Conflict and delegation

Ming Li[*] Sofia Moroni[†]

July, 2016
Very preliminary and incomplete, do not cite or circulate.

1. Introduction

In organizations or political institutions, it is commonplace that a person with the initial decision-making authority (“Principal") has to rely on the information held by another member of the organization, or in the case of political institutions, a politician or bureaucrat (“Agent"). Conventional wisdom has it that the principal would prefer that the agent have preferences that are congruent with her own. Viewed from two perspectives, it could be argued that this is even self-evident. From an information transmission perspective, if the principal keeps her decision-making authority and elicits advice from the agent, it is more difficult to trust reports provided by a more biased agent. From a delegation perspective, if the principal delegates the decision power to the agent, the latter will make decisions that are worse for the principal. In political writing, it is often argued that a democratic electoral system should aim at producing a political leader that best represents the preferences of the consituency that the political leader serves.

In this paper, we demonstrate the above arguments may have under-explored the principal’s ability to restrict the agent’s discretion and the multidimensional nature of typical political and economic environments. In particular, we address the following question. In scenarios where it is impossible to get perfect alignment of preferences on all dimensions between a principal and an agent, would it always be worthwhile to reduce misalignment in *one* dimension? We show that for a class of preferences, the answer is “No."

We model an environment in which the principal delegates a multi-dimensional decision to an agent. We generalize results found in the unidimensional literature and identify a condition under which delegation is valuable to the principal. In a setting with two-dimensional decisions where the ideal decisions for the

[*]Concordia University and CIREQ, 1455 Boulevard de Maisonneuve Ouest, Department of Economics, Montréal, Québec, CANADA H3G 1M8. Email: ming.li@concordia.ca.
[†]University of Pittsburgh, Department of Economics, 4714 Posvar Hall, 230 South Bouquet Street, Pittsburgh, PA 15260, USA. Email: moroni@pitt.edu.

principal and the agent are independent across dimensions, we show that, if their disagreement is fixed in one dimension, then as the disagreement widens, the principal's payoff is monotonically nondecreasing. Furthermore, as the disagreement becomes arbitrarily large, in the other dimension, the principal may ensure that the agent carry out her first best decision.

The intuition for our results is as follows. Suppose the principal and the agent have a systematic disagreement in one dimension, but their disagreement on the other dimension could be either nuanced or strong. Since the principal has the power to determine the set of actions allowable to the agent, she can limit the agent's choice on the first dimension by making sure it does not deviate too much from her own ideal. However, assuming that the agent has convex preferences, meaning his marginal loss from an action further away from his ideal is higher, the principal may manipulate the agent on the first dimension and ensure that the agent chooses the optimal action on the other dimension, even if they also have strong disagreement on the latter.

For a more concrete illustration, suppose that the principal is the writer of a constitution and that the agent is a political leader down the road. Political leaders tend to be more ambitious than the general public on social issues. On economic policy, however, even though their ideal decision may be quite different from the society's, depending on their ideology and financial interests. Given the constitution writer's ability to tie the political leader's hands, she actually prefers a political leader who is more ambitious on social issues and therefore gains more from realized ambitions and suffers more from thwarted ones. Given the higher effectiveness to thwart or facilitate the leader's ambition on social issues relative to the economic policy he adopts, the constitution writer is able to ensure the political leader's choice be closer to the ideal of society.

Related literature. Our paper belongs the literature on optimal delegation, which originated from Holmström (1977, 1984), and includes works byMelumad and Shibano (1991), Alonso and Matouschek (2008), Kováč and Mylovanov (2009), Martimort and Semenov (2006), and Goltsman et al. (2009) in the case of unidimensional delegation,[1] and Koessler and Martimort (2012) and Frankel (2015) in the case of multidimensional delegation.

Koessler and Martimort (2012) study multi-dimensional delegation where the agent has the same bias in each direction. In their model, the principal is always better off with a less biased agent. Frankel (2015) studies optimal multiple delegation when the principal faces an uncertain prior about the agent's preferences. He generalizes the "cap" formulation of the optimal delegation set in the unidimensional setting to a multidimensional one. Antić and Iaryczower (2016) analyze a delegation game where the principal not only chooses the set of actions, but also the scale at which each action can be implemented. They show that the principal may want to limit the scale of implementation for an agent biased towards too high an action but provide too much scale for an agent biased towards too low an action, which creates inefficiency. In the context of appointing bureaucracy chiefs, McCarty (2004) shows that when the appointment/removal power lies in different hands than the budgetary power, there could be inefficiency

[1] Krishna and Morgan (2008) study an environment where transfers are allowed.

due to the fact that the principal in charge of appointment/removal of the bureaucrat only appoints bureaucrats or replaces existing ones with those who have the same ideological positions as hers, neutralizing any possible opportunity to trade off between ideology and budget.

Levy and Razin (2007) identify the limits to strategic communication when the principal has to elicit multidimensional information from a single agent. Chakraborty and Harbaugh (2007, 2010) show that a limited amount of information transmission may be achieved in a class of preferences.

In the context of strategic information transmission, stronger incentive for information acquisition has often been identified as the reason why a principal may prefer a biased agent–Dewatripont and Tirole (1999) and Che and Kartik (2009) provide two such examples.

2. Model

A sender is informed about the realization of $\theta \in \Theta$ where $\Theta \subseteq \mathbb{R}^2$ is compact, convex set. The sender of type θ sends message $m(\theta)$, an element of a message space M, to the receiver and the receiver commits to an allocation $y(m)$ as a function of the message. We assume the message strategy of the sender $m : \Theta \to M$ is measurable. The receiver or principal's utility for action $y = (y_1, y_2) \in \mathbb{R}^2$ if the type of the sender is θ is given by

$$V(y,\theta) = \sum_{i=1}^{2} -\alpha_i (y_i^P(\theta) - y_i)^2$$

where $y_i^P : \Theta \to \mathbb{R}$ and $\alpha_i > 0$. Let A denote the diagonal matrix with $a_{ii} = \alpha_i$.

The receiver has a prior with distribution function F and continuous density f.

The utility of the sender of type θ from action y is given by[2]

$$U(y,\theta) = \sum_{i=1}^{2} -(\theta_i - y_i)^2.$$

If the strategy of the sender is a function $\sigma : \Theta \to \Delta M$. The receiver's expected utility from a contract $y : M \to \mathbb{R}^2$ is given by

$$V(y(\cdot)) = \mathbb{E}_{\sigma(\theta)} \left(\sum_{i=1}^{2} -\alpha_i (y_i^P(\theta) - y_i(m))^2 \right)$$

where the expectation is taken with respect to the distribution over types induced by the sender's strategy.

[2]These functional forms are more general than the preferences considered by Melumad and Shibano (1991) and less general than those considered by Alonso and Matouschek (2008) since we assume the sender has quadratic utility.

We next show that when searching for deterministic mechanisms, it is without loss to consider direct mechanisms.[3]

Lemma 1. *The principal's contracting problem can be stated as finding $y(\theta):\Theta\to\mathbb{R}^N$ such that*

$$\max_{y(\theta)} V(y)$$
$$\textit{subject to}$$
$$\theta \in argmax_{\theta'} \quad U(y(\theta'),\theta) \quad \textit{(IC-IP)}$$

Thus, $y(\theta)$ maximizes the expected utility of the receiver subject to the incentive compatibility of the informed party.

Proof. For every message space M the principal's problem can be written as

$$\max_{M,y(m),\sigma(\theta)} \mathbb{E}_{\sigma(\theta)}\left(\sum_{i=1}^{N} -\alpha_i (y_i^P(\theta)-y_i(m))^2\right)$$
$$U(y(m),\theta) \geq U(y(m'),\theta) \quad \forall m \in supp\,\sigma(\theta).$$

Let's now see that the solution of this program gives weakly less utility to the principal than the program in which, in addition to the message, the sender reveals her type to the receiver. The program in which the receiver reveals the type in addition to the original message can be written as,

$$\max_{M\times\Theta,y(m,\theta),\sigma(\theta)} \mathbb{E}_{\sigma(\theta)}\left(\sum_{i=1}^{N} -\alpha_i (y_i^P(\theta)-y_i(m,\theta))^2\right)$$
$$U(y(m,\theta),\theta) \geq U(y(m',\theta'),\theta) \quad \forall m \in supp\,\sigma(\theta).$$

The solution of the first program is always feasible in the second one since the principal can always choose to ignore the part of the message involving the sender's type. Thus, the second program must give the principal a weakly better payoff.

Suppose a contract $y(m,\theta)$ maximizes the receiver's utility.

Let $\sigma(\theta):\theta\to(\Delta M,\theta)$ denote the strategy of the sender. By IC a sender must be indifferent between all messages that are sent with positive probability. Also by the maximization of the receiver for each $(m,\theta),(m',\theta)\in supp\,\sigma(\theta)$ we must have $V(y(m,\theta),\theta)=V(y(m',\theta),\theta)$.

Note that the indifference curves of sender and receiver intersect in at most two points and thus the set $\{y(m,\theta)|(m,\theta)\in supp\,\sigma(\theta)\}$ contains at most two elements. We can assume without loss that $\sigma(\theta)$ has at most two elements in its support. Suppose $(1,0)$ is not the direction of the bias (if not we can make the same argument with direction $(0,1)$). Because $(1,0)$ is not the direction of the bias

[3] Alonso and Matouschek (2008) show a similar result in the one-dimensional setting.

the set $\text{argmax}_{(m,\theta)\in supp\,\sigma(\theta)} y_1(m,\theta)$ is a singleton. Let $(\bar{m},\theta) = \text{argmax}_{(m,\theta)\in supp\,\sigma(\theta)} y_1(m,\theta)$. Define $(y_1(\theta_1,\theta_2), y_2(\theta_1,\theta_2)) = (y_1(\bar{m},\theta), y_2(\bar{m},\theta))$. $(y_1(\theta_1,\theta_2), y_2(\theta_1,\theta_2))$ is Lebesgue measurable since $y_1(m,\theta)$ is measurable in the $M\times\Theta$ space. □

3. Valuable delegation

In this section we show that for a fairly general set of preferences the principal can always gain from delegating to an agent in 2-dimensional setting. Delegation will be valuable, in particular, even when the bliss points of the principal and the agent are arbitrarily distant. In contrast, in a one-dimensional setting delegation cannot be valuable when these bliss points are sufficiently apart. Our results suggest that it is strictly preferable to bundle two delegation decisions instead of delegating them independently. Furthermore, it also follows immediately from our results that delegation is valuable in an $N>2$ dimensional setting, since the principal can always bundle two decisions and allocated the rest independently.

In what follows, we provide a sufficient condition for arbitration to be valuable in the multi-dimensional setting. As a corollary, we show that when the principal weighs each dimension equally, the sufficient condition is satisfied as long as the receiver's preferred action in each direction increases with the preferred action of the sender in that direction. Thus, even if the preferred actions of sender and receiver are very distant from each other, that is for any arbitrarily large bias, as long as type dependence of their preferred actions is aligned, the principal can always profit from contracting with the agent when the dimension of the type space is at least 2.

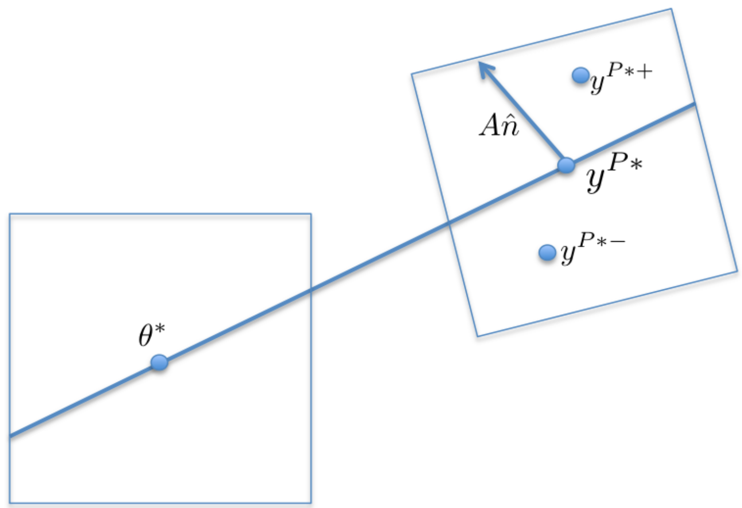
Let y^{P*} denote the expectation of y^P with respect to the prior on the type of the sender. We let y^{P*} denote the contract that takes action y^{P*} for every type. We say that *delegation is valuable* if there is a contract $y(\theta)$ such that $\Delta V(y) = (V(y) - V(y^{P*})) > 0$. That is, if there is contract that improves on the optimal uninformed action of the receiver.

Definition 1. We say that the preferences are *minimally aligned with respect to* $(\hat{n},\theta^*)\in S^N\times int\Theta$ if $\hat{n}\cdot y^{P*} = \hat{n}\cdot\theta^*$ and

$$\mathbb{E}((A\hat{n})\cdot y^P(z)|\hat{n}\cdot z\leq\hat{n}\cdot\theta^*) < (A\hat{n})\cdot y^{P*} < \mathbb{E}((A\hat{n})\cdot y^P(z)|\hat{n}\cdot z\geq\hat{n}\cdot\theta^*).$$

The condition of minimal alignment is a joint condition involving the weight matrix A and the preferred actions $y^P()$ of the receiver and is illustrated in Figure 1. The dot marked with y^{P*+} denotes $\mathbb{E}(y^P(z)|\hat{n}\cdot z\geq\hat{n}\cdot\theta^*)$ and y^{P*-} denotes $\mathbb{E}(y^P(z)|\hat{n}\cdot z\leq\hat{n}\cdot\theta^*)$. The condition requires that the angle between the vector that joins y^{P*+} and y^{P*} and the vector $A\hat{n}$ be less than $\frac{\pi}{2}$ while the angle between the vector that joins y^{P*-} and y^{P*} is greater than $\frac{\pi}{2}$.

For some intuition, consider the case in which A is the identity matrix. In this case, minimal alignment says that if the type of the sender is known to be above (below) a hyperplane then the optimal action



copy.png copy.png

Figure 1: Minimal alignment condition. The square on the left represents the agent's preferred actions while the tilted square on the right represents the principal's preferred actions.

of the receiver needs to be above (below) this hyperplane as well. Intuitively, if this is the case, the receiver can offer two actions, that separate types in the upper and lower half-spaces, which are closer to the optimal actions in the respective half-spaces. Corollary 1 below shows that when A is the identity, minimal alignment never fails under a mild condition over the dependence of receiver's preferred action $y^P(\cdot)$ on the type θ.

When A is not the identity the receiver experiences greater losses in one dimension than another, relative to the agent. In this case, the minimal alignment condition takes into account that the loss in the $\hat{n}$ direction have to be transformed into the receiver's units.

The minimal alignment condition is a generalization of the minimal alignment condition defined by Alonso and Matuschek (2008) for a one-dimensional environment.[4] In fact, when A is the identity matrix the condition can be understood as requiring minimal alignment in the one-dimensional sense in some direction. In the one dimensional setting minimal alignment is necessary and sufficient whereas in the 2-dimensional setting we can only show that the condition is sufficient.

Proposition 1. *If there is $(\hat{n}, \theta^*)$ such that the preferences are minimally aligned with respect to $(\hat{n}, \theta^*)$ arbitration is valuable for the principal.*

[4] Alonso and Matuschek (2008) define the preferences to be *minimally aligned* if there is a a state $\theta^* \in (0,1)$ such that $\mathbb{E}(y^P(z)|z \leq \theta^*) < \theta^* < \mathbb{E}(y^P(z)|z \geq \theta^*)$. That is, there must be a type θ^* such that θ^* is between the optimal action of the receiver conditional on the type being above θ^* and the optimal action conditional on the type being below θ^*. This condition may fail in the one dimensional setting, for example, whenever $y^P(z)$ is increasing and the bias is sufficiently large.

Proof. Consider the contract given by

$$y(\theta) = \begin{cases} (y_1^{P*} + \hat{n}_1\varepsilon, y_2^{P*} + \hat{n}_2\varepsilon) & \text{if } \hat{n}\cdot\theta \geq \hat{n}\cdot\theta^* \\ (y_1^{P*} - \hat{n}_1\varepsilon, y_2^{P*} - \hat{n}_2\varepsilon) & \text{if } \hat{n}\cdot\theta < \hat{n}\cdot\theta^* \end{cases}.$$

This contract satisfies the IC constraints for all types. In fact, senders types in the set $\{\theta | \hat{n}\cdot\theta = \hat{n}\cdot\theta^*\}$ are indifferent between $(y_1^{P*} + \hat{n}_1\varepsilon, y_2^{P*} + \hat{n}_2\varepsilon)$ and $(y_1^{P*} - \hat{n}_1\varepsilon, y_2^{P*} - \hat{n}_2\varepsilon)$ because the two actions are equidistant to their preferred action. In figure 1 these are the types on the straight line that divides the sender type space. Let's see that $y(\cdot)$ improves on y^{P*} for small enough ε.

$$\begin{aligned}
\Delta(V(y)) = \mathbb{E}_\theta &\left(-\alpha_1(y_1(\theta) - y_1^P(\theta))^2 - \alpha_2(y_2(\theta) - y_2^P(\theta))^2 + \alpha_1(y_1^{P*} - y_1^P(\theta))^2 + \alpha_2(y_2^{P*}(\theta) - y_2^P(\theta))^2\right) \\
= \mathbb{P}(\hat{n}\cdot\theta \geq \hat{n}\cdot\theta^*)&\Big(\alpha_1\left(2\varepsilon\hat{n}_1\left(\mathbb{E}\left(y_1^P(\theta)|\hat{n}\cdot\theta \geq \hat{n}\cdot\theta^*\right) - y_1^{P*}\right) - (\hat{n}_1\varepsilon)^2\right) + \\
&\alpha_2\left(2\varepsilon\hat{n}_2\left(\mathbb{E}\left(y_2^P(\theta)|\hat{n}\cdot\theta \geq \hat{n}\cdot\theta^*\right) - y_2^{P*}\right) - (\hat{n}_2\varepsilon)^2\right)\Big) + \\
\mathbb{P}(\hat{n}\cdot\theta \leq \hat{n}\cdot\theta^*)&\Big(\alpha_1\left(2\varepsilon\hat{n}_1\left(y_1^{P*} - \mathbb{E}\left(y_1^P(\theta)|\hat{n}\cdot\theta \leq \hat{n}\cdot\theta^*\right)\right) - (\hat{n}_1\varepsilon)^2\right) + \\
&\alpha_2\left(2\varepsilon\hat{n}_2\left(y_2^{P*} - \mathbb{E}\left(y_2^P(\theta)|\hat{n}\cdot\theta \leq \hat{n}\cdot\theta^*\right)\right) - (\hat{n}_2\varepsilon)^2\right)\Big).
\end{aligned}$$

Now, by the minimally aligned preferences condition

$$\alpha_1\hat{n}_1\left(\mathbb{E}\left(y_1^P(\theta)|\hat{n}\cdot\theta \geq \hat{n}\cdot\theta^*\right) - y_1^{P*}\right) + \alpha_2\hat{n}_2\mathbb{E}\left(y_2^P(\theta)|\hat{n}\cdot\theta \geq \hat{n}\cdot\theta^* - y_2^{P*}\right) > 0$$

and

$$\alpha_1\hat{n}_1\left(y_1^{P*} - \mathbb{E}\left(y_1^P(\theta)|\hat{n}\cdot\theta \leq \hat{n}\cdot\theta^*\right)\right) + \alpha_2\hat{n}_2\left(y_2^{P*} - \mathbb{E}\left(y_2^P(\theta)|\hat{n}\cdot\theta \leq \hat{n}\cdot\theta^*\right)\right) > 0$$

Thus, we have $\Delta V(y) > 0$ for small enough ε. □

In what follows, let d denote a vector with norm 1 and a denote a vector in $\mathbb{R}^2$, we denote a_d for $a\cdot d$. Thus, $y_d^P(\theta_1,\theta_2)$ denotes the receiver's optimal action in direction d given type $\theta = (\theta_1,\theta_2)$, and θ_d denotes the type in direction d. $\theta_{d^\perp}$ denotes the type θ in the direction perpendicular to d. The following Corollary shows that if the matrix A is the identity and y_d^P increases in θ_d for each d then the minimal alignment condition holds. This requirement amounts to monotonicity in in each direction d. In contrast, in the one-dimensional case monotonicity does not guarantee that delegation is valuable. In fact, in that case, delegation is never valuable for high enough bias.

Corollary 1. *If for each $d \in S^2$ y_d^P is increasing in θ_d and $A = I$, then delegation is valuable.*

Proof. For every $\theta^* \in int\Theta$ we can find $\hat{n}$ such that $\hat{n}\cdot\theta^* = \hat{n}\cdot y^{P*}$. As y_d^P is increasing in θ_d we obtain $\mathbb{E}(\hat{n}\cdot y^P(z)|\hat{n}\cdot z \leq \hat{n}\cdot\theta^*) < \hat{n}\cdot y^{P*} < \mathbb{E}(\hat{n}\cdot y^P(z)|\hat{n}\cdot z \geq \hat{n}\cdot\theta^*)$. □

Intuitively, consider Figure 1. When A is the identity, monotonicity guarantees that y^{P*+} is in the upper half-space and y^{P*-} is in the lower half-space and $A\hat{n} = \hat{n}$ is perpendicular to the vector $y^{P*} - \theta^*$. Therefore, minimal alignment holds.

Note that y_d^P is increasing with θ_d whenever $y^P(\theta) = M\theta + b$ where M is a positive definite matrix.

4. Discrete states

Now we consider a version of our model with a finite number of states.

The potential set of decisions is $Y = \mathbb{R}^2$. Let $\Theta = \left\{\theta^1, \theta^2, \ldots, \theta^K\right\}$ be the set of possible states of the world, with p_k the associated probability of each state ($k = 1, 2, \ldots, K$). Recall that the principal's preferences are

$$u^P(\theta, x) \quad = \quad -\alpha_1 \left[y_1^P(\theta) - x_1\right]^2 - \alpha_2 \left[y_2^P(\theta) - x_2\right]^2, \tag{1}$$

and the agent's be

$$u^A(\theta, x) \quad = \quad -\left[y_1^A(\theta) - x_1\right]^2 - \left[y_2^A(\theta) - x_2\right]^2, \tag{2}$$

where x is the action taken, $\theta \in \Theta$ is the state of the world, and $y_i^I(\theta)$ is agent I's ($I = A, P$) ideal action in state θ in the i-th ($i = 1, 2$) dimension.

For simplicity, we adopt the notation

$$y_i^{I,k} \equiv y_i^I(\theta^i),$$

where $i = 1, 2$, $k = 1, 2, \ldots, K$, and $I = A, P$. Thus, $y_1^{A,k}$ is the agent's ideal action on the first dimension in state k. We assume there is a minimal level of congruence between the preferences of the principal and the agent.

Assumption 1. The principal's preferences and the agent's preferences satisfy

$$y_i^{A,k} \geq y_i^{A,l} \text{ if and only if } y_i^{P,k} \geq y_i^{P,l}$$

for all $i = 1, 2$ and $k, l = 1, 2, \ldots, K$.

In other words, whenever the principal prefers a higher action in one state versus another, the agent does as well. This is satisfied, for example, when the agent has a fixed bias in each dimension versus the principal's ideal action.

Thus, the principal's optimization problem can be written as

$$\max_{x^1,x^2,\ldots,x^K} \sum_{k=1}^{K} -p_k \left[\alpha_1 \left(y_1^{P,k} - x_1^k\right)^2 + \alpha_2 \left(y_2^{P,k} - x_2^k\right)^2\right]$$

subject to the following IC constraints

$$-\left(y_1^{A,k} - x_1^k\right)^2 - \left(y_2^{A,k} - x_2^k\right)^2 \geq -\left(y_1^{A,k} - x_1^l\right)^2 - \left(y_2^{A,k} - x_2^l\right)^2,$$

for $k,l = 1,2,\ldots,n$.

Let $\mathscr{D}$ be the set of binding constraints. That is,

$$\mathscr{D} \equiv \left\{(k,l)|d(x^k, y^{A,k}) = d(x^l, y^{A,k})\right\}.$$

The Lagrangian of the problem can be written

$$\begin{aligned} \mathscr{L} = & \sum_{k=1}^{K} -p_k \left[\alpha_1 \left(y_1^{P,k} - x_1^k\right)^2 + \alpha_2 \left(y_2^{P,k} - x_2^k\right)^2\right] + \\ & \sum_{(k,l)\in\mathscr{D}} \mu_{k,l} \left[-\left(y_1^{A,k} - x_1^k\right)^2 - \left(y_2^{A,k} - x_2^k\right)^2 + \left(y_1^{A,k} - x_1^l\right)^2 + \left(y_2^{A,k} - x_2^l\right)^2\right]. \end{aligned}$$

Let $\mathscr{D}_k^1$ and $\mathscr{D}_k^2$ be defined respectivdely as

$$\begin{aligned} D_k^1 &\equiv \{l|(k,l) \in \mathscr{D}\}, \\ D_k^2 &\equiv \{l|(l,k) \in \mathscr{D}\}. \end{aligned}$$

Thus, the first order condition with respect to x_1^k and x_2^k can be written

$$2p_k \left(y_1^{P,k} - x_1^k\right) + \sum_{j\in D_i^1} \mu_{k,l} \left[2\left(y_1^{A,k} - x_1^k\right)\right] + \sum_{j\in D_i^2} \mu_{l,k} \left[-2\left(y_1^{A,l} - x_1^k\right)\right] = 0; \tag{3}$$

$$2p_k\lambda \left(y_2^{P,k} - x_2^k\right) + \sum_{j\in D_i^1} \mu_{k,l} \left[2\left(y_2^{A,k} - x_2^k\right)\right] + \sum_{j\in D_i^2} \mu_{l,k} \left[-2\left(y_2^{A,l} - x_2^k\right)\right] = 0. \tag{4}$$

From this point on, we focus our attention on a scenario where it is helpful to treat the two dimensions of decision differently. Think of the first dimension as an issue in which the agent has a bias versus the principal that is constant and independent of the state. To use the example in the introduction, it is conceivable that a politician could be more radical on cultural issues than the general public, but on matters of economic policy, his divergence from the society might be more nuanced, due to a combination of ideological and financial factors. From the perspective of the writer of a constitution as the principal, who wants to maximize the welfare of the general public, it is useful to think of ways to take advantage of these characteristics.

To highlight the distinctive features of this scenario, we make the following assumption.

Assumption 2. The agent's ideal action in the first dimension, $y_1^{A,k}$, is equal to 0 for all $i = 1, 2, \ldots, n$, and the principal's ideal action, $y_1^{P,k}$, is equal to $y_1^p \geq 0$ for all $i = 1, 2, \ldots, n$.[5]

Given this assumption, the first order condition regarding x_1^k becomes

$$2p_k \left(y_1^{P,k} - x_1^k \right) + \sum_{j \in D_i^1} \mu_{k,l} \cdot 2 \left(-x_1^k \right) + \sum_{j \in D_i^2} \mu_{l,k} \cdot (-2) \left(-x_1^k \right) \quad = \quad 0. \tag{5}$$

Note that under Assumption 2, in the optimal solution,

$$x_1^k \quad \geq \quad 0 \text{ for all } i = 1, \ldots, n. \tag{6}$$

The reason is that if $x_1^k < 0$ then by changing it to $x_1^{k\prime} = -x_1^k$, the principal's payoff is increased without affecting any IC constraints.

The following proposition, which is the main result of our analysis in this section, states that if we move the principal's and the agent's preferences further apart from each other in a systematic way, then the principal, in the outcome under optimal delegation, always becomes better off.

Proposition 1. *Let Assumptions 1 and 2 hold. Then, an increase in y_1^p will make the principal better off. Furthermore, as y_1^p goes to infinity, the principal's payoff converges to her full-information optimum in the second dimension, i.e., $x_2^k - y_2^{P,k}$ approaches zero.*

To see the meaning of the proposition, consider a two-dimensional environment as follows. In the first dimension, the agent's bias is fixed and independent of the state of the world; in the second one, the agent's bias may or may not vary with the state of the world. The proposition states that, fixing everything else, the principal always benefits from having an agent who is more biased in the first dimension. Due to the convexity of the agent's preferences, the principal can manipulate the choices available in the first dimension, in order to make sure he chooses actions that are closer to the principal's ideal in the second dimension. In particular, a deviation away from the agent's ideal action in the first dimension has a larger effect of deterrence if the agent has a larger bias in that dimension, given that the principal chooses the delegation set closer to her ideal actions in the first dimension. Furthermore, as the bias approaches to infinity, the principal will be able to achieve her ideal action in the second dimension in each state of the world.

Proof of Proposition 1. Let $V(y^p)$ be the principal's payoff under optimal delegation. Let us move y_1^P

[5] Alternatively and equivalently, we could assume $y_1^{P,k} = 0$ for all i and vary $y_1^{A,k}$, but it would make the notations more complicated.

marginally to the right, by the Envelope Theorem, we have

$$\frac{dV}{dy_1^P} = \sum_{k=1}^{K} -2p_k\left(y_1^P - x_1^k\right).$$

Using the first order condition with respect to x_1^k above and collecting terms, we obtain

$$\frac{dV}{dy_1^P} = \sum_{(k,l)\in\mathscr{D}} 2\mu_{k,l}\left(x_1^l - x_1^k\right).$$

Our goal is to show, under Assumptions 1 and 2, the following inequality holds.

$$\sum_{(k,l)\in\mathscr{D}} 2\mu_{k,l}\left(x_1^l - x_1^k\right) \geq 0. \tag{7}$$

Note that if $x_1^l \geq x_1^k$ for all $(k,l)\in\mathscr{D}$, the above condition is satisfied.

Lemma 1. *Let $y_2^{A,k} > y_2^{A,l}$ (which by assumption implies $y_2^{P,k} > y_{2,j}^P$) and $(k,l)\in\mathscr{D}$. Then, in the optimal solution, $x_2^k \geq x_2^l$.*

Proof of Lemma 1. A straightforward observation by summing the IC constraints (k,l) and (j,i) is that

$$x_1^k y_1^{A,k} + x_2^k y_2^{A,k} + x_1^l y_1^{A,l} + x_2^l y_2^{A,l} \geq x_1^k y_1^{A,l} + x_2^k y_2^{A,l} + x_1^l y_1^{A,k} + x_2^l y_2^{A,k},$$

or

$$\left(x_1^k - x_1^l\right)\left(y_1^{A,k} - y_1^{A,l}\right) + \left(x_2^k - x_2^l\right)\left(y_2^{A,k} - y_2^{A,l}\right) \geq 0. \tag{8}$$

The lemma is directly implied by (8). □

Lemma 2. *Let $y_2^{A,k} > y_2^{A,l}$ and $(k,l)\in\mathscr{D}$. Then, in the optimal solution, $(l,k)\notin\mathscr{D}$, which implies $\mu_{l,k}=0$.*

Proof of Lemma 2. Note that

$$\frac{\partial\left[d(y^{A,m},x^l)^2 - d(y^{A,m},x^k)^2\right]}{\partial y_2^{A,m}} = 2\left(x_2^k - x_2^l\right) \tag{9}$$

and that

$$d(y^{A,l},x^l)^2 - d(y^{A,l},x^k)^2 = d(y^{A,k},x^l)^2 - d(y^{A,k},x^k)^2 + \int_{y_2^{A,k}}^{y_2^{A,l}} \frac{\partial\left[d(y^{A,m},x^l)^2 - d(y^{A,m},x^k)^2\right]}{\partial y_2^{A,m}} dy_2^{A,m}.$$

By Lemma 1, we have

$$x_2^k - x_2^l \geq 0,$$

from which we may conclude

$$d(y^{A,l},x^l)^2 - d(y^{A,l},x^k)^2 \leq 0.$$

Hence, the constraint (j,i) is also satisfied and therefore redundant/slack, which implies $\mu_{l,k} = 0$. $\square$

Note that the assumption $y_2^{A,k} > y_2^{A,l}$ is without loss of generality, as the same argument can be applied to the case $y_2^{A,k} < y_2^{A,l}$.

Lemma 3. *Let $y_2^{A,k} > y_2^{A,l} > y_2^{A,k}$, then $(i,k) \notin \mathscr{D}$ and $(k,i) \notin \mathscr{D}$.*

Proof of Lemma 3. We show $(i,k) \notin \mathscr{D}$ (the argument for $(k,i) \notin \mathscr{D}$ is similar). Note that

$$d(y^{A,k},x^m)^2 - d(y^{A,k},x^k)^2 \quad = \quad d(y^{A,k},x^m)^2 - d(y^{A,k},x^l)^2 + d(y^{A,k},x^l)^2 - d(y^{A,k},x^k)^2. \tag{10}$$

In addition,

$$\begin{aligned} d(y^{A,k},x^l)^2 - d(y^{A,k},x^k)^2 &\geq 0, \\ d(y^{A,l},x^m)^2 - d(y^{A,l},x^l)^2 &\geq 0 \end{aligned}$$

due to the IC constraints. In addition, observe that

$$d(y^{A,k},x^m)^2 - d(y^{A,k},x^l)^2 \quad \geq \quad d(y^{A,l},x^m)^2 - d(y^{A,l},x^l)^2.$$

by (9) and Lemma 1. Combined with (10), we conclude the IC constraint (k,m) is automatically satisfied. $\square$

We want to show that (7) holds. Let $y_2^{A,k} > y_2^{A,k+1}$ for $k = 1,2,\ldots,K-1$ (and by assumption, $y_2^{P,k} > y_2^{P,k+1}$. Note that by Lemma 3, the only binding constraints are $(k,k+1)$ or $(k+1,k)$.

Let us define

$$\delta^k \quad \equiv \quad x_1^k - y_1^P.$$

From (5), we obtain that

$$p_k\delta^k \quad = \quad [\mu_{m,k} - \mu_{k,l}]x_1^k, \tag{11}$$

where

$$(k,l) \in \mathscr{D}, \;\; (m,k) \in \mathscr{D}.$$

In addition, by (6),

$$x_1^k \geq 0.$$

Thus, if

$$\delta^k \geq 0,$$

then

$$\begin{aligned} \mu_{m,k} - \mu_{k,l} &\geq 0; \\ [\mu_{m,k} - \mu_{k,l}] x_1^k &\geq [\mu_{m,k} - \mu_{k,l}] y_1^P. \end{aligned}$$

Similarly, if

$$\delta^k \leq 0,$$

would imply that

$$\begin{aligned} \mu_{m,k} - \mu_{k,l} &\leq 0; \\ [\mu_{m,k} - \mu_{k,l}] x_1^k &\geq [\mu_{m,k} - \mu_{k,l}] y_1^P. \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{dV}{dy_1^P} &= \sum_{k=1}^{K} 2 p_k \delta^k, \\ &\geq \sum_{(k,l) \in \mathscr{D}} [-\mu_{k,l} + \mu_{k,l}] y_1^P, \\ &= 0. \end{aligned}$$

To prove the second part of the proposition, note that as y_1^P approaches infinity, δ^k must remain bounded, because the principal can always guarantee herself a payoff bounded from below. So, x_1^k must approach infinity. By (11), $\mu_{m,k} - \mu_{k,l}$ converges to zero. Using it in (4), we conclude that $y_2^{P,k} - x_2^k$ must converge to zero. □

5. Continuous type space

In this section we consider a continuous type space. We can show that for some classes of utilities the implemented actions under delegation approach the first best in the direction opposite to the bias as the bias grows to infinity.

5.1 Constant bias

Next we consider the case in which the agent and the principal's bliss points differ by a constant bias. The principal's prior distribution over the type space is given by a continuous distribution $F(\theta_1, \theta_2)$ on Θ with density $f(\theta_1, \theta_2)$. Let $f_i(\theta_i)$ denote the density in θ_i.

The principal's preferred actions given the type are given by $y^P(\theta) = (\theta_1 - b_1, \theta_2 - b_2)$ and A is the identity matrix. Because the distribution is arbitrary it is without loss to assume that the bias vector is $(0, b)$.

Our main result in the setting is that the optimal contract converges in probability to a simple contract in which the action taken in the θ_1 direction is constant and equal to $\mathbb{E}(\theta_1) - b$ and in which there is perfect revelation in the dimension opposite to the bias. This result is stated formally in the following Theorem

Theorem 1 (Optimal contract as bias approaches infinity)**.** *If $y_1(\theta_1, \theta_2)$ is optimal then for each $\varepsilon > 0$ and $\delta > 0$ there is $\bar{b}$ such that $b \geq \bar{b}$*

$$Prob(|\mathbb{E}(\theta_1) - b - y_1(\theta_1, \theta_2)| \geq \varepsilon) < \delta$$

and,

$$Prob(|\theta_2 - y_2(\theta_1, \theta_2)| \geq \varepsilon) < \delta.$$

In order to prove Theorem 1 we first show that as b becomes sufficiently large the optimal contract becomes approximately constant in y_1. For large bias the receiver's preferred actions are very distant from the sender's preferred actions and any further loss in that direction has a great utility cost for the sender. In this case the principal is not able to induce revelation in the direction of the bias. However, as the bias grows large it becomes cheaper for the principal to induce revelation in the direction opposite to the bias: by distorting the allocation in the direction of the bias slightly the principal can harshly punish the choice of more desirable actions in the bias's orthogonal direction. Thus, as the bias grows large the action that is implemented in the direction of the bias converges in probability to a constant. The following Lemma formalizes this first result.

Lemma 2. *For every $\varepsilon > 0$ and $\delta > 0$ there is $\bar{b}$ such that for $b \geq \bar{b}$, the optimal contract $\hat{y}$ is such that*

$$Prob(|\hat{y}_1(\theta_1, \theta_2) - \hat{y}_1(\theta_1', \theta_2)| > \varepsilon) < \delta.$$

Proof. Let $\varepsilon > 0$ and $\delta > 0$. Since the mechanism $y_1(\theta_1, \theta_2) = \mathbb{E}(\theta_1) - b$, $y_2(\theta_1, \theta_2) = \theta_2$ satisfies IC the receiver's payoff is bounded below by $-\int(\mathbb{E}(\theta_1) - \theta_1)^2$. Thus for big enough b if $\hat{y}$ is optimal we have

$$\int(\mathbb{E}(\theta_1) - \theta_1)^2 \geq \int(\hat{y}_1(\theta_1, \theta_2) + b - \theta_1)^2 \geq Prob\left(|\hat{y}_1(\theta_1, \theta_2) + b - \theta_1| > \frac{b}{2}\right)\frac{b^2}{4}. \quad (12)$$

Thus, for b big enough $Prob\left(|\hat{y}_1(\theta_1,\theta_2)+b-\theta_1)|>\frac{b}{2}\right)\le\frac{\delta}{2}$. Analogously for every $a>0$,

$$\int(\mathbb{E}(\theta_1) \quad - \quad \theta_1)^2 \quad \geq \quad \int(\hat{y}_2(\theta_1,\theta_2) \quad - \quad \theta_2)^2 \quad \geq \quad Prob\,(|\hat{y}_2(\theta_1,\theta_2)-\theta_2|>a)\,a^2. \quad (13)$$

and, therefore, there is $\bar{a}$ such that for $a\ge\bar{a}$, $Prob\,(|\hat{y}_2(\theta_1,\theta_2)-\theta_2|>a)<\frac{\delta}{2}$. Now, suppose there are types θ_1,θ_1' and θ_2 such that $\hat{y}_1(\theta_1,\theta_2)+\tilde{\varepsilon}(\theta_1',\theta_2)=\hat{y}_1(\theta_1',\theta_2)$ with $\tilde{\varepsilon}(\theta_1',\theta_2)>\varepsilon>0$. From the IC constraint

$$(\hat{y}_1(\theta_1,\theta_2)-\theta_1)^2+(\hat{y}_2(\theta_1,\theta_2)-\theta_2)^2\le(\hat{y}_1(\theta_1',\theta_2)-\theta_1)^2+(\hat{y}_2(\theta_1',\theta_2)-\theta_2)^2 \implies$$
$$(\hat{y}_2(\theta_1',\theta_2)-\theta_2)^2\ge-2(\hat{y}_1(\theta_1,\theta_2)-\theta_1)\tilde{\varepsilon}(\theta_1,\theta_2)+\tilde{\varepsilon}(\theta_1,\theta_2)^2 \quad (14)$$

From equation (12) with probability $1-\delta/2$ we have $-(\hat{y}_1(\theta_1',\theta_2)-\theta_1)\ge\frac{b}{2}$. Let $\bar{b}$ be such that $\bar{b}\varepsilon\ge\bar{a}$. From equation (13), the last inequality in equation (14) is satisfied with probability at most $\delta/2$ when $b\ge\bar{b}$. Thus, we conclude

$$Prob(|\hat{y}_1(\theta_1,\theta_2)-\hat{y}_1(\theta_1',\theta_2)|>\varepsilon)<\delta,$$

for $b\ge\bar{b}$, implying that y_1 is nearly constant for large enough bias.

Proof of Theorem 1

Let $\hat{y}(b)$ denote the optimal contract for bias b. From Lemma 2 there is $\bar{y}_1(b)$, $\varepsilon(b)$ and $\delta(b)$ such that

$$Prob\,(|\hat{y}_1(b)(\theta_1,\theta_2)-\bar{y}_1(b)|>\varepsilon(b))<\delta(b),$$

where $\varepsilon(b)$ and $\delta(b)$ converging to zero as b converges to ∞. Define $\varepsilon(\theta_1,\theta_2)=\hat{y}_1(b)(\theta_1,\theta_2)-\bar{y}_1(b)$.

Let's see that for every ε' there is $\bar{b}$ such that for $b\ge\bar{b}$,

$$|\bar{y}_1(b)-\mathbb{E}\left(\theta_1\middle||\varepsilon(\theta_1,\theta_2)|\le\varepsilon(b)\right)+b|\le\varepsilon'$$

and

$$\int(\hat{y}_2(b)(\theta)-\theta_2)^2\le\varepsilon'.$$

Suppose not, then there is a sequence $b_n\to\infty$ such that either

$$|\bar{y}_1(b_n)-\mathbb{E}\left(\theta_1\middle||\varepsilon(\theta_1,\theta_2)|\le\varepsilon(b_n)\right)+b_n|\ge\alpha>0$$

or

$$\int(\hat{y}_2(b_n)(\theta)-\theta_2)^2\ge\alpha>0.$$

Consider the contract $\tilde{y}(b)$ defined as $\tilde{y}_1(b)(\theta_1,\theta_2)=\mathbb{E}(\theta_1\,|\,\varepsilon(\theta_1,\theta_2)\le\varepsilon(b))-b=\tilde{y}_1(b)$ and $\tilde{y}_2(b)(\theta_1,\theta_2)=$

θ_2. The contract $\tilde{y}(b)$ satisfies IC. Let $M = \max_{\theta_1,\theta_1'} |\theta_1' - \theta_1|$.[6]

$$
\begin{aligned}
V(\tilde{y}) - V(\hat{y}) &\geq \int \left(-(\tilde{y}_1(\theta) - \theta_1 + b_n)^2 + (\hat{y}_1(\theta) - \theta_1 + b_n)^2 - (\tilde{y}_2(\theta) - \theta_2)^2 + (\hat{y}_2(\theta) - \theta_2)^2\right) d\theta \\
&= \int_{|\varepsilon(\theta_1,\theta_2)| \leq \varepsilon(b_n)} \left(-(\tilde{y}_1(\theta) - \theta_1 + b_n)^2 + (\bar{y}_1 + \varepsilon(\theta_1,\theta_2) - \theta_1 + b_n)^2\right) d\theta \\
&\quad + \int_{|\varepsilon(\theta_1,\theta_2)| > \varepsilon(b_n)} \left(-(\tilde{y}_1(\theta) - \theta_1 + b_n)^2 + (\bar{y}_1 + \varepsilon(\theta_1,\theta_2) - \theta_1 + b_n)^2\right) d\theta \\
&\quad + \int (\hat{y}_2(\theta) - \theta_2)^2 \, d\theta \\
&\geq \int_{|\varepsilon(\theta_1,\theta_2)| \leq \varepsilon(b_n)} \left(-(\tilde{y}_1(\theta) - \theta_1 + b_n)^2 + (\bar{y}_1 + \varepsilon(\theta_1,\theta_2) - \theta_1 + b_n)^2\right) d\theta \\
&\quad + \int_{|\varepsilon(\theta_1,\theta_2)| > \varepsilon(b_n)} -(\tilde{y}_1(\theta) - \theta_1 + b_n)^2 \, d\theta + \int (\hat{y}_2(\theta) - \theta_2)^2 \, d\theta
\end{aligned}
$$

Writing $(\bar{y}_1 + \varepsilon(\theta_1,\theta_2) - \theta_1 + b_n)^2 = ((\bar{y}_1 + \varepsilon(\theta_1,\theta_2) - \tilde{y}_1) + (\tilde{y}_1 - \theta_1 + b_n))^2$ and noting that $\int_{|\varepsilon(\theta_1,\theta_2)| \leq \varepsilon(b_n)} (\tilde{y}_1 + b_n - \theta_1) \, d\theta = 0$ we obtain

$$
\begin{aligned}
V(\tilde{y}) - V(\hat{y}) &\geq \int_{|\varepsilon(\theta_1,\theta_2)| \leq \varepsilon(b_n)} \left((\bar{y}_1 - \tilde{y}_1 + \varepsilon(\theta_1,\theta_2))^2 + 2\varepsilon(\theta_1,\theta_2)(\tilde{y}_1 + b_n - \theta_1)\right) d\theta \\
&\quad + M^2\delta + \int (\hat{y}_2(\theta) - \theta_2)^2
\end{aligned}
$$

If $|\bar{y}_1(b_n) - \tilde{y}_1(b_n)| \geq \alpha$ or $\int (\hat{y}_2(b_n)(\theta) - \theta_2)^2 \geq \alpha$, for big enough b_n the last expression is positive, contradicting the optimality of $\hat{y}$. Finally, since the type space has a continuous density $\mathbb{E}(\theta_1 \,\big|\, |\varepsilon(\theta_1,\theta_2)| \leq \varepsilon(b))$ converges to $\mathbb{E}(\theta_1)$ in probability. Since the limit in probability is unique almost surely $\mathbb{E}(\theta_1) - \bar{y}_1 + b$ converges to zero in probability. Convergence of $\int (\hat{y}_2(b)(\theta) - \theta_2)^2$ to zero implies convergence of $\hat{y}_2(b)(\theta_2) - \theta_2$ to zero in probability. □

5.2 Type-dependent bias

In the previous section we considered a model in which the agent and the principal's bliss points differed only by a constant bias. However, the main result is not restricted to the constant bias case. In this section, we provide an example with type dependent bias in which as the bias converges to infinity the principal attains the first best in the direction opposite to the bias. We show that we can construct a lower bound on the payoffs of the principal that increases with the disagreement between the sender and receiver.

Suppose now that the receiver's utility is

[6] In what follows, we omit the dependence on b_n for ease of notation.

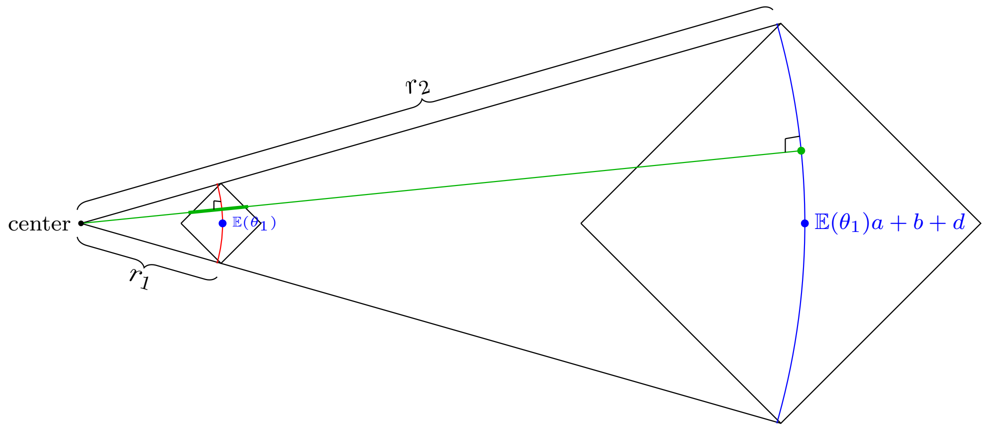


Figure 2: Sets of the sender and receiver's preferred actions. The left set represents the sender's preferred actions and the right set the receiver's.

$$V(y,\theta) = \sum_{i=1}^{N} -(a\theta_i + b_i - y_i)^2,$$

and the utility of the sender is

$$U(y,\theta) = \sum_{i=1}^{N} -(\theta_i - y_i)^2$$

as before.

The set of the agent and the principal's preferred actions are illustrated in Figure 2. The tilted square on the left represents the preferred actions of each type of agent. The tilted square on the right represent the preferred action of the principal as a function of the state. For example, when the preferred action of the agent is on the top corner of the left square the preferred action principal is on the top corner of the right square. More generally, when the state (θ_1,θ_2) the preferred action of the principal is $(b_1+a\theta_1, b_2+a\theta_2)$.

Given a delegation set, the agent chooses the action in the delegation set that is closest in Euclidian distance to his preferred action given his privately known type. The blue and red arcs drawn in Figure 2 are part of circles with the same center. The green line goes from the center of the arcs to the furthermost arc and is, thus, perpendicular to the arcs. If the delegation set is given by the blue arc in Figure 2 then the types of sender that are in the intersection of the green line and the set of sender types (the bold segment in the figure) choose the action in the delegation set corresponding to the intersection between the green line and the blue arc (the green dot on the blue arc).

The delegation set corresponding to the blue arc in Figure 2 can be parametrized as $\{\alpha \in (\underline{\alpha}, \bar{\alpha}) : (r_2 \cos\alpha - d_1 + b_1, r_2 \sin\alpha - d_2 + b_2\}$ for some angles $\underline{\alpha}$, $\bar{\alpha}$, radius $r_2 > 0$ and vector (d_1, d_2). Let's see that by appropriate choice of $\underline{\alpha}$, $\bar{\alpha}$, r_2 and (d_1,d_2) the principal can offer a delegation set for each bias b such that

$$\lim_{\|b\|\to\infty} \mathbb{E}\left((\theta_{b^\perp} a - y_{b^\perp}(\theta))^2\right) = 0. \tag{15}$$

That is, the action taken by the agent in the direction that is perpendicular to the bias converges to the preferred action of the principal in that direction.

Assume without loss that the bias is $(b,0)$. Suppose $a>1$ as in Figure 2. Let d be the distance between the center and the leftmost corner of the square. To find r_1 and r_2 that parametrize the sets in Figure 1 we need to solve the following conditions

$$r_1 = d + \mathbb{E}(\theta_1), \quad r_2 = a\mathbb{E}(\theta_1) + b + d \text{ and } ar_1 \sin\alpha = r_2,$$

which imply $r_2 = \mathbb{E}(\theta_1)a + \frac{a}{a-1}b$, $r_1 = \frac{b}{a-1} + \mathbb{E}(\theta_1)$, $\bar{\alpha} = \sin^{-1}\left(\frac{a\max\{\theta_2|\theta_2\in\Theta\}}{r_2}\right)$, $\underline{\alpha} = \sin^{-1}\left(\frac{a\min\{\theta_2|\theta_2\in\Theta\}}{r_2}\right)$ and $d = (\mathbb{E}(\theta_1)a + \frac{b}{a-1}, 0)$. Two observations are relevant at this point (1) since $ar_1\sin\alpha = r_2\sin\alpha$ the types on the red arc in Figure 1 choose an action in direction $(0,1)$ that exactly matches the principal's preferred action in that direction. (2) The angles $\bar{\alpha},\underline{\alpha}\to 0$ as $b\to\infty$, thus, as the bias increases the types assigned to each point in the delegation set differ less and less direction perpendicular to the bias. Observations (1) and (2) imply that as b increases the assignment in direction perpendicular to the bias converges to the principal preferred action in that direction and condition (15) follows.

The geometric intuition for this result is simple. As the bias grows large the arcs of the circles become closer and closer to a vertical line. Thus, in the limit types in every horizontal line in the agent's type set will be induced to choose an action that matches the principal preferred in the direction opposite to the bias and coincides with the expectation of the principal's preferred action in the direction of the bias.

References

Alonso, R. and N. Matouschek (2008). Optimal delegation. *The Review of Economic Studies 75*(1), 259–293.

Antić, N. and M. Iaryczower (2016). Delegation with strings attached: Underfunding a biased agent. Technical report.

Chakraborty, A. and R. Harbaugh (2007). Comparative cheap talk. *Journal of Economic Theory 132*(1), 70–94.

Chakraborty, A. and R. Harbaugh (2010). Persuasion by cheap talk. *The American Economic Review 100*(5), 2361–2382.

Che, Y.-K. and N. Kartik (2009). Opinions as incentives. *Journal of Political Economy 117*(5), 815–860.

Dewatripont, M. and J. Tirole (1999, February). Advocates. *Journal of Political Economy 107*(1), 1–39.

Frankel, A. (2015). Delegating Multiple Decisions. *American Economic Journal: Microeconomics* (Forthcoming).

Goltsman, M., J. Hörner, G. Pavlov, and F. Squintani (2009, July). Arbitration, mediation, and cheap talk. *Journal of Economic Theory 144*(4), 1397–1420.

Holmström, B. R. (1977). *On incentives and control in organizations*. Ph.D. dissertation, Stanford University.

Holmström, B. R. (1984). On the theory of delegation. In M. Boyer and R. Kihlstrom (Eds.), *Bayesian Models in Economic Theory*, pp. 115–141. New York: North-Holland.

Koessler, F. and D. Martimort (2012). Optimal delegation with multi-dimensional decisions. *Journal of Economic Theory 147*(5), 1850 – 1881.

Kováč, E. and T. Mylovanov (2009). Stochastic mechanisms in settings without monetary transfers: The regular case. *Journal of Economic Theory 144*(4), 1373–1395.

Krishna, V. and J. Morgan (2008). Contracting for information under imperfect commitment. *RAND Journal of Economics 39*(4), 905–925.

Levy, G. and R. Razin (2007). On the limits of communication in multidimensional cheap talk: A comment. *Econometrica 75*(3), 885–893.

Martimort, D. and A. Semenov (2006). Continuity in mechanism design without transfers. *Economics Letters 93*(2), 182 – 189.

McCarty, N. (2004). The appointments dilemma. *American Journal of Political Science 48*(3), 413–428.

Melumad, N. D. and T. Shibano (1991). Communication in settings with no transfers. *The RAND Journal of Economics 22*(2), 173–198.