FEDERAL RESERVE BANK OF BOSTON™

RISK AND POLICY ANALYSIS UNIT

Working Paper
RPA 18-01
March 30, 2018

Optimal Delegation under Unknown Bias: The Role of Concavity

Noam Tanner

Risk and Policy Analysis (RPA) Working Papers present economic, financial and policy-related research conducted by staff in the Federal Reserve Bank of Boston's Risk and Policy Analysis Unit. RPA Working Papers can be downloaded without charge at: http://www.bostonfed.org/bankinfo/qau/wp/

Optimal Delegation Under Unknown Bias: The Role of Concavity

Noam Tanner*

Federal Reserve Bank of Boston†

March 30, 2018

Abstract

A principal is uncertain of an agent's preferences and cannot provide monetary transfers. The principal, however, does control the discretion granted to the agent. In this paper, we provide a simple characterization of when it is optimal for the principal to screen by offering different terms of discretion to the agent. When the principal's utility is sufficiently concave, it is optimal for the principal to pool and to offer all agents the same discretion. Thus, for any number of agents and any distribution over agent preferences, the optimal contract is simple: the principal sets a cap and forbids actions above this cap (interval delegation). For less concave preferences, it is optimal for the principal to screen. The principal benefits by providing agents a choice between interval delegation and gap delegation, which allows for more extreme actions but prohibits intermediate actions. Moreover, we provide new intuition for the optimality of interval delegation when the principal knows the agent's preferences: the payoff distributions generated by sets containing gaps are mean-preserving spreads of those generated by intervals.

*noam.tanner@bos.frb.org; Federal Reserve Bank of Boston, Risk and Policy Analysis Unit.
The author gratefully acknowledges the help of Johannes Hörner and Larry Samuelson for their valuable comments, discussions, and support throughout this entire project. The author would also like to thank Dirk Bergemann, Tibor Heumann, Michal Kowalik, Daniel Krähmer, Claire Labonne, Tymofiy Mylovanov, Esteban Peralta, and Matthew Pritsker for helpful discussions.

†The views expressed herein are those of the author and do not necessarily represent those of the Federal Reserve Bank of Boston or the Federal Reserve system.

1 Introduction

1.1 Motivation

Principals frequently delegate decisions to an agent with greater knowledge of the impact of these decisions on their intended outcomes. This agent could be a standing committee drafting a bill for a legislative body (Gilligan and Krehbiel, 1987), a financial supervisor engaged in rulemaking (Melumad and Shibano, 1994), or a monopolist setting prices at the behest of a regulatory authority (Amador and Bagwell, 2016). The standing committee may have expert knowledge relevant to the bill being crafted. The supervisor may have greater knowledge of the institutional detail of financial markets. Finally, the monopolist may know the precise costs of production. This knowledge is directly relevant to the objectives of the principal. As the less informed principal, the legislature would like to make use of the expert and institutional knowledge of the committee to craft an effective bill or rule. In the monopoly example, the regulator may wish to balance firm profits and consumer welfare.

In these settings, the preferences of the principals may differ from those of the agents. For example, a committee of climate scientists may have greater concern for environmental harm than the legislature. For a given environmental state of the world, the committee would prefer a more aggressive bill targeting climate change than the legislature. If the committee is an industry group, the committee may prefer a *less aggressive* response than the legislature. Hence, we say that, in either case, the committee is *biased* away from the principal's ideal choice given the state of the world. Thus, the principal faces a trade-off. The more discretion the principal allows the agent, the better the principal can make use of the agent's information; however, the more leeway given to the agent, the more room there is for the agent to take actions undesirable to the principal. When the principal is uncertain of the degree of the agent's bias, this trade-off is more difficult to solve.

While monetary transfers may help the principal to elicit information to better solve this tradeoff (Laffont and Martimort, 2002), they may be infeasible in many settings. The legislature may be unable to provide transfers to the standing committee. The budget for the financial supervisor may be set outside the standard appropriations framework. Thus, the legislature cannot reward (or fine) it for its rulemaking. It may also be illegal for the monopolist to provide monetary transfers to the regulator.

In the setting studied here, the agent first learns his preferences and then learns the payoff-relevant state of the world. For example, the standing committee's views on the gravity of climate change have generally been developed before learning of the impact of a particular bill on curbing carbon emissions. While the principal cannot elicit information using transfers, the principal can elicit information in a manner similar to that in the transfer setting: by screening. By *screening* we mean providing the agent with a menu of different options such that agents with different preferences will select different options from the menu. By appropriately designing this menu, the principal can incentivize each preference type of agent to select actions that better align with the principal's preferences.

In this setting without transfers, this menu will consist of ranges of possible decisions. Once the agent selects an option from the menu, the agent will be restricted to make a decision within the selected range. For example, the regulator could allow the monopolist to choose between multiple regulatory frameworks, but must price following the chosen framework. For example, Sappington (2002) documents that in telecommunication regulation, the regional Bell Operating Companies were provided a choice between different earnings sharing plans and price cap regulation. Similarly, the legislature can set different guidelines on rulemaking for the supervisor and allow the supervisor to choose between the different guidelines.

We characterize when it is optimal to screen. In a *pooling* contract, the agent only has one option. The key result in this paper is that screening may be optimal even when the agent's (initial) information regarding its preferences is unrelated to its (later) information regarding the payoff-relevant state. Thus, even if different standing committees have different biases from the legislature, but have access to the *same* scientific and institutional information, the legislature may benefit by screening between the two types of committees. In the setting studied in this paper, the feature that determines the structure of the optimal contract is the concavity of the loss function.

We show that pooling is optimal when the loss function of the principal is sufficiently concave and screening is optimal otherwise. Thus, for sufficiently risk-averse principals, pooling is always optimal. In this setting of sufficient risk aversion, while the set of possible contracts is complex, we prove that for any number of agents and for all distributions over agent preferences, an optimal contract is simple: the principal offers a menu of identical intervals. Surprisingly, in this setting, there is no benefit to the principal from providing the agent a choice over decision rules. What is important from the principal's standpoint is the design of this single decision rule. This result is in contrast with much of the standard results on mechanism design with transfers. In this mechanism design literature, the principal benefits by providing the agents with a choice over decisions (Laffont and Martimort, 2002). More formally, pooling contracts are only optimal on a non-generic subset of parameters. In addition, this optimal decision rule takes a simple form: the principal sets a cap and allows the agent to take any decision below this cap (i.e. threshold delegation is optimal).

In contrast, for less concave preferences, the principal benefits from screening. The principal benefits by offering less biased agents a set of extreme options (a set which contains gaps), while offering the more biased agents a set of only moderate options (an interval of options). A legislature that wishes to screen between a more and less ideological (biased) committee would offer the committee a choice between guidelines that allow for more extreme bills and guidelines that only allow moderate bills. Counterintuitively, the less ideological committee would select the guidelines with extreme options, while the more ideological committee would select the more moderate guidelines.

In order to understand the intuition when screening is optimal, notice that a decision rule is a set on the real line. Thus, when designing this set, the principal has three instruments at his disposal:

1. The choice of *left boundary*, the smallest action the agent can take.
2. The choice of *right boundary*, the largest action the agent can take.
3. Inserting *gaps*, prohibiting the choice of intermediate actions.

With regard to the choice of the left boundary, optimality reinforces incentive compatibility. The principal prefers to allow less biased agents to take lower actions. In addition, less biased agents will prefer sets with a lower left boundary, while more biased agents will prefer sets with higher boundaries. Yet, since more biased agents will be disinclined to select low actions, there is no loss to the principal from offering *all* agents the opportunity to select these low actions. Thus, tailoring the left boundary alone will not be enough for the principal to generate additional payoff from screening beyond that achieved by pooling.

With regard to the right boundary, optimality and incentive-compatibility clash. The principal would prefer to offer *more biased* agents *lower right boundaries.* In contrast, we will see that incentive compatibility will require the principal to offer the *more biased* agents *higher right boundaries* since they prefer higher actions. Hence, modifying the left and right boundaries alone will also not yield the principal addition payoff from screening.

Thus, in order to effectively screen and allow the less biased agents to take higher actions, the principal must make their delegation sets less appealing to the more biased types. The principal can achieve this by inserting gaps into the delegation sets of the less biased agents. Doing so will discourage the more biased agents from choosing a set with larger actions because they will be prohibited from taking intermediate actions. However, notice that this gap forces the agent to choose between extreme actions. The less biased agents still select the extreme option set because they prefer the discretion to select low outcomes (the more biased agents do not value these low outcomes). Returning to the example, the less ideological committee may prefer the extreme option guidelines since they are willing to craft bills at both ends of the spectrum. In contrast, while the more ideological committee may be more inclined to craft bills at one end of the spectrum, it is strongly disinclined to craft bills at the other end. Thus, the more *ideological* committee may select the more *moderate* guidelines.

In this way, the principal can use gaps to screen and it remains to determine when inserting these gaps is beneficial to the principal. We will show that a gap increases the variance over agent choices (and yields a mean-preserving spread of an interval delegation set). Hence, the more risk-averse the principal, the more costly will be the insertion of a gap. Hence, for sufficiently risk-averse principals (those most sensitive to variance) inserting a gap will be suboptimal. The gain from providing the less biased type the discretion to take higher actions (and preventing the more biased type from taking excessively high actions) is offset by the uncertainty generated by the extreme option set. Thus, pooling is optimal for these risk-averse principals. In contrast, for less risk-averse principals, inserting a gap will not be costly enough to discourage its use[1]. Thus, for these less risk-averse principals, screening will be optimal.

We would like to stress that screening is not driven by differing knowledge over the future payoff-relevant state of the world. In this paper, we show that even when the information the agent learns is independent over time, the principal still benefits

[1] In fact, for a principal with absolute value loss, inserting a gap brings no loss to the principal.

from screening. What drives the optimality of screening is the degree of risk-aversion of the principal. When the principal's utility is sufficiently concave and the principal is sufficiently risk-averse, the principal does not benefit from screening. When the principal is not as risk-averse, it is optimal for the principal to screen. We note that screening is optimal for concave functions and is not the result of risk-neutrality.

There are multiple interpretations of the main result. First, this result shows when threshold delegation, the optimal delegation strategy in the setting of known preferences, is robust to preference uncertainty. In addition, this paper provides insight into when screening may be observed in regulatory settings. In settings where the principal can diversify (or is less risk averse) over outcomes, choices among regulatory frameworks may be observed. In addition, the optimality of screening has important implications for political science. Mylovanov (2008) has shown that veto rules with default decisions may be used to implement the optimal contract without transfers. In the setting where screening is optimal, such a veto rule does not implement the optimal contract. Thus, if the principal is a legislature delegating a decision to a committee with an unknown political inclination, a veto rule may not be optimal. Finally, when a principal is more risk-neutral, the choice of discretion that allows for extreme actions may signal that the agent is less ideological.

Following a literature review, we present the model in section 2. Section 3 provides new intuition for the original result of Melumad and Shibano (1991) in settings of *preference certainty* when threshold delegation is optimal. This new intuition is leveraged in the proof of the main result in our setting of *preference uncertainty.* Section 3 provides graphical examples that illustrate the main results of this paper. In section 4, we set the stage for the proof of the main result by analyzing two simple settings: the case of known preference and the case where preferences are unknown, but the principal is restricted to offering menus of intervals with no gaps. The proof that intervals are optimal for known preferences uses only elementary techniques from information economics. It shows that the principal benefits from filling in gaps (regardless of the endpoints). This result is useful for characterizing the optimal pooling contract. We then characterize the optimal menu when the principal is restricted to offer only interval delegation sets. We show that incentive compatibility requirements drastically limit the allowable set of contracts and that a pooling contract can outperform all of the remaining contracts. In section 5, we use the results from section 4 and prove the main results. We show that pooling is optimal for concave enough loss functions and screening is optimal for all other loss functions. We provide comparative statics in section 6. In section 7, we show that even when pooling is optimal, the optimality of pooling hinges on the no transfer restriction. Section 8 concludes the paper.

1.2 Related Literature

Of the delegation literature, the two most important papers for this project are those of Holmström (1984) and Melumad and Shibano (1991) (abbreviated as MS). Szalay (2005) studies a setting where the agent must exert costly effort to acquire information. Martimort and Semenov (2006) provide conditions for when the optimal delegation set is an interval (but still consider the case when the preferences of the agent are known). Mylovanov (2008) studies veto-based delegation and Kováč and Mylovanov (2009) also study when stochastic mechanisms yield optimal payoffs to the principal. In a fundamental paper, Alonso and Matouschek (2008) study the optimal delegation problem under more general preferences and distributions over the state space. Amador and Bagwell (2013) provide even more general results and introduce innovative techniques to analyzing the delegation problem. Ambrus and Egorov (2017) provide a superb analysis of the setting of delegation with and without monetary transfers (including the impact of money burning). All these papers differ from the present paper in that they assume known preferences of the agent.

Amador and Bagwell (2012, 2013) provide applications of the theory of delegation to tariff caps. Amador, Werning, and Angeletos (2006) apply the theory of delegation to study commitment and flexibility in saving rules. Carrasco and Fuchs (2009) consider a setting of implementing a decision with agents who have different preferences. Once again, these papers assume that the preferences of the agent are known by the principal. Armstrong (1995) and Frankel (2014) consider the case of preference uncertainty but do not study the optimal screening contracts. Lewis and Sappington (1989) study a model of regulatory options, but consider a setting where transfers are allowed, unlike the setting of this paper.

Another related literature is that of sequential screening, such as Courty and Li (2000). In that paper, the functional form of the agents' utility is monotonic. In the present paper, the utilities are not monotonic. Another difference is that transfers are allowed in Courty and Li (2000), but are prohibited in the present paper. The closest work to this paper was the innovative study of sequential delegation by Kováč and Krähmer (2016). This paper complements their analysis by studying an alternative preference structure and finding a new setting where screening is optimal (screening in this paper is analogous to sequential delegation in their paper). Like Kováč and Krähmer (2016), we find that screening may be optimal and that inserting gaps to delegation sets may also be optimal. But in our setting we find an additional driver for the optimality of screening: the degree of risk-aversion of the principal. In this environment, screening may be optimal when agent types have *different* biases, but the *same* knowledge over future payoff-relevant states. Thus, bias may matter for screening in addition to knowledge.

This paper also generalizes results from Tanner (2015) by extending the class of functions where pooling is optimal. In addition, it also characterizes when screening is optimal. In Tanner (2015), pooling was found to be optimal in all settings explored. Hence, this paper provides a simple characterization regarding when pooling is optimal.

2 The Model

2.1 Preferences

The setting is similar to that of Melumad and Shibano (1991), except that we introduce uncertainty in the bias between the agent and principal. Thus, the payoff functions of the agent depend on the state of nature ($s \in [0,1]$), the action implemented by the principal ($x \in \mathbb{R}$), and the bias of the agent ($k \in \mathbb{R}_+$). Let $U : \mathbb{R} \to \mathbb{R}$, where U is a symmetric, differentiable, strictly concave function, that is maximized at zero (without loss of generality, we normalize U so that $U(0) = 0$)[2]. The utility function of the principal is $U^P(x,s) = U(x-s)$. Let $0 \leq k_1 < k_2 < \cdots < k_N$. The utility function of agent i is $U^i(x,s) = U(x - s - k_i)$, where $i \in \{1, 2, \ldots, N\} = \mathcal{N}$. k_i and s are random and statistically independent, where s is distributed uniformly over $[0,1]$. The probability that bias k_i is chosen is denoted by p_i.

2.2 Actions, Timing, and Solution Concept

Before describing the timing of the game, we first define a delegation set. A *delegation set* is a set of actions, D, that the agent will be restricted to take. D must be compact and we denote the set of compact subsets of the real line by $\mathcal{D}$. (We only need to restrict attention to closed sets since for every unbounded closed set, there is a bounded closed set that produces the same outcome and provides identical incentives). By the taxation principle, the principal will offer the agents a menu of sets, $m =$

[2]These conditions also imply that $U(\cdot)$ is strictly decreasing over $\mathbb{R}_+$ and strictly increasing over $\mathbb{R}_-$.

$\{D_1, \ldots, D_N\}$ (We will describe the timing of the game below). Let $\mathcal{D}^N = \mathcal{M}$ be the set of all menus of delegation sets. We call a menu, $m \in \mathcal{M}$, a *convex menu* if every $D \in m$ is a convex set. In other words, all the delegation sets in a convex menu are convex. We call a menu, $m \in \mathcal{M}$, a *nonconvex menu* if there exists a nonconvex set, $D' \in m$.

The timing of the game is as follows:

Time 0: Nature chooses k_i (bias) for the agent. The agent observes this value, but the principal does not.

Time 1: The principal offers the agent a menu of delegation sets, $m = \{D_i\}_{i \in \mathcal{N}} \in \mathcal{M}$.

Time 2: The agent selects one of the sets, D_i, and this selection is observed by the principal.

Time 3: The state of the world, s, is chosen by nature. It is observed by the the agent but not the principal.

Time 4: The agent picks a *final action* $d \in D_i$, which is observed by the principal. Payoffs are then determined.

The interpretation of this formulation is that the final, payoff-relevant action of the agent is restricted by the principal. The action chosen by the agent must be an element of his chosen delegation set, which was designed by the principal. Hence, the principal's strategy is an element $\{D_i\}_{i \in \mathcal{N}} \in \mathcal{M} = \mathcal{D}^N$.

The agent's strategy is an action at each information set in the game tree. Thus, the agent's strategy is σ, where

$$\sigma : \mathcal{N} \times \mathcal{M} \times [0,1] \to \mathcal{D} \times \mathbb{R}, \tag{2.1}$$

$$\sigma(i, m, s) = (\sigma_{\mathcal{M}}(i, m), \sigma_{\mathcal{D}}(i, m, s)) \tag{2.2}$$

$$\sigma_{\mathcal{M}}(i, m) \in m, \tag{2.3}$$

$$\sigma_{\mathcal{D}}(i,m,s) \in \sigma_{\mathcal{M}}(i,m), \tag{2.4}$$

and

$$i \in \mathcal{N} = \{1,\ldots,N\}, s \in [0,1]. \tag{2.5}$$

$\sigma_{\mathcal{M}}(i,m)$ represents the delegation set agent i chooses from the menu offered by the principal (implied by condition 2.3). $\sigma_{\mathcal{D}}(i,m,s)$ represents the final action chosen by the agent after observing the state of nature. Notice that the final action must be an element of the delegation set chosen (implied by condition 2.4).

The solution concept used throughout this paper is *Perfect Bayes-Nash Equilibrium.* The principal chooses $m = \{D_1,\ldots,D_N\}$ to maximize ex-ante expected utility:

$$\max_{m\in\mathcal{M}} \sum_{i=1}^{n} p_i\bigg(\int_0^1 U^P\big(\sigma_{\mathcal{D}}(i,m,s),s\big)ds\bigg). \tag{2.6}$$

Each type of agent ($i \in \{1,\ldots,N\}$), chooses the final action, $\sigma_{\mathcal{D}}(i,m,s)$ to maximize ex-post utility conditional on the original choice of delegation set from the menu, $\sigma_{\mathcal{M}}(i,m)$:

$$\sigma_{\mathcal{D}}(i,m,s) \in \operatorname*{argmax}_{x\in\sigma_{\mathcal{M}}(i,m)} U^i(x,s) = \operatorname*{argmax}_{x\in\sigma_{\mathcal{M}}(i,m)} U(x-s-k_i). \tag{2.7}$$

Their are two points to notice. First, notice that $\sigma_{\mathcal{D}}$ is determined by $\sigma_{\mathcal{M}}$. We call the value of $\sigma_{\mathcal{M}}(i,m)$, agent i's choice of delegation set. Second, we do not need to assume that the menu contains only bounded closed sets. In other words, $\sigma_{\mathcal{M}}(i,m)$ need not be compact. Since the loss function is symmetric and decreasing in the distance from $s+k_i$, we know that there exists a $Q(i,m)$ such that $|\sigma_{\mathcal{D}}(i,m,s)| \leq Q(i,m), \forall s \in [0,1]$. Thus, we know that even for menus containing closed (but not bounded sets), $\sigma_{\mathcal{D}}$ is well-defined. Each agent type, chooses the delegation set, D_j, from the menu, $m = \{D_1,\ldots,D_N\}$ in order to maximize interim expected utility given his future final actions, $\mathbb{E}^i D_j$, where:

$$\mathbb{E}^i D_j = \int_0^1 U^i\big(\sigma_{\mathcal{D}}(i,m,s),s\big)ds, \tag{2.8}$$

where $\sigma_{\mathcal{D}}(i,m,s) \in D_j$.

Before deriving results about the optimal convex and pooling menus, we need to derive some properties about the equilibrium strategies of the agents. We will state the useful properties of the equilibrium best responses in the next subsection (we will characterize $\sigma_{\mathcal{D}}(i,m,s)$ and prove some useful lemmas about this function). Melumad and Shibano (1991) prove the relevant properties of these best response functions for a general class of utility functions. We will show below that the utility functions assumed here satisfy the properties necessary for Melumad and Shibano's proof. Hence, their results apply in this setting.

2.3 Characterizing the Agent's Best Response

The agent type's ($i \in \mathcal{N}$) behavior is very simple conditional on the choice of a delegation set, $\sigma_{\mathcal{M}}(i,m) = D \in m$. Agent type i will choose, for each state s, the element in D closest to $s + k_i$. We call this point $x_i^D(s)$. More formally, for every compact set $D \subseteq \mathbb{R}$ (we just need the set to be closed, but we assume compactness for a smoother description):

$$x_i^D(s) \in \operatorname*{argmax}_{x \in D} U^i(x,s) = \operatorname*{argmax}_{x \in D} U(x - s - k_i). \tag{2.9}$$

Comparing (2.7) and (2.9) we define

$$x_i^D(s) := \sigma_{\mathcal{D}}(i,m,s), \tag{2.10}$$

when $\sigma_{\mathcal{M}}(i,m) = D$.

We call x_i^D the *delegation schedule generated by* D for type i. This function maps the current state, s, to an optimal action of the agent within his chosen delegation set, D. Let $x_i^O(s) = x_i^{\mathbb{R}}(s) = \text{argmax}_{x \in \mathbb{R}} U(x - s - k_i) = s + k_i$. x_i^O is the optimal delegation schedule for type i: for each state, s, x_i^O yields the best possible final action for the agent type. Call the range of x_i^O the *set of ideal actions for type i*. Notice that the range of $x_i^O = [k_i, 1 + k_i]$. In addition, the ideal action set of the principal is $[0, 1]$. We denote the principal's optimal delegation schedule by $x_P(s) = s$. The properties of $x_i^D(\cdot)$ are listed in Appendix A. The most important result is that the delegation schedule is an increasing function of the state. Figures 1 and 2 present plots of delegation schedules of convex and nonconvex sets. They provide intuition for the behavior of agents given different choices of delegation sets. The left graph plots the choice of outcome for two interval delegation sets as a function of the state. Notice that the agent selects the lowest element of each set for a range of low states, selects his ideal action for an intermediate range, and then selects the highest element of each delegation set for the remaining interval of states. The right graph plots the delegation schedule for a set with a gap. The key difference is that the delegation schedule is discontinuous. The delegation schedule "jumps" from the left endpoint of the gap to the right endpoint of the gap. This jump will introduce variance to the principal.

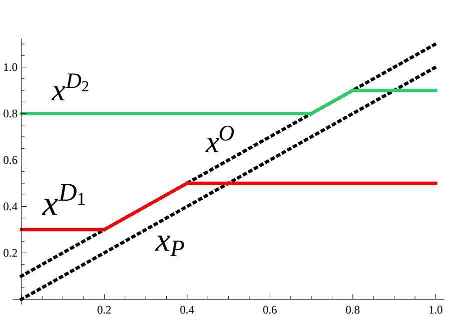

Figure 1: Delegation schedules of k_1 type

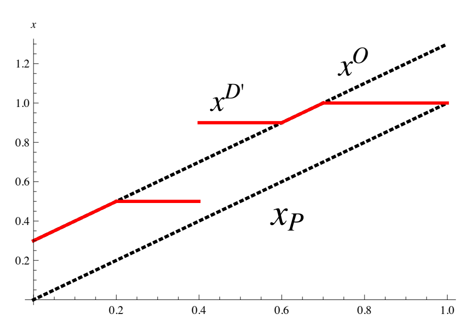

Figure 2: Schedule for a set with a gap.

A delegation set D is *nonredundant for agent type* i if every action in set D is taken by player i for a particular state. Stated more formally, a delegation set D is nonredundant for agent type i if $I(x_i^D) := Image(x_i^D) = D$. Corollary 9.7 in Appendix A shows that for every closed set D, there is a nonredundant set $D' = I(x_i^D)$ for player i (which we know is compact). Thus, we can restrict attention to nonredundant menus. Corollary 9.7 (the nonredundancy result) will simplify analysis when studying screening menus (defined below).

The interim expected utility to the agent of type i after choosing the delegation set D is:

$$\mathbb{E}^i D = \int_0^1 U\big(x_i^D(s) - s - k_i\big)ds.$$

Thus, by the Taxation Principle, the principal's optimization program is :

$$\max_{\{D_1,\dots,D_n\}\in\mathcal{M}} \sum_{i=1}^{n} p_i \left(\int_0^1 U\big(x_i^{D_i}(s) - s\big)ds\right) \tag{2.11}$$

subject to the type incentive constraint for delegation sets (IC_k^i):

$$\mathbb{E}^i D_i \geq \mathbb{E}^i D_j, \forall i, j.$$

We denote $\mathbb{E}_i^P D_i := \int_0^1 U(x_i^{D_i}(s) - s)ds$ as the *expected payoff to the principal from type* i. A set D' *improves upon* D *for type* i if $\mathbb{E}_i^P D' > \mathbb{E}_i^P D$ and $\mathbb{E}^i D' \geq \mathbb{E}^i D$. Notice that the menu that improves upon another only satisfies the IC_j^i constraints. Thus, it may not be incentive compatible since another type, k, may prefer D' to D_k.

A menu is *convex* if all sets in it are convex. A menu $m \in \mathcal{D}^N$ is *pooling* if all sets in it are identical ($m = \{D_1, \dots, D_N\}$, where $D_1 = D_2 = \dots = D_N = D$). We call D the *pooling set*. A menu $m \in \mathcal{D}^N$ is *screening* if contains at least two nonredundant sets that are different. Notice that the restriction to nonredundant sets is not such a demanding requirement since $I(x_i^{D_i}) \subseteq D_i$, and, therefore,

$$\mathbb{E}^i\bigg(I(x_i^{D_i})\bigg) = \mathbb{E}^i(D_i) \geq \mathbb{E}^i(D_j) \geq \mathbb{E}^i\bigg(I(x_j^{D_j})\bigg),$$

$\forall i, j \in \mathcal{N}$. Hence, the expected payoff and incentive constraints are preserved.

3 Examples to Illustrate Main Results

We illustrate the main results with graphical examples. In our first example, we show that screening between agents with different biases is feasible. In addition, it is feasible using a menu of intervals. In the next example, we show that when intervals are used, incentive-compatibility conflicts with optimality for the principal. We then illustrate how gaps may help and harm the principal. In example 3, we show how inserting gaps into the delegation set may lower the principals payoff. Given the lesson from example 3, we show an example where the gap may not hurt the principal (example 4), and inserting a gap may improve the payoff of the principal by allowing him to screen. In example 5, we provide intuition for the role of concavity of the principal's loss function in determining the benefits from screening. In the final example we illustrate, for sufficiently risk-averse principals, why pooling is optimal. We now show the feasibility of screening with a menu of intervals.

3.1 Example 1: Feasibility of Screening Menus Consisting Only of Intervals

First, we show that the problem is feasible (even under the restriction to convex menus). We show that there exist menus that are convex, incentive compatible, and screening. Let $U^P(x, s) = -(x - s)^2$, $U^1(x, s) = -(x - s - 0.1)^2$, and $U^2(x, s) = -(x -$

$s - 0.4)^2$. (Hence, $k_1 = 0.1$ and $k_2 = 0.4$.) Let $m = \{D_1, D_2\}$, where $D_1 = [0.3, 0.5]$ and $D_2 = [0.8, 0.9]$. Figure 1 plots the delegation schedule for type k_1 (the delegation schedule for type k_2 is analogous), while figures 3 and 4 plot the loss for the various delegation sets. The shaded area under each plot represents the expected utility loss from each set.

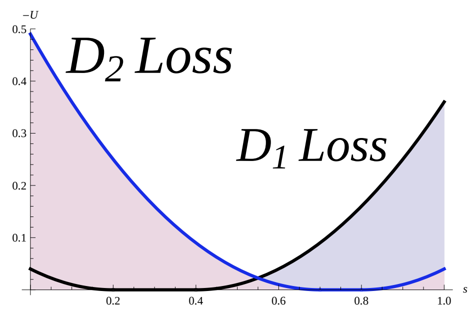

Figure 3: Loss to k_1 type

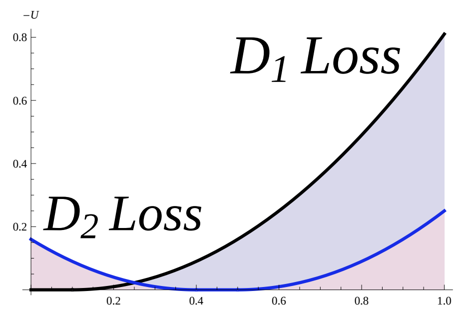

Figure 4: Loss to k_2 type

For each type, D_1 yields the agent smaller loss in lower states and larger loss in higher states than D_2. As can be seen from figure 3 by comparing the shaded areas, type k_1 strictly prefers D_1 to D_2 and type k_2 strictly prefers D_2 to D_1: $\mathbb{E}^1 D_1 > \mathbb{E}^1 D_2$ and $\mathbb{E}^2 D_2 > \mathbb{E}^2 D_1$. Notice that menu $D_1 = [0.3, 0.5]$ is "lower" than menu $D_2 = [0.8, 0.9]$. Under these delegation sets, type k_1 is more sensitive to the losses from lower states than the losses from the higher states. Hence, type k_1 prefers D_1 to D_2 (a similar logic applies to the k_2 type). This sensitivity will be described more formally below by a single-crossing lemma (Lemma 4.3). This lemma implies that, under incentive compatibility, the delegation sets in a screening menu must "increase" with bias (the delegation set chosen by type i, D_i, must be lower than the delegation set chosen by type j, D_j, for $k_j > k_i$). We now show why, with interval menus, incentive-compatibility conflicts with optimality for the principal. Thus, in order for screening to be optimal, menus with sets containing gaps will have to be used.

3.2 Example 2: Interval IC Conflicts with Optimality

In this subsection, we provide an example of how to improve upon an incentive compatible menu of intervals with an interval pooling menu (a menu consisting of a single interval). The menu $m = \{D_1, D_2\}$ in Example 1 is incentive compatible, but is not optimal. Notice that the left endpoint of D_1, 0.3, is higher than $k_1 = 0.1$ and that the left endpoint of D_2 is also higher than $k_2 = 0.4$. By thickening each set to k_i (replacing D_1 with $D_1' = D_1 \cup [0.1, 0.3]$ and D_2 with $D_2' = D_2 \cup [0.4, 0.8]$) the principal could increase expected payoff from each type. For example, by thickening D_1, the principal would have a constant loss of $U(k_1) = -(0.1)^2$ over the interval $[0.1, 0.3]$, but this would be an improvement in expected utility. The shaded area in Figure 5 illustrates the gain from thickening[3]. In addition, notice that the gain to the principal from thickening the set to k_1 is greater than the gain from the agent (the shaded area is greater than the area below the curve of the agent's loss). This will be useful in proving that menus consisting only of intervals are optimal. The reasoning for thickening D_2 is similar.

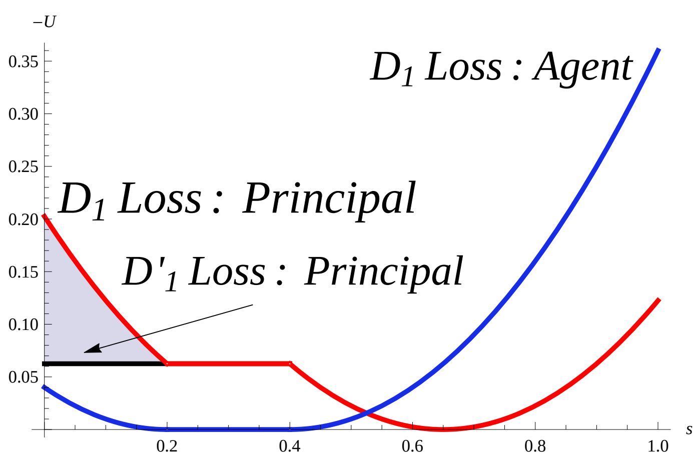

Figure 5: Gain from thickening to k_i.

[3]The plot in Figure 5 is for different parameter values in order to provide a clearer figure, but expected gain is similar for the parameters introduced earlier in this section.

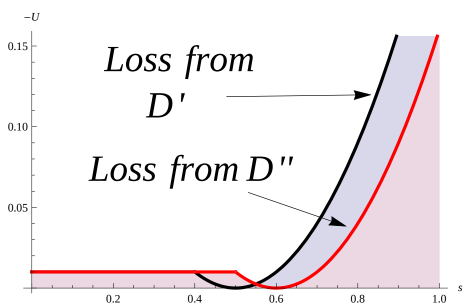

Figure 6: Loss to k_1 type

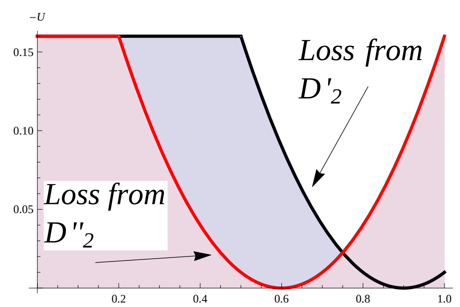

Figure 7: Loss to k_2 type

While the menu $m' = \{D_1' = [0.1, 0.5], D_2' = [0.4, 0.9]\}$ yields the principal a higher expected payoff, it is not incentive compatible (type k_1 prefers D_2' to D_1'). However, we can find a pooling menu that yields the principal strictly higher expected payoff from menu $m'' = \{[0.3, 0.6]\}$. $D_2' = [0.4, 0.9]$ is too high: thinning it from the right (replacing it with $D_2'' = [0.4, 0.6]$, illustrated in Figure 7) would increase the expected payoff of the principal. In addition, $D_1' = [0.1, 0.5]$ is too low: thickening it from the right (replacing it with $D'' = [0.1, 0.6]$, illustrated in Figure 6), would also increase the expected payoff of the principal. Hence, the menu $m'' = \{[0.1, 0.6]\}$ would yield the principal strictly higher expected payoff than the menu m', and, therefore, m. In section 4, we prove that this holds for all incentive compatible menus of intervals. Thus, using menus containing only intervals, screening will not be optimal. In order for screening to be optimal, we will need to introduce menus of sets containing gaps. The next example illustrates the impact of gaps on the principal's payoff.

3.3 Example 3: New Intuition for Interval Delegation Under Known Preferences

Example 2 restricted attention to a convex menus. The main result of this paper shows when this restriction is without loss of generality. In order to prove this result, we leverage intuition regarding the optimality of interval delegation in the setting of *known preferences*. In other words, we provide new intuition for the result of Melumad and Shibano (1991) regarding interval optimality. Figures 8-10 below provide a graphical illustration of this proof. In fact, the proof provides a generalization of Melumad and Shibano (1991). In order to discuss this intuition, we review some terms and variables.

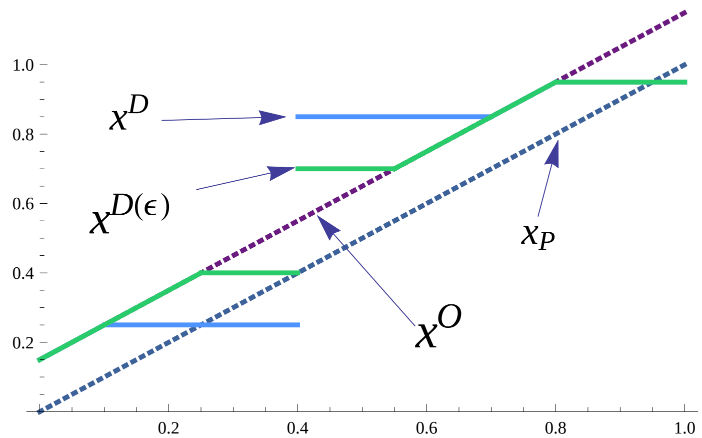

Figure 8: Thickening each set to k_i.

Let $x_P(s) = s$ denote the ideal action of the principal at state s (recall that the principal's bias is normalized to 0). Let $x^D(s)$ denote the delegation schedule for an agent with bias $k > 0$ under set delegation set D. Let $x^O(s) = s + k$ denote the ideal action of the agent at state s (we do not include a subscript i since there will only be one agent in this example). Let $V^P(s) = x^D(s) - s$ denote the *deviation* of the agent from the principal's ideal action at state s when the agent is restricted to delegation set D. Assume that D has a gap, $G = (l, h) \subseteq [k, 1 + k]$. Let $D(\epsilon)$ denote the set $D \cup (l, l + \epsilon] \cup [h - \epsilon, h)$. Figures 9 and 10 plot[4] the deviation and the distribution of the deviation (conditional on $s \in [l - k, h - k]$ for two sets: D and $D(\epsilon)$.

[4]In these plots, $D = [0.15, 0.25] \cup [0.85, 0.95]$, $G = (0.25, 0.85)$, $k = 0.15$, $\epsilon = 0.15$, $D(\epsilon) = [0.15, 0.40] \cup [0.7, 0.95]$.

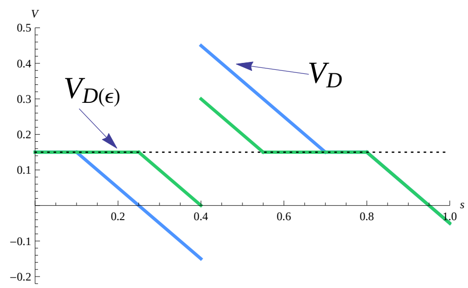

Figure 9: Plots of deviation

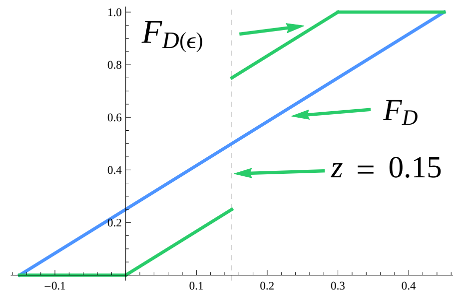

Figure 10: Plots of cdfs

Since $x_{D(\epsilon)}(s) = x_D(s)$ for all $s \notin [l-k, h-k]$, we condition the distribution on falling in this particular interval for the plot. Notice that the deviation of the agent under D is a mean-preserving spread of the deviation under $D(\epsilon)$. Hence, a principal with any strictly concave utility function would prefer to restrict the agent to $D(\epsilon)$ over D. In fact, the principal would prefer to completely fill in the gap in D (set $\epsilon = \frac{h-l}{2}$). Hence, under known bias, the optimal delegation set is convex. This is a result of Melumad and Shibano (1991). In fact, their proof of this result relied on quadratic loss utility. This proof holds for any strictly concave utility function.

This mean-preserving spread argument can be extended to the case of preference uncertainty. Notice that the mean-preserving spread argument holds independent of the bias k. Thus, in a setting with multiple types of agents, filling in the gap would increase expected payoff to the principal from *all types* of the agent. Hence, if a pooling menu had gaps, the expected payoff to the principal from each type of agent (k_i) could be increased by filling in the gaps of the delegation set (see Appendix D). In this way, we see that the optimal pooling menu is interval. In addition, because the introduction of a gap yields a mean-preserving spread, we see that the harm to the principal from gaps is the result of risk aversion. In the next example, we show that, with a principal with absolute value loss, screening is optimal. In addition, the principal receives a higher payoff by using a menu containing a set with a gap.

3.4 Example 4: When Screening is Optimal

In this subsection, we provide a simple example where screening is optimal. Notice that there is no difference between the types with regard to the information relevant to the payoff-relevant states. When the contract is offered, both types of agents have the same information of the future state. They have the same beliefs over the state, s. (In fact, the principal has the same beliefs regarding the payoff-relevant state). They differ only in their bias. We show that screening is optimal in this very simple setting of preference heterogeneity (but no payoff-relevant informational heterogeneity amongst the agents).

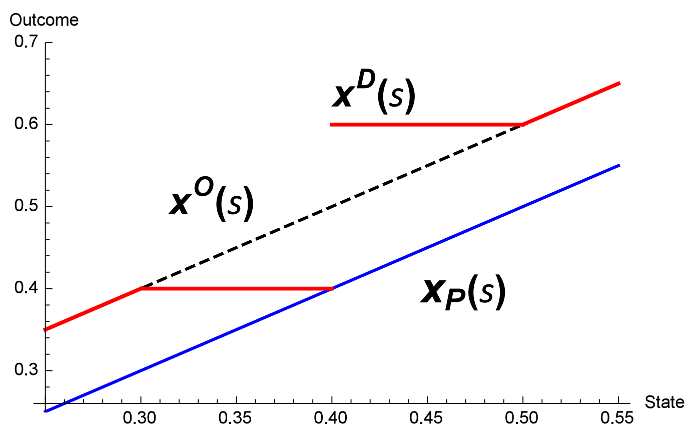

Figure 11: Delegation schedule with gap $(0.4, 0.6)$.

Let the biases be the same as before ($k_1 = 0.1$, $k_2 = 0.4$) and be equally likely (with probability $= 0.5$). But let the loss function be absolute value instead of quadratic. Thus, $U^P(x, s) = -|x - s|$, $U^1(x, s) = -|x - s - 0.1|$, and $U^2(x, s) = -|x - s - 0.4|$. Figure 11 (above) plots the delegation schedule generated by inserting the gap over $(0.4, 0.6)$. It shows the outcome chosen by the agent with low bias as a function of the state. Notice that, due to the gap, there is a jump in the delegation schedule. This will generate variance in the loss of the principal. Figure 12 plots this loss to the principal from inserting a gap over $[0.4, 0.6]$. Figure 13 plots the loss from filling in this gap. Notice that for absolute value loss, *the principal is indifferent between adding a gap and filling it in.* This will allow the principal to improve upon the optimal payoff obtained using a pooling menu.

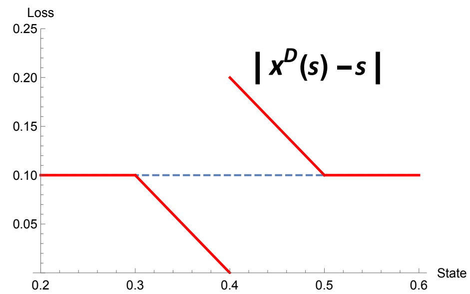

Figure 12: Plot with Gap

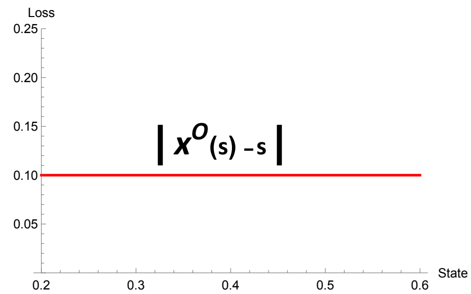

Figure 13: Plot with Filled Gap

The optimal payoff to the principal from a pooling menu is achieved by offering the delegation set $D^* = [0.1, 0.7]$ (this will follow from our alternative derivation of the Melumad and Shibano result). The principal would prefer to offer $[0.1, .8]$ to the less biased agent and $[0.4, .6]$ to the more biased agent, but this is not incentive compatible (since the more biased agent's set is strictly contained in the less biased agent's set). Hence, notice that the principal wishes to lower the highest action of the set selected by the more biased type. In addition, the payoff to the principal could be improved by offering the less biased agent $D_1 = [0.1, 0.4] \cup [0.6, 0.7]$ and the more biased agent $D_2 = [0.4, 0.69]$. The principal gains since the gap does not impact his expected payoff (see Figures 12 and 13) and since it lowers the highest action taken by the more biased agent. The gap induces the more biased agent to select the set with slightly lower highest action (which strictly increases the expected payoff of the principal). Finally, the exclusion of all actions in the interval $[0.1, 0.4)$ induces the less biased agent to select D_1 even though it has a gap. Thus, since a screening menu yields strictly higher payoff than that achieved by the optimal pooling menu. The optimal menu is screening and contains a set with a gap. In the next example, we show why this argument may hold for a range of more risk-averse principals, but may fail for more risk-averse principals.

3.5 Example 5: The Role of Concavity

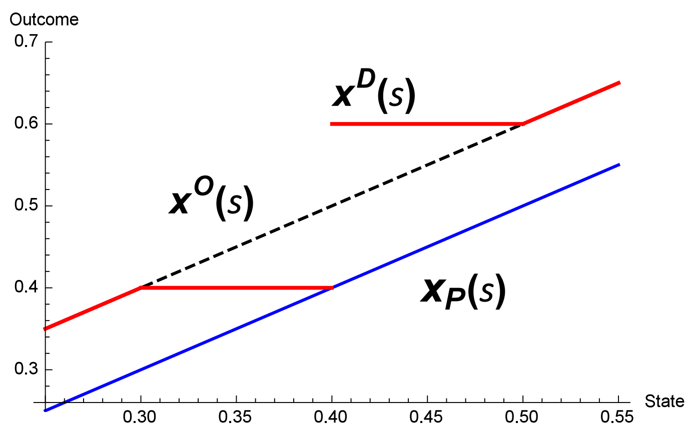

Figure 14: Delegation schedule with gap $(0.4, 0.6)$.

In the previous example, we saw why a screening menu with a gap can yield the principal higher payoff than the optimal pooling menu for absolute value loss. In this section we illustrate why the concavity of the loss function impacts the structure of the optimal contract. Figure 14 plots the outcomes chosen by an agent given a set with a gap over $(0.4, 0.6)$. It shows that the gap induces the agent to select its left endpoint for a fraction of the states, while selecting the right endpoint for another fraction. Thus, the gap introduces *variance.* Figures 15 and 16 illustrate the loss to the principal from the agent's choices given this set with a gap. For the states where the agent selects the left endpoints, the loss is lower. For states where the agent selects the right endpoint, the gap is higher. As preferences become more concave, the loss from the right endpoint states outweighs the gain from the left endpoint states. For absolute value, the loss is equivalent to the gain and the principal is indifferent between filling in (small) gaps and keeping them. For more concave losses, like quadratic (illustrated in Figure 16) the principal prefers to fill in losses.

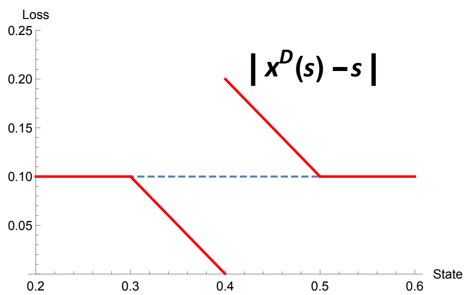

Figure 15: Gap Plot: Absolute Loss

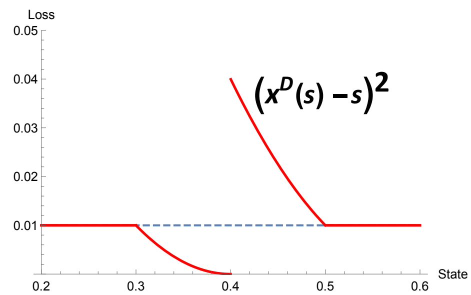

Figure 16: Gap Plot: Quadratic Loss

In the final example, we provide an illustration of why pooling is optimal when the principal is sufficiently risk averse. We show how to take a menu containing a set with a gap and turn it into a pooling menu that yields the principal a higher expected payoff.

3.6 Example 6: Improving on Nonconvex Menus

Let $U^P(x,s) = -(x-s)^2$ be the utility of the principal, $U^1(x,s) = -(x-s-0.2)^2$ be the utility of the less biased type, and $U^2(x,s) = -(x-s-.4)^2$ be the utility of the biased type. Let $m = \{D_1, D_2\}$, where $D_1 = [0.2, 0.3] \cup \{1\}$ and $D_2 = [0.7, 0.95]$. This menu is incentive compatible ($\mathbb{E}^1 D_1 > \mathbb{E}^1 D_2$, $\mathbb{E}^2 D_2 > \mathbb{E}^2 D_1$), see Figures 17 and 18. In addition, the largest point in D_1 is higher than that in D_2 and the lowest point in D_1 is lower than that in D_2. Hence, unlike the case of convex menus, incentive compatibility does not guarantee monotonicity in sets.

Notice that D_1 has a gap, $G = (0.3, 1)$. We illustrate that we can find a new incentive compatible menu, m', consisting only of intervals (a convex menu) that yields the principal higher expected payoff than m. Hence, as illustrated in the previous section, there would be a pooling menu that yields the principal higher

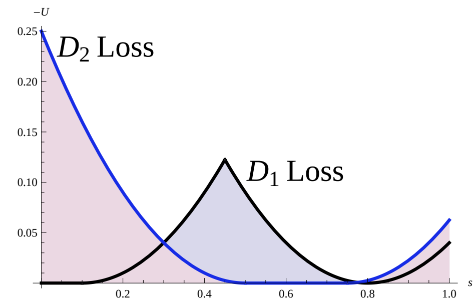

Figure 17: Loss to k_1 type

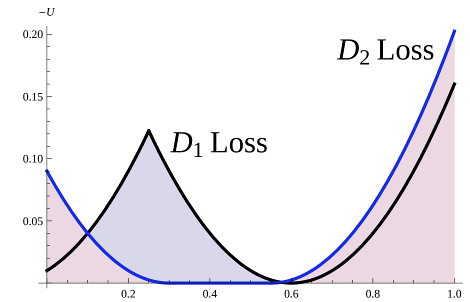

Figure 18: Loss to k_2 type

expected payoff than m' (and, hence, m). In order to do this, we fill in the gap in set D_1 (replace D_1 with $F = [0.2, 0.3] \cup \{1\} \cup (0.3, 1) = D_1 \cup G = [0.2, 1]$), and then thin it from the right (replace F with a new set, $D' = [.2, a]$, where $a < 1$) in such a way that:

- The unbiased agent is indifferent between the new (convex set), $D' = [0, a]$, and the original set, D_1: $\mathbb{E}^1 D' = \mathbb{E}^1 D_1$. Notice that Area A and Area B in Figure 19 are equal. In this case, $D' = [0.2, a]$, where $a \approx 0.746$.

- The new set D' yields lower expected payoff to the other agent than the original set, D_1: $\mathbb{E}^2 D' < \mathbb{E}^2 D_1 < \mathbb{E}^2 D_2$. Notice that Area C is *less* than Area D in Figure 20.

- The expected utility to the principal is higher from the new set than the original set: $\mathbb{E}_1^P D' > \mathbb{E}_1^P D_1$. See Figure 22.

Hence the new menu, $m' = \{D' = [0.2, a], D_2\}$, is incentive compatible and convex and yields the principal higher expected payoff. Thus, we have reduced the example to a case of convex menus, which can be improved upon by a pooling menu, as illustrated in Example 2.

The reason this menu yields the principal strictly higher expected payoff is that the *gain* in expected payoff to the principal from filling in a gap is at least as high as

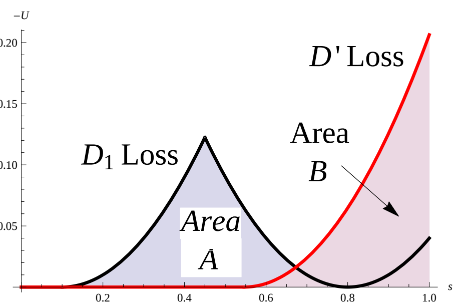

Figure 19: Loss to k_1 type

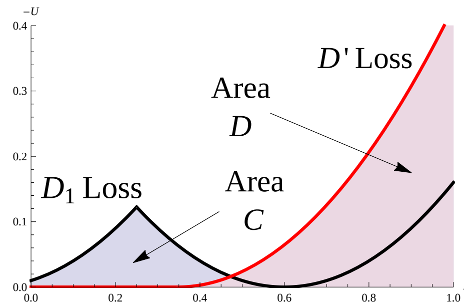

Figure 20: Loss to k_2 type

the gain in expected payoff of the agent (type 1) from filling in a gap. To illustrate this graphically, the union of Area A and Area A' in Figure 21 represents the gain to the agent from filling in a gap. The difference between areas H and G in Figure 22 represents the gain to the principal from filling in the gap. However, the loss in expected payoff from thinning the filled in set is lower for the principal than the agent (compare Figures 21 and 22). Thus, the "filled-in and thinned" set keeps type 1 at the same expected payoff (and preserves incentive compatibility), but strictly benefits the principal. The "fill-in and thin" variational argument will be used to prove that the restriction to convex menus is without loss for sufficiently risk-averse principals. Yet, the "fill-in and thin" argument will not hold for preferences that are less concave than quadratic loss. For absolute value preferences, the principal does not gain any expected utility from filling in gaps (and this holds for power loss functions whose exponent is strictly less than 2). Thus, pooling is not necessarily optimal when preferences are not sufficiently concave.

Note that this "fill-in and thin" modification works because the gap introduces a mean-preserving spread. Notice that, unlike the absolute value loss case, the principal was not indifferent between a small gap and no gap. The increase in concavity of the loss function made the mean-preserving spread more costly to the principal. Observe that in Figure 17, Area G (the region where the principal gains from introducing a

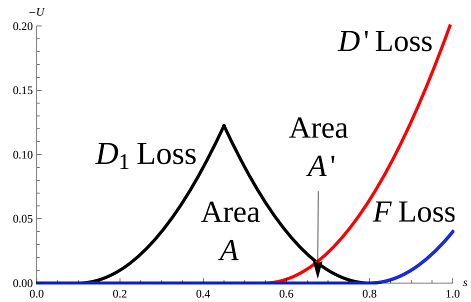

Figure 21: Loss to k_1 type.

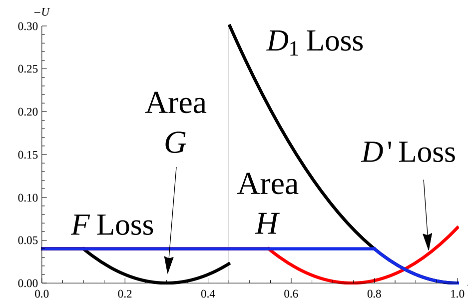

Figure 22: Loss to principal.

gap) is dwarfed by Area H (the region where the principal loses from introducing a gap). Recall that, for quadratic loss, these two areas are equal. The more concave the loss function, the greater is the difference between the areas and more costly is a gap to the principal. Thus, there is a parameter above which gaps are no longer effective. For exponents above this parameter, pooling is optimal. In the setting studied in this paper, this knife-edge case is quadratic loss. We now state the results precisely, beginning with those for convex menus.

4 Convex Menus

We first prove that if the principal is restricted to offer convex menus, then the optimal menu will be a pooling menu. The proof uses a variational argument. We show that for all incentive compatible screening menus, there is a pooling menu that yields the principal higher payoff. In order to find this pooling menu we present three lemmas regarding convex delegation sets. The first establishes that filling in the gap between zero and the lowest point of the set can only help the principal (no matter what type of agent). The second lemma will tell us how to select the maximal point of the pooling delegation set. The final lemma is a version of single-crossing and will allow us to order the sets in the proposed delegation menu. We use this order to construct a pooling menu that yields the principal higher expected payoff.

Recall that the principal's program is to maximize:

$$\mathbb{E}^P(\mathcal{M}) = \max_{m=\{D_1,\ldots,D_N\}} \mathbb{E}^P(m) = \max_{m=\{D_1,\ldots,D_N\}} \sum_{i=1}^{N} p_i \left(\int_0^1 U(x_i^{D_i}(s) - s) ds \right)$$

subject to the type incentive constraint for delegation sets (IC_k^i):

$$\mathbb{E}^i D_i \geq \mathbb{E}^i D_j, \forall\ i, j.$$

Note that $\mathbb{E}^P(m = \{D_1, \ldots, D_N\}) = \sum_{i=1}^N p_i \left(\int_0^1 U(x_i^{D_i}(s) - s) ds \right)$. Recall that

$$\mathbb{E}_i^P D_i = \int_0^1 U(x_i^{D_i}(s) - s) ds$$

is the *expected payoff to the principal from type i.* Since we first restrict attention to the case when all the D_i are convex we may let $D_i = [a_i, b_i]$. We have the following lemma:

Lemma 4.1. *(Down to k_i Lemma) Let $D = [a, b]$, where $a > k_i \geq 0$, then $\mathbb{E}_i^P D < \mathbb{E}_i^P D'$, when $D' = [k_i, b]$.*

Proof. See Appendix B. □

Thus, if the agent's type is known, Lemma 4.1 implies that an optimal convex delegation set is of the form $D = [k_i, b]$. Letting $\Psi_i(b) := \mathbb{E}_i^P[k_i, b]$, the expected payoff to the principal from the set $[k_i, b]$ when $1 + k_i \geq b \geq k_i$ would be:

$$\int_0^{b-k_i} U(k_i)ds+\int_{b-k_i}^1 U(b-s)ds = U(k)(b-k_i)+\int_{-k_i}^{\min\{1-b,0\}} U(s)ds+\int_0^{\max\{1-b,0\}} U(s)ds. \tag{4.1}$$

For $b < k_i$, $\Psi_i(b) = \int_0^1 U(b-s)ds$. Thus, we have the following lemma:

Lemma 4.2. *(Known Bias Optimum Lemma) Let the bias, k_i, of the agent be known by the principal. If $k_i \geq \frac{1}{2}$, then an optimal convex delegation set is $D_i^* = \{\frac{1}{2}\}$. If $k_i < \frac{1}{2}$, then an optimal convex delegation set is $D_i^* = [0, 1-b_i^*]$, where $b_i^* = 1-k_i$. For this case, $\Psi_i(\cdot)$ is strictly increasing from $[0, 1-k_i]$ and strictly decreasing from $[1-k_i, 1+k_i]$. Thus, $D_i^* = [0, q_i]$, where $q_i = \max\{\frac{1}{2}, 1-k_i\}$.*

Proof. See Appendix B. □

Lemmas 4.1 and 4.2 characterize the delegation sets the principal would choose if the bias of the agent was known. The forthcoming lemma is a single-crossing result that will characterize the menu the principal must offer when the bias of the agent is unknown. In order to derive this result, we first introduce some notation. If we have two intervals, one contained in the other, then any agent would always select the larger interval and the incentive constraints would fail. Hence, we define an ordering, $\succsim$, on non-nested sets. In this way, if $D_1 = [a_1, b_1]$ and $D_2 = [a_2, b_2]$, then:

(A) $a_1 < a_2$ and $b_1 < b_2$

(B) $a_1 > a_2$ and $b_1 > b_2$.

Note that, in a sense, under case (A), D_1 is "lower" than D_2. Thus, if (A) holds, we say that $D_1 \precsim D_2$. If (B) holds, we say that $D_1 \succsim D_2$.

Lemma 4.3. *(Single-Crossing Lemma) Let $D_i, D_j \in m$ such that m is IC_k and $k_i < k_j$. If $D_i \succeq D_j$ then $\mathbb{E}^i D_i \geq \mathbb{E}^i D_j$ then $\mathbb{E}^j D_i > \mathbb{E}^j D_j$, which violates the IC_k condition.*

Proof. See Appendix B. □

Lemma 4.3 states that if a lower bias agent prefers a higher interval to a lower interval, then a higher bias agent prefers the higher interval, as well.

Armed with these three lemmas, we are ready to outline the proof that the optimal convex menu is pooling (Proposition 4.4). If the bias of all types is $\geq \frac{1}{2}$, then by the Known Bias Optimum Lemma (Lemma 4.2), the optimal convex menu is $[0, \frac{1}{2}]$. This yields the principal the optimal expected utility from each type. Thus, we assume that at least one type has bias less than $\frac{1}{2}$ ($k_1 < \frac{1}{2}$).

The goal is to find a pooling menu that yields the principal higher expected payoff. We construct this pooling menu in three steps. First, we know that an IC_k menu must be of the form $m = \{D_1 = [a_1, b_1], \ldots, D_N = [a_N, b_N]\}$, where $b_i \leq b_j$ for all $j > i$. By Lemma 4.1, we know that if we were to replace each D_i by $D_i^0 = [k_i, b_i]$, we would improve the principal's expected utility. There is one problem though, the menu of D_i^0's may not be IC_k. However, we can replace this menu with a pooling menu (which is trivially incentive compatible). From single-crossing, we know that the b_i are increasing in i. In addition, from Lemma 4.2 (Known Bias Optimum Lemma), we know that the $q_i = 1 - k_i$ (the optimal end points under known bias) are decreasing in i. Thus, if some $b_i > q_i$, then $b_j > q_j$ for all $j > i$. In words, if type i's delegation is too large, all higher types' delegation sets are too large. Thus, by shrinking the delegation sets of all such types (and expanding the delegation sets of the types whose sets are not too large) we can achieve a pooling delegation set that yields the principal higher utility than the original menu m. We now state the result:

Proposition 4.4 (No Need to Screen: Convex Menus)**.** *If menus in a delegation set are restricted to contain only convex sets, then there exists an optimal pooling menu,* $m = \{P^*, P^*, \ldots, P^*\} = \{P^*\}$ *that is optimal.*

Proof. See Appendix B. □

Hence, if an organization is restricted to offering a convex menu of guidelines, then the organization should set the same guidelines for each member. We also note that incentive compatibility was only used to show that the sets are ordered ($b_i < b_{i+1}$ for all i). Once the sets are ordered, the argument does not use incentive compatibility. Thus, let $d_i^* = \max_{d \in D_i} d$. In this way, if we are given a menu $m = \{D_1, \ldots, D_N\}$ where (i) each D_i is convex and (ii) $d_i^* \leq d_{i+1}^*$ for all i (and strict inequality holds for, at least, $i = 1$), then we can use the argument in the proof to find a (convex) pooling menu that yields the principal strictly higher expected utility. Call a menu satisfying (i) and (ii) a *nice menu* (notice that a nice menu may not be IC_k). Hence, we have the following lemma:

Lemma 4.5. *Given a nice menu, m, then there exists a pooling menu (m') with convex delegation set $D' = [k_1, \gamma]$ ($\hat{m} = \{[k_1, \gamma]\}$) such that:*

$$\mathbb{E}^P(\hat{m}) = \sum_{i=1}^{N} p_i \mathbb{E}_i^P [k_1, \gamma] > \sum_{i=1}^{N} p_i \mathbb{E}_i^P D_i = \mathbb{E}^P(m). \tag{4.2}$$

Proof. Proof follows immediately from the proof of Proposition 4.4. □

This lemma will prove useful in the next section. Suppose an incentive-compatible menu, $m = \{D_1, \ldots, D_N\}$, yields the principal less expected payoff than a nice menu, $m^n = \{D_1^n, \ldots, D_N^n\}$:

$$\mathbb{E}^P(m^n) = \sum_{i=1}^{N} p_i \mathbb{E}_i^P D_i^n > \sum_{i=1}^{N} p_i \mathbb{E}_i^P D_i = \mathbb{E}^P(m). \tag{4.3}$$

Thus, Lemma 4.5 shows that we can find a pooling menu $\hat{m} = \{\hat{D}\}$ such that :

$$\mathbb{E}^P(\hat{m}) = \sum_{i=1}^{N} p_i \mathbb{E}_i^P \hat{D} > \sum_{i=1}^{N} p_i \mathbb{E}_i^P D_i^n = \mathbb{E}^P(m^n) > \mathbb{E}^P(m). \tag{4.4}$$

We state this as the following corollary:

Corollary 4.6. *Let m be an incentive-compatible menu. If there is a nice menu (that is not necessarily incentive-compatible), m^n, that yields the principal higher expected payoff as in equation (4.3), then there is a convex, pooling menu (a singleton delegation set composed of a convex set) that yields the principal strictly higher expected payoff as in equation (4.4).*

In this section, we restricted the analysis to convex menus. In the next section, we provide conditions for when this analysis is without loss. When it is without loss, we will do so by showing that, for each incentive compatible menu, there is a nice menu that yields the principal strictly higher expected payoff. Hence, by Corollary 4.6, there is a convex pooling menu that yields the principal strictly higher expected payoff.

5 Main Result

In this section, we characterize when it is optimal for the principal to pool. In other words, we show when the restricted analysis to convex menus is without loss. Note that when it is with loss, screening is optimal and the menu may include sets with gaps. We examine the setting where the principal and agent have power loss functions[5] of degree ≥ 1: $U(\cdot) = -|\cdot|^{\ell}$, where $\ell \in \mathbb{R}$ and $\ell \geq 1$. Notice that if $\ell = 2$, then loss functions are quadratic. In order to analyze nonconvex menus, we will need to analyze nonconvex sets. Hence, we will analyze sets with gaps. A *gap* in a delegation set, D, is an interval of the form (l, h) where $D \cap (l, h) = \emptyset$ and $l, h \in D$. It is shown in Appendix C that a gap exists in every nonconvex delegation set.

[5]It is also simple to show that the optimal contract under absolute value loss functions is similar to that of loss functions of degree strictly less than 2.

In this way, we can define the expected payoff increase from filling in a gap. This will be used in the proofs to characterize the optimal delegation menu (Proposition 5.4 and Proposition 5.5). This proof for Proposition 5.4 will require us to fill in gaps to convert (nonconvex) incentive-compatible menus into nice menus. We are now ready to outline the proof of the main result.

Our goal is to show that, for loss functions of power ≥ 2, for every incentive compatible menu $m = \{D_1, \ldots, D_N\}$, there is a nice menu yielding the principal higher expected payoff. Thus, by Corollary 4.6, we know that there is a convex pooling menu that yields the principal strictly higher payoff than m. In order to construct a nice menu, we first find an interval delegation set of the form $I_1 = [k_1, a_1^*]$ that gives the least biased agent the same payoff as in the original contract:

$$\mathbb{E}^1 I_1 = \mathbb{E}^1 D_1. \tag{5.1}$$

We then show that equation (5.1) implies that I_1 also satisfies the following inequalities:

$$\mathbb{E}^P I_1 \geq \mathbb{E}^P D_1. \tag{5.2}$$

and

$$\mathbb{E}^j I_1 \leq \mathbb{E}^j D_1, \tag{5.3}$$

for all $j > i$. Equation (5.2) states that this modified delegation set, I_1, yields the principal higher expected payoff than the original delegation set for agent 1, D_1. Equation (5.3) states that all higher types, $j > 1$, prefer their original delegation set, D_i, to the modified delegation set I_1 (and, hence, their original delegation sets, D_j, to I_1). We then repeat this construction of replacing the original delegation set with an interval starting at k_1 for each higher type ($j > 1$). This construction yields a nice menu that improves upon the original expected payoff of the principal. This holds because each modification of the menu improves upon the original expected payoff of the principal. Second, all the sets in the new delegation menu are intervals. Finally, by incentive compatibility and equation (5.3) we know that each type prefers their interval delegation set to those of all lower types. Hence, we have that $a_j^* \geq a_i^*$, for all $j > i$. In this way, the menu $\{I_1, \ldots, I_N\}$ is a nice menu and we have the following result:

Proposition 5.1. *If $\ell \geq 2$, for every nonconvex screening menu, there exists a convex pooling menu that yields the principal greater expected utility.*

Proof. See Appendix C. □

What is critical to the proof of this result is that gains to principal from filling in a gap are greater than gains to the agent from filling in a gap. In addition, in all settings, the principal also weakly gains expected payoff from filling in a gap. If $\ell \geq 2$, the expected gain to the principal from filling in a gap is greater than the expected gain to an agent from filling in a gap. This is formalized as Lemma 11.3 in Appendix C. In the quadratic loss setting ($\ell = 2$), the gain to the principal is exactly equal to the gain to the agent. Thus, for sufficiently concave loss functions, we find a nice menu that yields the principal higher expected payoff than the original menu. In this way, by Corollary 4.6, we know that there is a pooling menu that yields the principal higher expected payoff than the original menu. However, we still need to show the optimal pooling menu is convex.

In order to do this, we first show how to improve upon an arbitrary set with a gap. We shall show that the same modification strictly improves utility, *independent of the bias of the agent*. Thus, if there is a pooling menu with a gap, we can use this particular modification to raise the principal's expected payoff. The modification used in this section will completely fill in the gap. In other words, if D contains a gap, (l, h), the set $D' = D \cup (l, h)$ will yield the principal strictly higher expected utility. We call this modification *gap filling*. We state this in the following lemma:

Lemma 5.2. *(Gap Filling Lemma) Let D be a set with gap, $G = (l, h)$. Let $D(\epsilon, l, h) := D \cup [l, l+\epsilon) \cup (h-\epsilon, h]$, where $\epsilon \leq \frac{h-l}{2}$. Then for all $\epsilon \in [0, \frac{h-l}{2}]$, we have*

$$\mathbb{E}_i^P D = \int_0^1 U(x_i^D(s) - s)ds < \int_0^1 U(x_i^{D(\epsilon,l,h)}(s) - s)ds = \mathbb{E}_i^P D(\epsilon, l, h).$$

Hence, completely filling in the gap (replacing D with $D' = D \cup G$) would yield the principle higher utility.

Proof. See Appendix D. □

The intuition for this result is simple. The distribution of loss to the principal generated by a gap is a mean-preserving spread of distribution of the loss generated by a (partially) filled in gap. Thus, we have further intuition for Melumad and Shibano's (1991) result about the optimality of intervals. Gaps in delegation sets generate "riskier" lotteries for the principal than those generated by intervals. Hence, intervals are optimal.

In this way, as argued in the previous paragraph we have the following proposition:

Proposition 5.3 (Convex Optimal Pooling)**.** *If $\ell \geq 2$, given any distribution over N types of agents, their exists a convex set D_P^* that could serve as an optimal pooling set. All other optimal pooling menus differ from D_P^* on a set that will be played with probability zero in equilibrium.*

Proof. See Appendix D. □

Thus, Proposition 5.1 shows that for every nonconvex screening menu there is a pooling menu that yields the principal higher expected utility. Proposition 5.3 shows that, of the pooling menus, convex menus are optimal. Hence, convex menus are optimal and the result from section 4 is without loss for $\ell \geq 2$:

Proposition 5.4. *For $\ell \geq 2$, the restriction to convex sets is without loss: the optimal delegation menu is pooling and convex.*

Proof. This result is immediate from Propositions 5.1 and 5.3. $\square$

In the next result, we show when screening is optimal.

Proposition 5.5. *If $\ell < 2$ and biases are strictly less than $\frac{1}{2}$, the optimal contract is screening and certain agents receive gaps.*

We first provide intuition for the proof. In the optimal pooling menu, we know that the right endpoint is too high: the highest type is allowed to choose actions that are excessively high given this type's bias. We will introduce a gap in the optimal pooling set. We will then thin the high type's set from the right and thin it from the left (raise the left endpoint up to k_N). We will show that incentive compatibility is still preserved and the principal benefits from this variation. The principal will benefit since the gain from thinning the high type's set from the right[6] is "higher order" than the loss from introducing a gap to the delegation sets of all other types.

Proof. From above, we know the optimal pooling menu is achieved by an interval contract $D_{POOL} = [k_1, a^*]$. In addition, we know that $a^* > 1 - k_N$, where $1 - k_N$ is the optimal right boundary if the principal knew for sure that the agent had bias, k_N[7]. We show that there are $\epsilon, \delta > 0$ such that the following menu yields strictly higher payoff to the principal: $\{D'_L, D'_H\}$, where $D'_L = [k_1, k_N] \cup [k_N + \delta, a^*]$ and $D'_H = [k_N, a^* - \epsilon]$. δ is chosen such that

[6]The principal does not lose any payoff from thinning from the left in this case.

[7]If $a^* \leq 1 - k_N$, then the menu would not be optimal. The principal could strictly increase payoff by thickening the D_{POOL} from the right by a small amount.

$$\mathbb{E}^N D'_H = \mathbb{E}^N D'_L. \tag{5.4}$$

Thus, we can write δ as a function of ϵ, $\delta(\epsilon)$. Also, notice that for small enough ϵ, $\delta(\epsilon)$ is a strictly increasing function of epsilon[8]. Thus, types 1 though $N-1$ choose D'_1 and the highest type chooses D'_N.

In addition, notice that, in addition, to the lower right boundary, the less biased agents would lose an additional $\int_0^{k_N-k_i} U(s)ds$. Thus, for small enough ϵ, incentive compatibility is satisfied. It remains to show that the principal benefits.

Using calculations similar to those used in the proof of Proposition 5.4, the loss to the principal from inserting a gap from type i is:

$$PG_i(\epsilon) = \int_{k_i-\delta(\epsilon)}^{k_i+\delta(\epsilon)} U(s)ds - 2\delta(\epsilon)U(k_i). \tag{5.5}$$

The gain to the principle from lowering the endpoint of type N is:

$$T_N(\epsilon) = -\epsilon U(k_N) + \int_{1-a^*}^{1-a^*+\epsilon} U(s)ds. \tag{5.6}$$

Note that $T_N(\epsilon)$ is positive for small epsilon since $a^* > 1-k_N$, the right boundary of the pooling menu is higher than that chosen facing an agent with known bias k_N.

Thus, the screening menu improves upon the pooling menu if:

$$\sum_{i=1}^{N-1} p_i PG_i(\epsilon) + p_N T_N(\epsilon) > 0, \tag{5.7}$$

which holds if the derivative at ϵ is strictly positive for all small enough ϵ:

$$\sum_{i=1}^{N-1} p_i PG'_i(\epsilon) + p_N T'_N(\epsilon) > 0, \tag{5.8}$$

[8]And is continuous by the Implicit Function Theorem.

$$\iff 1 > \frac{\sum_{i=1}^{N-1} p_i P G_i'(\epsilon)}{-p_N T_N'(\epsilon)} = \sum_{i=1}^{N-1} \frac{p_i P G_i'(\epsilon)}{-G_N'(\epsilon)} \frac{-G_N'(\epsilon)}{-p_N T_N'(\epsilon)}, \tag{5.9}$$

where $G_N(\epsilon)$ is the loss[9] from a gap to type N and is given by:

$$G_N(\epsilon) = 2\int_0^{\delta(\epsilon)} U(s)ds. \tag{5.10}$$

Note that from equation (5.4), the loss from the gap to the most biased type must exactly equal the gain from a higher right boundary. Thus, by calculations used in the proof of Proposition 5.4, we have that:

$$G_N(\epsilon) = 2\int_0^{\delta(\epsilon)} U(s)ds = \int_{1+k_N-a^*}^{1+k_N-a^*+\epsilon} U(s)ds = L_N(\epsilon), \tag{5.11}$$

for all small enough ϵ. Hence, for all small enough ϵ,

$$G_N'(\epsilon) = L_N'(\epsilon). \tag{5.12}$$

Plugging in the above equality into equation (5.9) we get:

$$\sum_{i=1}^{N-1} \frac{p_i P G_i'(\epsilon)}{-G_N'(\epsilon)} \frac{-L_N'(\epsilon)}{-p_N T_N'(\epsilon)}. \tag{5.13}$$

Notice that both $L_N(\epsilon) = \int_{1+k_N-a^*}^{1+k_N-a^*+\epsilon} U(s)ds$ and $T_N(\epsilon) = -\epsilon U(k_N) + \int_{1-k_N}^{1-k_N+\epsilon} U(s)ds$ are bounded and nonzero for all ϵ small. Thus, the ratio:

$$\frac{-L_N'(\epsilon)}{p_N T_N'(\epsilon)} = \frac{-U(1+k_N-a^*+\epsilon)}{p_N\big(-U(k_N)+U(1-a^*+\epsilon)\big)} \tag{5.14}$$

is bounded for all ϵ small enough.

[9]Hence we multiplied by -1 to multiply and divide by a positive term.

Thus, it remains to show that $\frac{p_i P G_i'(\epsilon)}{G_N'(\epsilon)}$ converges to zero as ϵ converges to zero.

By L'Hopital's Rule (WLOG ignoring the constant P_N):

$$\lim_{\epsilon\to 0}\frac{PG_i'(\epsilon)}{G_N'(\epsilon)}=\lim_{\epsilon\to 0}\frac{PG_i(\epsilon)}{G_N(\epsilon)}=\lim_{\epsilon\to 0}\frac{\int_{k_i-\delta(\epsilon)}^{k_i+\delta(\epsilon)}U(s)ds-2\delta(\epsilon)U(k_i)}{2\int_0^{\delta(\epsilon)}U(s)ds}=\lim_{\delta\to 0}\frac{\int_{k_i-\delta}^{k_i+\delta}U(s)ds-2\delta U(k_i)}{2\int_0^{\delta}U(s)ds}, \tag{5.15}$$

where the last equality follows since $\delta(\epsilon)$ is strictly increasing and continuous in ϵ. In this way, with some abuse of notation, we can treat δ as a variable and not a function.

By repeated application of L'Hopital's rule we have:

$$\lim_{\delta\to 0}\frac{\int_{k_i-\delta}^{k_i+\delta}U(s)ds-2\delta U(k_i)}{2\int_0^{\delta}U(s)ds}=\lim_{\delta\to 0}\frac{U(k_i+\delta)+U(k_i-\delta)-2U(k_i)}{2U(\delta)} \tag{5.16}$$

$$=\lim_{\delta\to 0}\frac{U'(k_i+\delta)-U'(k_i-\delta)}{2U'(\delta)}=\lim_{\delta\to 0}\frac{U''(k_i+\delta)+U''(k_i-\delta)}{2U''(\delta)} \tag{5.17}$$

$$=\frac{\ell(\ell-1)\Big((k_i+\delta)^{\ell-2}+(k_i-\delta)^{\ell-2}\Big)}{\ell(\ell-1))(\delta)^{\ell-2}}. \tag{5.18}$$

But notice that:

$$\lim_{\delta\to 0}(\delta)^{\ell-2}=\infty, \tag{5.19}$$

since $1<\ell<2$. Therefore, since $\ell(\ell-1)\Big((k_i+\delta)^{\ell-2}+(k_i-\delta)^{\ell-2}\Big)$ is bounded as $\delta\to 0$,

$$\lim_{\delta\to 0}\frac{\ell(\ell-1)\Big((k_i+\delta)^{\ell-2}+(k_i-\delta)^{\ell-2}\Big)}{\ell(\ell-1))(\delta)^{\ell-2}}=0. \tag{5.20}$$

Hence, from equation (5.15)

$$\lim_{\epsilon \to 0} \frac{PG_i'(\epsilon)}{G_N'(\epsilon)} = 0. \tag{5.21}$$

The proof is done for ℓ such that $1 < \ell < 2$. For absolute value loss ($\ell = 1$), the above proof is simpler since $PG_i(\epsilon) = 0$ for all ϵ small ($\epsilon < k_1$) by linearity of the loss (and the result that gaps introduce a mean-preserving spread- recall Example 4 in section 3). Thus, since $T_N(\epsilon) > 0$, equation (5.7) holds and screening is optimal. $\square$

Notice that it was the relative gains from filling in a gap that drove the optimality result for $\ell \geq 2$. For less concave preferences, the principal can begin at the optimal pooling contact, introduce a small gap to the less biased types, lower the upper threshold for the most biased type, while preserving the IC constraints. Thus, less biased types may take higher actions than more biased types in equilibrium. The principal screens by placing a gap in the less biased types delegation sets. This discourages the more biased types from choosing these sets with higher actions. In order to incentivize the less biased types to select sets with gaps, the principal raises the smallest action of the more biased agents. To the more biased agents, this restriction of discretion over low actions is much less costly than it is to the less biased agents.

The next subsection describes the comparative statics of the optimal pooling menu. For $\ell \geq 2$, pooling is optimal. Thus, for these parameters, the next subsection describes the comparative statics of the optimal menu.

6 Comparative Statics of the Optimal Pooling Menu

Let $\mathbf{p}_N := (p_1, \ldots, p_N)$, where $\sum_{i=1}^{N} p_i \leq 1$ (the subscript N denotes the dimension of the vector- $\mathbf{p}_L$ would be an L-tuple). and $\mathbf{k}_N := (k_1, \ldots, k_N)$ (again, the subscript N denotes the dimension of the vector- $\mathbf{k}_L$ would be an L-tuple). Denote the optimal nonredundant pooling delegation set by $D^*(\mathbf{p}_N, \mathbf{k}_N)$. We have the following result:

Proposition 6.1. *Fix $\boldsymbol{p}_N$. The optimal pooling delegation set, $D^*(\boldsymbol{p}_N, \boldsymbol{k}_N)$ is weakly decreasing in $\boldsymbol{k}_N$. Formally,*

$$\boldsymbol{k}'_N \geq \boldsymbol{k}_N \Rightarrow D^*(\boldsymbol{p}_N, \boldsymbol{k}'_N) \subseteq D^*(\boldsymbol{p}_N, \boldsymbol{k}_N), \tag{6.1}$$

where $\boldsymbol{k}'_N \geq \boldsymbol{k}_N$ iff $k'_i \geq k_i$ for all $i \in \{1, \ldots, N\}$.

Proof. See Appendix E. □

Now, let $F_{(\mathbf{p}_N, \mathbf{k}_N)}(z)$ denote the cumulative density function of the distribution of the biases $\mathbf{k}_N$ under the probability distribution $\mathbf{p}_N$. We can extend the previous result:

Proposition 6.2. *The optimal pooling delegation set, $D^*(\boldsymbol{p}_N, \boldsymbol{k}_N)$ is weakly decreasing in first-order stochastic dominance ($\succsim_{1st}$). Formally,*

$$(\boldsymbol{p}'_N, \boldsymbol{k}'_N) \succsim_{1st} (\boldsymbol{p}_L, \boldsymbol{k}_L) \Rightarrow D^*(\boldsymbol{p}'_N, \boldsymbol{k}'_N) \subseteq D^*(\boldsymbol{p}_L, \boldsymbol{k}_L), \tag{6.2}$$

where $(\boldsymbol{p}'_N, \boldsymbol{k}'_N) \succsim_{1st} (\boldsymbol{p}_L, \boldsymbol{k}_L)$ iff $F_{(\boldsymbol{p}'_N, \boldsymbol{k}'_N)}(z) \leq F_{(\boldsymbol{p}_L, \boldsymbol{k}_L)}(z)$ for all $z \in \mathbb{R}$.

Proof. We provide a complete proof in Appendix E. The intuition of the proof is to show that if one lottery, a, over types (first-order) stochastically dominates another, b, then we can convert a into b through a sequences of monotonic adjustments to the bias and to the probabilities. □

Thus, if one draw of types is "more biased" (according to first-order stochastic dominance), then the principal will offer the riskier draw a smaller delegation set.

7 Example with Transfers

In order to show one of the consequences of the no transfers assumption, we provide an example of delegation with transfers and show a broad class of settings where the optimal menu is not pooling. The utility functions in this section are a special case of those in Ambrus and Egorov (2017). The utility of the prinicpal is $U^P(x, s, T) = -(x-s)^2 - T$. The utility of agent type i is $U^i(x, s, T) = -(x - s - k_i)^2 + T$. In this example, there are two types of agents. For agent 1, $k_1 = 0$ and we denote this agent as Agent U (unbiased). For the other agent, $0 < k < \frac{1}{2}$, we denote this agent as Agent B (biased). In this case, the principal offers the agent a menu[10] of delegation set and transfer pairs $m_T = \{(D_U, T_U), (D_B, T_B)\}$, where $(D_j, T_j) \in \mathcal{D} \times \mathbb{R}$. Let p denote the probability of an agent of type U (with $1 - p$ denoting the probability of type B). The timing of the game is the same as in Section 3.

Since there are no restrictions on transfers, the Principal's problem is different in this setting. In particular, there will now be individual rationality (IR) constraints. Thus, the Principal's problem is to maximize:

$$\max_{\{(D_U,T_U),(D_B,T_B)\}} p\big(\mathbb{E}^P_U D_U - T_U\big) + (1-p)\big(\mathbb{E}^P_B D_B - T_B\big), \tag{7.1}$$

subject to the IC constraints

$$\mathbb{E}^U D_U + T_U \geq \mathbb{E}^U D_B + T_B,\ \mathbb{E}^B D_B + T_B \geq \mathbb{E}^B D_U + T_U \tag{7.2}$$

and IR constraints

$$\mathbb{E}^U D_U + T_U \geq \bar{w}_U, \mathbb{E}^B D_B + T_B \geq \bar{w}_B, \tag{7.3}$$

where $\bar{w}_U, \bar{w}_B \in \mathbb{R}$.

[10] Once again, we can restrict attention to two-option menus by the Taxation Principle.

We can now find a broad class of settings where the optimal menu is not pooling.

Proposition 7.1. *Denote* $\tilde{w} := \int_0^k -s^2 ds$. *Let* $\bar{w}_B < \tilde{w}$. *Denote* $\mu := \tilde{w} - \bar{w}_B$ *and let* $0 > w_U > -\mu$. *For every optimal IC and IR compatible pooling menu, there is an IC and IR compatible separating menu that yields the principal strictly higher expected payoff.*

Proof. The proof of this result is provided in the online appendix. □

The intuition for the argument is that any noninterval pooling delegation set can be improved upon by an interval pooling delegation set. Thus, like the case without transfers, the optimal pooling menu is interval. In addition, the marginal expected payoff from lengthening the interval from the right to the biased type is higher than that to the low type. Finally, the marginal expected payoff gain from lengthening the interval from the right to the biased type has a higher magnitude than the expected marginal change to the principal. Thus, the principal can charge the biased type their expected return from lengthening the interval. This variation would preserve the biased type's utility (and, hence, IC and IR constraints), would preserve the IC (and IR) constraint of the unbiased type, and would yield the principal strictly higher expected payoff. Thus, the optimality of pooling is a consequence of the lack of transfers and not the nonmonotonicity of the utility functions.

Notice that the proof does not rule out optimality of pooling for certain parameter values[11]. If the outside option utilities are too high, the optimal menu could be pooling. However, the proof does show that even if the optimal menu is pooling, the delegation set is drastically different from the optimal set in the case without transfers. Either an optimal menu is not pooling, or it cannot be strictly contained in the unit interval.

[11] The conditions on the outside utility options ensure that the optimal pooling menu is strictly contained in the interval $[0, 1)$.

8 Conclusion

In this paper we have shown that under uncertainty over the preferences of agents, the optimal contract for delegation depends on the risk-aversion of the principal. When the principal is more risk-averse, we showed that there is no benefit to screening. In contrast, when the principal is less risk averse (or can diversify risk over outcomes), there is a benefit to screening. Surprisingly, under the optimal screening menu, less biased types choose a set of extreme actions. They do so because they strongly value the option to take small actions in addition to large actions. In other words, the less ideological actions select the extreme option set since they place a greater value on outcomes across the entire spectrum. The more biased agents are discouraged by the extreme options since they do not value small actions. In addition, screening is beneficial to the principal since the lower risk-aversion allows the principal to tolerate the additional variance from the extreme actions chosen by the less biased types. This paper provides a very simple characterization for when screening is beneficial. The key insight is that the benefits of screening rely on the degree of risk-aversion of the principal's preferences. For less risk-averse principals, screening may be optimal. Thus, the optimality of screening is not just a result of the knowledge of future payoff-relevant states. It may also occur when the agents have different biases and have *identical* knowledge of future states. Thus, in addition to knowledge, bias matters for screening.

The results in this paper suggest several directions for future work. First, it remains to generalize the results of this paper to settings where the states are not distributed uniformly, the bias is not constant, and the utility is not necessarily quadratic. One can also explore whether stochastic mechanisms yield the principal higher expected payoff than deterministic mechanisms. In addition, one may generalize the results of this paper to the case of more than 1 agent (potentially to analyze hierarchies). In addition, as in Amador and Bagwell (2012), one may consider alternate quasilinear utilities to incorporate the possibility of money burning in delegation. One may further consider (finite) repeated interaction to see if there is a screening menu that yields the principal strictly higher expected payoff than the optimal (repeated pooling) delegation menu.

9 Appendix A: Proofs for Results in Section 3

In this Appendix, we show that we can replace any set in a menu with a nonredundant set (Corollary 9.7). In order prove this, we will need knowledge of the properties of the delegation schedule (Lemmas 9.1-9.4) and a result that, given a compact delegation set, the set of outcomes chosen by the agent is compact (Lemma 9.5). These will directly imply the nonredundancy result.

Lemmas 9.1-9.4 are the delegation schedule analogs of *delegation rule* lemmas in Proposition 1 of Melumad and Shibano (1991). Notice that $U^i = U^P$, for $k_i = 0$. In addition, U^i is single-peaked (for each s, there is an x such that $\frac{\partial U^i}{\partial x}\big(x - s - k_i\big) = 0$), $\frac{\partial^2 U^i}{\partial x^2}\big(x - s - k_i\big) < 0$, and $\frac{\partial^2 U^i}{\partial x \partial s}\big(x - s - k_i\big) > 0$. These are the conditions on the utility function for Proposition 1 of Melumad and Shibano (1991). Hence, we can cite a few results of Proposition 1 from their paper:

Lemma 9.1 (Delegation schedules are weakly increasing)**.** *For all $i \in \mathcal{N}$, $x_i^D(s)$ is weakly increasing in s and the only discontinuities of it are jump discontinuities.*

Thus, we know that for all D and i, $x_i^D(s)$ is weakly increasing and, hence, has only jump discontinuities. Let $x_i^{D+}(s) = \lim_{r \to s^+} x_i^D(s)$ and $x_i^{D-}(s) = \lim_{r \to s^-} x_i^D(s)$. By part (iii) of Proposition 1 of Melumad and Shibano (1991) we have:

Lemma 9.2. *At a point of discontinuity, $\tau \in [0, 1]$, of x_i^D, we have that: (a) $|x_i^{D+}(\tau) - \tau - k_i| = |x_i^{D-}(\tau) - \tau - k_i|$. (b) $x_i^D(\tau) \in \{x_i^{D-}(\tau), x_i^{D+}(\tau)\}$.*

In addition, we have the following corollary:

Corollary 9.3. *If τ is a point of discontinuity of the delegation schedule,*

$$x_i^{D-}(\tau) < \tau + k_i < x_i^{D+}(\tau). \tag{9.1}$$

Finally, we know that the function x_i^D achieves both the right and left hand limits at a point of discontinuity (though not at the same point):

Lemma 9.4. *If τ is a point of discontinuity, there exist $s_1, s_2 \in [0,1]$ such that $x_i^D(s_1) = x_i^{D+}(\tau)$ and $x_i^D(s_2) = x_i^{D-}(\tau)$.*

Proof. By Lemma 9.2, we know that $x_i^D(\tau) \in \{x_i^{D-}(\tau), x_i^{D+}(\tau)\}$. W.L.O.G. assume $x_i^D(\tau) = x_i^{D-}(\tau)$. By Corollary 9.3, we know that there is s_1 close enough to τ that $\tau + k_i < s_1 + k_i < x_i^{D+}(\tau)$. In addition, we know that $x_i^{D+}(\tau) \in D$, by compactness (and, therefore, closure) of D. Also, we know that $\Big(x_i^{D-}(\tau), x_i^{D+}(\tau)\Big) \cap D = \emptyset$. Otherwise, $x_i^D(\tau)$ is not optimal. By Lemma 9.2, $|x_i^{D-}(\tau) - \tau - k_i| = |x_i^{D+}(\tau) - \tau - k_i|$. Thus, $|x_i^{D-}(\tau) - s_1 - k_i| > |x_i^{D+}(\tau) - s_1 - k_i|$. Hence, $x_i^{D+}(\tau) = x_i^D(s_1)$. $\square$

Lemma 9.5. *For all $D \in \mathcal{D}$ and $i \in \mathcal{N}$, $I(x_i^D)$ is compact.*

Proof. We prove this by showing that the set $I(x_i^D)$ is bounded and closed. Hence, by the Heine-Borel Theorem it is compact.

Step 1: $I(x_i^D)$ is bounded.

Proof. First, since x_i^D weakly-increasing by Lemma 9.1. Thus, its range is bounded: $x_i^D(0) \leq x_i^D(s) \leq x_i^D(1)$ for all $s \in [0,1]$. $\square$

Step 2: Admissible sets are closed.

Proof. We prove this step by showing that the complement of $I(x_i^D)$ is open.

Let $q \in \mathbb{R}$, where q is in the complement of $I(x_i^D)$, $I(x_i^D)^C$. Then, $\exists \epsilon > 0$ such that $(q - \epsilon, q + \epsilon) \subseteq I(x_i^D)^C$. Otherwise, q would be a right or left hand limit of x_i^D by Lemma 9.1. But then, by Lemma 9.4, $q \in I(x_i^D)$. But this is a contradiction. Thus, $I(x_i^D)^C$ is open and $I(x_i^D)$ is closed. $\square$

Since $I(x_i^D)$ is closed and bounded, by the Heine-Borel Theorem it is compact. $\square$

Corollary 9.6. *For all $D \in \mathcal{D}$ and $i \in \mathcal{N}$, $I(x_i^{I(x_i^D)}) = I(x_i^D)$.*

The following result combines this lemma and corollary together and allows us to reduce attention to nonredundant sets:

Corollary 9.7. *For every set $D \in \mathcal{D}$ and $i \in \mathcal{N}$, there is a nonredundant set $D' = I(x^D) \in \mathcal{D}$ such that $I(x_i^{D'}) = I(x_i^D)$ and $x_i^{D'}(s) = x_i^D(s), \forall s \in [0,1]$.*

10 Appendix B: Proofs for Results in Section 4

Proof of Lemma 4.1 (Down to k_i Lemma):

Proof. If $a \geq 1 + k_i$, then $U(x_i^D(s) - s) < U(k_i)$ for all $s \in [0,1]$. Thus,

$$\mathbb{E}_i^P D = \int_0^1 U(x_i^D(s) - s)ds < \int_0^1 U(k_i)ds = \mathbb{E}_i^P D'. \tag{10.1}$$

If $a < 1 + k_i$, then

$$E_i^P D = \int_0^1 U(x_i^D(s) - s)ds = \int_0^{a-k_i} U(a_i - s)ds + \int_{a-k_i}^1 U(x_i^D(s) - s)ds \tag{10.2}$$

$$= \int_{k_i}^a U(s)ds + \int_{a-k_i}^1 U(x_i^D(s) - s)ds < \int_0^{a-k_i} U(k_i)ds + \int_{a-k_i}^1 U(x_i^D(s) - s)ds = \mathbb{E}_i^P D' \tag{10.3}$$

since $a > k_i$. Thus, D is not optimal. □

Proof of Lemma 4.2:

Proof. Notice that if $k_i \geq \frac{1}{2}$, then $\Psi_i(k_i) > \Psi_i(b), \forall\, b > k_i \geq \frac{1}{2}$. In addition, for strictly concave $U(\cdot)$, $\text{argmax}_{b \in [0,1]} \int_0^1 U(b - s)ds = \frac{1}{2}$. Thus, for $k_i \geq \frac{1}{2}$, an optimal convex delegation set is $[0, \frac{1}{2}]$.

If $k_i < \frac{1}{2}$ and $b < k_i$, then $\Psi_i(k_i) > \Psi_i(b)$ because the agent only takes one action for $b \leq k_i$.

But over $b \in (k_i, 1]$, $\Psi(b)$ is differentiable and strictly concave. Hence, $\Psi_i(b)$ is maximized by b^* such that $\Psi_i'(b^*) = U(k) - U(1-b^*) = 0$. Hence, $b^* = 1 - k_i$ (and the FOC is sufficient because of the strict negative sign of $\Psi_i''(b)$). Thus, we know that $\Psi_i(\cdot)$ is strictly increasing until b^* and decreasing after. $\square$

Proof of Lemma 4.3 (Single-Crossing Lemma):

Proof. Assume that $a_i > a_j$ and $b_i > b_j$. WLOG we can assume (by the nonredundancy result) that $b_i \leq 1 + k_i$ and $a_j \geq k_j$.

$$\mathbb{E}^j D_j = \int_0^{a_j - k_j} U(a_j - s - k_j)ds + \int_{b_j - k_j}^{1} U(b_j - s - k_j)ds = \int_0^{a_j - k_j} U(s)ds + \int_0^{1 + k_j - b_j} U(s)ds. \tag{10.4}$$

In addition, we have:

$$\mathbb{E}^j D_i = \int_0^{a_i - k_j} U(a_i - s - k_j)ds + \int_{b_i - k_j}^{1} U(b_i - s - k_j)ds = \int_0^{a_i - k_j} U(s)ds + \int_0^{1 + k_j - b_i} U(s)ds. \tag{10.5}$$

The IC_k condition implies that $\mathbb{E}^j D_j \geq \mathbb{E}^j D_i$

$$\iff \int_{1+k_j-b_i}^{1+k_j-b_j} U(s)ds \geq \int_{a_j-k_j}^{a_i-k_j} U(s)ds \iff \int_{1+k_i-b_i}^{1+k_i-b_j} U(s)ds > \int_{a_j-k_i}^{a_i-k_i} U(s)ds, \tag{10.6}$$

since $k_i < k_j$ and $U(\cdot)$ is strictly concave. Thus, $\mathbb{E}^i D_j > \mathbb{E}^i D_i$ and the IC_k condition is violated. $\square$

Proof of the No Need to Screen Result for Convex Menus (Proposition 4.4):

Proof. If there is no type with bias strictly less than $\frac{1}{2}$, then the principal can achieve optimal expected payoff by offering the menu $m = \{\frac{1}{2}\}$. If there is a type with bias strictly less than $\frac{1}{2}$, then by the Single-Crossing Lemma (Lemma 4.3), we know that if a menu $m = \{D_1 = [a_1, b_1], \dots, D_N = [a_N, b_N]\}$ satisfies the IC_k constraints, we need

$$b_1 \leq b_2 \leq \cdots \leq b_N. \tag{10.7}$$

By Lemma 4.1 (Down to k Lemma), we know that if we were to replace each D_i by $D_i^0 = [k_i, b_i]$ (forming the menu m^0), then $\mathbb{E}^P D_i^0 \geq \mathbb{E}^P D_i$ and therefore:

$$\mathbb{E}^P(m^0) = \sum_{i=1}^{N} p_i \mathbb{E}^P D_i^0 \geq \sum_{i=1}^{N} p_i \mathbb{E}^P D_i = \mathbb{E}^P(m). \tag{10.8}$$

Thus, while m^0 may yield the principal a higher expected payoff, it may not be incentive compatible. Thus, we will modify this menu further (making it both a incentive compatible and a pooling menu).

By the Known Bias Optimum Lemma (Lemma 4.2) we know that the optimal complete information (over types) delegation set for type i is equal to $D^{i*} = [0, q_i]$, where $q_i = \max\{\frac{1}{2}, 1 - k_i\}$. Thus, we have:

$$q_1 \geq q_2 \geq \cdots \geq q_N. \tag{10.9}$$

Roughly, these equations state that the optimal delegation sets (under complete information) are decreasing[12]. In contrast, the sets in a non-pooling, but IC_k menu, must be increasing. We will use this contrast to achieve a contradiction.

If $b_1 \geq q_1$, then the pooling menu with pooling set $[0, q_1]$ is IC_k (trivially) and yields the principal strictly higher expected payoff than m^0 (and m) from the Known

[12] Also, at least one of the inequalities must hold strictly. Otherwise, the principal can maximize expected payoff with $m = \{D\}$, where $D = \{\frac{1}{2}\}$.

Bias Optimum Lemma[13]. If $b_N \leq q_N$, then the pooling menu with pooling set $[0, q_N]$ is IC_k (trivially) and yields the principal strictly higher expected payoff than m^0 (and m) by the Known Bias Optimum Lemma. If $b_1 < q_1$ and $b_N > q_N$ we define the *Turning Point Type*, i^*:

$$i^* = \max\{i \in \mathcal{N} | b_i \leq q_i\}. \tag{10.10}$$

By equations (10.7) and (10.9) we know that i^* is well-defined because for all $i < i^*$, $b_i \leq b_{i^*} \leq q_{i^*} \leq q_i$ and for all $j > i^* + 1$, $b_j \geq b_{i^*+1} > q_{i^*+1} \geq q_j$.

WLOG assume $b_{i^*} < q_{i^*+1} \leq q_{i^*} < b_{i^*+1}$. Let the pooling delegation set be $D^* = [0, q_{i^*+1}]$. From the Known Bias Optimum Lemma, we know the pooling menu yields the principal strictly higher expected payoff than menu m^0 (and m). In addition, it satisfies IC_k (trivially). Hence, we have shown by contradiction that the optimal convex menu must be pooling. $\square$

11 Appendix C: Proofs of Results in Section 5

11.1 A Note on Gaps

In the class of games studied in this paper, the principal offers the agent a menu of sets. In order to discuss the types of sets the principal may find optimal to offer, we introduce a useful definition. We define carefully the definition of a *gap* in a delegation set. Assume that a delegation set, D, is not convex. Thus, if there exists a point y such that $x, z \in D$ and $x < y < z$, let $G_D^+(y) = [y, u_y^D)$, where

$$u_y^D = \sup_t \{t \in \mathbb{R} | t > y, [y, t) \cap D = \emptyset\}.$$

[13]Since the expected payoff to the principal is decreasing when the delegation set is too large. Thus, there is a gain from shrinking each D_i^0.

In addition, let $G_D^-(y) = [d_y^D, y)$, where

$$d_y^D = \inf_t \{t \in \mathbb{R} | y > t, (t, y] \cap D = \emptyset\}.$$

Thus, define $G_D(y) := G_D^-(y) \cup G_D^+(y)$, where $G_D(y)$ is the largest gap containing y.

We state one more lemma which will be useful later.

Lemma 11.1. *$d_y^D, u_y^D \in D$.*

Proof. This follows from the openness of the complement of a compact set (D is compact). Since compact sets are closed, these points must be contained in a compact set, D. $\square$

11.2 Proof of Proposition 5.1

In order to simplify the exposition of the proof, we make two assumptions in this Appendix: (1) that D_i is contained in $(-\infty, 1 + k_i]$ and (2) that the gap, G, is contained in $[k_i, 1+k_i]$. In Tanner (2018) we show that these assumptions are WLOG.

The proof will modify the set D_i twice. The first modification will fill in all of the gaps inside of D_i to obtain a new set, D_i''. The first modification will increase the expected payoff of the agent and increase the expected utility of the principal, but will effect the incentive compatibility conditions. In order to preserve these conditions, we will modify the set again. We will "thin" this set from the right so as to preserve the indifference of agent i, creating set, I_i. We will then show that all agents $j > i$ prefer there original sets D_j to I_i. Then we will demonstrate that I_i yields the principal a higher expected payoff than D_i. By performing these modifications for all types of agent, we will then create a nice menu. By Lemma 4.5, there will then exist a pooling menu that improves upon this nice menu. Thus, there will exist a pooling menu that improves on the original menu (which was not necessarily nice).

11.3 Aligned Thinning Lemma

Before we construct an I_i that satisfies equations (5.1), (5.2), and (5.3), we first prove a useful lemma. In order to prove this lemma, we introduce some notation. Let A be a closed (and bounded) set. Denote $\max A := \max_{x \in A} x$. We call a closed (and bounded) delegation set D *thick at the top* if there exists an $\epsilon > 0$ such that $[\max D - \epsilon, \max D] \subseteq D$. We call the set $D^-(\delta)$ a *δ-thinning of D* if $D^-(\delta) = D \cap (-\infty, \max D - \delta]$ and $\delta \geq 0$ is chosen so that $[\max D - \delta, \max D] \subseteq D$.

Lemma 11.2. *(Aligned Thinning Lemma) Let D be thick at the top, let $\delta > 0$, and let $D^-(\delta)$ be a δ-thinning of D. We have the following inequality:*

$$\Delta^j(D^-(\delta), D) < \Delta^i(D^-(\delta), D) < \Delta_i^P(D^-(\delta), D), \tag{11.1}$$

for all $j > i$.

In words, equation (11.1) states that thinning certain thick at the top delegation sets causes the least expected utility loss to the principal (it may even be a gain) and causes more expected utility losses for higher-bias types. Thus, thinning sets will prove to be a powerful variation that preserves incentive compatibility while maintaining expected utility gains to the principal.

Proof. We let $d^* = \max D$. We break this lemma into two results: (i) $\Delta^j(D^-(\delta), D) < \Delta^i(D^-(\delta), D)$ for all $j > i$ and (ii) $\Delta^i(D^-(\delta), D) < \Delta_i^P(D^-(\delta), D)$.

Proof of (i): We first show that $\Delta^j(D^-(\delta), D) < \Delta^i(D^-(\delta), D)$ for all $j > i$. We will prove this relation by writing $\Delta^j(D^-(\delta), D)$ as a function of k_j and show that it is decreasing in k_j: $\Delta^j(D^-(\delta), D) = \mathbb{E}^i D^-(\delta) - \mathbb{E}^i D$

$$= \int_{\max\{0, d^* - \delta - k_j\}}^{1} U(d^* - \delta - s - k_j) ds - \int_{\max\{0, d^* - k_j\}}^{1} U(d^* - s - k_j) ds \tag{11.2}$$

$\big($since $x_j^{D^-(\delta)}(s) = x_j^D(s)$ for all $s \in [0, \max\{0, d^* - \delta - k_j\}]$ and $U(0) = 0$ $\big)$

$$= \int_{1+k_j-d^*}^{1+k_j-d^*+\delta} U(s)ds - \int_{\max\{k_j-d^*,0\}}^{\max\{k_j+\delta-d^*,0\}} U(s)ds > \int_{1+k_l-d^*}^{1+k_l-d^*+\delta} U(s)ds - \int_{\max\{k_l-d^*,0\}}^{\max\{k_l+\delta-d^*,0\}} U(s)ds = \mathbb{E}^l D^-(\delta) - \mathbb{E}^l D, \tag{11.3}$$

for all $k_l > k_j$ since one of three cases holds: (A) $k_j + \delta - d^* < 0$ (B) $k_j - d^* < 0 \leq k_j + \delta - d^*$ (C) $k_j - d^* \geq 0$.

If (A) holds, then $\mathbb{E}^j D^-(\delta) - \mathbb{E}^j D = \int_{1+k_j-d^*}^{1+k_j-d^*+\delta} U(s)ds$, which is strictly decreasing in k_j. If (B) holds then $\mathbb{E}^j D^-(\delta) - \mathbb{E}^j D = \int_{1+k_j-d^*}^{1+k_j-d^*+\delta} U(s)ds - \int_0^{k_j+\delta-d^*} U(s)ds$. Differentiating with respect to k_j we get

$$U(1+k_j-d^*+\delta) - U(1+k_j-d^*) - \big(U(k_j+\delta-d^*) - U(0)\big) < U(1+k_j-d^*+\delta) - U(1+k_j-d^*) - \big(U(\delta) - U(0)\big) < 0 \tag{11.4}$$

since $k_j - d^* \leq 0 < k_j - d^* + \delta$, $U(0) = 0$, $U(\cdot)$ is strictly decreasing over $\mathbb{R}_+$, and $U(\cdot)$ is strictly concave. If (C) holds, the result follows by an argument identical that in (B). $\square$

Proof of (ii): $\Delta^i(D^-(\delta), D) < \Delta_i^P(D^-(\delta), D)$

We first reduce the analysis to a specific D and $\delta(D)$ (where the thinning will be a function of the set D). The reason we can do this is because if D is thick at the top and $D^-(\delta)$ is a δ-thinning, then $D^-(\delta)$ is also a closed and bounded set. In addition, recall that $\Delta_i^P(D^-(\delta), D) = \mathbb{E}_i^P D^-(\delta) - E_j^P D$. Hence, letting $\delta_0 = 0$, $\sum_{r=1}^L \delta_r = \delta$, and $S_r = \sum_{h=0}^r \delta_h$, then:

$$\Delta_i^P(D^-(\delta), D) = \Delta_i^P(D^-(\sum_{r=1}^L \delta_r), D) = \sum_{r=1}^L \Delta_i^P\Bigg(D^-(\delta_r + S_{r-1}), D^-(S_{r-1})\Bigg). \tag{11.5}$$

In this way, if $\Delta_i^P\Bigg(D^-(\delta_r + S_{r-1}), D^-(S_{r-1})\Bigg) > \Delta^i\Bigg(D^-(\delta_r + S_{r-1}), D^-(S_{r-1})\Bigg)$, for all $r \in \{1, \ldots, L\}$ then $\Delta_j^P(D^-(\delta), D) > \Delta^j(D^-(\delta), D)$.

Cases: (I) $\max D > 1, \max D - 1 \geq \delta > 0$. (II) $1 \geq \max D \geq k_i, \max D - k_i \geq \delta > 0$.

Proof for Case (I), $\max D > 1, \max D - 1 \geq \delta > 0$*:* In this case, by removing an interval of length δ, the principal's expected utility is strictly increased while the agent's expected utility is strictly decreased. It is increased for the principal since the principal prefers a choice of $d' < d$ for all $d, d' > 1$. For the agent, the preferences are reversed. Hence, for case (I), $\Delta_i^P(D^-(\delta), D) > \Delta^i(D^-(\delta), D)$. $\square$

Proof for Case (II), $1 \geq \max D > k_i, \max D - k_i \geq \delta > 0$

$$\Delta_i^P(D^-(\delta), D) = \mathbb{E}_i^P D^-(\delta) - \mathbb{E}_i^P D = \int_{1-d^*}^{1+\delta-d^*} U(s)ds - \delta U(k_i) \tag{11.6}$$

(since $x_i^{D^-(\delta)}(s) = x_i^D(s)$ for all $s \in [0, d^* - \delta - k_i]$, and D is thick at the top so $x_i^D(s) = s + k_i$ for all $s \in [d^* - \delta - k_i, d^* - k_i]$)

$$> \int_{1+k_i-d^*}^{1+k_i+\delta-d^*} U(s)ds = \int_{d^*-\delta-k_i}^{1} U(d^*-\delta-s-k_i)ds - \int_{d^*-k_j}^{1} U(d^*-s-k_i)ds = \mathbb{E}^i D^-(\delta) - \mathbb{E}^i D = \Delta^i(D^-(\delta), D). \tag{11.7}$$

Hence, for case (II), $\Delta_i^P(D^-(\delta), D) > \Delta^i(D^-(\delta), D)$. $\square$

Thus, from cases (I) and (II) and the argument at the beginning of case (ii) we know that for all D that are thick at the top, for any δ-thinning we have: $\Delta_i^P(D^-(\delta), D) > \Delta^i(D^-(\delta), D)$.

From case (i) we know that $\Delta^i(D^-(\delta), D) > \Delta^j(D^-(\delta), D)$ for all $j > i$ and from case (ii) we know $\Delta_i^P(D^-(\delta), D) > \Delta^i(D^-(\delta), D)$. Hence, we have completed the proof for the Aligned Thinning Lemma (Lemma 11.2). $\square$

Now we construct D_i'' and I_i.

11.4 Construction of I_i

Let $a_i = \max_{d \in D_i} d$ and $D_i'' = [k_i, a_i]$. In words, we filled in all of the gaps in D_i and thickened D_i from the left to k_i. This replacement strictly raises agent i's expected payoff. We will then replace D_i'' with the interval, $I_i = [k_i, a_i^*]$, such that agent i is indifferent between I_i and D_i (thus, $a_i^* < a_i$). Such an a_i^* exists by the Intermediate Value Theorem. Hence, I_i satisfies equation (5.1). It remains to show that it satisfies equations (5.2) and (5.3).

Equation (5.3) is the easier relation to prove. First, note that D_i'' fills in all gaps in D_i and thickens D_i to k_i if necessary. $\mathbb{E}^j D_i'' - \mathbb{E}^j D_i$ is just the expected gain to type $j > i$ from filling in all the gaps in D_i. The gain from a gap $G \subseteq [k_j, 1 + k_j]$ is the same to type j as it is to type i since $U(\cdot)$ is a function of the absolute value of distance from action d to $s + k_j$. Thus, the only nontrivial case where the change in expected payoff from filling in a gap is different is for the case when $G = (l, h)$, $h > k_j$, and $l < k_j$. In this case, the loss to the agent j over the gap is

$$\int_0^{\frac{h+l}{2}-k_j} U(l-s-k_j)ds+\int_{\frac{h+l}{2}-k_j}^{h-k_j} U(h-s-k_j)ds = \int_{k_j-l}^{\frac{h-l}{2}} U(s)ds+\int_0^{\frac{h-l}{2}} U(s)ds. \tag{11.8}$$

For type i, the loss over the gap is

$$\int_{\max\{0,l-k_i\}}^{\frac{h+l}{2}-k_i} U(l-s-k_i)ds + \int_{\frac{h+l}{2}-k_i}^{h-k_i} U(h-s-k_i)ds \tag{11.9}$$

$$= \int_{\max\{k_i-l,0\}}^{\frac{h-l}{2}} U(s)ds + \int_0^{\frac{h-l}{2}} U(s)ds < \int_{k_j-l}^{\frac{h-l}{2}} U(s)ds + \int_0^{\frac{h-l}{2}} U(s)ds. \tag{11.10}$$

Thus, the loss to type i is greater from this kind of gap. Hence, filling it in will increase the expected payoff of type i by more than that of type $j > i$. Hence, we have that $\Delta^i(D_i'', D_i) > \Delta^j(D_i'', D_i)$.

In addition $\Delta^i(I_i, D_i'') > \Delta^j(I_i, D_i'')$ by the Aligned Thinning Lemma (Lemma 11.2). Thus,

$$\mathbb{E}^j I_i - \mathbb{E}^j D_i = \Delta^j(I_i, D_i'') + \Delta^j(D_i'', D_i) < \Delta^i(I_i, D_i'') + \Delta^i(D_i'', D_i) = \mathbb{E}^i I_i - \mathbb{E}^i D_i = 0. \tag{11.11}$$

So $\mathbb{E}^j I_i < \mathbb{E}^j D_i$ for all $j > i$. This shows that equation (5.3)is satisfied by I_i. It remains to show that I_i satisfies equation (5.2).

Let $L_i^j(G)$ denote the loss over a gap, G of D, where $G = (l, h) \subseteq [k_i, 1 + k_i]$, to an agent j or a principal with agent i. Thus, $j \in \{P, 1, 2, \ldots, N\}$ and if $j \neq P$, then $j = i$:

$$\mathbb{E}_i^P D = \int_0^{l-k_i} U(x_i^D(s)-s)ds + \int_{l-k_i}^{\frac{h+l}{2}-k_i} U(l-s)ds + \int_{\frac{h+l}{2}-k_i}^{h-k_i} U(h-s)ds + \int_{h-k_i}^{1} U(x_t^D(s)-s)ds \tag{11.12}$$

$$= \int_0^{l-k_i} U(x_i^D(s) - s)ds + \int_{-k_i}^{-k_i+\frac{h-l}{2}} U(s)ds + \int_{-k_i-\frac{h-l}{2}}^{-k_i} U(s)ds + \int_{h-k_i}^{1} U(x_t^D(s) - s)ds \tag{11.13}$$

$$= \int_0^{l-k_i} U(x_i^D(s) - s)ds + \int_{-k_i-\frac{h-l}{2}}^{-k_i+\frac{h-l}{2}} U(s)ds + \int_{h-k_i}^{1} U(x_t^D(s) - s)ds \tag{11.14}$$

$$= \int_0^{l-k_i} U(x_i^D(s) - s)ds + \int_{k_i-\frac{h-l}{2}}^{k_i+\frac{h-l}{2}} U(s)ds + \int_{h-k_i}^{1} U(x_t^D(s) - s)ds. \tag{11.15}$$

When the gap is filled in, D becomes $D \cup G$. Thus, the loss to the principal is:

$$\mathbb{E}_i^P D \cup G = \int_0^{l-k_i} U(x_i^D(s)-s)ds + \int_{l-k_i}^{\frac{h+l}{2}-k_i} U(k_i)ds + \int_{\frac{h+l}{2}-k_i}^{h-k_i} U(k_i)ds + \int_{h-k_i}^{1} U(x_t^D(s)-s)ds \tag{11.16}$$

$$= \int_0^{l-k_i} U(x_i^D(s)-s)ds + (h-l)U(k_i) + \int_{h-k_i}^{1} U(x_t^D(s)-s)ds. \tag{11.17}$$

When the gap is filled in, D becomes $D \cup G$. By equations (11.15) and (11.17), the gain from filling in the gap to the principal, $\Delta_i^P(D \cup G, D)$ is:

$$\Delta_i^P(D \cup G, D) = \mathbb{E}_i^P D \cup G - \mathbb{E}_i^P(D) = (h-l)U(k_i) - \int_{k_i - \frac{h-l}{2}}^{k_i + \frac{h-l}{2}} U(s)ds. \tag{11.18}$$

The payoff to an agent with bias k_i is:

$$\mathbb{E}^i D = \int_0^{l-k_i} U(x_i^D(s)-s-k_i)ds + \int_{l-k_i}^{\frac{h+l}{2}-k_i} U(l-s-k_i)ds \tag{11.19}$$

$$+ \int_{\frac{h+l}{2}-k_i}^{h-k_i} U(h-s-k_i)ds + \int_{h-k_i}^{1} U(x_t^D(s)-s-k_i)ds$$

$$= \int_0^{l-k_i} U(x_i^D(s)-s-k_i)ds + \int_0^{\frac{h-l}{2}} U(s)ds + \int_{-\frac{h-l}{2}}^{0} U(s)ds + \int_{h-k_i}^{1} U(x_t^D(s)-s-k_i)ds \tag{11.20}$$

$$= \int_0^{l-k_i} U(x_i^D(s)-s-k_i)ds + 2\int_0^{\frac{h-l}{2}} U(s)ds + \int_{h-k_i}^{1} U(x_t^D(s)-s-k_i)ds \tag{11.21}$$

$$\mathbb{E}^i D \cup G = \int_0^{l-k_i} U(x_i^D(s) - s - k_i) ds + \int_{l-k_i}^{\frac{h+l}{2}-k_i} U(0) ds + \int_{\frac{h+l}{2}-k_i}^{h-k_i} U(0) ds + \int_{h-k_i}^{1} U(x_t^D(s) - s - k_i) ds \tag{11.22}$$

$$= \int_0^{l-k_i} U(x_i^D(s) - s - k_i) ds + \int_{h-k_i}^{1} U(x_t^D(s) - s - k_i) ds. \tag{11.23}$$

By equations (11.21) and (11.23), for agent type i, the gain from filling in the gap $\Delta^i(D \cup G, D)$ is:

$$\Delta^i(D \cup G, D) = \mathbb{E}^i D \cup G - \mathbb{E}^i D = -\int_{-\frac{h-l}{2}}^{\frac{h-l}{2}} U(s) ds. \tag{11.24}$$

Thus, the gain is clearly positive since $U(\cdot) \leq 0$. We now show that for power loss functions $\big(U(\cdot) = -|\cdot|^\ell$, where $\ell \in \mathbb{N}\big)$ and $\ell \geq 2$, the gain to the principal is *at least* as high as the gain to the agent. We state this in the next lemma:

Lemma 11.3. *If $U(\cdot) = -|\cdot|^\ell$, where $\ell \in \mathbb{R}$ and $\ell \geq 2$, then for $\Phi = \frac{h-l}{2}$:*

$$\Delta_i^P(D \cup G, D) \geq \Delta^i(D \cup G, D). \tag{11.25}$$

Proof. Notice that the gain to the principal from filling in a gap minus the gain to agent from filling in a gap is:

$$f(\Phi) := -2\Phi U(k) + \int_{k-\Phi}^{k+\Phi} U(s) ds - 2\int_0^{\Phi} U(s) ds \tag{11.26}$$

Notice that, $f(0) = f'(0) = 0$. Yet,

$$f''(\Phi) = U'(k + \Phi) - U'(k - \Phi) - 2U'(\Phi). \tag{11.27}$$

Hence,

$$f(\Phi) = f''(\xi), \tag{11.28}$$

for some $\xi \in (0, \Phi)$. Hence, $f(\Phi) > 0 \iff U'''(\Phi) < 0$.

Notice that this result relies on the fact that $U'(0) = 0$. It can be shown that screening may be optimal for absolute value loss functions.

□

For quadratic functions ($U(s) = -s^2$), it turns out that these two gains are equal since, (letting $\Phi = \frac{h-l}{2}$):

$$\Delta^i(D \cup G, D) = \int_{-\Phi}^{\Phi} s^2 ds = \frac{2}{3}\Phi^3. \tag{11.29}$$

In contrast, by equation (11.18) and simple calculation:

$$\Delta_i^P(D \cup G, D) = -(2\Phi)k_i^2 + \int_{k_i-\Phi}^{k_i+\Phi} s^2 ds = \frac{2}{3}\Phi^3 = \Delta^i(D \cup G, D). \tag{11.30}$$

Recalling from equation (11.24) that $\Delta^i(D \cup G, D) > 0$, equation (11.25) implies that $\Delta_i^P(D \cup G, D) > 0$.

Lemma 11.3 held for filling in one gap. But if we fill in multiple (all) gaps, the result would clearly hold. Thus, we have:

$$\Delta^P(D_i'', D_i) \geq \Delta^i(D_i'', D_i). \tag{11.31}$$

We also know from the Aligned Thinning Lemma (Lemma 11.2) that:

$$\Delta^P(I_i, D_i'') > \Delta^i(I_i, D_i''). \tag{11.32}$$

Hence by equations (11.31) and (11.32),

$$\mathbb{E}^P I_i - \mathbb{E}^P D_i = \Delta^P(I_i, D_i'') + \Delta^P(D_i'', D_i) \tag{11.33}$$

$$> \Delta^i(I_i, D_i'') + \Delta^i(D_i'', D_i) = \mathbb{E}^i I_i - \mathbb{E}^i D_i = 0.$$

Thus, I_i satisfies equation (5.2) and the proof of Proposition 5.1 is complete.

12 Appendix D: Proof of Proposition 5.3

Proof of Lemma 5.2 (Gap Filling Lemma):

There are two possible cases for the gap $G = (l, h)$: (a) $l \geq k_1$ or (b) $l < k_1$. We prove the lemma for case (a) here. The proof of the lemma for case (b) is contained in Tanner (2018).

In each case we show that the deviation generated by the unfilled set is a mean-preserving spread of the (partially) filled in set.

12.1 Proof of Case (a)

We track the deviation of the delegation schedule from the ideal point of the principal. In order to assist the exposition of this section, we again show Figures 23 and 24 from Section 3.3:

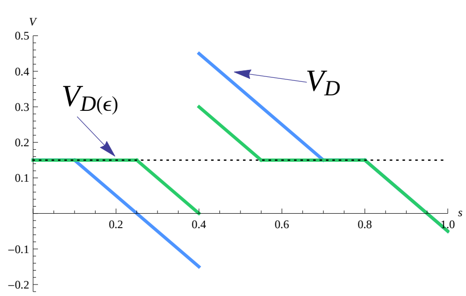

Figure 23: Plots of deviation

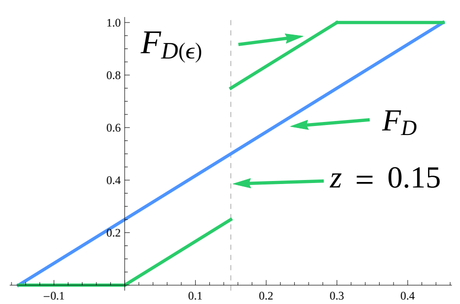

Figure 24: Plots of cdfs

The deviations are defined in section 3. These deviations are graphed in figure 23.

$$V_i^D(s) = x_i^D(s) - s = \begin{cases} l - s & \text{when } s \in [l - k_i, \frac{h+l}{2} - k_i) , \\ h - s & \text{when } s \in (\frac{h+l}{2} - k_i, h - k_i] . \end{cases} \tag{12.1}$$

Let $\epsilon \leq \frac{h-l}{2}$, $D(\epsilon, l, h) = D \cup [l, l+\epsilon] \cup [h-\epsilon, h]$. First, notice that

$$x_i^D(s) - s = x_i^{D(\epsilon,l,h)}(s) - s, \tag{12.2}$$

when $s \in [0, l-k_i] \cup [h-k_i, 1]$. When $s \in [l-k_i, h-k_i]$,

$$V_i^{D(\epsilon,l,h)}(s) = x_i^{D(\epsilon,l,h)}(s) - s, \tag{12.3}$$

where

$$V_i^{D(\epsilon,l,h)}(s) = \begin{cases} k_i & \text{when } s \in [l-k_i, l+\epsilon-k_i]\ , \\ l+\epsilon-s & \text{when } s \in [l+\epsilon-k_i, \frac{h+l}{2}-k_i)\ , \\ h-\epsilon-s & \text{when } s \in (\frac{h+l}{2}-k_i, h-\epsilon-k_i]\ , \\ k_i & \text{when } s \in [h-\epsilon-k_i, h-k_i]\ . \end{cases} \tag{12.4}$$

We argue that V_i^D is a mean-preserving spread of $V_{D(\epsilon,l,h)}(s)$, when s is uniformly distributed between the interval $[l-k_i, h-k_i]$.

Lemma 12.1. *$V_i^D(s)$ is a mean-preserving spread of $V_i^{D(\epsilon,l,h)}(s)$, when s is uniformly distributed between the interval $[l-k_i, h-k_i]$.*

Proof. For this proof, we fix l and h. Thus, we denote $D(\epsilon, l, h)$ by $D(\epsilon)$.

Part A: We first show that $V_i^D(s)$ and $V_i^{D(\epsilon)}(s)$ have the same mean.

Let $\eta = h - l$.

$$\mathbb{E}V_i^D = \frac{1}{\eta}\left(\int_{l-k_i}^{\frac{h+l}{2}-k_i}(l-s)ds + \int_{\frac{h+l}{2}-k_i}^{h-k_i}(h-s)ds\right) = \frac{1}{\eta}\left(\int_0^{\frac{h-l}{2}}(k_i-t)dt + \int_{-\frac{h-l}{2}}^{0}(k_i-t)dt\right), \tag{12.5}$$

where equality was obtained by a change of variables $s = t + l - k_i$ in the first integral and $s = t + h - k_i$ in the second integral.

But by (12.5) we have that $\mathbb{E}V_i^D = k_i$ and, thus, $\mathbb{E}V_i^D = \mathbb{E}V_i^{D(\epsilon)} = k_i$ since

$$\mathbb{E}V_i^{D(\epsilon)} = \frac{1}{\eta}\Big(\int_{l-k_i}^{l+\epsilon-k_i} k_i ds + \int_{l+\epsilon-k_i}^{\frac{h+l}{2}-k_i} (l+\epsilon-s)ds + \int_{\frac{h+l}{2}-k_i}^{h-\epsilon-k_i} (h-\epsilon-s)ds + \int_{h-\epsilon-k_i}^{h-k_i} k_i ds\Big) \tag{12.6}$$

$$= \frac{1}{\eta}\left(2\epsilon k_i + \int_0^{\frac{h-l}{2}-\epsilon} (k_i - s)ds + \int_{-\frac{h-l}{2}+\epsilon}^{0} (k_i - s)ds\right) = k_i, \tag{12.7}$$

where the equality between equations (12.6) and (12.7) follows a change of variables similar to that in (12.5).

Part B: We now show that for all $t \in \mathbb{R}$, *we have:*

$$\int_{-\infty}^{t} \big(F_D(s) - F_{D(\epsilon)}(s)\big)ds \geq 0.$$

(and the inequality holds strictly for $s \in (k_i - \frac{h-l}{2}, k_i + \frac{h-l}{2})$.

Let $F_D(x)$ denote the cdf of V_i^D and let $F_{D(\epsilon)}(x)$ denote the cdf of $V_{D(\epsilon,l,h)}$.

$$F_D(x) = \begin{cases} 0 & \text{when } s \in (\infty, k_i - \frac{h-l}{2}] \ , \\ \frac{1}{\eta}(s - (k_i - \frac{h-l}{2})) & \text{when } s \in [k_i, \frac{h-l}{2} + k_i] \ , \\ 1 & \text{when } s \in [\frac{h-l}{2} + k_i, \infty) \ . \end{cases} \tag{12.8}$$

$$F_{D(\epsilon)}(x) = \begin{cases} 0 & \text{when } s \in (-\infty, k_i - \frac{h-l}{2} + \epsilon] \ , \\ \frac{1}{\eta}(s - (k_i - \frac{h-l}{2} + \epsilon)) & \text{when } s \in [k_i - \frac{h-l}{2} + \epsilon, k_i) \ , \\ \frac{2\epsilon}{\eta} + \frac{1}{\eta}(s - (k_i - \frac{h-l}{2} + \epsilon)) & \text{when } s \in [k_i, k_i + \frac{h-l}{2} - \epsilon) \ , \\ 1 & \text{when } s \in [k_i + \frac{h-l}{2} - \epsilon, \infty) \ . \end{cases} \tag{12.9}$$

Notice that $\int_{-\infty}^{t}\big(F_D(s) - F_{D(\epsilon)}(s)\big)ds \geq 0$, for all $t \in (-\infty, k_i)$ since $F_D(s) \geq F_{D(\epsilon)}(s)$ for all such s (and the inequality is strict from $[k_i - \frac{h-l}{2}, k_i)$).

Define $\psi(\cdot)$ such that:

$$\psi(s) := F_D(k_i + s) - F_{D(\epsilon)}(k_i + s).$$

Notice that for $s \in [-\frac{h-l}{2}, \frac{h-l}{2}]$, we have $\psi(s) = -\psi(s)$. Thus, since $F_D(s) = F_{D(\epsilon)}(s)$ for $s \in (-\infty, k_i - \frac{h-l}{2}] \cup [k_i + \frac{h-l}{2}, \infty)$ and since $F_D(s) > F_{D(\epsilon)}(s)$ for all such $s \in [k_i - \frac{h-l}{2}, k_i)$, equations (12.8) and (12.9) we have that $\int_{-\infty}^{t} \big(F_D(s) - F_{D(\epsilon)}(s)\big)ds \geq 0$ for all $s \in \mathbb{R}$ (and the inequality is strict for $s \in [k_i - \frac{h-l}{2}, k_i + \frac{h-l}{2})$. □

Lemma 12.1 shows that filling in a gap (by any amount) strictly increases the utility of the principal (and, of course, the agent), *independent of the agent's bias!* This is summarized in the following lemma:

Lemma 12.2. *(Gap Filling Lemma- Case (a), $G \subseteq [k_i, 1 + k_i]$) Let $k_i \geq 0$. Let D be a set with gap, $G = (l, h)$, such that $G \subseteq [k_i, 1 + k_i]$. Then we know that $\mathbb{E}_i^P D = \int_0^1 U(x_i^D(s) - s)ds < \int_0^1 U(x_i^{D(\epsilon)}(s) - s)ds = \mathbb{E}_i^P D(\epsilon)$. Hence, completely filling in the gap (replacing D with $D' = D \cup G$) would yield the principle strictly higher utility:*

$$\mathbb{E}_i^P D < \mathbb{E}_i^P D' = \mathbb{E}_i^P (D \cup G). \tag{12.10}$$

Proof. We know that $x_i^D(s) = x_i^{D(\epsilon)}(s)$ for all $s \in [0, l] \cup [h, 1]$. Thus, we just need to show that

$$\int_l^h U(x_i^D(s) - s)ds < \int_l^h U(x_i^{D(\epsilon)}(s) - s)ds. \tag{12.11}$$

But from Lemma 12.1 that $V_i^D = x_i^D(s) - s$ is a mean-preserving spread of $V_i^{D(\epsilon)} = x_i^{D(\epsilon)}(s) - s$. So since $U(\cdot)$ is strictly concave we have that:

$$\frac{1}{\eta}\int_l^h U(x_i^D(s) - s)ds < \frac{1}{\eta}\int_l^h U(x_i^{D(\epsilon)}(s) - s)ds, \tag{12.12}$$

which is equivalent to equation (12.11).

Equation (12.10) is obtained by replacing D with $D(\epsilon)$, where $\epsilon = \frac{h-l}{2}$ (since $D(\frac{h-l}{2}) = D \cup G$). $\square$

13 Appendix E: Proof of Proposition 6.2

Proof of Proposition 6.2:

Proof. Let $I(u) = [0, u]$.

$$V(k_i, u) := V_i(u) = \mathbb{E}_i^P I(u) = \begin{cases} \int_0^1 U(u-s)ds & \text{when } u \in [0, k_i] \ , \\ (u-k_i)U(k_i) + \int_{u-k_i}^1 U(u-s)ds & \text{when } u \in [k_i, 1+k_i] \ , \\ U(k_i) & \text{when } u \geq 1+k_i. \end{cases} \tag{13.1}$$

Differentiating we get:

$$\frac{\partial V}{\partial u}(k_i, u) := \frac{d}{du}\big(\mathbb{E}_i^P I(u)\big) = \begin{cases} \int_0^1 U'(u-s)ds = U(u) - U(u-1) & \text{when } u \in [0, k_i] \ , \\ \int_{u-k_i}^1 U'(u-s)ds = U(k_i) - U(u-1) & \text{when } u \in [k_i, 1+k_i] \ , \\ 0 & \text{when } u \geq 1+k_i. \end{cases} \tag{13.2}$$

Notice that for all k_i, $V_i(u)$ is strictly increasing on $[0, \frac{1}{2}]$. Recall that if $k_1 \geq \frac{1}{2}$ the optimal delegation menu is $\{\frac{1}{2}\}$. Thus, we can restrict attention to the case where $k_1 < \frac{1}{2}$. Under this restriction, we know that $V_i(u)$ is strictly decreasing (for all i) for $u \in [1-k_1, 1+k_1]$ and is weakly decreasing for $s \in [1-k_1, \infty)$. Hence, we know that a u^* which optimizes $\sum_{i=1}^N p_i V_i(u)$ must be contained in $[k_1, 1-k_1]$. For each k_i, $V_i(u)$ is single-peaked. In addition, it is strictly quasi-concave over this interval. Hence, an optimal $u^* \in [k_1, 1-k_1]$ is unique. In addition, while it may not be strictly concave over $\mathbb{R}_+$, it is strictly concave over this interval. Thus, first-order conditions are sufficient to determine optimality. Thus, letting $m_u = \{I(u)\}$. $\sum_{i=1}^{N} p_i V(k_i, u) = \mathbb{E}^P(m_u)$, then u^* satisfies

$$\sum_{i=1}^{N} p_i \frac{\partial V}{\partial u}(k_i, u^*) = 0. \tag{13.3}$$

Notice that $\frac{\partial V}{\partial u}(k_i, u)$ is nonincreasing in k_i: if $k_i' \geq k_i$, then $\frac{\partial V}{\partial u}(k_i', u) \leq \frac{\partial V}{\partial u}(k_i, u)$. In addition, if the p_i are held fixed, and if $k_i' \geq k_i$, then the fact that $\frac{\partial V}{\partial u}(k_i, u)$ is nonincreasing in k_i' gives us:

$$\sum_{i=1}^{N} p_i \frac{\partial V}{\partial u}(k_i, u^*) \leq 0, \tag{13.4}$$

where u^* maximized the expression when biases where k_i and not k_i'. Thus, the optimal u is nonincreasing in k_i. The nonincreasing partial derivative in k_i also implies that if an $\epsilon > 0$ of probability is shifted from i to j ($p_i' = p_i + \epsilon$ and $p'j = p_j - \epsilon$) where $k_i > k_j$ ($k_j < k_i$), then the optimal u is nonincreasing (nondecreasing) in ϵ.

Hence, let $\mathbf{p}_N := (p_1, \ldots, p_N)$, where $\sum_{i=1}^{N} p_i \leq 1$ (the subscript N denotes the dimension of the vector- $\mathbf{p}_L$ would be an L-tuple). and $\mathbf{k}_N := (k_1, \ldots, k_N)$ (again, the subscript N denotes the dimension of the vector- $\mathbf{k}_L$ would be an L-tuple). Denote the optimal pooling delegation set by $D^*(\mathbf{p}_N, \mathbf{k}_N)$.

Let k_L be the highest value in the support of $F_{(\mathbf{p}_L, \mathbf{k}_L)}(z)$. First-order stochastic dominance implies that there exist $k_{N-j+1}', k_{N-j+2}', \ldots, k_N' \geq k_L$ such that:

$$F_{(\mathbf{p}_N', \mathbf{k}_N')}(k_{N-j+1}') \geq p_N. \tag{13.5}$$

(Notice that the rest of the points in the support of $F_{(\mathbf{p}_N', \mathbf{k}_N')}$ are $< k_L$.) Hold $\mathbf{p}_N'$ fixed, but replace $\mathbf{k}_N'$ with $\mathbf{k}_N^2$ (the 2 is a superscript not an exponent) such that $k_i^2 = k_i'$ for $i < N - j + 1$ and $k_i^2 = k_L$ for all $i \geq N - j + 1$. By Proposition 6.1 we know that $D^*(\mathbf{p}'_N, \mathbf{k}'_N) \subseteq D^*(\mathbf{p}'_N, \mathbf{k}^2_N)$.

Notice that $(\mathbf{p}'_N, \mathbf{k}^2_N) = (\mathbf{p}^2_{N-j+1}, \mathbf{k}^2_{N-j+1})$, where $p^2_i = p'_i$ if $i < N - j + 1$ and $p^2_{N-j+1} = \sum_{h=0}^{j-1} p_{N-h}$ and $k^2_i = k'_i$ if $i < N - j + 1$ and $k^2_{N-j+1} = k_L$. Then fix $\mathbf{k}^2_{N-j+1}$ but replace $\mathbf{p}^2_{N-j+1}$ with $\hat{\mathbf{p}}^2_{N-j+1}$ such that $\hat{p}^2_{N-j+1} = p^2_{N-j+1}$ if $N-j+1$ is the smallest index i such that $k^2_i \geq k_{L-1}$ (and denote this index by i^*_2). If not, let $\hat{p}^2_{N-j+1} = p_L$ and let $\hat{p}^2_{i^*_2} = p^2_{i^*_2} + \sum_{h=0}^{j-1} p_{N-h}$. For all other i, let $\hat{p}^2_i = p^2_i$. Thus, we just transferred probability from the highest value of of bias in the support $F_{(\mathbf{p}'_N, \mathbf{k}^2_N)}$ and transferred it to a lower value and left all other probabilities fixed. Remembering that $\frac{\partial V}{\partial u}(k_i, u)$ is nonincreasing we can conclude that:

$$D^*_P(\mathbf{p}^2_{N-j+1}, \mathbf{k}^2_{N-j+1}) \subseteq D^*_P(\hat{\mathbf{p}}^2_{N-j+1}, \mathbf{k}^2_{N-j+1}). \tag{13.6}$$

Continuing in this fashion, we have a finite sequence of $(\hat{\mathbf{p}}^2_{N-j+1}, \mathbf{k}^2_{N-j+1}), (\hat{\mathbf{p}}^3_{N-j+1}, \mathbf{k}^3_{N-j+1}), \ldots,$ $(\hat{\mathbf{p}}^q_{N-j+1}, \mathbf{k}^q_{N-j+1}) = (\mathbf{p}_L, \mathbf{k}_L)$, where

$$D^*_P(\mathbf{p}^r_{N-j+1}, \mathbf{k}^r_{N-j+1}) \subseteq D^*_P(\hat{\mathbf{p}}^s_{N-j+1}, \mathbf{k}^s_{N-j+1}), \tag{13.7}$$

if $s > r$. $\square$

References

[1] Alonso, N. and N. Matouschek. (2008). "Optimal Delegation", *Review of Economic Studies*, **75**: 259-293.

[2] Amador, M. and K. Bagwell. (2012). "Tariff Revenue and Tariff Caps", *American Economic Review: Papers and Proceedings*, **102**: 459-465.

[3] Amador, M. and K. Bagwell. (2013). "The Theory of Delegation with an Application to Tariff Caps," *Econometrica*, **81**(4): 1541-600.

[4] Amadaor, M. and K. Bagwell. (2016). “Regulating a Monopolist with Uncertain Costs without Transfers." Working Paper.

[5] Amador, M., Werning, I., and G-M. Angeletos. (2006). “Commitment vs. Flexibility", *Econometrica*, **74**: 365-396.

[6] Ambrus, A. and G. Egorov. (2017). “Delegation and Nonmonetary Incentives", *Journal of Economic Theory*, **171**: 101-135.

[7] Armstrong, M. (1995). “Delegation and Discretion", Discussion paper no. 9421, Munich Personal RePEc Archive, Munich.

[8] Avery, C. and M. Meyer. (2011). “Designing Promotion and Hiring Procedures with Biased Evaluators" (mimeo).

[9] Bertrand, M. and S. Mullainathan. (1999). “Is There a Discretion in Wage Setting? A Test Using Takeover Legislation", *RAND Journal of Economics* **30**, 535-554.

[10] Carrasco, v. and W. Fuchs. (2009). “Dividing and Discarding: A Procedure for Taking Decisions with Non-transferable Utility" (mimeo).

[11] Courty, P. and H. Li. (2000) “Sequential Screening", *Review of Economic Studies*, **67**, 697-717.

[12] Frankel, A. (2014). “Aligned Delegation," *American Economic Review*, **104**(1):66-83.

[13] Gilligan, T. and K. Krehbiel. (1987). "Collective Decisionmaking and Standing Committees: An Informational Rationale for Restrictive Amendment Procedures", *Journal of Law, Economics, and Organization* **3**: 287-335.

[14] Holmström, B. (1984). "On the Theory of Delegation", in M. Boyer and R. Kihlstrom (eds.) *Bayesian Models in Economic Theory* (New York: North-Holland).

[15] Jensen, M. (1986). "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers", *American Economic Review* **76**, 323-329.

[16] Kováč, E. and D. Krähmer. (2016). "Optimal Sequential Delegation", *Journal of Economic Theory* **163**: 849-888.

[17] Kováč, E. and T. Mylovanov. (2009). "Stochastic Mechanisms in Settings Without Monetary Transfers: The Regular Case", *Journal of Economic Theory* **144**: 1373-1395.

[18] Laffont, J-J. and D. Martimort (2002). *The Theory of Incentives: The Principal-Agent Model*, Princeton and Oxford: Princeton University Press.

[19] Lewis, T.R. and D.E.M. Sappington (1989). "Regulatory Options and Price-Cap Regulation", *The RAND Journal of Economics* **20(3)**: 405-416.

[20] Martimort, D. and A. Semenov. (2006). "Continuity in Mechanism Design Without Transfers", *Economics Letters* **93**: 182-189.

[21] Melumad, N. and T. Shibano. (1991). "Communication in Settings with No Transfers", *The RAND Journal of Economics* **22**: 173-98.

[22] Melumad, N. and T. Shibano. (1994). "The Securities and Exchange Commission and the Financial Accounting Standards Board: Regulation Through veto-Based Delegation", *Journal of Accounting Research* **32**: 1-37.

[23] Mylovanov, T. (2008). "Veto-Based Delegation", *Journal of Economic Theory*, **138**, 297-307.

[24] Sappington, D. (2002). “Price Regulation", in M. Cave, S. Majumdar, andI. Vogelsang, (eds.) *The Handbook of Telecommunications Economics, Volume I: Structure, Regulation, Competition.* (Amsterdam: Elsevier Science Publications).

[25] Szalay, Dezsö. (2005). “The Economics of Clear Advice and Extreme Options," *Review of Economic Studies*, **72**, 1173-1198.

[26] Tanner, Noam. (2015). "Optimal Delegation Under Uncertain Bias." PhD Dissertation.

[27] Tanner, Noam. (2018). "Optimal Delegation Under Uncertain Bias: Online Appendix." PhD Dissertation.

[28] Wolinsky, A. (2003). “Information Transmission When the agent's Preferences are Uncertain." *Games and Economic Behavior* **42**: 319:326.