Experimentation with Delegation*

Yingni Guo
Department of Economics, Yale University

1 Introduction

We consider the dynamic interaction between a principal (she) who owns a project of unknown quality and an agent (he) who has the expertise to work on the project. The agent is better informed but biased. In each period, the principal has the authority to assign a delegation set and let the agent to select any action from this set. Yet, contingent transfers are infeasible and the principal cannot commit to a dynamic delegation rule across time. We are interested in how dynamic delegation without commitment improves on decision-making, what the optimal delegation strategy is, and what the limitations are compared with the situations where the principal can commit to future actions and/or make transfers.

This paper is closely related to the literature on delegation. Contrasting with most of the mechanism-design literature, the delegation literature assumes that contingent transfers are not feasible. Contingent transfers may be infeasible in numerous settings. For example, legal rules prohibit contingent transfers between a regulator and a regulated monopolist firm. Holmström (1977) first defines and analyzes the delegation problem. He addresses the question when delegation mechanism is optimal. Amador and Bagwell (2012) consider a general representation of the delegation problem and provide conditions under which interval delegation sets are optimal. Mylovanov (2009) examines veto-based delegation rule where the agent makes a proposal and the principal can either approve it or implement a default decision. He shows that the principal can implement an optimal outcome with a properly chosen default decision.

The delegation literature mainly focuses on static settings. Our model is different in the sense that we examine the optimal delegation rule in a dynamic environment, assuming that the principal cannot commit to future actions. More specifically, the principal cannot commit to a contingent dynamic delegation plan. Lack of commitment across time is a prevailing phenomenon, especially when two parties cannot specify a contract that covers the entire time horizon of the relationship. Take the regulation story as an example. The regulator cannot commit to future regulations. Instead, it learns from past history and optimally updates its regulatory terms across time.

Lack of commitment and contingent transfers distinguishes our work from the literature on optimal contracts for experimentation with adverse selection and dynamic moral hazard. Halac, Kartik and Liu (2012) consider contracting at period zero with full commitment power

*I would like to thank Dirk Bergemann, Eduardo Faingold, Johannes Hörner, Larry Samuelson, and Juuso Välimäki for helpful discussion.

from the principal. Since there is hidden information at the time of contracting, without loss of generality, the principal's problem is to offer the agent a menu of dynamic contracts. A dynamic contract specifies a sequence of transfers as a function of the publicly observable history. Hörner and Samuelson (2012) also considers situations where the principal cannot commit to future actions. Yet, instead of assigning permissible action sets, the principal in their model makes offers regarding how to split the surplus from successful experimentation.

The analysis of our benchmark setting is built on the exponential bandit model introduced by Keller, Rady and Cripps (2005). We characterize the optimal dynamic rule that gives the principal the highest expected payoff and shows that the basic tradeoff is between the benefit from inducing the agent's private information and the cost of excessive experimentation. Moreover, the principal is weakly better off when she can commit to a dynamic delegation plan.

2 The model

Experimentation and payoffs. Time $t \in [0, \infty)$ is continuous. The principal and the agent face a two-armed bandit problem. One arm S is "safe" and yields a known deterministic flow payoff s. The other arm R is "risky" and can be either "good" or "bad". If it is bad, then it always yields 0. If it is good, then it yields lump-sum payoffs after exponentially distributed random times. If R is pulled over an interval of time $[t, t + dt)$, the probability that a lump-sum payoff realizes at some point in the interval is $\lambda\theta dt$, where $\lambda > 0$ is a constant known to all players, $\theta = 1$ if R is good, and $\theta = 0$ if R is bad. Thus, the arrival of the first lump sum reveals that the risky arm is a good one. The magnitude of these lump sums is h. Let $g = \lambda h$ be the flow payoff to the principal if the risky arm is good.

Agent's bias. The agent faces the same bandit as the principal does and receives flow payoff s if S is pulled. However, he gets different payoffs from the principal if the risky arm is pulled. The agent's flow payoff from the risky arm is $g + b$ if it is good and 0 if bad. The constant b measures the agent's bias towards (or against) the risky arm and is commonly known[1]. We assume that $0 < s < g$ and $0 < s < g + b$, so both players strictly prefer R, if it is good, to S, and strictly prefer S to R, if it is bad. At time 0, players do not know the risky arm's type. They start with a common prior p_0 that the risky arm is good and share a common discount rate $r > 0$.

Agent's private information. Throughout the game, both players observe the delegation set assigned by the principal, the action chosen by the agent and the outcome of the risky arm if it was pulled. The only information asymmetry comes from the agent's private information regarding the risky arm's type. Before the game starts, the agent privately observe a binary signal $\theta \in \Theta \equiv \{\theta_l, \theta_h\}$ which endows him a better knowledge of the risky arm's type. If $\theta = \theta_h$, the agent updates his prior to $\bar{p}_0 > p_0$ that the risky arm is a good one. If $\theta = \theta_l$, the agent updates his prior to $\underline{p}_0 < p_0$. We assume that the agent gets signal θ_h with probability q_0. Therefore, we have

$$p_0 = q_0 \bar{p}_0 + (1 - q_0)\underline{p}_0.$$

In particular, we assume that $\underline{p}_0 \in (0, 1)$ and $\bar{p}_0 = 1$, implying that a high signal perfectly reveals the risky arm's type. The interpretation of this information structure is that the agent

[1] We analyse and present the case where b is positive as a benchmark.

gets a chance to pull the risky arm costlessly for some time before the game starts. If the agent ever observes any lump sum payoff, he updates his belief to $\bar{p}_0 = 1$ and is called the high type agent. If not, the agent is a low type one and assigns probability $\underline{p}_0$ to the risky arm being good at time 0. Let q_0 be the probability of high type agent at time 0. Then we have $p_0 = q_0 + (1 - q_0)\underline{p}_0$.

Delegation. At time t, the principal determines a non-empty permissible set $D_t \subset \{R, S\}$ and then the agent chooses an action $a_t \in D_t$. The discrete analogue is that at time $t \in \{0\Delta, 1\Delta, 2\Delta, ...\}$, the principal assigns a non-empty set D_t from which the agent chooses an action a_t which is played in $[t, t + \Delta)$. The principal would like to utilize the agent's private information while restricting the expression of his bias. The equilibrium concept is Markov perfect equilibrium and we characterize the equilibria as $\Delta \to 0$ that gives the principal the highest expected payoff.

3 Main Results

Faced with a biased yet better informed agent, the principal can use her delegation authority to achieve a higher expected payoff than the case where the principal always assigns singleton delegation sets and 'takes over' the task of experimentation. One useful tool is that the principal can let the agent choose either to stop experimenting or to experiment perpetually. This menu does not allow the agent to switch back to the safe arm when he gets sufficiently pessimistic and force the low type agent to reveal his type earlier than he would like to. Another useful tool to induce information is through possibly *excessive* experimentation. Here, excessive experimentation means that the principal first assigns $\{R\}$ for a period longer than what she would like to if she had seen the low type agent's signal. The purpose of excessive experimentation is to let the low type agent become sufficiently pessimistic to separate from the high type agent. The optimal delegation rule, shaped by the tradeoff between the benefit from separating two types of agents and the cost of excessive experimentation, depends on the degree of the agent's bias.

When the bias term b is sufficiently small, the principal always offers the delegation set $\{R\}$ until the low type agent's belief hits the principal's optimal stopping time (had the principal seen a low signal at time 0). Then, by offering either to stop experimentation or to experiment perpetually, the principal is able to separate two types of agents and gets the expected payoff as if she had seen the agent's signal.

When the misalignment of preferences is moderate, the principal always offers the delegation set $\{R\}$ until the low type agent prefers to play S forever than to play R perpetually. The principal gains from separating two types of agents but has to pay the cost of excessive experimentation.

As b gets sufficiently large, the cost of excessive experimentation outweighs the informational gain from the separation of two types of agents. The principal always assigns singleton delegation sets and 'takes over' the task of experimentation.

4 Conclusion

We characterize the optimal delegation rule by a principal when she faces a better informed and biased agent. Without being able to commit to future actions, the principal can not separate two types of agents at time 0. Instead, the principal has to assign delegation set $\{R\}$ and 'experiment' for a period sufficiently long before the low type agent is willing to separate from a high type one. The principal finds it optimal to induce the agent's private information when bias is small or moderate and prefers to 'take over' the task of experimentation when the bias is sufficiently large. We also show that the principal is weakly better off if she can commit to a contingent delegation plan based on the agent's reports.

References

[1] Amador, Manuel and Kyle Bagwell. 2012. The Theory Of Optimal Delegation With An Application to Tariff Caps. *Econometrica*, forthcoming.

[2] Armstrong, Mark and John Vickers. 2010. A Model of Delegated Project Choice. *Econometrica*, 78(1): 213-244.

[3] Bergemann, Dirk and Jusso Välimäki. 1996. Learning and Strategic Pricing. *Econometrica*, 64: 1125-49.

[4] Halac, Marina, Navin Kartik and Qingmin Liu. 2012. Optimal Contracts for Experimentation. *Working paper*.

[5] Holmström, Bengt. 1977. On Incentives and Control in Organizations. *PhD dissertation,* Stanford University.

[6] Hörner, Johaness, and Larry Samuelson. 2012. Incentives for Experimenting Agent. *Working paper*.

[7] Keller, Godfrey, Sven Rady and Martin Cripps. 2005. Strategic Experimentation with Exponential Bandits. *Econometrica*, 73(1): 39-68.

[8] Mylovanov, Tymofiy. 2008. Veto-Based Delegation. *Journal of Economic Theory*, 138: 297-307.

[9] Rosenberg, Dinah, Eilon Solan and Nicolas Vieille. 2007. Social Learning in One-arm Bandit Problems. *Econometrica*, 75(6): 1591-1611.