to maximize her payoff. There are several payoff functions of the decision maker that are commonly considered in the literature.*

- *matching utilities: in this case, the states space and the action space are finite, and $\Omega = A = \{1, \ldots, n\}$. the agent gains positive utility if the chosen action matches the state, i.e., the utility of the agent is $u(a, \omega; \theta) = \theta \cdot \mathbf{1}[a = \omega]$, where*

[6] In this case, the agent solves an optimal stopping problem for acquiring additional information. We can also have a continuous time version for acquiring information when the agent has access to signals following a Brownian motion [Georgiadis and Szentes, 2020] or a Poisson process [Zhong, 2017].

$\mathbf{1}[\cdot]$ *is the indicator function and* θ *is the private type of the agent.*[7] *Given belief* G*, when the agent chooses the action optimally, the expected utility of the agent is* $V(G,\theta) = \theta \cdot \max_{\omega \in \Omega} G(\omega)$*. Thus, by letting* $v(G) = \max_{\omega \in \Omega} G(\omega)$*, the valuation of the agent is linear and* $V(G,\theta) = v(G) \cdot \theta$ *for any posterior* G *and type* θ*.*

- *error minimization: in the case,* $\Omega = A \subseteq \mathbb{R}$*, and the agent minimizes the square error between the chosen action and the true state, i.e., the utility of the agent is* $u(a, \omega; \theta) = -\theta \cdot (a - \omega)^2$*. Given belief* G*, the optimal choice of the agent is* $\mathbf{E}[\omega]$*, and the expected utility of the agent is* $V(G,\theta) = -\theta \cdot \mathrm{Var}(G)$*, where* $\mathrm{Var}(G)$ *is the variance of distribution* G*. Thus, by letting* $v(G) = -\mathrm{Var}(G)$*, the valuation of the agent is linear and* $V(G,\theta) = v(G) \cdot \theta$ *for any posterior* G *and type* θ*.*

Example 4.1.4. *Consider the model of monopoly auction in Mussa and Rosen [1978]. In this example, the agent is a firm selling a product to a consumer with private value for different quality levels of the product. The state space* $\Omega = \mathbb{R}_+$ *represents the space of valuations of the consumers. The firm has private cost parameter* c*, and the cost for producing the product with quality* q *is* $c \cdot q^2$*.*[8] *Let* F_G *and* f_G *be the cumulative function and density function given posterior belief* G*. Assuming that the distribution* G *is regular, i.e., the virtual value function* $\phi_G(z) = z - \frac{1-F_G(z)}{f_G(z)}$ *is non-decreasing in* z*,*[9] *the optimal mechanism of the firm with cost* c *is to provide the product with quality* $q(z) = \frac{\max\{0, \phi_G(z)\}}{2c}$

[7] In the special case that $\theta = 1$ with probability 1, this utility function is the matching utility considered in Bergemann, Bonatti, and Smolin [2018]. Note that in Bergemann, Bonatti, and Smolin [2018], the agent has an exogenous private signal that is informative about the state, while in our model that private signal is assumed to be endogenous.

[8] Yang [2020] considers a similar setting with linear cost function $c \cdot q$.

[9] Note that the assumption on regularity is not essential for this example. For any distribution G that is not regular, we can apply the ironing technique in Myerson [1981b] to show that the valuation of the agent, i.e., the profit of the firm, is still a linear function by substituting the virtual value with ironed virtual value.

to the agent with value z, and the expected profit of the firm is

$$\int_{\mathbb{R}_+} \frac{\max\{0, \phi_G(z)\}^2}{4c} \,\mathrm{d}G(z).$$

Let $\theta = \frac{1}{c}$ *be the private type of the firm, and let* $v(G) = \frac{1}{4}\int_{\mathbb{R}_+} \max\{0, \phi_G(z)\}^2 \,\mathrm{d}G(z)$*. The valuation function is* $V(G,\theta) = v(G)\cdot\theta$ *given any type* θ *and any belief* G*, which satisfies the linearity condition.*

4.2. Menu Complexity of the Optimal Mechanisms

In this section, we will provide characterizations for the revenue optimal mechanism without any assumption on the valuation function.

Theorem 4.2.1. *In the revenue optimal mechanism, the following two properties hold.*

(1) The agent does not acquire costly information under equilibrium. That is,

$$\mathbf{E}_{\theta\sim F}\Big[\mathbf{E}_{G\sim x(\theta)|D}\big[C^A(\hat{\sigma}^*_{G,\theta}, G)\big]\Big] = 0$$

where $\hat{\sigma}^*_{G,\theta} \in \arg\max_{\hat{\sigma}\in\hat{\Sigma}} \mathbf{E}_{\hat{G}\sim\hat{\sigma}|G}\Big[V(\hat{G},\theta)\Big] - C^A(\hat{\sigma}, G)$.[10]

(2) Revealing full information is in the menu of the optimal mechanism, i.e., there exists a type $\theta \in \Theta$ *such that* $x(\theta) = \sigma^F$*.*

Before the formal proof, we first illustrate its implications and interpretations. By taxation principal, any mechanism can be represented as offering a menu of experiment-payment pairs to the agent, and the agent chooses the utility maximization one. The menu complexity of the mechanism is defined as the minimum number of menu entries

[10]Since $C^A(\hat{\sigma}, G) \geq 0$ for any $\hat{\sigma}$ and G, the agent only acquires costly information for a set with measure zero.

required to represent the mechanism. By Theorem 4.2.1, it is easy to see the the menu complexity of the optimal mechanism can be unbounded.

Corollary 4.2.2. *Over the worst case of the cost functions and type distributions, the menu complexity of the optimal mechanism is unbounded.*

Next we interpret the results. The second statement of Theorem 4.2.1 is the standard no distortion at the top result. As we made no assumption on the type space here, we cannot pin down the type that receives full information. In Section 4.2.1, we will provide more structural properties when there is a natural order on the type space, and in that case the highest type will receive full information. What is more interesting is the first statement, where the theorem states that under equilibrium, the agent never (except for a set with measure zero) has incentive to acquire additional costly information after receiving the signal from the data broker. This holds because if the agent with type θ acquires additional information by conducting experiment $\hat{\sigma}_\theta$ with positive cost, the data broker can directly supply this experiment to the agent in the information structure, and increases the payment of type θ by the cost of the experiment $\hat{\sigma}_\theta$. The new mechanism increases the expected revenue of the data broker, and eliminates the incentives for the agent with type θ to further acquire any additional information. In the following proof, we will formally show that this new mechanism is also incentive compatible and individual rational.[11]

[11]Note that this result relies crucially on Assumption 2. If Assumption 2 is violated, it is possible that the agent has strict incentive to acquire additional costly information under equilibrium.

Proof of Statement 1 of Theorem 4.2.1. Let $\mathcal{M} = (x, p)$ be the optimal mechanism. Let $\kappa_{\theta,\sigma}$ be the optimal choice of experiments for agent with type θ when she receives the realized experiment σ from mechanism $\mathcal{M}$. By contradiction, let $\hat{\Theta}$ be the set of types with positive measure such that for any $\hat{\theta} \in \hat{\Theta}$, the cost for additional experiments given optimal best response strategy $\kappa_{\hat{\theta},\sigma}$ for agent with type $\hat{\theta}$ is positive, i.e.,

$$\int_{\Sigma} \int_{\Omega} \int_{S} C^A(\kappa_{\hat{\theta},\sigma}(s), \hat{G}_{\sigma,s,D}) \,\mathrm{d}\sigma(s|\omega) \,\mathrm{d}D(\omega) \,\mathrm{d}x(\sigma|\hat{\theta}) > 0,$$

where $\hat{G}_{\sigma,s,D}$ is the posterior given experiment σ and signal s, assuming the prior is D. Let $\hat{x}$ and $\hat{p}$ be the allocation and payment rule such that

- for any $\theta \notin \hat{\Theta}$, $\hat{x}(\theta) = x(\theta)$ and $\hat{p}(\theta) = p(\theta)$;
- for any $\hat{\theta} \in \hat{\Theta}$, for any $\sigma \in \Sigma$, $\hat{x}(\sigma|\hat{\theta}) = x(\kappa_{\hat{\theta},\sigma} \circ \sigma|\hat{\theta})$,[12] and

$$\hat{p}(\hat{\theta}) = p(\hat{\theta}) + \int_{\Sigma} \int_{\Omega} \int_{S} C^A(\kappa_{\hat{\theta},\sigma}(s), \hat{G}_{\sigma,s,D}) \,\mathrm{d}\sigma(s|\omega) \,\mathrm{d}D(\omega) \,\mathrm{d}x(\sigma|\hat{\theta}).$$

Let $\hat{\mathcal{M}} = (\hat{x}, \hat{p})$. It is easy to verify that

$$\begin{aligned}
&\mathrm{Rev}(\hat{\mathcal{M}}) = \int_{\Theta} \hat{p}(\theta) \,\mathrm{d}F(\theta) = \int_{\Theta \backslash \hat{\Theta}} \hat{p}(\theta) \,\mathrm{d}F(\theta) + \int_{\hat{\Theta}} \hat{p}(\theta) \,\mathrm{d}F(\theta) \\
&= \int_{\Theta \backslash \hat{\Theta}} p(\theta) \,\mathrm{d}F(\theta) \\
&\quad + \int_{\hat{\Theta}} \left(p(\theta) + \int_{\Sigma} \int_{\Omega} \int_{S} C^A(\kappa_{\hat{\theta},\sigma}(s), \hat{G}_{\sigma,s,D}) \,\mathrm{d}\sigma(s|\omega) \,\mathrm{d}D(\omega) \,\mathrm{d}x(\sigma|\hat{\theta}) \right) \mathrm{d}F(\theta) \\
&< \int_{\Theta \backslash \hat{\Theta}} p(\theta) \,\mathrm{d}F(\theta) + \int_{\hat{\Theta}} p(\theta) \,\mathrm{d}F(\theta) = \mathrm{Rev}(\mathcal{M}).
\end{aligned}$$

[12] Note that under this new sequential experiment $\hat{x}(\hat{\theta})$, the signals generated by experiments in all stages will be revealed to the agent.

The inequality holds because the types in set $\hat{\Theta}$ occur with positive measure. Thus the revenue of mechanism $\hat{\mathcal{M}}$ is strictly higher. Moreover, in mechanism $\hat{\mathcal{M}}$, the utility of the agent has at least the same expected utility compared to mechanism $\mathcal{M}$ by not acquiring any additional information upon receiving the signal. Therefore, mechanism $\hat{\mathcal{M}}$ is individual rational. Next it is sufficient to show that mechanism $\hat{\mathcal{M}}$ is incentive compatible. It is easy to verify that the agent with any type θ has no incentive to deviate to type $\hat{\theta} \notin \hat{\Theta}$ since her utility for reporting truthfully weakly increases, while her utility for misreporting $\hat{\theta}$ remains the same. Finally, for any type θ, under mechanism $\hat{\mathcal{M}}$, the utility for deviating the report from type θ to type $\hat{\theta} \in \hat{\Theta}$ is

$$\begin{aligned}
U(\theta; \hat{\theta}, \hat{\mathcal{M}}) &= \mathbf{E}_{G\sim\hat{x}(\hat{\theta})|D}\Big[V(G, \hat{\Sigma}, \theta)\Big] - \hat{p}(\hat{\theta}) \\
&= \mathbf{E}_{G\sim\hat{x}(\hat{\theta})|D}\Big[V(G, \hat{\Sigma}, \theta)\Big] - p(\hat{\theta}) - \int_{\Sigma}\int_{\Omega}\int_{S} C^A(\kappa_{\hat{\theta},\sigma}(s), \hat{G}_{\sigma,s,D}) \,\mathrm{d}\sigma(s|\omega) \,\mathrm{d}D(\omega) \,\mathrm{d}x(\sigma|\hat{\theta}) \\
&= \int_{\Sigma}\int_{\Omega}\int_{S} \Big(\mathbf{E}_{\hat{G}\sim\kappa_{\hat{\theta},\sigma}(s)|\hat{G}_{\sigma,s,D}}\Big[V(\hat{G}, \hat{\Sigma}, \theta)\Big] - C^A(\kappa_{\hat{\theta},\sigma}(s), \hat{G}_{\sigma,s,D})\Big) \,\mathrm{d}\sigma(s|\omega) \,\mathrm{d}D(\omega) \,\mathrm{d}x(\sigma|\hat{\theta}) \\
&\quad - p(\hat{\theta}) \\
&\le \int_{\Sigma}\int_{\Omega}\int_{S} V(\hat{G}_{\sigma,s,D}, \hat{\Sigma}, \theta) \,\mathrm{d}\sigma(s|\omega) \,\mathrm{d}D(\omega) \,\mathrm{d}x(\sigma|\hat{\theta}) - \hat{p}(\hat{\theta}) \\
&= U(\theta; \hat{\theta}, \mathcal{M}).
\end{aligned}$$

The inequality holds because by Assumption 2, given any posterior $\hat{G}_{\sigma,s,D}$, a feasible choice for the agent is to choose $\kappa_{\hat{\theta},\sigma}(s) \in \hat{\Sigma}$, pay cost $C^A(\kappa_{\hat{\theta},\sigma}(s), \hat{G}_{\sigma,s,D})$, and then choose additional experiments optimally given the realized signal. The utility of this choice is upper bounded by directly choosing the optimal experiment from $\hat{\Sigma}$, which induces value

$V(\hat{G}_{\sigma,s,D}, \hat{\Sigma}, \theta)$ for the agent. Thus, we have

$$U(\theta; \hat{\mathcal{M}}) - U(\theta; \hat{\theta}, \hat{\mathcal{M}}) \geq U(\theta; \mathcal{M}) - U(\theta; \hat{\theta}, \mathcal{M}) \geq 0,$$

and mechanism $\hat{\mathcal{M}}$ is incentive compatible. □

PROOF OF STATEMENT 2 OF THEOREM 4.2.1. For any mechanism $\mathcal{M} = (x, p)$, let $\bar{p} = \sup_\theta p(\theta)$. By adding the choice $(\sigma^F, \bar{p})$ into the menu of mechanism $\mathcal{M}$, the revenue of the data broker only increases. □

Note that although under equilibrium, the agent has no incentive to acquire additional information. The optimal revenue is not equal to the case when the agent cannot acquire additional information. In fact, the ability to potentially acquire additional information distorts the incentives of the agent, and decreases the revenue the seller can extract from the agent. Letting $\mathrm{OPT}\left(F, \hat{\Sigma}\right)$ be the optimal revenue when the type distribution is F and the set of possible experiments for the agent is $\hat{\Sigma}$, we have the following characterization for the optimal revenue of the data broker.

Proposition 4.2.3. *For any set of experiments $\hat{\Sigma} \subseteq \hat{\Sigma}' \subseteq \Sigma$, any type distribution F, we have*

$$\mathrm{OPT}\left(F, \hat{\Sigma}'\right) \leq \mathrm{OPT}\left(F, \hat{\Sigma}\right).$$

Proof. For any set of experiments $\hat{\Sigma} \subseteq \hat{\Sigma}' \subseteq \Sigma$, for any type distribution F, let $\mathcal{M}$ be the optimal mechanism when the type distribution is F, and the agent can conduct additional experiments in $\hat{\Sigma}'$. Next we show that mechanism $\mathcal{M}$ is incentive compatible and individual rational when the set of additional experiments for the agent is $\hat{\Sigma}$. By

Theorem 4.2.1, for any type θ, when the set of experiments is $\hat{\Sigma}'$, the agent has no incentive to acquire additional information on equilibrium path. Thus when the set of additional experiments is $\hat{\Sigma}$, by reporting the type truthfully, the utility of the agent remains the same. Therefore, mechanism $\mathcal{M}$ is individual rational. In addition, since $\hat{\Sigma} \subseteq \hat{\Sigma}'$, in mechanism $\mathcal{M}$, the utility of deviating to any other type is weakly smaller when the set of additional experiments is $\hat{\Sigma}$. Thus mechanism $\mathcal{M}$ is incentive compatible as well given the set $\hat{\Sigma}$. Therefore, we have

$$\mathrm{OPT}\left(F, \hat{\Sigma}'\right) = \mathrm{Rev}(\mathcal{M}, \hat{\Sigma}') = \mathrm{Rev}(\mathcal{M}, \hat{\Sigma}) \leq \mathrm{OPT}\left(F, \hat{\Sigma}\right). \qquad \square$$

An immediate implication of Proposition 4.2.3 is that the revenue of the data broker is maximized when $\hat{\Sigma} = \{\sigma^N\}$ i.e., the agent cannot acquire additional information.

4.2.1. Linear Valuation

In this section, we will obtain more structure results on the optimal mechanism by restricting the type space and the family of valuation functions. We assume that the type space is single dimensional, i.e., $\Theta = [\underline{\theta}, \overline{\theta}] \subseteq \mathbb{R}$, and the valuation function of the agent is linear. The valuation functions illustrated in Example 4.1.3 and 4.1.4 satisfy the required assumptions.

Let $\phi(\theta) = \theta - \frac{1-F(\theta)}{f(\theta)}$ be the virtual value function of the agent. Let $\theta^* = \inf_\theta\{\phi(\theta) \geq 0\}$ be the lowest type with virtual value 0. We introduce the following regularity assumption on the type distribution to simplify the exposition in the chapter. This assumption is widely adopted in the auction design literature since Myerson [1981b].

Assumption 3. *The distribution F is* regular, *i.e., the corresponding virtual value function $\phi(\theta)$ is monotone non-decreasing in θ.*

In this section, we characterize the optimal revenue of the data broker using the Envelope Theorem [Milgrom and Segal, 2002]. For agent with private type θ, the interim utility given revelation mechanism $\mathcal{M} = (x, p)$ is

$$U(\theta) = \mathbf{E}_{G\sim x(\theta)|D}\Big[V(G, \hat{\Sigma}, \theta)\Big] - p(\theta).$$

The derivative of the utility is

$$U'(\theta) = \mathbf{E}_{G\sim x(\theta)|D}\Big[V_3(G, \hat{\Sigma}, \theta)\Big] = \mathbf{E}_{G\sim x(\theta)|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{\theta,G}|G}\Big[v(\hat{G})\Big]\Big]$$

where $\hat{\sigma}_{\theta,G} \in \hat{\Sigma}$ is the optimal experiment for the agent given private type θ and belief G, and $V_3(G, \hat{\Sigma}, \theta)$ is the partial derivative on the third coordinate. Thus, we have

$$U(\theta) = \int_{\underline{\theta}}^{\theta} \mathbf{E}_{G\sim x(z)|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\Big[v(\hat{G})\Big]\Big] \,\mathrm{d}z + U(\underline{\theta}).$$

Then the revenue of the data broker is

$$\begin{aligned}
\mathrm{Rev}(\mathcal{M}) &= \mathbf{E}_{\theta\sim F}\Big[\mathbf{E}_{G\sim x(\theta)|D}\Big[V(G, \hat{\Sigma}, \theta)\Big] - \int_{\underline{\theta}}^{\theta} \mathbf{E}_{G\sim x(z)|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\Big[v(\hat{G})\Big]\Big] \,\mathrm{d}z - U(\underline{\theta})\Big] \\
&= \mathbf{E}_{\theta\sim F}\Big[\mathbf{E}_{G\sim x(\theta)|D}\Big[V(G, \hat{\Sigma}, \theta) - \frac{1-F(\theta)}{f(\theta)} \cdot \mathbf{E}_{\hat{G}\sim\hat{\sigma}_{\theta,G}|G}\Big[v(\hat{G})\Big]\Big]\Big] - U(\underline{\theta}) \\
&= \mathbf{E}_{\theta\sim F}\Big[\mathbf{E}_{G\sim x(\theta)|D}\Big[\phi(\theta) \cdot \mathbf{E}_{\hat{G}\sim\hat{\sigma}_{\theta,G}|G}\Big[v(\hat{G})\Big] - C^A(\hat{\sigma}_{\theta,G}, G)\Big]\Big] - U(\underline{\theta}), \qquad (4.1)
\end{aligned}$$

where the second equality holds by integration by parts. The next lemma provides sufficient and necessary conditions on the allocations such that the resulting mechanism is incentive compatible and individual rational.

Lemma 4.2.4. *An allocation rule x can be implemented by an incentive compatible and individual rational mechanism if and only if for any $\theta, \theta' \in \Theta$,*[13]

$$\text{(IC)} \qquad \int_{\theta'}^{\theta} \mathbf{E}_{G\sim x(z)|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\Big[v(\hat{G})\Big]\Big] - \mathbf{E}_{G\sim x(\theta')|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\Big[v(\hat{G})\Big]\Big]\,\mathrm{d}z \geq 0,$$

$$\text{(IR)} \qquad \int_{\underline{\theta}}^{\theta} \mathbf{E}_{G\sim x(z)|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\Big[v(\hat{G})\Big]\Big]\,\mathrm{d}z + U(\underline{\theta}) \geq V(D, \hat{\Sigma}, \theta).$$

The proof of Lemma 4.2.4 is deferred in Appendix C.1. The incentive constraint on allocation is similar to the integral monotonicity provided in Yang [2020], where the author considers selling data to an agent without any ability to further acquire information. Note that it is not sufficiently to consider experiments that are Blackwell monotone for designing incentive compatible and individual rational mechanisms.[14] This illustrates a distinction between our model and the classical single-item auction where in the latter case the monotonicity of the interim allocation is sufficient to ensure the incentive compatibility of the mechanisms.

When agents can acquire endogenous information, the incentive constraints cannot be simplified to the monotonicity constraint, and the individual rational constraint may bind for types higher than the lowest type $\underline{\theta}$. Thus the point-wise optimization method

[13] If $\theta < \theta'$, we use $\int_{\theta'}^{\theta}$ to represent $-\int_{\theta}^{\theta'}$.

[14] Sinander [2019] shows that under some regularity conditions, experiments that are monotone in Blackwell order can be implemented by incentive compatible mechanism. However, those conditions are violated in our model and Blackwell monotone experiments may not be implementable. This issue of implementation with Blackwell monotone experiments for selling information has also been observed in the model of Yang [2020].

for classical auction design cannot be applied when there is endogenous information. In the following theorem, we provide a full characterization of the optimal mechanism under Assumption 2 and 3 by directly tackling the constraints on the integration of allocations. Note that the regularity assumption (Assumption 3) is only made to simplify the exposition. The same characterization holds for irregular distributions by adopting the ironing techniques in Toikka [2011]. The detailed proof of Theorem 4.2.5 is provided in Appendix C.1.

Theorem 4.2.5. *For linear valuations, under Assumption 2 and 3, there exists an optimal mechanism* $\widehat{\mathcal{M}}$ *with allocation rule* $\hat{x}$ *such that,*[15]

- *for any type* $\theta \geq \theta^*$*, the data broker reveals full information, i.e.,* $\hat{x}(\theta) = \sigma^F$*;*
- *for any type* $\theta < \theta^*$*, the data broker commits to information structure*

$$\hat{x}(\theta) = \arg\max_{\hat{\sigma} \in \hat{\Sigma}} \mathbf{E}_{G \sim \hat{\sigma}|D}[V(G, \theta)] - C^A(\hat{\sigma}, D)$$

  where ties are broken by maximizing the cost $C^A(\hat{\sigma}, D)$*;*
- $U(\underline{\theta}) = V(D, \hat{\Sigma}, \underline{\theta})$.

Theorem 4.2.5 implies that there is no distortion at the top in the optimal mechanism. Intuitively, when the agent has sufficiently high type, i.e., $\phi(\theta) > 0$, by fully revealing the information to the agent, the expected virtual value is maximized since $\mathbf{E}_{G \sim \sigma|D}\left[\mathbf{E}_{\hat{G} \sim \hat{\sigma}_{\theta,G}|G}\left[v(\hat{G})\right]\right]$ is maximized when the signal σ fully reveals the state. Moreover, with fully revealed state, the posterior belief of the agent is a singleton, and hence the cost of the endogenous information is zero since there is no additional information

[15]The characterization on allocation actually holds in any optimal mechanism except for a set of types with measure zero.

available. By Equation (4.1), this allocation maximizes the virtual surplus, and hence the expected revenue of the data broker.

According to the characterization in Theorem 4.2.5, for any type $\theta < \theta^*$, the utility of the agent in the optimal mechanism $\widehat{\mathcal{M}}$ is

$$\begin{aligned} U(\theta) &= \int_{\underline{\theta}}^{\theta} \mathbf{E}_{G\sim\hat{x}(z)|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\Big[v(\hat{G})\Big]\Big] \, \mathrm{d}z + U(\underline{\theta}) \\ &= \int_{\underline{\theta}}^{\theta} \mathbf{E}_{G\sim\hat{\sigma}_{\theta,D}|D}[v(G)] \, \mathrm{d}z + V(D, \hat{\Sigma}, \underline{\theta}) = V(D, \hat{\Sigma}, \theta). \end{aligned}$$

Thus the individual rational constraint does not only bind for the lowest type, but also for all types below the monopoly type θ^*. Note that this is different from the Myerson's auction design problem or selling information when the agent cannot acquire additional information. In those cases, the utility of the low type agents coincide with the outside option because the seller chooses not to sell to these agents. Here the data broker provide valuable information with positive payment to the agent such that the low type agents are exactly indifferent between participation and choosing the outside option.

4.3. Pricing for Full Information

Note that although the optimal mechanism may require a complex price discrimination scheme against different types when the agent can acquire additional costly information, in the remaining part of this section, we will provide sufficient conditions on the prior distribution, the cost function or the type distribution such that pricing for revealing full information is optimal or approximately optimal.

Proposition 4.3.1. *For any cost function C^A and any prior D, if $\sigma^N \in \arg\max_{\hat{\sigma}\in\hat{\Sigma}} \mathbf{E}_{\hat{G}\sim\hat{\sigma}|D}\left[V(\hat{G},\theta^*)\right] - C^A(\hat{\sigma}, D)$, the optimal mechanism is to post a price for revealing full information.*

In Appendix C.1, we will show that when the monopoly type θ^* find it optimal to not acquire costly information, all types below θ^* will have strictly incentives to not acquire costly information. Combining this observation with the characterization in Theorem 4.2.5, we directly obtain the result that the optimal mechanism is pricing for revealing full information.

Note that the condition in Proposition 4.3.1 is also necessary for pricing for revealing full information to be revenue optimal when $\theta^* > \underline{\theta}$ and the type distribution is continuous with positive density everywhere. This is because if the monopoly type θ^* has strict incentive to acquire costly information given the prior, there exists a positive measure of types below θ^* that also have strict incentive to acquire costly information given the prior. Hence the data broker can have strictly revenue increase by price discriminating those low types.

There are two interpretations for Proposition 4.3.1. Fixing the cost function C^A, we say the prior D is *sufficiently informative* if $\sigma^N \in \arg\max_{\hat{\sigma}\in\hat{\Sigma}} \mathbf{E}_{\hat{G}\sim\hat{\sigma}|D}\left[V(\hat{G},\theta^*)\right] - C^A(\hat{\sigma}, D)$. This is intuitive since when the prior is sufficiently close to the degenerate pointmass distribution, the marginal cost for additional information is sufficiently high while the marginal benefit of additional information is bounded. Thus it is not beneficial for the agent to not acquiring any additional information. We formalize the intuition in Appendix C.1.

An alternative interpretation for Proposition 4.3.1 is that by holding the prior D as fixed, when the cost C^A of acquiring additional information is sufficiently high, for agent with type θ^*, she has no incentive to acquire additional information given the prior. Note that this can be achieved by scaling any cost function up by a sufficiently large constant. Thus posting a deterministic price for revealing full information is also optimal when the information acquisition is sufficiently costly. This is a generalization for the case where cost of information is infinite for any experiment except σ^N.

So far we have shown the optimality of revealing full information with conditions on the prior or the cost of acquiring additional information. Without any such assumptions, the optimal mechanism may contain a continuum of menus, which discriminate different types of the agent by offering experiments with increasing level of informativeness. However, we show that the additional benefit of price discrimination is limited, as posting a deterministic price for revealing full information is approximately optimal for revenue maximization given the same set of assumptions as in Theorem 4.2.5. The proof of Theorem 4.3.2 is provided in Appendix C.1.

Theorem 4.3.2. *For linear valuations, under Assumption 2 and 3, for any prior D and any cost function C^A, posting a deterministic price for revealing full information achieves at least half of the optimal revenue.*

Theorem 4.3.2 is shown by identifying the worst case type distribution and cost function that maximize the multiplicative gap between the optimal revenue and the revenue from posted pricing, and then directly proving that the gap in the worst case is 2. In comparison, Bergemann et al. [2021] showed that even when the valuation distribution is a singleton, if the informative signal the agent receives is exogenous instead of endogenous, for any constant $c > 1$ and $m > 1$, there exists a valuation function of the agent and a signal structure such that any mechanism that is a c-approximation ratio to the optimal revenue has menu complexity at least m. This distinction shows that to approximate the optimal revenue, complex mechanisms are necessary for the exogenous information setting, while simple mechanisms are sufficient for the endogenous information setting.

In Proposition 4.2.3 we have shown that the ability of acquiring additional information distorts the incentives of the agent, and reduces the optimal revenue of the data broker. In the following proposition, we discuss the implication of endogenous information acquisition on the social welfare and the expected utility of the agent. Recall that $|\hat{\Sigma}| = 1$ is equivalent to the setting where the agent cannot acquire additional information.

Proposition 4.3.3. *For any valuation function V, any prior D, and any type distribution F, in the revenue optimal mechanism,*

- *the social welfare is minimized when $|\hat{\Sigma}| = 1$;*[16]
- *the utility of the agent is minimized when $|\hat{\Sigma}| = 1$.*

The proof of statement 1 in Proposition 4.3.3 is provided in Appendix C.1, and the second statement is implied by the first statement and Proposition 4.2.3. Intuitively, when the agent cannot acquire additional information, in the optimal mechanism, the data broker will not provide information to lower types to reduce the information rent from higher types, which minimizes the social welfare.

[16]Our result actually implies that the expected value for each type of the agent is minimized when $|\hat{\Sigma}| = 1$.

4.4. Applications

In this section, we apply the characterizations of the optimal mechanisms to several leading examples of selling information.

Error Minimization. Here we consider the model where the agent is a decision maker trying to minimize the square error of the chosen action. That is, let the state space and action space be $\Omega = A \subseteq \mathbb{R}$, and the agent minimizes the square error between the chosen action and the true state, i.e., the utility of the agent is $u(a, \omega; \theta) = -\theta \cdot (a - \omega)^2$. This is one of the models illustrated in Example 4.1.3.

Recall that in Example 4.1.3, we show that the valuation function of the agent is linear with the form $V(G, \theta) = \theta \cdot v(G)$, where $v(G) = -\mathrm{Var}(G)$ is the variance of distribution G. Let F be the distribution over the types and let θ^* be the monopoly type in distribution F. We assume that the prior distribution D over states is a Gaussian distribution $\mathcal{N}(0, \eta^2)$ with variance η^2. The agent can repeatedly pay a unit cost c to observe a Gaussian signal $s = \omega + \epsilon$ where $\epsilon \sim \mathcal{N}(0, 1)$.

Next we illustrate the optimal mechanism in this setting by applying Theorem 4.2.5.

- For any $\theta \geq \theta^*$, the data broker reveals the states to the firm with price $p = \theta^* \cdot \mathrm{Var}(D)$.
- For any $\theta < \theta^*$, the optimal allocation solves the following Bayesian persuasion problem

$$\hat{x}(\theta) = \arg\max_{\hat{\sigma} \in \hat{\Sigma}} \mathbf{E}_{G \sim \hat{\sigma}|D}[V(G, \theta)] - C^A(\hat{\sigma}, D).$$

Note that in this example, the agent can only decide the number of Gaussian signals to observe, and with k signals, the cost is k, and the variance of the posterior is $\frac{\eta^2}{1+k\eta^2}$ regardless of the realized sequence of the observed signals. Thus, letting

$$k_\theta = \text{argmax}_k -\theta \cdot \frac{\eta^2}{1+k\eta^2} - kc,$$

in the optimal mechanism, the data broker commits to a signal structure that is Blackwell equivalent to revealing k_θ Gaussian signals with unit variance, and charges the agent with price $k_\theta \cdot c$.

Note that for any $\theta < \theta^*$, the optimal number of signals revealed to the agent k_θ is weakly increasing in θ and η, and is weakly decreasing in c. Moreover, fixing distribution F and correspondingly the monopoly type θ^*, when η is sufficiently small or when c is sufficiently large, $k_\theta = 0$ for any $\theta < \theta^*$, and the optimal mechanism reduces to posted pricing mechanism.

Monopoly auction. Here we consider the monopoly auction model introduced in Mussa and Rosen [1978]. This is introduced in Example 4.1.4. We consider a simple case that the state space $\Omega = \{\omega_1, \omega_2\}$ is binary, where $\omega_i \in \mathbb{R}$ represents the value of the consumer and $0 < \omega_1 < \omega_2$. In this case, given posterior belief G of the firm, the virtual value of the consumer simplifies to

$$\phi_G(\omega_1) = \omega_1 - \frac{G(\omega_2)(\omega_2 - \omega_1)}{G(\omega_1)};$$

$$\phi_G(\omega_2) = \omega_2.$$

According to Mussa and Rosen [1978], the optimal mechanism of the firm with cost c is to provide the product with quality $q(\omega_i) = \frac{\max\{0, \phi_G(\omega_i)\}}{2c}$ to the agent with value ω_i, and the expected profit of the firm is $\frac{1}{c} \cdot v(G)$ where

$$v(G) \triangleq G(\omega_1) \cdot \frac{\max\{0, \phi_G(\omega_1)\}^2}{4} + G(\omega_2) \cdot \frac{\omega_2^2}{4}.$$

Suppose the cost c is the private information of the firm, and let $\theta = 1/c$. Recall that F is the distribution over the types and D is the prior over states. Let θ^* be the monopoly type in distribution F. We assume that the firm can flexibly design any experiment, i.e., $\hat{\Sigma}$ contains all possible experiments. In addition, for any $\hat{\sigma} \in \hat{\Sigma}$, the cost is the reduction in entropy, i.e., $C^A(\hat{\sigma}, G) = H(G) - \mathbf{E}_{\hat{G} \sim \hat{\sigma}|G}\left[H(\hat{G})\right]$ for any posterior G where $H(G) = -\sum_i G(\omega_i) \log G(\omega_i)$ is the entropy function.

Since the valuation function of the firm is linear, next we illustrate the optimal mechanism in this setting by applying Theorem 4.2.5.

- For any $\theta \geq \theta^*$, or equivalently for any $c \leq 1/\theta^*$, the data broker reveals full information to the firm with price

$$p = \theta^* \cdot (\mathbf{E}_{\omega \sim D}[v(G_\omega)] - v(D)),$$

where G_ω is the pointmass distribution on state ω.

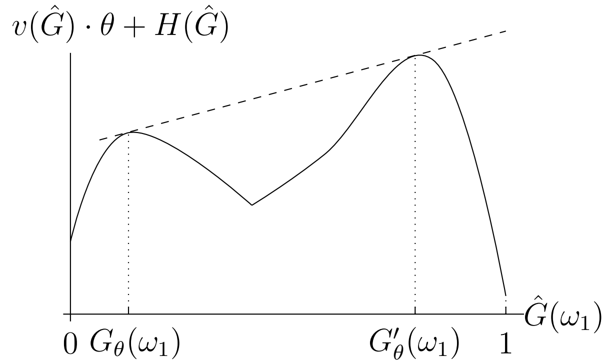

Figure 4.1. The figure is the value of $v(\hat{G})\cdot\theta+H(\hat{G})$ as a function of $\hat{G}(\omega_1)$.

- For any $\theta < \theta^*$, or equivalently for any $c > 1/\theta^*$, the optimal allocation solves the following Bayesian persuasion problem

$$\begin{aligned}\hat{x}(\theta) &= \arg\max_{\hat{\sigma}\in\hat{\Sigma}} \mathbf{E}_{G\sim\hat{\sigma}|D}[V(G,\theta)] - C^A(\hat{\sigma}, D) \\ &= \arg\max_{\hat{\sigma}\in\hat{\Sigma}} \mathbf{E}_{G\sim\hat{\sigma}|D}[\theta \cdot v(G) + H(G)] - H(D).\end{aligned}$$

  By the concavification approach in Kamenica and Gentzkow [2011], the optimal signal structure has signal space of size 2. As illustrated in Fig. 4.1, if the prior satisfies $G_\theta(\omega_1) < D(\omega_1) < G'_\theta(\omega_1)$, the data broker induces posterior either G_θ or G'_θ for type θ. Otherwise, the data broker reveals no information to the firm.

References

Jacob D Abernethy and Rafael M Frongillo. A characterization of scoring rules for linear properties. In *Conference on Learning Theory*, pages 27–1, 2012.

Zoë Abrams. Revenue maximization when bidders have budgets. In *Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm*, pages 1074–1082, 2006.

Anat R Admati and Paul Pfleiderer. A monopolistic market for information. *Journal of Economic Theory*, 39(2):400–438, 1986.

Anat R Admati and Paul Pfleiderer. Direct and indirect sale of information. *Econometrica: Journal of the Econometric Society*, pages 901–928, 1990.

Mohammad Akbarpour, Scott Duke Kominers, Shengwu Li, and Paul R Milgrom. Investment incentives in near-optimal mechanisms. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 26–26, 2021.

Saeed Alaei, Hu Fu, Nima Haghpanah, Jason Hartline, and Azarakhsh Malekian. Bayesian optimal auctions via multi-to single-agent reduction. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, page 17, 2012.

Saeed Alaei, Hu Fu, Nima Haghpanah, and Jason Hartline. The simple economics of approximately optimal auctions. In *Proc. 54th IEEE Symp. on Foundations of Computer Science*, pages 628–637. IEEE, 2013.

Saeed Alaei, Jason Hartline, Rad Niazadeh, Emmanouil Pountourakis, and Yang Yuan. Optimal auctions vs. anonymous pricing. *Games and Economic Behavior*, 2018.

Mark Armstrong. Multiproduct Nonlinear Pricing. *Econometrica*, 64(1):51–75, 1996.

Brian Baisa. Auction design without quasilinear preferences. *Theoretical Economics*, 12 (1):53–78, 2017.

Richard E Barlow and Albert W Marshall. Tables of bounds for distributions with monotone hazard rate. *Journal of the American Statistical Association*, 60(311):872–890, 1965.

Xiaohui Bei, Nick Gravin, Pinyan Lu, and Zhihao Gavin Tang. Correlation-robust analysis of single item auction. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 193–208. SIAM, 2019.

Dirk Bergemann, Alessandro Bonatti, and Alex Smolin. The design and price of information. *American economic review*, 108(1):1–48, 2018.

Dirk Bergemann, Yang Cai, Grigoris Velegkas, and Mingfei Zhao. Is selling complete information (approximately) optimal? *working paper*, 2021.

Alexander W Bloedel and Weijie Zhong. The cost of optimally-acquired information. Technical report, Technical report, Working paper, Stanford University, 2020.

Kim C Border. Implementation of reduced form auctions: A geometric approach. *Econometrica: Journal of the Econometric Society*, pages 1175–1187, 1991.

Alexei Boulatov and Sergei Severinov. Optimal mechanism with budget constrained buyers. Technical report, Mimeo, October, 2018.

J. Bulow and J. Roberts. The simple economics of optimal auctions. *The Journal of Political Economy*, 97:1060–90, 1989.

Yang Cai, Constantinos Daskalakis, and Christos Papadimitriou. Optimum statistical estimation with strategic data sources. In *Conference on Learning Theory*, pages 280–296, 2015.

Yang Cai, Nikhil R Devanur, and S Matthew Weinberg. A duality-based unified approach to bayesian mechanism design. *ACM SIGecom Exchanges*, 15(1):71–77, 2016.

Andrew Caplin, Mark Dean, and John Leahy. Rational inattention, optimal consideration sets, and stochastic choice. *The Review of Economic Studies*, 86(3):1061–1094, 2019.

Gabriel Carroll. Robustness and separation in multidimensional screening. *Econometrica*, 85(2):453–488, 2017.

Gorkem Celik and Okan Yilankaya. Optimal auctions with simultaneous and costly participation. *The BE Journal of Theoretical Economics*, 9(1), 2009.

Shuchi Chawla, Jason D Hartline, David L Malec, and Balasubramanian Sivan. Multi-parameter mechanism design and sequential posted pricing. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 311–320. ACM, 2010.

Shuchi Chawla, David L Malec, and Azarakhsh Malekian. Bayesian mechanism design for budget-constrained agents. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 253–262. ACM, 2011.

Yeon-Koo Che and Ian Gale. The optimal mechanism for selling to a budget-constrained buyer. *Journal of Economic theory*, 92(2):198–233, 2000.

Yeon-Koo Che and Ian Gale. Revenue comparisons for auctions when bidders have arbitrary types. *Theoretical Economics*, 1(1):95–118, 2006.

Yeon-Koo Che, Jinwoo Kim, and Konrad Mierendorff. Generalized reduced-form auctions: A network-flow approach. *Econometrica*, 81(6):2487–2520, 2013.

Yiling Chen and Fang-Yi Yu. Optimal scoring rule design. *arXiv preprint arXiv:2107.07420*, 2021.

Yiling Chen and Shuran Zheng. Prior-free data acquisition for accurate statistical estimation. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 659–677, 2019.

Yiling Chen, Nicole Immorlica, Brendan Lucier, Vasilis Syrgkanis, and Juba Ziani. Optimal data acquisition for statistical estimation. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 27–44, 2018.

Yu Cheng, Nick Gravin, Kamesh Munagala, and Kangning Wang. A simple mechanism for a budget-constrained buyer. In *International Conference on Web and Internet Economics*, pages 96–110. Springer, 2018.

Jacques Crémer and Fahad Khalil. Gathering information before signing a contract. *The American Economic Review*, pages 566–578, 1992.

Nikhil R Devanur and S Matthew Weinberg. The optimal mechanism for selling to a budget constrained buyer: The general case. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 39–40. ACM, 2017a.

Nikhil R Devanur and S Matthew Weinberg. The optimal mechanism for selling to a budget constrained buyer: The general case. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 39–40, 2017b.

Peerapong Dhangwatnotai, Tim Roughgarden, and Qiqi Yan. Revenue maximization with a single sample. *Games and Economic Behavior*, 91:318–333, 2015.

Ky Fan and Georg Gunther Lorentz. An integral inequality. *The American Mathematical Monthly*, 61(9):626–631, 1954.

Moran Feldman, Ola Svensson, and Rico Zenklusen. Online contention resolution schemes. In *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms*, pages 1014–1033. SIAM, 2016.

Yiding Feng and Jason D Hartline. An end-to-end argument in mechanism design (prior-independent auctions for budgeted agents). In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 404–415. IEEE, 2018.

Yiding Feng, Jason D Hartline, and Yingkai Li. Optimal auctions vs. anonymous pricing: Beyond linear utility. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 885–886, 2019.

Yiding Feng, Jason Hartline, and Yingkai Li. Simple mechanisms for non-linear agents. *arXiv preprint arXiv:2003.00545*, 2020.

Rafael Frongillo and Jens Witkowski. A geometric perspective on minimal peer prediction. *ACM Transactions on Economics and Computation*, 5(3):1–27, 2017.

Hu Fu, Jason Hartline, and Darrell Hoy. Prior-independent auctions for risk-averse agents. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 471–488, 2013.

George Georgiadis and Balazs Szentes. Optimal monitoring design. *Econometrica*, 88(5): 2075–2107, 2020.

Alex Gershkov, Benny Moldovanu, Philipp Strack, and Mengxi Zhang. A theory of auctions with endogenous valuations. *Journal of Political Economy*, 129(4):1011–1051, 2021a.

Alex Gershkov, Benny Moldovanu, Philipp Strack, and Mengxi Zhang. Optimal auctions: Non-expected utility and constant risk aversion. *Accepted at Review of Economic Studies*, 2021b.

Nima Haghpanah and Jason Hartline. Reverse mechanism design. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 757–758. ACM, 2015.

Jason Hartline, Yingkai Li, Liren Shan, and Yifan Wu. Optimal scoring rules for multi-dimensional effort. *working paper*, 2021a.

Jason Hartline, Yingkai Li, Liren Shan, and Yifan Wu. Optimal scoring rule with adverse selection. *working paper*, 2021b.

Jason D Hartline. Approximation in mechanism design. *American Economic Review*, 102 (3):330–36, 2012.

Jason D Hartline, Yingkai Li, Liren Shan, and Yifan Wu. Optimization of scoring rules. *arXiv preprint arXiv:2007.02905*, 2020.

Charles A Holt Jr. Competitive bidding for contracts under alternative auction procedures. *Journal of Political Economy*, 88(3):433–445, 1980.

Yaonan Jin, Pinyan Lu, Qi Qi, Zhihao Gavin Tang, and Tao Xiao. Tight approximation ratio of anonymous pricing. In *Proceedings of the 51th Annual ACM Symposium on Theory of Computing, to appear*, 2019.

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

Ian King, Linda Welling, and R Preston McAfee. Investment decisions under first and second price auctions. *Economics Letters*, 39(3):289–293, 1992.

Yuqing Kong. Optimizing multi-task peer prediction. *arXiv preprint arXiv:2103.02214*, 2021.

Jean-Jacques Laffont and Jacques Robert. Optimal auction with financially constrained buyers. *Economics Letters*, 52(2):181–186, 1996.

Nicolas S Lambert. Elicitation and evaluation of statistical forecasts. *working paper*, 2011.

Yingkai Li. Selling data to an agent with endogenous information. *arXiv preprint arXiv:2103.05788*, 2021.

Shuze Liu, Weiran Shen, and Haifeng Xu. Optimal pricing of information. *Proceedings of the 22nd ACM Conference on Economics and Computation*, 2021.

Brendan Lucier. An economic view of prophet inequalities. *ACM SIGecom Exchanges*, 16(1):24–47, 2017.

Eric Maskin. Auctions, development, and privatization: Efficient auctions with liquidity-constrained buyers. *European Economic Review*, 44(4–6):667–681, 2000.

Eric Maskin and John Riley. Optimal auctions with risk averse buyers. *Econometrica*, pages 1473–1518, 1984.

Steven A Matthews. Selling to risk averse buyers with unobservable tastes. *Journal of Economic Theory*, 30(2):370–400, 1983.

John McCarthy. Measures of the value of information. *Proceedings of the National Academy of Sciences of the United States of America*, 42(9):654, 1956.

Flavio M Menezes and Paulo K Monteiro. Auctions with endogenous participation. *Review of Economic Design*, 5(1):71–89, 2000.

Jeffrey Mensch. Screening inattentive buyers. *working paper*, 2021.

Paul Milgrom and Ilya Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601, 2002.

Oskar Morgenstern and John von Neumann. *Theory of games and economic behavior*. Princeton university press, 1953.

Michael Mussa and Sherwin Rosen. Monopoly and product quality. *Journal of Economic theory*, 18(2):301–317, 1978.

R. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6:58–73, 1981a.

Roger B Myerson. Optimal auction design. *Mathematics of operations research*, 6(1): 58–73, 1981b.

Eric Neyman, Georgy Noarov, and S Matthew Weinberg. Binary scoring rules that incentivize precision. *Proceedings of the 22nd ACM Conference on Economics and Computation*, 2021.

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

Kent Osband and Stefan Reichelstein. Information-eliciting compensation schemes. *Journal of Public Economics*, 27(1):107–115, 1985.

Mallesh M Pai and Rakesh Vohra. Optimal auctions with financially constrained buyers. *Journal of Economic Theory*, 150:383–425, 2014.

Michael Richter. Mechanism design with budget constraints and a population of agents. *Games and Economic Behavior*, 115:30–47, 2019.

John Riley and Richard Zeckhauser. Optimal selling strategies: When to haggle, when to hold firm. *The Quarterly Journal of Economics*, pages 267–289, 1983.

Tim Roughgarden, Vasilis Syrgkanis, and Eva Tardos. The price of anarchy in auctions. *Journal of Artificial Intelligence Research*, 59:59–101, 2017.

Xianwen Shi. Optimal auctions with information acquisition. *Games and Economic Behavior*, 74(2):666–686, 2012.

Christopher A Sims. Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690, 2003.

Ludvig Sinander. The converse envelope theorem. *arXiv preprint arXiv:1909.11219*, 2019.

Alex Smolin. Disclosure and pricing of attributes. *Available at SSRN 3318957*, 2020.

Guofu Tan. Entry and R&D in procurement contracting. *Journal of Economic Theory*, 58(1):41–60, 1992.

Juuso Toikka. Ironing without control. *Journal of Economic Theory*, 146(6):2510–2526, 2011.

Q. Yan. Mechanism design via correlation gap. In *Proc. 22nd ACM Symp. on Discrete Algorithms*, pages 710–719. SIAM, 2011.

Kai Hao Yang. Selling consumer data for profit: Optimal market-segmentation design and its consequences. *working paper*, 2020.

Luis Zermeno. A principal-expert model and the value of menus. *Job market paper, Massachusetts Institute of Technology*, 2011.

Weijie Zhong. Optimal dynamic information acquisition. *Available at SSRN 3024275*, 2017.

APPENDIX A

Appendix to Chapter 2

A.1. Other Posted Pricing Mechanisms

A.1.1. Oblivious Posted Pricing

For oblivious posted pricing mechanisms [e.g. Chawla et al., 2010], we show how to apply resemblant property between the ironed price-posting payoff curve and optimal payoff curve to obtain approximation results for agents with general utility. Similar to sequential posted pricing, we will define the oblivious posted price in quantile space.

Definition A.1.1. *An* oblivious posted pricing mechanism *is* $(\{q_i\}_{i\in N})$ *where the adversary chooses an ordering* $\{o_i\}_{i\in N}$ *of the agents, and* $\{q_i\}_{i\in N}$ *denotes the quantile corresponding to the per-unit prices to be offered to agents at the time they are considered according to the order* $\{o_i\}_{i\in N}$ *if the item is not sold to previous agents. Note that quantiles* $\{q_i\}_{i\in N}$ *can be dynamic and depends on both the order and realization of the past agents.*

Given the definition of the oblivious quantile pricing mechanism, we denote the payoff of the oblivious quantile pricing mechanism $(\{q_i\}_{i\in N})$ for agents with a collection of price-posting payoff curves $\{P_i\}_{i\in N}$ by $\text{OPP}(\{P_i\}_{i\in N}, \{q_i\}_{i\in N})$, and the optimal payoff for the oblivious quantile pricing mechanism is

$$\text{OPP}(\{P_i\}_{i\in N}) = \max_{\{q_i\}_{i\in N}} \text{OPP}(\{P_i\}_{i\in N}, \{q_i\}_{i\in N}).$$

Similar to Theorem 2.2.2, we have the following reduction framework for oblivious posted pricing for non-linear agents. The proof is identical to Theorem 2.2.2, hence omitted here.

Theorem A.1.1. *Fix any set of (non-linear) agents with price-posting payoff curves* $\{P_i\}_{i\in N}$ *that are* ζ*-resemblant to their optimal payoff curves* $\{R_i\}_{i\in N}$*. If there exists an oblivious posted pricing mechanism* $(\{q_i\}_{i\in N})$ *that is a* γ*-approximation to the ex ante relaxation for linear agents analog with price-posting payoff curves* $\{P_i\}_{i\in N}$*, i.e.,* $\mathrm{OPP}(\{P_i\}_{i\in N}, \{q_i\}_{i\in N}) \geq {}^1/_\gamma \cdot \mathrm{EAR}(\{\bar{P}_i\}_{i\in N})$*, then this mechanism is also a* $\gamma\zeta$*-approximation to the ex ante relaxation for non-linear agents, i.e.,* $\mathrm{OPP}(\{P_i\}_{i\in N}, \{q_i\}_{i\in N}) \geq {}^1/_{\gamma\zeta} \cdot \mathrm{EAR}(\{R_i\}_{i\in N})$.

For the single item setting, there exists an oblivious posted pricing mechanism that is a 2-approximation to the ex ante relaxation for linear agents [Feldman et al., 2016]. In addition, if the price-posting payoff curves are the same for all gents, the approximation ratio is improved to ${}^1/_{(1 - {}^1/_{\sqrt{2\pi}})}$ [Yan, 2011].

A.1.2. Anonymous Pricing

A desirable property for the multi-agent setting is anonymity. This requires that the price posted to all agents are the same. Note that for welfare maximization, although anonymous pricing guarantees 2-approximation for linear agents [Lucier, 2017], it may lead to huge welfare loss for non-linear agents. This is illustrated in the following example.

Example A.1.2. *Consider the single-item setting with two budget agents. Let* v *be a sufficiently large number. Agent* 1 *has value* v *and no budget constraint while agent* 2 *has* *value v^2 and budget 1. The welfare optimal mechanism allocates the item to agent 2, with welfare v^2. However, if the anonymous price is at most v, then agent 1 will buy the whole item and if the anonymous price is larger than v, the item is sold with probability at most $\frac{1}{v}$. Thus anonymous pricing can guarantee welfare at most v, with approximation factor at least v, which is unbounded.*

Thus we will focus on revenue maximization for anonymous pricing. Alaei et al. [2018] showed that for linear agents, the central assumption for constant approximation of anonymous pricing is concavity of the price posting revenue curves. Next we provide a general reduction framework for anonymous pricing for non-linear agents. Note that $\mathrm{AP}(\{P_i\}_{i\in N})$ is the optimal revenue from anonymous pricing when the price-posting payoff curves are $\{P_i\}_{i\in N}$.

Theorem A.1.2. *Fix any set of (non-linear) agents with price-posting payoff curves $\{P_i\}_{i\in N}$ that are ζ-resemblant to their optimal payoff curves $\{R_i\}_{i\in N}$. If the price-posting payoff curves are concave, then anonymous pricing is a ζe-approximation to the ex ante relaxation on the optimal payoff curves, i.e.,* $\mathrm{AP}(\{P_i\}_{i\in N}) \geq {}^1/_{\zeta e} \cdot \mathrm{EAR}(\{R_i\}_{i\in N})$.

Proof. Let $\{q_i\}_{i\in N}$ be the optimal ex ante relaxation for ex ante revenue curves $\{R_i\}_{i\in N}$, and let $q_i^\dagger$ be the quantile assumed to exist by ζ-resemblance such that $q_i^\dagger \leq q_i$ and $\bar{P}_i(q_i^\dagger) \geq \frac{1}{\zeta} R_i(q_i)$ for each i. Since the price-posting payoff curves are concave, we have $\{P_i\}_{i\in N} = \{\bar{P}_i\}_{i\in N}$, and

$$\mathrm{EAR}(\{P_i\}_{i\in N}) = \mathrm{EAR}(\{\bar{P}_i\}_{i\in N}) \geq \sum\nolimits_i \bar{P}_i(q_i^\dagger) \geq \frac{1}{\zeta} \sum\nolimits_i R(q_i) = \frac{1}{\zeta} \mathrm{EAR}(\{R_i\}_{i\in N}).$$

By Alaei et al. [2018], $e \cdot \mathrm{AP}(\{P_i\}_{i\in N}) \geq \mathrm{EAR}(\{P_i\}_{i\in N})$ if the price-posting payoff curves $\{P_i\}_{i\in N}$ are concave. Combining the inequalities, we have

$$\zeta e \cdot \mathrm{AP}(\{P_i\}_{i\in N}) \geq \mathrm{EAR}(\{R_i\}_{i\in N}). \qquad \square$$

As instantiation of the reduction framework in Theorem A.1.2, we can show that agents are 1-resemblant and have concave price posting revenue curve when they have public budget and regular valuation distributions, and they are $(2 + \ln {}^{\bar{v}}/_{C})$-resemblant and have concave price posting revenue curve when they have capacitated risk averse utility with maximum value $\bar{v}$, capacity $C \leq \bar{v}$, and regular valuation distributions.

A.2. Public Budget Agent

Theorem 2.3.5. *An agent with public budget and regular valuation distribution has the ironed price-posting revenue curve $\bar{P}$ that equals to (i.e. 1-resemblant) her optimal revenue curve R.*

Proof. For an agent with public budget w, the $\hat{q}$ ex ante optimal mechanism is the solution of the following program,

$$\begin{aligned} \max_{(x,p)} \quad & \mathbf{E}_v[p(v)] \\ s.t. \quad & (x,p) \text{ are IC, IR}, \\ & \mathbf{E}_v[x(v)] = \hat{q}, \\ & p(\bar{v}) \leq w. \end{aligned} \tag{A.1}$$

where $\bar{v}$ is the highest possible value of the agent. Consider the Lagrangian relaxation of the budget constraint in (A.1),

$$\begin{aligned}
&\min_{\lambda\geq 0}\max_{(x,p)} \quad && \mathbf{E}_v[p(v)] + \lambda w - \lambda p(\bar{v}) \\
\text{(A.2)} \qquad &s.t. && (x,p) \text{ are IC, IR,} \\
& && \mathbf{E}_v[x(v)] = \hat{q}.
\end{aligned}$$

Let λ^* be the optimal solution in program (A.2). If we fix $\lambda = \lambda^*$ in program (A.2), its inner maximization program can be thought as a $\hat{q}$ ex ante optimal mechanism design for a linear agent with Lagrangian objective function $\mathbf{E}_v[p(v)] - \lambda^* p(\bar{v})$. Thus, we define the Lagrangian price-posting revenue curve $P_{\lambda^*}(\cdot)$ where $P_{\lambda^*}(q)$ is the maximum value of the Lagrangian objective $\mathbf{E}_v[p(v)] - \lambda^* p(\bar{v})$ in price-posting mechanism with per-unit price $V(q)$. For any $q \in (0, 1]$, by the definition, $P_{\lambda^*}(q) = qV(q) - \lambda^* V(q)$. For $q = 0$, notice that the agent with $\bar{v}$ is indifferent between purchasing or not purchasing. Thus, by the definition, $P_{\lambda^*}(q) = 0$ if $q = 0$.

Now, we consider the concave hull of the Lagrangian price-posting revenue curve $P_{\lambda^*}(\cdot)$ which we denote as $\hat{P}_{\lambda^*}(\cdot)$. Let $q^\dagger$ be the smallest solution of equation $P_{\lambda^*}(q) = qP'_{\lambda^*}(q)$. Since $P_{\lambda^*}(0) \leq 0$, $P_{\lambda^*}(1) = 0$ and $P_{\lambda^*}(\cdot)$ is continuous, $q^\dagger$ always exists. Then, for any $q \leq q^\dagger$, $\hat{P}_{\lambda^*}(q) = q\,P'_{\lambda^*}(q^\dagger)$. For any $q \geq q^\dagger$, we show $\hat{P}_{\lambda^*}(q) = P_{\lambda^*}(q)$ by the following arguments. First notice that $P_{\lambda^*}(q^\dagger) \geq 0$, and hence $q^\dagger \geq \lambda^*$. Consider $P''_{\lambda^*}(q) = V''(q)(q - \lambda^*) + 2V'(q)$. Clearly, $V'(q) \leq 0$. If $V''(q) \leq 0$, then $P''_{\lambda^*}(q) \leq 0$. If $V''(q) > 0$, then $P''_{\lambda^*}(q) = V''(q)(q - \lambda^*) + 2V'(q) \leq qV''(q) + 2V'(q) \leq 0$, where $qV''(q) + 2V'(q)$ is non-positive due to the regularity of the valuation distribution.

To summarize, $\hat{P}_{\lambda^*}(\cdot)$, the concave hull of the Lagrangian price-posting revenue curve satisfies

$$\hat{P}_{\lambda^*}(q) = \begin{cases} q\, P'_{\lambda^*}(q^\dagger) & \text{if } q \in [0, q^\dagger] \\ P_{\lambda^*}(q) & \text{if } q \in [q^\dagger, 1] \end{cases}$$

Therefore, use the similar ironing technique based on the revenue curves for linear agents with irregular valuation distribution [e.g. Myerson, 1981a, Bulow and Roberts, 1989, Alaei et al., 2013], Lemma A.2.1 (stated below) suggests that the $\hat{q}$ ex ante optimal mechanism irons quantiles between $[0, q^\dagger]$ under $\hat{q}$ ex ante constraint, which is still a posted-pricing mechanism. □

Lemma A.2.1 (Alaei et al., 2013)**.** *For incentive compatible and individual rational mechanism $(x(\cdot), p(\cdot))$ and an agent with any Lagrangian price-posting revenue curve $P_{\lambda^*}(q)$, the expected Lagrangian objective of the agent is upper-bounded by her expected marginal Lagrangian objective of the same allocation rule, i.e.,*

$$\mathbf{E}_v[p(v)] + \lambda^* p(\bar{v}) \leq \mathbf{E}_q\left[\hat{P}'_{\lambda^*}(q) \cdot x(V(q))\right].$$

Furthermore, this inequality holds with equality if the allocation rule $x(\cdot)$ is constant all intervals of values $V(q)$ where $\hat{P}_{\lambda^*}(q) > P_{\lambda^*}(q)$.

APPENDIX B

Appendix to Chapter 3

B.1. Canonical Scoring Rules

In this section, we will formally prove Theorem 3.2.4. In the subsequent discussion, the boundary of the report space is denoted by ∂R and the interior of the report space by $\text{relint}(R) = R \setminus \partial R$.

Lemma B.1.1 (Abernethy and Frongillo, 2012)**.** *Any proper and μ-differentiable scoring rule for eliciting the mean S coincides with a canonical scoring rule (defined by u, ξ, and κ) at reports in the relative interior of the report space, i.e., it satisfies equation* (3.2) *for all $r \in \text{relint}(R)$.*

The main new results need to show that canonical scoring rules are without loss for Program (3.1) are extensions of Lemma B.1.1 to the boundary of the report space ∂R. The form of scoring rules considered enters the program in two places: the objective and the boundedness constraint. The two lemmas below show that canonical scoring rules are without loss in these two places in the program.

Lemma B.1.2. *Any μ-differentiable, bounded, and proper scoring rule S for eliciting the mean is equal in expectation of truthful reports to a canonical scoring rule (defined by u, ξ, and κ), i.e., it satisfies equation* (3.3).

Lemma B.1.3. *For any μ-differentiable and proper scoring rule S for eliciting the mean that induces utility function u (via Lemma B.1.2) and satisfies score bounded in $[0, B]$, there is a canonical scoring rule defined by u (and some ξ and κ) that satisfies the same score bound, i.e., it satisfies equation* (3.4).

Note that Lemma B.1.2 implies that the utility function u corresponding to any μ-differentiable scoring rule S can be identified (via the equivalent cannonical scoring rule); thus, the assumption of Lemma B.1.3 is well defined. Lemma B.1.2 and Lemma B.1.3 combine to imply that Program (3.1) and Program (3.6) are equivalent.

Next, we will formally prove Lemma B.1.2 and B.1.3. First we show that when the scoring rule is bounded, the corresponding functions $u(r), \xi(r), \kappa(\omega)$ in the characterization of Lemma B.1.1 are bounded in the interior as well.

Lemma B.1.4. *For any bounded scoring rule S, there exist convex function $u : R \to \mathbb{R}$ and function $\kappa : \Omega \to \mathbb{R}$ such that for any report $r \in \text{relint}(R)$ and any state $\omega \in \Omega$,*

$$S(r, \omega) = u(r) + \xi(r) \cdot (\omega - r) + \kappa(\omega)$$

where $\xi(r) \in \partial u(r)$ is a subgradient of u, and functions $u(r), \xi(r), \kappa(\omega)$ are bounded for any report $r \in \text{relint}(R)$ and any state $\omega \in \Omega$.

Proof. Since scoring rule S is bounded, let $\bar{B}_\omega = \sup_{r \in \text{relint}(R)} S(r, \omega)$ and $\underline{B}_\omega = \inf_{r \in \text{relint}(R)} S(r, \omega)$. Let $\hat{r} \in \text{relint}(R)$ be a report in the interior such that both $u(\hat{r})$ and $\xi(\hat{r})$ are finite. Note that for any state $\omega \in \Omega$, state ω locate on the boundary of the report space, i.e., $\omega \in \partial R$, and the report space is a linear combination of the state space.

For any report $r \in \text{relint}(R)$, by the convexity of function u, we have

$$u(r) \geq u(\hat{r}) - \xi(\hat{r}) \cdot (r - \hat{r})$$

and hence $u(r)$ is bounded below.

Next we show that $u(r)$ is bounded above for any report $r \in \text{relint}(R)$. We first show that fixing any state ω, any report r which is a linear combination of ω and $\hat{r}$ has bounded utility $u(r)$. If $u(r) \leq u(\hat{r})$, then naturally $u(r)$ is bounded above. Otherwise, note that

$$\begin{aligned}
&\bar{B}_\omega - \underline{B}_\omega \geq S(r, \omega) - S(\hat{r}, \omega) = u(r) + \xi(r) \cdot (\omega - r) - u(\hat{r}) - \xi(\hat{r}) \cdot (\omega - \hat{r}) \\
&\geq (u(r) - u(\hat{r})) \cdot \frac{\|\omega - \hat{r}\|}{\|\hat{r} - r\|} + u(\hat{r}) - u(\hat{r}) - \xi(\hat{r}) \cdot (\omega - \hat{r}) \geq u(r) - u(\hat{r}) - \xi(\hat{r}) \cdot (\omega - \hat{r}),
\end{aligned}$$

where the first inequality holds because the scoring rule is bounded. The second inequality holds because the convex function u projected on line $(\omega, \hat{r})$ is still a convex function. The last inequality holds because report r lies in between ω and $\hat{r}$. Therefore, we have that $u(r)$ is bounded above for report r lies in between ω and $\hat{r}$. For any state $\omega \in \Omega$, let $\hat{u}(\omega) = \lim_{k\to\infty} u(r^k)$ where $\{r^k\}_{k=1}^{\infty}$ is a sequence of report on line $(\omega, \hat{r})$ that converges to ω. Since $u(r^k)$ are bounded for any r^k, we have that $\hat{u}(\omega)$ is bounded as well. Since the report space is a subset of the convex hull of the state space, we have that for any report $r \in \text{relint}(R)$, $u(r)$ is upper bounded by the convex combination of $\hat{u}(\omega)$, which is also bounded by above.

For any state $\omega \in \Omega$, we have

$$S(\hat{r}, \omega) = u(\hat{r}) + \xi(\hat{r}) \cdot (\omega - \hat{r}) + \kappa(\omega),$$

which implies $\kappa(\omega)$ is bounded since all other terms are bounded.

Finally, for any report $r \in \text{relint}(R)$ and any state $\omega \in \Omega$,

$$S(r,\omega) = u(r) + \xi(r)\cdot(\omega - r) + \kappa(\omega),$$

which implies $\xi(r)\cdot(\omega - r)$ is bounded. Since the boundedness holds for all directions, the subgradient $\xi(r)$ must also be bounded. □

Lemma B.1.5. *Given any state space Ω and report space R with non-empty interior, for any distribution $G \in \Delta(\Omega)$ with mean μ_G, there exists a sequence of posteriors $\{G^k\}$ such that for any bounded function $\phi(\omega)$ in space Ω, we have $\{\mathbf{E}_{\omega\sim G^k}[\phi(\omega)]\}$ converges to $\mathbf{E}_{\omega\sim G}[\phi(\omega)]$.*

Proof. Since space R has a non-empty interior, let $\widetilde{G}$ be a distribution with mean $\mu_{\widetilde{G}}$ in the interior of R. Let the sequence of posteriors $G^k = (1 - 1/k)\cdot G + 1/k \cdot \widetilde{G}$. For any bounded function $\phi(\omega)$ in space Ω, we have

$$\begin{aligned}&\lim_{k\to\infty} \mathbf{E}_{\omega\sim G^k}[\phi(\omega)]\\ &= \lim_{k\to\infty}[(1 - 1/k)\cdot \mathbf{E}_{\omega\sim G}[\phi(\omega)] + 1/k \cdot \mathbf{E}_{\omega\sim \widetilde{G}}[\phi(\omega)]] \to \mathbf{E}_{\omega\sim G}[\phi(\omega)]\,.\end{aligned}$$ □

Proof of Lemma B.1.2. By Lemma B.1.1, for μ-differentiable proper scoring rule S, there exists convex function $u : R \to \mathbb{R}$ and function $\kappa : \Omega \to \mathbb{R}$ such that for any report $r \in \text{relint}(R)$ and any state $\omega \in \Omega$, we have

$$S(r,\omega) = u(r) + \xi(r)\cdot(\omega - r) + \kappa(\omega)$$

where $\xi(r) \in \nabla u(r)$ is a subgradient of u. By Lemma B.1.4, since the scoring rule is bounded, function u is convex and bounded and hence continuous in the interior. Thus, we can well define the value of u on the boundary as its limit from the interior, i.e., set $u(r) = \lim_{k\to\infty} u(r^k)$ for any r on the boundary of the report space R and $\{r^k\}_{k=1}^{\infty}$ as a sequence of interior reports converging to r. Thus we can replace the convex function u with continuous and convex function u for bounded scoring rules and the characterization still holds in the interior.

For any bounded proper scoring rule, we have that $u(r)$ is bounded for any report $r \in \text{relint}(R)$ and $\kappa(\omega)$ is bounded for any state $\omega \in \Omega$. Given any posterior G such that $\mu_G \in \partial R$, let $\{G^k\}$ be the sequence of posteriors constructed in Lemma B.1.5.

(1) The identity function $\phi(\omega) = \omega$ is bounded. Therefore, the mean of the posteriors converges, i.e., $\lim_{k\to\infty} \mu_{G^k} = \mu_G$. And all means $\{\mu_{G^k}\}$ are in the interior of R.

(2) Function $\kappa(\omega)$ is bounded. Therefore, the expected value for function κ converges. That is, $\lim_{k\to\infty} \mathbf{E}_{\omega\sim G^k}[\kappa(\omega)] = \mathbf{E}_{\omega\sim G}[\kappa(\omega)]$.

(3) The ex post score $S(r,\omega)$ is bounded. Therefore, the expected score for reporting μ_G converges, i.e., $\lim_{k\to\infty} \mathbf{E}_{\omega\sim G^k}[S(\mu_G,\omega)] = \mathbf{E}_{\omega\sim G}[S(\mu_G,\omega)]$.

Moreover, considering the sequence of expected score for reporting μ_{G^k} with distribution G, we have

$$\lim_{k\to\infty} \mathbf{E}_{\omega\sim G}[S(\mu_{G^k},\omega)] = \lim_{k\to\infty}[u(\mu_{G^k}) + \mathbf{E}_{\omega\sim G}[\xi(\mu_{G^k})\cdot(\omega - \mu_{G^k})] + \mathbf{E}_{\omega\sim G}[\kappa(\omega)]]$$

$$= \lim_{k\to\infty}[u(\mu_{G^k}) + \mathbf{E}_{\omega\sim G^k}[\kappa(\omega)]] = \lim_{k\to\infty}[\mathbf{E}_{\omega\sim G^k}[S(\mu_{G^k},\omega)]$$

where the second equality holds because $\lim_{k\to\infty} \mathbf{E}_{\omega\sim G^k}[\kappa(\omega)] = \mathbf{E}_{\omega\sim G}[\kappa(\omega)]$ and $\lim_{k\to\infty} \mu_{G^k} = \mu_G$. Combining the equalities, we have

$$\mathbf{E}_{\omega\sim G}[S(\mu_G,\omega)] = \lim_{k\to\infty} \mathbf{E}_{\omega\sim G^k}[S(\mu_G,\omega)] \leq \lim_{k\to\infty} \mathbf{E}_{\omega\sim G^k}[S(\mu_{G^k},\omega)]$$
$$= \lim_{k\to\infty} \mathbf{E}_{\omega\sim G^k}[S(\mu_{G^k},\omega)] = \lim_{k\to\infty} \mathbf{E}_{\omega\sim G}[S(\mu_{G^k},\omega)] \leq \mathbf{E}_{\omega\sim G}[S(\mu_G,\omega)]$$

where the inequalities holds by the properness of the scoring rule. Therefore, all inequalities must be equalities, and hence

$$\mathbf{E}_{\omega\sim G}[S(\mu_G,\omega)] = \lim_{k\to\infty} \mathbf{E}_{\omega\sim G^k}[S(\mu_{G^k},\omega)]$$
$$= \lim_{k\to\infty} \mathbf{E}_{\omega\sim G^k}[u(\mu_{G^k}) + \kappa(\omega)] = u(\mu_G) + \mathbf{E}_{\omega\sim G}[\kappa(\omega)]\,.$$

where the last equality hold since function u is continuous.

Finally, given any bounded, continuous and convex function u with bounded subgradients and any bounded function κ, the corresponding canonical scoring rule is proper, bounded, and the expected score coincides. □

Proof of Lemma B.1.3. If a proper scoring rule S is induced by function u and bounded by B in space Ω, by Lemma B.1.1, there exists function $\kappa : \Omega \to \mathbb{R}$ such that for any report $r \in \text{relint}(R)$ and any state $\omega \in \Omega$,

$$S(r,\omega) = u(r) + \xi(r)\cdot(\omega - r) + \kappa(\omega)$$

where $\xi(r) \in \nabla u(r)$ is a subgradient of u. Moreover, the score $S(r,\omega) \in [0, B]$ for any report and state $r \in R, \omega \in \Omega$. Thus, it holds that for any report and state $r \in$

relint$(R), \omega \in \Omega$

$$S(\omega,\omega) - S(r,\omega) = u(\omega) - u(r) - \xi(r)(\omega - r) \leq B.$$

For any report $R \in \partial R$, there exists a sequence of reports r_i such that $\{r_k\}$ converges to r and $\xi(r) = \lim_{k\to\infty} \xi(r_k)$ is a subgradient at report r. Thus, it holds that for any report $r \in \partial R$ and state $\omega \in \Omega$,

$$S(\omega,\omega) - S(r,\omega) = u(\omega) - u(r) - \lim_{k\to\infty} \xi(r_k)(\omega - r) \leq B.$$

Therefore, the canonical scoring rule defined by u with the same function κ is proper and bounded in $[0, B]$. □

B.2. Proofs of Lemma 3.4.7-Lemma 3.4.10

Proof of Lemma 3.4.7. This result follows because the extended distribution is symmetric on the extended state space, thus, its optimal scoring rule is max-over-separate (Corollary 3.4.4). This scoring rule can be applied to the original space where it is still max-over-separate. The optimal max-over-separate scoring rule for the original space is no worse. □

Proof of Lemma 3.4.8. Let $\tilde{u}$ be the optimal utility function corresponding to OPT$(\tilde{f}, B, \widetilde{\Omega})$. Since the distribution $\tilde{f}$ is center symmetric, by Theorem 3.4.2, the utility function $\tilde{u}$ is symmetric V-shaped. Thus, we have

$$\begin{aligned}\mathrm{OPT}(\tilde{f}, B, \widetilde{\Omega}) &= \int_{\tilde{R}} \tilde{u}(r)\,\tilde{f}(r)\,\mathrm{d}r \\ &= \frac{1}{2}\int_{R} \tilde{u}(r)\,f(r)\,\mathrm{d}r + \frac{1}{2}\int_{R} \tilde{u}(2\mu_D - r)\,f(r)\,\mathrm{d}r \\ &= \int_{R} \tilde{u}(r)\,f(r)\,\mathrm{d}r = \mathrm{Obj}(\tilde{u}, f).\end{aligned}$$

□

Proof of Lemma 3.4.9. Let $\hat{u}$ be the optimal solution of Program (3.6) with distribution f and state space $\widetilde{\Omega}$, i.e., $\mathrm{Obj}(\hat{u}, f) = \mathrm{OPT}(f, B, \widetilde{\Omega})$. On the other hand, utility function $\hat{u}$ may not be optimal for distribution $\tilde{f}$, thus, $\mathrm{OPT}(\tilde{f}, B, \widetilde{\Omega}) \geq \mathrm{Obj}(\hat{u}, \tilde{f})$. We have,

$$\begin{aligned}\mathrm{OPT}(\tilde{f}, B, \widetilde{\Omega}) \geq \mathrm{Obj}(\hat{u}, \tilde{f}) &= \int_{\tilde{R}} \hat{u}(r)\,\tilde{f}(r)\ \mathrm{d}r = \frac{1}{2}\int_{R} \tilde{u}(r)\,f(r)\,\mathrm{d}r + \frac{1}{2}\int_{R} \tilde{u}(2\mu_D - r)\,f(r)\,\mathrm{d}r \\ &\geq \frac{1}{2}\int_{R} \tilde{u}(r)\,f(r)\,\mathrm{d}r = \frac{1}{2}\,\mathrm{OPT}(f, B, \widetilde{\Omega})\end{aligned}$$

where the final inequality follows from convexity of $\hat{u}$, $\int_R (2\mu_D - r)\,f(r)\,\mathrm{d}r = \mu_D$, Jensen's Inequality, and $\hat{u}(\mu_D) = 0$. □

The approach to proving Lemma 3.4.10, i.e., $\mathrm{OPT}(f, B, \widetilde{\Omega}) \geq \frac{1}{4}\,\mathrm{OPT}(f, B, \Omega)$, is as follows. Let u be the optimal utility corresponding to $\mathrm{OPT}(f, B, \Omega)$. We construct $\tilde{u}$ that (a) exceeds u at all point $r \in R$ and (b) is feasible for $\mathrm{OPT}(f, 4B, \widetilde{\Omega})$. The utility function $\tilde{u}/4$, thus, has objective value at least $\frac{1}{4}\,\mathrm{OPT}(f, B, \Omega)$ and is feasible for $\mathrm{OPT}(f, B, \widetilde{\Omega})$. The optimal utility is only better.

The proof of the lemma introduces the following constructs.

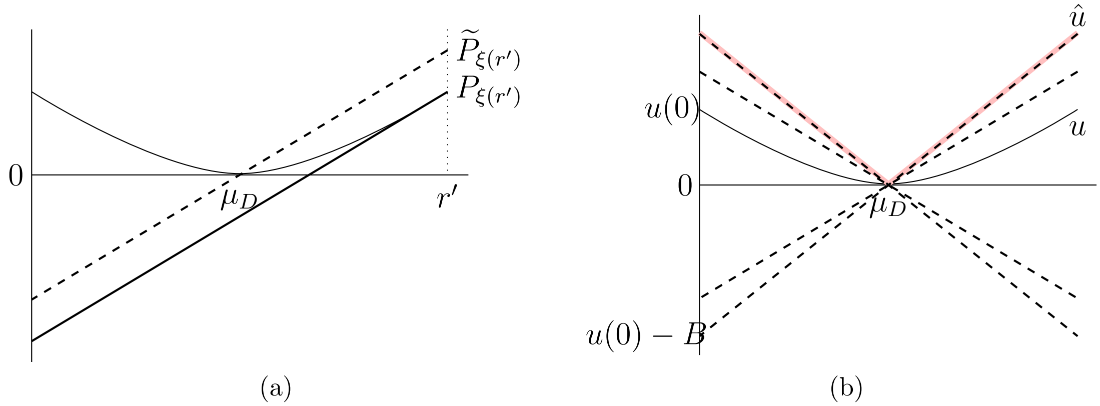

Figure B.1. The figure on the left hand side illustrates a hyperplane for report r' on the boundary of the report space, which is shifted from a tangent plane of u at the boundary r'. The figure on the right hand side illustrates the extended utility function $\tilde{u}$ that takes the supremum over all hyperplanes shifted from the feasible tangent planes to intersect with the $(\mu_D, 0)$ point.

- The *extended utility function* $\tilde{u}$ for program $\mathrm{OPT}(f, 4B, \widetilde{\Omega})$ given utility function u for the program $\mathrm{OPT}(f, B, \Omega)$ is defined as follows.

  Feasibility of u for Program (3.6) defines subgradients $\{\xi(r) : r \in R\}$ that satisfy the boundedness condition. Let $\mathcal{G}_u$ be the set of all subgradients of u that satisfy the boundedness constraint. Clearly the latter set contains the former set. Define the extended utility function $\tilde{u}$ as the convex function defined by the supremum of the supporting hyperplanes given by the subgradients $\mathcal{G}_u$ shifted to intersect with the $(\mu_D, 0)$ point. See Figure B.1.

  Convexity of u implies that its supporting hyperplane at r with subgradient $\xi(r)$ is below $u(\mu_D) = 0$ at μ_D. Thus, relative to the supporting hyperplanes of u these supporting hyperplanes of $\tilde{u}$ are shifted upwards.

The extended utility function $\tilde{u}$ is *convex-conical* as it is defined by supporting hyperplanes that all contain point $(\mu_D, 0)$.

- The *extended state spaces* are $\Omega \subset \widetilde{\Omega}' \subset \widetilde{\Omega}'' \subset \widetilde{\Omega}$. State space $\widetilde{\Omega}'$ is the union of the original state space and its point reflection about μ_D as $\widetilde{\Omega}' = \Omega \cup \{2\mu_D - \omega : \omega \in \Omega\}$, state space $\widetilde{\Omega}''$ is the convex hull of $\widetilde{\Omega}'$, and state space $\widetilde{\Omega}$ (as previously defined) is the extended rectangular state space containing $\widetilde{\Omega}''$.

Lemma 3.4.10, i.e., $\mathrm{OPT}(f, 4B, \widetilde{\Omega}) \geq \mathrm{OPT}(f, B, \Omega)$, follows by combining the following lemmas.

Lemma B.2.1. *For any feasible solution u for Program (3.6), the extended utility function $\tilde{u}$ is at least u, i.e., $\tilde{u}(r) \geq u(r)$ for any report $r \in R$.*

Lemma B.2.2. *For any feasible solution u for Program (3.6) with score bound B and state space Ω, the extended utility function $\tilde{u}$ is a feasible solution of Program (3.6) with score bound $2B$ and state space Ω.*

Lemma B.2.3. *Any convex-conical utility function $\tilde{u}$ that is a feasible solution of Program (3.6) with score bound $2B$ and state space Ω is a feasible solution to Program (3.6) with bound $2B$ and state space $\widetilde{\Omega}'$.*

Lemma B.2.4. *Any convex-conical utility function $\tilde{u}$ that is a feasible solution of Program (3.6) with score bound $2B$ and state space $\widetilde{\Omega}'$ is a feasible solution to Program (3.6) with bound $2B$ and state space $\widetilde{\Omega}'' = \mathrm{conv}(\widetilde{\Omega}')$.*

Lemma B.2.5. *Any convex-conical utility function $\tilde{u}$ that is a feasible solution of Program (3.6) with score bound $2B$ and state space $\widetilde{\Omega}''$ is a feasible solution to Program (3.6) with bound $4B$ and state space $\widetilde{\Omega}$.*

Proof of Lemma B.2.1. Since the supporting hyperplanes of $\tilde{u}$ are shifted upwards relative to u, we have $\tilde{u}(r) \geq u(r)$ at all $r \in R$. Thus, $\tilde{u}$ obtains at least the objective value of u, i.e., $\mathrm{Obj}(f, \tilde{u}) \geq \mathrm{Obj}(f, \tilde{u})$. □

Proof of Lemma B.2.2. First, the subgradients of $\tilde{u}$ are a subset of the subgradients of u that satisfy the boundedness constraint. Lemma B.2.6 (stated and proved at the end of this subsection) shows that the set of subgradients $\mathcal{G}_u$ of u that satisfy the boundedness constraint is closed. As $\tilde{u}$ is defined the supremum over these hyperplanes, closure of the set implies that the supremum at any report $r \in R$ is attained on one of these hyperplanes.

Now observe that in the construction of $\tilde{u}$, the supporting hyperplanes of u are shifted up by at most B. The boundedness constraint corresponding to state μ_D and the report r with subgradient $\xi(r) \in \nabla u(r)$ implies that the supporting hyperplane corresponding to $\xi(r)$ at r has value at least $-B$ at μ_D. Thus, in the construction of the extended utility function $\tilde{u}$, the hyperplane corresponding to $\xi(r)$ is shifted up by at most B and, at any state $\omega \in \Omega$, $\tilde{u}(\omega) \leq u(\omega) + B$.

Finlly, the boundedness constraint is the difference between the utility at a given state and the value of any supporting hyperplane of the utility evaluated at that state. From u to $\tilde{u}$ the former has increased by at most B and the latter is no smaller; thus, $\tilde{u}$ satisfies the boundedness constraint on state space Ω with bound $2B$. □

Proof of Lemma B.2.3. The lemma follows by the geometries of the boundedness constraint and convex cones. The boundedness constraint requires a bounded difference between the utility at any state (in the state space) and the value at that state on any supporting hyperplane of the utility function (corresponding to any report in the report space). For convex-conical utility functions, the supporting hyperplanes are also supporting hyperplanes of the cone defined by the point reflection of the utility function around its vertex $(\mu_D, 0)$, henceforth, the reflected cone. Thus, the boundedness constraint for convex-conical utility function requires that the difference between the original cone and the reflected cone be bounded at all states in the state space.

The original space Ω and the reflected state space $\{2\mu_D - \omega : \omega \in \Omega\}$ are symmetric with respect to the original cone and the reflected cone. Thus, if states in the original state space are bounded, by comparing a state on the cone to the same state on the reflected cone; then states in the reflected state space are bounded by comparing its reflected state (in the original state space) on the reflected cone to its reflected state on the original cone.

Thus, if a boundedness constraint holds on Ω it also holds on the reflected state space $\{2\mu_D - \omega : \omega \in \Omega\}$ and their union. □

Proof of Lemma B.2.4. Consider the cone and reflected cone defined in the proof of Lemma B.2.3 and the geometry of the boundedness constraint. Notice that, by convexity of the cone defining the utility function $\tilde{u}$ and concavity of the reflected cone, the convex combination of the bounds, i.e., the difference of values of states on these two cones, of any set of states is at least the bound of the convex combination of the states. Hence, if the boundedness constraint holds on state space $\widetilde{\Omega}'$, then it holds on its convex hull $\widetilde{\Omega}'' = \mathrm{conv}(\widetilde{\Omega}')$. □

Proof of Lemma B.2.5. Consider any ray from μ_D. Since the utility $\tilde{u}$ is a convex cone, the utility on this ray is a linear function of the distance from μ_D. The same holds for this ray evaluated on the point reflection of the utility at μ_D. The difference between these utilities is also linear. Thus, by the geometry of the boundedness constraint for convex-conical utility functions, on any ray from μ_D, the bound is linear. Considering the state space $\widetilde{\Omega}''$ and $\widetilde{\Omega}$, if the former is scaled by a factor of two around μ_D, then it contains the latter (by simple geometry, see Figure 3.2). Thus, if the convex-conical utility function $\tilde{u}$ satisfies bound $2B$ on state space $\widetilde{\Omega}''$ it satisfies bound $4B$ on state space $\widetilde{\Omega}$. □

Lemma B.2.6. *For any feasible solution u for Program (3.6), the set $\mathcal{G}_u$ of all subgradients of u satisfying the bounded constraints is a closed set.*

Proof. By Lemma B.1.2, any feasible solution u for Program (3.6) is convex, bounded and continuous with bounded subgradients. For any convex, bounded and continuous function u, let $\{\xi^k(r^k)\}_{k=1}^{\infty} \subseteq \mathcal{G}_u$ be a convergent sequence of subgradients in set $\mathcal{G}_u$, where r^k is the report corresponds to the k^{th} subgradient. Let $\xi^* = \lim_{k\to\infty} \xi^k(r^k)$ be the limit of the subgradients. Since the report space is a closed and bounded space, there exists a subsequence of reports $\{r^{k_j}\}_{j=1}^{\infty} \subseteq \{r_k\}_{k=1}^{\infty}$ such that $\{r^{k_j}\}_{j=1}^{\infty}$ converges. Letting report $r = \lim_{j\to\infty} r^{k_j}$, we have report r is in the report space, i.e., $r \in R$. Moreover, we have $\lim_{j\to\infty} \xi^{k_j}(r^{k_j}) = \lim_{k\to\infty} \xi^k(r^k) = \xi^*$. Next we show that ξ^* is a subgradient for some report $r \in R$ such that the bounded constraints of the induced scoring rule are satisfied for any state $\omega \in \Omega$, i.e., $\xi^* \in \mathcal{G}_{u,r}$.

First for any state ω, we have

$$u(r) + \xi^* \cdot (\omega - r) = \lim_{j\to\infty} [u(r^{k_j}) + \xi^* \cdot (\omega - r^{k_j})]$$

$$= \lim_{j\to\infty} [u(r^{k_j}) + \xi^{k_j}(r^{k_j}) \cdot (\omega - r^{k_j})] \leq u(\omega),$$

where the first equality holds because function u and function $\xi^* \cdot r$ are continuous and bounded in reports. The inequality holds because $\xi^{k_j}(r^{k_j})$ is a subgradient for report r^{k_j}. Thus ξ^* is subgradient for report r. Next we show that the scoring rule induced by subgradient ξ^* is bounded for report r. For any state ω, we have

$$u(\omega) - u(r) - \xi^* \cdot (\omega - r) = u(\omega) - \lim_{j\to\infty} [u(r^{k_j}) + \xi^{k_j}(r^{k_j}) \cdot (\omega - r^{k_j})]$$

$$\leq u(\omega) - (u(\omega) - B) = B,$$

where the inequality holds because the subgradient $\xi^{k_j}(r^{k_j})$ satisfies the bounded constraint for report r^{k_j} at state ω, i.e., $\xi^{k_j}(r^{k_j}) \in \mathcal{G}_{u,r^{k_j}}$ and $u(r^{k_j}) + \xi^{k_j}(r^{k_j}) \cdot (\omega - r^{k_j}) \geq u(\omega) - B$. Therefore, $\xi^* \in \mathcal{G}_{u,r} \subset \mathcal{G}_u$, which implies the set $\mathcal{G}_u$ is a closed set. $\square$

APPENDIX C

Appendix to Chapter 4

C.1. Linear Valuation

Before the proof of the theorems in Section 4.2.1, we first present the following lemma showing that experiment σ^F that reveals full information is the most valuable for the agent. Recall that $\hat{\sigma}_{\theta,G} \in \hat{\Sigma}$ is the optimal experiment the agent chooses when her type is θ and her posterior belief after receiving the signal from the data broker is G.

Lemma C.1.1. *Let σ^F be the experiment that reveals full information. For any experiment $\sigma \in \Sigma$, any prior D, and any type θ, we have*

$$\mathbf{E}_{G\sim\sigma^F|D}\left[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{\theta,G}|G}\left[v(\hat{G})\right]\right] \geq \mathbf{E}_{G\sim\sigma|D}\left[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{\theta,G}|G}\left[v(\hat{G})\right]\right].$$

Proof. For the fully informative experiment σ^F, for any experiment σ, any prior D, and any type θ, we have

$$\mathbf{E}_{G\sim\sigma^F|D}\left[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{\theta,G}|G}\left[v(\hat{G})\right]\right] = \mathbf{E}_{\omega\sim D}[v(\omega)] \geq \mathbf{E}_{G\sim\sigma|D}\left[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{\theta,G}|G}\left[v(\hat{G})\right]\right].$$

The inequality holds since v is convex in G and σ^F fully reveals the states. □

Lemma C.1.2. *An allocation rule x can be implemented by an incentive compatible and individual rational mechanism if and only if for any $\theta, \theta' \in \Theta$,*[1]

$$\text{(IC)} \qquad \int_{\theta'}^{\theta} \mathbf{E}_{G\sim x(z)|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\big[v(\hat{G})\big]\Big] - \mathbf{E}_{G\sim x(\theta')|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\big[v(\hat{G})\big]\Big]\ \mathrm{d}z \geq 0,$$

$$\text{(IR)} \qquad \int_{\underline{\theta}}^{\theta} \mathbf{E}_{G\sim x(z)|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\big[v(\hat{G})\big]\Big]\ \mathrm{d}z + U(\underline{\theta}) \geq V(D, \hat{\Sigma}, \theta).$$

Proof of Lemma 4.2.4. Given allocation rule x, by the envelope theorem, for any incentive compatible mechanism $\mathcal{M}$, the interim utility $U(\theta)$ is convex in θ and

$$U(\theta) = \int_{\underline{\theta}}^{\theta} \mathbf{E}_{G\sim x(z)|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\big[v(\hat{G})\big]\Big]\ \mathrm{d}z + U(\underline{\theta}).$$

Note that the mechanism $\mathcal{M} = (x, p)$ is individual rational if and only if $U(\theta) \geq V(D, \hat{\Sigma}, \theta)$ for any type θ, i.e.,

$$\int_{\underline{\theta}}^{\theta} \mathbf{E}_{G\sim x(z)|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\big[v(\hat{G})\big]\Big]\ \mathrm{d}z + U(\underline{\theta}) \geq V(D, \hat{\Sigma}, \theta).$$

Moreover, the corresponding payment rule for mechanism $\mathcal{M}$ is

$$p(\theta) = \mathbf{E}_{G\sim x(\theta)|D}\left[V(G, \hat{\Sigma}, \theta) - \int_{\underline{\theta}}^{\theta} \mathbf{E}_{G\sim x(z)|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\big[v(\hat{G})\big]\Big]\ \mathrm{d}z\right] - U(\underline{\theta}).$$

Next we verify the incentive constraints of the given mechanism. Note that for any $\theta, \theta' \in \Theta$, letting $U(\theta; \theta')$ be the utility of the agent with type θ when she reports θ' in

[1] If $\theta < \theta'$, we use $\int_{\theta'}^{\theta}$ to represent $-\int_{\theta}^{\theta'}$.

mechanism $\mathcal{M}$, we have

$$
\begin{aligned}
&U(\theta) - U(\theta;\theta') \\
&= U(\theta) - U(\theta') - \mathbf{E}_{G\sim x(\theta')|D}\Big[V(G,\hat{\Sigma},\theta)\Big] + \mathbf{E}_{G\sim x(\theta')|D}\Big[V(G,\hat{\Sigma},\theta')\Big] \\
&= \int_{\theta'}^{\theta} \mathbf{E}_{G\sim x(z)|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\Big[v(\hat{G})\Big]\Big]\,\mathrm{d}z - \mathbf{E}_{G\sim x(\theta')|D}\Big[\int_{\theta'}^{\theta} V_3(G,\hat{\Sigma},z)\,\mathrm{d}z\Big] \\
&= \int_{\theta'}^{\theta} \Big(\mathbf{E}_{G\sim x(z)|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\Big[v(\hat{G})\Big]\Big] - \mathbf{E}_{G\sim x(\theta')|D}\Big[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\Big[v(\hat{G})\Big]\Big]\Big)\,\mathrm{d}z.
\end{aligned}
$$

Thus $U(\theta) - U(\theta;\theta') \geq 0$ if and only if the integral constraint in the statement of Lemma 4.2.4 is satisfied. □

Theorem 4.2.5. *For linear valuations, under Assumption 2 and 3, there exists an optimal mechanism* $\widehat{\mathcal{M}}$ *with allocation rule* $\hat{x}$ *such that,*[2]

- *for any type* $\theta \geq \theta^*$*, the data broker reveals full information, i.e.,* $\hat{x}(\theta) = \sigma^F$*;*
- *for any type* $\theta < \theta^*$*, the data broker commits to information structure*

$$\hat{x}(\theta) = \arg\max_{\hat{\sigma}\in\hat{\Sigma}} \mathbf{E}_{G\sim\hat{\sigma}|D}[V(G,\theta)] - C^A(\hat{\sigma}, D)$$

  where ties are broken by maximizing the cost $C^A(\hat{\sigma}, D)$*;*
- $U(\underline{\theta}) = V(D,\hat{\Sigma},\underline{\theta})$.

Proof of Theorem 4.2.5. We first show that allocation rule $\hat{x}$ combined with $U(\underline{\theta}) = V(D,\hat{\Sigma},\underline{\theta})$ can be implemented as an incentive compatible and individual rational mechanism. One way to prove this is to verify the constraints specified in Lemma 4.2.4.

[2]The characterization on allocation actually holds in any optimal mechanism except for a set of types with measure zero.

However, directly verifying the incentive constraints in Lemma 4.2.4 for allocation $\hat{x}$ might be challenging as we impose little structure on the information costs.[3] Thus we adopt an alternative approach by explicitly constructing an incentive compatible and individual rational mechanism $\widehat{\mathcal{M}}$. Then we show that the constructed mechanism has allocation $\hat{x}$ and utility for the lowest type $U(\underline{\theta}) = V(D, \hat{\Sigma}, \underline{\theta})$.

First consider a mechanism $\mathcal{M}'$ that post a deterministic price p for revealing full information. The price p is chosen such that the agent purchases information from the seller if and only if $\theta \geq \theta^*$. Note that given mechanism $\mathcal{M}'$, for agent with type $\theta < \theta^*$, she will choose not to participate the auction, and then subsequently conduct experiment

$$\hat{\sigma}_\theta = \arg\max_{\hat{\sigma} \in \hat{\Sigma}} \mathbf{E}_{G \sim \hat{\sigma}|D}[V(G, \theta)] - C^A(\hat{\sigma}, D).$$

We assume that the agent breaks tie by maximizing the cost $C^A(\hat{\sigma}, D)$. Now let $\widehat{\mathcal{M}}$ be the mechanism that reveals full information for types $\theta \geq \theta^*$ with price p, and commits to information structure $\hat{\sigma}_\theta$ for types $\theta < \theta^*$ with price $C^A(\hat{\sigma}_\theta, D)$. It is easy to verify that $\widehat{\mathcal{M}}$ has allocation rule $\hat{x}$ and the utility of the lowest type $\underline{\theta}$ in $\widehat{\mathcal{M}}$ is $V(D, \hat{\Sigma}, \underline{\theta})$. Moreover, by the proof of Theorem 4.2.1, mechanism $\widehat{\mathcal{M}}$ is incentive compatible and individual rational.

Note that when the posterior G is in the support of $\sigma^F|D$, the agent will not acquire additional costly information since σ^F fully reveals the state. Moreover, when the posterior G is in the support of $\hat{\sigma}_\theta|D$, by Assumption 2, the agent will not acquire additional costly information because otherwise $\hat{\sigma}_\theta$ is not the utility maximization information structure given prior D. Combining the observations, we have that $C^A(\hat{\sigma}_{\theta,G}, G) = 0$ for G in the

[3]Without additional structures on the costs, it is hard to characterize the optimal strategy $\hat{\sigma}_{\theta,G}$ given any type θ and posterior G.

support of $\hat{x}(\theta)|D$, and hence by Equation (4.1), the revenue of mechanism $\widehat{\mathcal{M}}$ is

$$\begin{aligned}\mathrm{Rev}(\widehat{\mathcal{M}}) &= \mathbf{E}_{\theta\sim F}\left[\mathbf{E}_{G\sim\hat{x}(\theta)|D}\left[\phi(\theta)\cdot\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{\theta,G}|G}\left[v(\hat{G})\right] - C^A(\hat{\sigma}_{\theta,G}, G)\right]\right] - U(\underline{\theta})\\ &= \mathbf{E}_{\theta\sim F}\left[\phi(\theta)\cdot\mathbf{E}_{G\sim\hat{x}(\theta)|D}\left[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{\theta,G}|G}\left[v(\hat{G})\right]\right]\right] - V(D,\hat{\Sigma},\underline{\theta})\\ &= \int_{\theta^*}^{\bar{\theta}}\phi(\theta)\cdot\mathbf{E}_{G\sim\sigma^F|D}[v(G)]\ \mathrm{d}\theta + \int_{\underline{\theta}}^{\theta^*}\phi(\theta)\cdot\mathbf{E}_{G\sim\hat{\sigma}_{\theta,D}|D}[v(G)]\ \mathrm{d}\theta - V(D,\hat{\Sigma},\underline{\theta}).\end{aligned} \tag{C.1}$$

Now consider any incentive compatible and individual rational mechanism $\mathcal{M}$ with allocation x, again by Equation (4.1), the revenue of mechanism $\mathcal{M}$ is

$$\begin{aligned}\mathrm{Rev}(\mathcal{M}) &= \mathbf{E}_{\theta\sim F}\left[\mathbf{E}_{G\sim x(\theta)|D}\left[\phi(\theta)\cdot\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{\theta,G}|G}\left[v(\hat{G})\right] - C^A(\hat{\sigma}_{\theta,G}, G)\right]\right] - U(\underline{\theta})\\ &\leq \mathbf{E}_{\theta\sim F}\left[\phi(\theta)\cdot\mathbf{E}_{G\sim x(\theta)|D}\left[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{\theta,G}|G}\left[v(\hat{G})\right]\right]\right] - U(\underline{\theta}),\end{aligned}$$

where the inequality holds since $C^A(\hat{\sigma}_{\theta,G}, G) \geq 0$ for any posterior G. For any type $\theta \geq \theta^*$, i.e., $\phi(\theta) \geq 0$, by applying Lemma C.1.1, the contribution of revenue from type θ is

$$\begin{aligned}\mathrm{Rev}(\mathcal{M};\theta) &\triangleq \phi(\theta)\cdot\mathbf{E}_{G\sim x(\theta)|D}\left[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{\theta,G}|G}\left[v(\hat{G})\right]\right]\\ &\leq \phi(\theta)\cdot\mathbf{E}_{G\sim\sigma^F|D}\left[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{\theta,G}|G}\left[v(\hat{G})\right]\right]\\ &= \phi(\theta)\cdot\mathbf{E}_{G\sim\sigma^F|D}[v(G)]\,.\end{aligned} \tag{C.2}$$

Next we bound the revenue contribution from types $\theta < \theta^*$, i.e., $\phi(\theta) < 0$.

$$
\begin{aligned}
&\mathbf{E}_{\theta\sim F}[\mathrm{Rev}(\mathcal{M};\theta)\cdot \mathbf{1}\left[\theta < \theta^*\right]] - U(\underline{\theta}) \\
&= \int_{\underline{\theta}}^{\theta^*} f(\theta)\cdot\phi(\theta)\cdot \mathbf{E}_{G\sim x(\theta)|D}\left[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{\theta,G}|G}\left[v(\hat{G})\right]\right]\,\mathrm{d}\theta - U(\underline{\theta}) \\
&= -\int_{\underline{\theta}}^{\theta^*} (f(\theta)\cdot\phi(\theta))' \int_{\underline{\theta}}^{\theta} \mathbf{E}_{G\sim x(z)|D}\left[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\left[v(\hat{G})\right]\right]\,\mathrm{d}z\,\mathrm{d}\theta - U(\underline{\theta}) \\
&\le -\int_{\underline{\theta}}^{\theta^*} (f(\theta)\cdot\phi(\theta))'\cdot(V(D,\hat{\Sigma},\theta) - U(\underline{\theta}))\,\mathrm{d}\theta - U(\underline{\theta}) \\
&= -\int_{\underline{\theta}}^{\theta^*} (f(\theta)\cdot\phi(\theta))'\cdot V(D,\hat{\Sigma},\theta)\,\mathrm{d}\theta - U(\underline{\theta})(f(\underline{\theta})\cdot\phi(\underline{\theta})+1) \\
&\le -\int_{\underline{\theta}}^{\theta^*} (f(\theta)\cdot\phi(\theta))'\cdot V(D,\hat{\Sigma},\theta)\,\mathrm{d}\theta - V(D,\hat{\Sigma},\underline{\theta})(f(\underline{\theta})\cdot\phi(\underline{\theta})+1) \\
&= \int_{\underline{\theta}}^{\theta^*} \phi(\theta)\cdot\mathbf{E}_{G\sim\hat{\sigma}_{\theta,D}|D}[v(G)]\,\mathrm{d}\theta - V(D,\hat{\Sigma},\underline{\theta}). \qquad \text{(C.3)}
\end{aligned}
$$

The second equality holds by integration by parts. The first inequality holds by (1) $(f(\theta)\cdot\phi(\theta))'$ is non-negative for $\theta \le \theta^*$ under Assumption 3 [c.f., Devanur and Weinberg, 2017b]; and (2) $\int_{\underline{\theta}}^{\theta} \mathbf{E}_{G\sim x(z)|D}\left[\mathbf{E}_{\hat{G}\sim\hat{\sigma}_{z,G}|G}\left[v(\hat{G})\right]\right]\,\mathrm{d}z \ge V(D,\hat{\Sigma},\theta) - U(\underline{\theta})$ according to the individual rational constraints in Lemma 4.2.4. The last inequality holds since $U(\underline{\theta}) \ge V(D,\hat{\Sigma},\underline{\theta})$ and $f(\underline{\theta})\cdot\phi(\underline{\theta}) + 1 = f(\underline{\theta})\cdot\underline{\theta} + F(\underline{\theta}) \ge 0$. Finally, the last equality holds by integration by parts and the facts that $\phi(\theta^*) = 0$ and

$$
V(D,\hat{\Sigma},\theta) = V(D,\hat{\Sigma},\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \mathbf{E}_{G\sim\hat{\sigma}_{z,D}|D}[v(G)]\,\mathrm{d}z.
$$

Combining Equations (C.1) to (C.3), we have

$$\begin{aligned}\mathrm{Rev}(\mathcal{M}) &\leq \mathbf{E}_{\theta\sim F}[\mathrm{Rev}(\mathcal{M};\theta)\cdot \mathbf{1}\left[\theta \geq \theta^*\right]] + \mathbf{E}_{\theta\sim F}[\mathrm{Rev}(\mathcal{M};\theta)\cdot \mathbf{1}\left[\theta < \theta^*\right]] - U(\underline{\theta}) \\ &\leq \int_{\theta^*}^{\bar{\theta}} \phi(\theta)\cdot \mathbf{E}_{G\sim\sigma^F|D}[v(G)]\ \mathrm{d}\theta + \int_{\underline{\theta}}^{\theta^*} \phi(\theta)\cdot \mathbf{E}_{G\sim\hat{\sigma}_{\theta,D}|D}[v(G)]\ \mathrm{d}\theta - V(D,\hat{\Sigma},\underline{\theta}) = \mathrm{Rev}(\widehat{\mathcal{M}}).\end{aligned}$$

Thus mechanism $\widehat{\mathcal{M}}$ is revenue optimal. □

Proposition C.1.3. *For any cost function C^A and any prior D, if $\sigma^N \in \arg\max_{\hat{\sigma}\in\hat{\Sigma}} \mathbf{E}_{\hat{G}\sim\hat{\sigma}|D}\left[V(\hat{G},\theta^*)\right] - C^A(\hat{\sigma},D)$, the optimal mechanism is to post a price for revealing full information.*

Proof. By Theorem 4.2.5, it is sufficient to show that if

$$\sigma^N \in \arg\max_{\hat{\sigma}\in\hat{\Sigma}} \mathbf{E}_{\hat{G}\sim\hat{\sigma}|D}\left[V(\hat{G},\theta^*)\right] - C^A(\hat{\sigma},D),$$

then $\hat{\sigma}_{\theta,D} = \sigma^N$ for any $\theta < \theta^*$. Suppose by contradiction that there exists $\theta < \theta^*$ such that $C^A(\hat{\sigma}_{\theta,D}, D) > 0$, i.e.,

$$\mathbf{E}_{G\sim\hat{\sigma}_{\theta,D}|D}[v(G)]\cdot\theta - C^A(\hat{\sigma}_{\theta,D},D) \geq \mathbf{E}_{G\sim\sigma^N|D}[v(G)]\cdot\theta.$$

Since $\theta^* > \theta$, we have that

$$\begin{aligned}&\mathbf{E}_{G\sim\hat{\sigma}_{\theta,D}|D}[v(G)]\cdot\theta^* - C^A(\hat{\sigma}_{\theta,D},D) - \mathbf{E}_{G\sim\sigma^N|D}[v(G)]\cdot\theta^* \\ &> \left(\mathbf{E}_{G\sim\hat{\sigma}_{\theta,D}|D}[v(G)] - \mathbf{E}_{G\sim\sigma^N|D}[v(G)]\right)\cdot\theta - C^A(\hat{\sigma}_{\theta,D},D) \geq 0,\end{aligned}$$

contradicting to the assumption that σ^N is one of the optimal choices for type θ^* given the prior D. □

In the following lemma, in the case that the state space Ω is finite, we formalize the intuition that when the prior is sufficiently close to the degenerate pointmass distribution, the marginal cost for additional information exceeds the marginal benefit of additional information.

Definition C.1.1. *The set of possible experiments* $\hat{\Sigma}$ is finitely generated *if it is generated by* σ^N *and a finite set* $\hat{\Sigma}'$ *through sequential learning, where* σ^N *is the one that always reveals no additional information with zero cost, and any* $\hat{\sigma} \in |\hat{\Sigma}'|$ *is an experiment that provides an informative signal about the state with fixed cost* $c_{\hat{\sigma}} > 0$.

Lemma C.1.4. *Suppose* Ω *is finite and* $\hat{\Sigma}$ *is finitely generated. Suppose that there exists* $\bar{v} < \infty$ *such that* $\max_{\omega \in \Omega} v(\omega) \leq \bar{v}$ *and* $v(G) \geq \min_{\omega \in \Omega} G(\omega) \cdot v(\omega)$ *for any* G*. Then there exists* $\epsilon > 0$ *such that any prior* D *satisfying* $D(\omega) > 1 - \epsilon$ *for some* $\omega \in \Omega$ *is sufficiently informative.*[4]

Proof. Let $c_m = \min_{\hat{\sigma} \in \hat{\Sigma}'} > 0$. By construction, for any experiment $\hat{\sigma} \in \hat{\Sigma}$, we have $C^A(\hat{\sigma}, G) \geq c_m$ for any G. Let ω^* be the state such that $D(\omega^*) > 1 - \epsilon$. Given prior D, the utility increase of type θ^* for additional information is at most

$$\theta^* \cdot \left(\sum_{\omega \in \Omega} D(\omega) v(\omega) - v(D) \right) \leq \theta^* \cdot \left(\sum_{\omega \neq \omega^*} D(\omega) v(\omega) \right) < \theta^* \cdot \epsilon \cdot \bar{v}.$$

The first inequality holds since $v(D) \geq D(\omega^*) v(\omega^*)$, and the second inequality holds since $v(\omega) \leq \bar{v}$ and $\sum_{\omega \neq \omega^*} D(\omega) < \epsilon$. Thus, when $\epsilon = \frac{c_m}{\theta^* \cdot \bar{v}}$, the cost of information is always

[4] We can have similar results when $\hat{\Sigma}$ is not finitely generated. For example, when the cost function is the reduction in entropy, by applying the techniques in Caplin, Dean, and Leahy [2019], for any valuation function v, there exists $\epsilon > 0$ such that any prior D satisfying $D(\omega) > 1 - \epsilon$ for some $\omega \in \Omega$ is sufficiently informative.

higher than the benefit of information, and the agent with type θ^* will never acquire any additional information given prior D. □

Before the proof of Theorem 4.3.2, we first introduce the definition of quantiles and revenue curves, which are helpful for bounding the approximation ratio. For any distribution F, let $q_F(\theta) \triangleq \Pr_{z\sim F}[z \geq \theta]$ be the quantile corresponding to type θ. Accordingly, we can define $\theta(q)$ as the type corresponds to quantile q. The revenue curve as a function of the quantile is defined as $R_F(q) \triangleq q \cdot \theta(q)$. Note that the regularity condition in Assumption 3 is equivalent to the concavity assumption for the revenue curve.

Lemma C.1.5 (Myerson, 1981b). *A distribution F is regular if and only if $R_F(q)$ is concave in q.*

Theorem 4.3.2. *For linear valuations, under Assumption 2 and 3, for any prior D and any cost function C^A, posting a deterministic price for revealing full information achieves at least half of the optimal revenue.*

Proof of Theorem 4.3.2. We first normalize the primitives such that $\theta^* \cdot q(\theta^*) = 1$. For any type θ, let $c(\theta) \triangleq V(D, \hat{\Sigma}, \theta)$ be the value of the agent for not participating the auction. It is easy to verify that $c(\theta)$ is convex in θ. Let $\bar{x} \triangleq \mathbf{E}_{\omega\sim D}[v(\omega)]$ be the maximum possible allocation. According to Theorem 4.2.5, if the distribution F is regular, in the revenue optimal mechanism, the expected utility of the agent is $c(\theta)$ for any $\theta < \theta^*$ and is $(\theta - \theta^*) \cdot \bar{x} + c(\theta^*)$ for any $\theta \geq \theta^*$.

Suppose $\hat{\theta}$ is the cutoff type that participates the auction in the optimal price posting mechanism for distribution F. It is easy to verify that $\hat{\theta} \leq \theta^*$ since revealing full information to any type above the monopoly type only increases the expected revenue. Moreover,

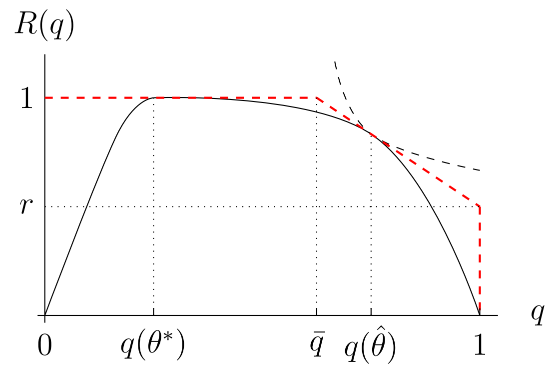

Figure C.1. The figure illustrates the reduction on the type distribution that maximizes the approximation ratio between the optimal revenue and the price posting revenue. The black solid curve is the revenue curve for distribution F and the red dashed curve is the revenue curve for distribution $\hat{F}$. The black dashed curve is the revenue curve $\bar{F}$ such that the seller is indifferent at deterministically selling at any prices with negative virtual value.

for any type $\theta < \theta^*$, since the payment that inducing $\hat{\theta}$ to be the cutoff type is $\hat{\theta} \cdot \bar{x} - c(\hat{\theta})$, we have that

$$(\hat{\theta} \cdot \bar{x} - c(\hat{\theta})) \cdot q_F(\hat{\theta}) \geq (\theta \cdot \bar{x} - c(\theta)) \cdot q_F(\theta).$$

That is, any type $\theta < \theta^*$,

$$q_F(\theta) \leq \frac{(\hat{\theta} \cdot \bar{x} - c(\hat{\theta})) \cdot q_F(\hat{\theta})}{\theta \cdot \bar{x} - c(\theta)}.$$

Let $\bar{F}$ be the distribution such that

$$q_{\bar{F}}(\theta) = \frac{(\hat{\theta} \cdot \bar{x} - c(\hat{\theta})) \cdot q_F(\hat{\theta})}{\theta \cdot \bar{x} - c(\theta)}$$

for any type θ. Thus the virtual value function $\bar{\phi}(\theta)$ for distribution $\bar{F}$ is

$$\bar{\phi}(\theta) = \theta - \frac{\theta \cdot \bar{x} - c(\theta)}{\bar{x} - c'(\theta)} \leq 0.$$

Moreover,

$$\bar{\phi}'(\theta) = \frac{c''(\theta) \cdot (c(\theta) - \theta \cdot \bar{x})}{(\bar{x} - c'(\theta))^2} \leq 0.$$

Thus the revenue curve such that the seller is indifferent at selling at any price is convex. Let $\hat{F}$ be the distribution with piecewise linear revenue curve illustrated in Figure C.1. Thus we have that

$$q_{\hat{F}}(\theta) = \begin{cases} \frac{1}{\theta} & \theta \geq \frac{1}{\bar{q}}, \\ \frac{1 - r\bar{q}}{\theta(1-\bar{q}) + 1 - r} & \theta < \frac{1}{\bar{q}}. \end{cases}$$

Let $p(\theta) \triangleq \theta \cdot \bar{x} - c(\theta) \geq 0$. First note that distribution $\hat{F}$ is first order stochastically dominated by $\bar{F}$, the optimal revenue from posted pricing is weakly smaller for distribution $\hat{F}$. Moreover, both distributions achieve the same price posting revenue by choosing the price $p(\hat{\theta})$ such that the cutoff type is $\hat{\theta}$. Thus the optimal price posting revenue for distribution $\hat{F}$ is attained by choosing price $p(\hat{\theta})$. This further indicates that optimal price posting revenue is the same for distribution F and $\hat{F}$, i.e., $\mathrm{PP}(F, c) = \mathrm{PP}(\hat{F}, c)$. Secondly, since distribution F is first order stochastically dominated by $\hat{F}$, it is easy to verify that $\mathrm{OPT}\,(F, c) \leq \mathrm{PP}(\hat{F}, c)$. Therefore, the ratio between the price posting revenue and the optimal revenue is minimized when the type distribution is $\hat{F}$.

For distribution $\hat{F}$, since the optimal price is $p(\hat{\theta})$, we have

$$(\hat{\theta} \cdot \bar{x} - c(\hat{\theta})) \cdot q_{\hat{F}}(\hat{\theta}) \geq (\theta \cdot \bar{x} - c(\theta)) \cdot q_{\hat{F}}(\theta).$$

Let $\zeta = (\hat{\theta} \cdot \bar{x} - c(\hat{\theta})) \cdot q_{\hat{F}}(\hat{\theta})$ and let

$$\hat{c}(\theta) = \theta \cdot \bar{x} - \frac{\zeta}{q_{\hat{F}}(\theta)}.$$

It is easy to verify that $\mathrm{PP}(\hat{F}, c) = \mathrm{PP}(\hat{F}, \hat{c}) = \zeta$. Moreover, $\hat{c}(\theta)$ is convex and $c(\theta) \geq \hat{c}(\theta)$ for any θ, which implies that any feasible mechanism for c is also feasible for $\hat{c}$, and hence $\mathrm{OPT}\left(\hat{F}, c\right) \leq \mathrm{OPT}\left(\hat{F}, \hat{c}\right)$. Thus to prove Theorem 4.3.2, it is sufficient to bound $\frac{\mathrm{PP}(\hat{F},\hat{c})}{\mathrm{OPT}\left(\hat{F},\hat{c}\right)}$. Note that by construction, the monopoly type for distribution $\hat{F}$ is $\frac{1}{\bar{q}}$. Hence the optimal revenue is

$$\begin{aligned}
\mathrm{OPT}\left(\hat{F}, \hat{c}\right) &= \int_r^{\frac{1}{\bar{q}}} \hat{f}(\theta)(\theta \cdot \hat{c}'(\theta) - \hat{c}(\theta))\,\mathrm{d}\theta + \left(\frac{1}{\bar{q}} \cdot \bar{x} - \hat{c}(\frac{1}{\bar{q}})\right) \cdot \bar{q} \\
&= \int_r^{\frac{1}{\bar{q}}} \hat{f}(\theta) \cdot \zeta \cdot \frac{1-r}{1-r\bar{q}}\,\mathrm{d}\theta + \zeta \\
&= \zeta \cdot \left(\frac{(1-\bar{q})(1-r)}{1-r\bar{q}} + 1\right) \leq 2\zeta,
\end{aligned}$$

where the inequality is tight if $\bar{q} = r = 0$. Combining the observations, for any distribution F and any function c induced by the set of experiments $\hat{\Sigma}$, we have

$$\frac{\mathrm{PP}(F, c)}{\mathrm{OPT}\,(F, c)} \geq \frac{\mathrm{PP}(\hat{F}, \hat{c})}{\mathrm{OPT}\left(\hat{F}, \hat{c}\right)} \geq \frac{1}{2}.$$

□

Proposition C.1.6. *For any valuation function V, any prior D, and any type distribution F, in the revenue optimal mechanism,*

- *the social welfare is minimized when $|\hat{\Sigma}| = 1$;*[5]
- *the utility of the agent is minimized when $|\hat{\Sigma}| = 1$.*

Proof for Statement 1 of Proposition 4.3.3. Recall that $\theta^* = \inf_\theta\{\phi(\theta) \geq 0\}$. Note that for any type $\theta \geq \theta^*$, the agent receives full information regardless of the set of possible experiments $\hat{\Sigma}$ for the agent. For any type $\theta < \theta^*$, it is easy to verify that the agent receives no information when $|\hat{\Sigma}'| = 1$. Since no information is the least preferred allocation for the agent, the social welfare is minimized when $|\hat{\Sigma}'| = 1$. □

[5] Our result actually implies that the expected value for each type of the agent is minimized when $|\hat{\Sigma}| = 1$.