Theoretical Economics 16 (2021), 381–402 1555-7561/20210381

On selecting the right agent

Geoffroy de Clippel
Department of Economics, Brown University

Kfir Eliaz
Department of Economics, Tel-Aviv University and Department of Finance, University of Utah

Daniel Fershtman
Department of Economics, Tel-Aviv University

Kareen Rozen
Department of Economics, Brown University

Each period, a principal must assign one of two agents to a new task. Each agent privately learns whether he is qualified for the task. An agent wishes to be chosen independently of qualification and chooses whether to apply for the task. The principal wishes to appoint the most qualified agent and chooses which agent to assign as a function of the public history of profits. We fully characterize when the principal's first-best payoff is attainable in equilibrium and identify a simple strategy profile achieving this first-best whenever feasible. Additionally, we provide a partial characterization of the case with many agents and discuss how our analysis extends to other variations of the game.

Keywords. Dynamic allocation without transfers, repeated games with imperfect monitoring.

JEL classification. C73, D23.

1. Introduction

How should a principal dynamically assign tasks to the agents who are most qualified to complete them when agents hope to be selected regardless of their qualification and when the principal cannot observe qualification or use contingent monetary transfers?

Geoffroy de Clippel: declippel@brown.edu
Kfir Eliaz: kfire@tauex.tau.ac.il
Daniel Fershtman: danielfer@tauex.tau.ac.il
Kareen Rozen: kareen_rozen@brown.edu

This paper greatly benefitted from helpful comments by two anonymous referees. Special thanks are also due to Nageeb Ali, Heski Bar-Isaac, Eddie Dekel, Yingni Guo, Johannes Hörner, Ehud Kalai, Jin Li, Benny Moldovanu, Alessandro Pavan, Michael Powell, Phil Reny, Ariel Rubinstein, Yuval Salant, Larry Samuelson, Lones Smith, Bruno Strulovici, Asher Wolinsky, and various seminar audiences for their helpful suggestions. This paper builds on and merges ideas from earlier works, previously circulated as "The Silent Treatment" (de Clippel, Eliaz, and Rozen) and "Dynamic Delegation: Specialization and Favoritism" (Fershtman). Geoffroy de Clippel and Kfir Eliaz thank the BSF for financial support under Award 003624. Daniel Fershtman gratefully acknowledges financial support from the German Research Foundation through CRC TR 224.

 https://doi.org/10.3982/TE4027

This question is pertinent to many economically relevant situations. Consider a manager who must decide which employee to assign to a new project or client, or a politician in office who needs to designate a staffer in charge of new legislation, or an organization that needs someone to direct a new initiative. Oftentimes, such employees receive a monthly salary or fixed payment per task. Interested employees may be required to communicate their availability, provide some evidence of serious intention, or pitch their vision for the project at hand. Alternatively, one can think of situations where the agents propose "ideas" to a decision-maker. For instance, think tanks and researchers submit proposals for a grant; engineers suggest directions for new versions of a product. The problem can also be interpreted as a stylized representation of a median voter choosing between office-driven politicians in each election.

To address the above question, we analyze an infinitely repeated game between a principal and two agents. The key features of the game are the following. Every period, a new task arrives and each agent privately learns whether he is qualified for it, where the probability of being qualified (denoted by θ) is commonly known. The agents simultaneously decide whether to apply for the task and the principal selects at most one applicant. Each agent cares only about being selected. The principal, however, cares about the profit from a completed task, which is either high or low (an unassigned task generates no profit), such that a high profit is more likely for a qualified agent. It follows that the principal's *first-best* outcome is to pick the most qualified agent in every period. The question is, under what conditions can the first-best be achieved in a perfect public equilibrium (PPE) and with what strategies?

Our first main result answers this question by characterizing the full set of parameter values (the probability of being qualified, the probabilities of a high profit for a qualified and unqualified agent, and the common discount factor) for which the first-best is attainable in PPE. In addition, we identify a simple strategy profile, dubbed the *Markovian last resort* (MLR), that achieves the first-best whenever it is feasible; that is, over the entire set of parameter values for which first-best is attainable.

The MLR strategy profile can be described as follows. At each history, one agent is designated as the agent of *last resort*, and the remaining agent is designated as *discerning*. The agent of last resort proposes himself regardless of whether he is qualified, while the discerning agent proposes himself if and only if he is qualified. The principal selects the agent of last resort if he is the only one available and otherwise picks the discerning agent. The first agent of last resort is chosen arbitrarily and he remains in that role as long as all of the principal's past profits were high. Otherwise, the agent of last resort is the most recent agent who generated a low profit. This profile has the following appealing features. First, it requires players to keep track of very little information: they need only know who was the last agent to generate low profit. Second, it does not require the agents to punish the principal to ensure she follows the strategy: MLR remains an equilibrium even when the principal's discount factor is zero. Furthermore, the MLR strategy profile is also an ex post PPE with respect to agents' qualifications: taking expectations over the future path of play, each agent's proposal decision remains optimal regardless of his belief about the other agent's privately observed qualification.

In Section 3, we turn to analyze the challenging case of more than two agents. We generalize the MLR strategy profile in a natural way by having $n-1$ discerning agents and the principal choosing at random from among discerning proposers whenever possible. Clearly, the MLR profile delivers the first-best outcome for the principal and the only question remaining is when it constitutes an equilibrium. We first note that it is impossible to attain the principal's first-best in PPE (or even in Nash equilibrium) if agents' abilities are below $1-\sqrt[n-1]{\frac{1}{n}}$. We then characterize the set of parameter values for which the MLR profile constitutes a PPE. We show that when ability is strictly above $1-\sqrt[n-1]{\frac{1}{n}}$, the MLR is an equilibrium when agents are patient enough and realized profits are sufficiently informative of qualification. In this sense, we obtain a characterization of the widest range of *abilities* for which first-best is achievable and show that it is achievable by the MLR.[1]

This leaves open the question of whether another strategy profile attains the principal's first-best in PPE for a wider range of parameters than the MLR. To at least partially address this question, we compare the performance of the MLR with an intuitive class of strategy profiles, which we call *hierarchical*. In a hierarchical strategy profile, agents are assigned priorities, the lowest-priority agent serves as last resort while all other agents are discerning, the principal picks the proposing agent with the highest priority, and a discerning agent moves down the ranking if he generates a low profit, with the ranking of agents with a higher priority than him unaffected. The MLR profile can be thought of as a "flat" hierarchy with only two tiers: the last resort is at the bottom and everyone else has the same priority. Would more tiers help attain the principal's first-best in PPE for a *wider* range of parameter values? We show that (i) no hierarchical strategy profile "dominates" MLR in the sense of attaining the first-best in PPE whenever MLR does and (ii) MLR dominates any hierarchical profile that sends a "failing" agent to the bottom of the ranking.

Our paper provides a thorough analysis of a common strategic dilemma: how should one select the "right" expert (idea, candidate) when the supply side mainly cares about being chosen and possesses private information pertinent for identifying the right choice? While we naturally abstract from many details present in real-life situations, many of these often share a few key features with our stylized model: the decision-maker repeatedly faces the same group of individuals who want to be selected, she cannot credibly commit to a decision rule, and cannot make contingent transfers. Our analysis identifies a simple and intuitive strategy profile that attains the decision-maker's first-best payoff whenever this is feasible. Its structure is independent of the parameters and is reminiscent of the tendency to avoid—whenever possible—choosing the most recent individual to generate a disappointing result.

[1] It remains an open question whether the set of parameters where MLR is a PPE corresponds to the widest set of *all parameters* for which first-best is achievable. This is in contrast to the two-agent case, in which this is indeed the case. The difficulty stems from the fact that, unlike in the two-agent case, the shape of the set of PPE payoffs is unknown. In particular, it is unclear whether it is feasible to bring more than one agent to the lowest PPE payoff. Indeed, we are not aware of any work that fully characterizes the set of PPE payoffs in a setting with incomplete information, no transfers, and more than two players.

The remainder of the paper is organized as follows. The next section introduces the model. Section 2 presents our main results for the two-agent case and Section 3 analyzes the case of more than two agents. The related literature is discussed in Section 4 and Section 5 discusses various extensions.

2. A model

There is one principal and two agents, 1 and 2. Each period $t = 0, 1, 2, \ldots$ there is a new task (or project) available and the principal can choose at most one agent to carry it out. The principal's profit from a period-t project is $y_t \in \{0, 1\}$ and is determined stochastically depending on whether the agent assigned to carry it out is qualified to do so. An unassigned project generates zero profits. A qualified agent has probability $\alpha \in (0, 1)$ of generating high profit for the principal, while an unqualified agent generates high profit with a strictly smaller probability $\beta \in [0, \alpha)$. Given our normalization for profits, $\beta \geq 0$ implies the principal prefers to hire an unqualified agent over hiring no one. In each period t, the probability that each agent i is qualified for the current project is constant and equal to θ. Thus, the parameter θ captures the common ability of the agents.[2] Each agent privately observes whether he is qualified for the specific project at hand, but the agents' ability level θ is commonly known.

In every period, the stage game unfolds as follows. Each agent privately observes whether he is qualified for the current project and decides whether to submit a proposal to the principal. The principal then decides which agent, if any, to select among the proposers.

Agent i gets a positive payoff in period t if the principal picks him in that period. We normalize this payoff to 1 (having a different payoff for each agent has no effect on our analysis). Agent i's objective is then to maximize the expectation of the discounted sum $\sum_{t=0}^{\infty} \delta^t 1\{x_t = i\}$, where δ is each agent's discount factor, $1\{\cdot\}$ is the indicator function, and $x_t \in \{1, 2\} \cup \{\varnothing\}$ is the identity of the agent that the principal picks in period t, if any. That is, each agent simply wants to be selected regardless of the end profit from the project.[3]

The principal's profit in a given period is zero if she does not choose any agent and is otherwise equal to the realized profit from the project. Her objective is to maximize the expectation of the discounted sum $\sum_{t=0}^{\infty} \delta_0^t y_t$, where δ_0 is the principal's discount factor and $y_t \in \{0, 1\}$ is her period-t profit.

The agents' proposal decisions, the agent chosen by the principal (if any), and the realized profit are all publicly observed.[4] We define a *public history* at any period t as

[2] We later discuss how the results extend to the case of individual-specific abilities.

[3] The assumption that agents simply want to be selected regardless of the end profit from the project captures situations where agents want to accumulate experience, build a resume, or obtain certain resources associated with carrying out a project, and where the principal's payoff from a project cannot be verified by an outside party. Our analysis would not change (but it would be more tedious) if each agent also received some fixed bonus when profits are high. We refer the reader to our working paper for a treatment of this case.

[4] As we will show, our results would not change if players could only observe the identity of the last agent who generated a low profit for the principal.

the sequence $h^t = ((x_{t'}, y_{t'}, S_{t'}))_{t'=0}^{t-1}$, where $S_\tau \subseteq \{1,2\} \cup \{\varnothing\}$ is the set of agents who made a proposal in each period $\tau < t$, and, as defined above, x_τ and y_τ denote the chosen agent and the profit he generated. A *public strategy for agent i* determines, for each period t, the probability with which he makes a proposal to the principal as a function of his current qualification and the public history of the game. A *public strategy for the principal* determines, for each period t, a lottery over which agent to select (if any) from among the set of agents who propose, given that set of proposers and the public history of the game. We apply the notion of *perfect public equilibrium* (PPE), that is, sequential equilibria where players use public strategies.

Discussion. There are three key features in our model. First, the principal is better off selecting some agent than not selecting any, which fits situations where the loss from not performing a task outweighs the loss from not doing it well. This assumption pins down an important property of the first-best: at every history, one agent must propose himself if and only if he is qualified, while the other agent must propose regardless of his qualifications. As is evident from the proof of our main result, this implication facilitates the derivation of the necessary conditions for attaining the first-best. In our concluding remarks, we briefly discuss the case in which an unqualified agent leads to an expected loss.

Second, the principal cannot sign complete contracts with the agents that specify transfers as a function of profits. This feature captures situations where either the principal's payoff cannot be verified by an outside party (e.g., it may include intangible elements such as perceived reputation) or because of institutional constraints that preclude such contracts (as in most public organizations where subordinates, who receive a constant wage, may propose themselves to an executive decision-maker).

Third, the principal cannot pick an agent who has not submitted a proposal. This captures situations where either institutional norms or explicit rules require an agent to give tangible evidence for his ability to take on the project and to lay out his plans explicitly. Allowing agents' messages to be cheap talk (in the sense that a principal can still pick an agent who declares himself unqualified or unavailable) significantly complicates the derivation of necessary and sufficient conditions for attaining the first-best in PPE. It, therefore, remains an open question whether there exists a single strategy profile that implements the first-best in PPE for the widest range of parameters.[5]

3. Main result

A strategy profile achieves the principal's first-best if a qualified agent is chosen in every period where at least one agent is qualified and some agent is chosen in all other periods.[6] We say that an agent is discerning when he applies if and only if he is qualified.

[5]While the MLR strategy profile described in the Introduction still attains the first-best in PPE, we do not know if it does so for the largest set of parameters. The source of complication is that under cheap-talk messages, the first-best allocation (who should be assigned the task at each history) is not pinned down. Consequently, there are many continuation payoffs that need to be considered in deriving the necessary conditions for attaining the first-best in PPE.

[6]Of course, the principal would prefer to pick only high-profit proposals when possible, but no one knows at the selection stage whether high profit will be realized.

Clearly, a strategy profile implements the first-best in PPE only if every period at least one agent is discerning.

Our main result consists of two parts. First, it provides a complete characterization of the parameter values for which the principal can attain the first-best in *any* PPE. Second, it shows that a simple strategy profile, which we next introduce, attains the first-best PPE payoff over the entire region of parameters for which a first-best PPE exists.

DEFINITION 1 (The Markovian Last-Resort Strategy Profile). At each history, one agent is designated as the agent of last resort and the remaining agent is designated as discerning. The agent of last resort proposes himself independently of his qualification, while the discerning agent proposes himself if and only if he is qualified. The principal selects the agent of last resort if he is the only one available, and otherwise picks the discerning agent. The identity of the initial agent of last resort is chosen arbitrarily, and he remains in that role as long as all the principal's past profits were high. Otherwise, the agent of last resort is the most recent agent who generated low profit for the principal.

Clearly, the principal achieves her first-best if she and the agents follow the MLR strategy profile. She is sure to select an agent each period and will select a qualified agent whenever one exists. The question then is, under what conditions is this profile a PPE?

PROPOSITION 1. *(a) A PPE that attains the principal's first-best exists if and only if*

$$\delta \geq \frac{1}{\beta + 2\theta(\alpha - \beta)}. \tag{1}$$

(b) The MLR strategy profile is a PPE if and only if (1) holds. Hence, there is a strategy profile that attains first-best in PPE if and only if the MLR profile is a PPE.

This result implies that the first-best is attainable in equilibrium when agents are patient enough if and only if the agents are sufficiently able, in the sense that

$$2\theta > \frac{1-\beta}{\alpha-\beta}, \tag{2}$$

where the inverse of the right-hand side, $\frac{\alpha-\beta}{1-\beta} = 1 - \frac{1-\alpha}{1-\beta} < 1$, measures how informative low profits are of qualification. Since the right-hand side is greater than 1, this inequality holds only if $\theta > \frac{1}{2}$. Thus, the first-best is attainable in PPE only if the agents are more likely to be qualified than not.

The key incentive constraint, which generates condition (1), is the one facing an unqualified discerning agent. To express this constraint, let V_1^{D} and V_1^{LR} represent agent 1's average discounted payoff (prior to learning his qualification status) under the MLR strategy profile when he is discerning and when he is last resort, respectively. These are defined as

$$V_1^{\mathrm{D}} = \theta\big((1-\delta)\cdot 1 + \alpha\delta V_1^{\mathrm{D}} + (1-\alpha)\delta V_1^{\mathrm{LR}}\big) + (1-\theta)\delta V_1^{\mathrm{D}},$$
$$V_1^{\mathrm{LR}} = (1-\theta)\big((1-\delta)\cdot 1 + \delta V_1^{\mathrm{LR}}\big) + \theta\big(\alpha\delta V_1^{\mathrm{LR}} + (1-\alpha)\delta V_1^{\mathrm{D}}\big).$$

To see why, consider first the continuation payoff when agent 1 is discerning. With probability θ, he is qualified, and, hence, proposes himself and is chosen, receiving an immediate payoff of 1. If he succeeds—an event with probability α—he remains discerning in the next period. Otherwise, he becomes last resort. With probability $1-\theta$, the discerning agent is not qualified and does not propose, leading to the selection of the last-resort agent. The continuation payoff of the last-resort agent is derived in a similar way. The only difference is that the last-resort agent is chosen only when the discerning agent is unqualified (which happens with probability $1-\theta$), and he switches roles when the discerning agent is picked and fails (which happens with probability $\theta(1-\alpha)$).

The incentive compatibility (IC) constraint for an unqualified discerning agent not to propose is given by

$$\delta V_1^{\mathrm{D}} \geq 1-\delta+\beta\delta V_1^{\mathrm{D}}+(1-\beta)\delta V_1^{\mathrm{LR}},$$

which can be rewritten as

$$V_1^{\mathrm{D}}-V_1^{\mathrm{LR}} \geq \frac{1-\delta}{\delta(1-\beta)}. \tag{3}$$

Solving explicitly for V_1^{D} and V_1^{LR} yields

$$V_1^{\mathrm{D}}-V_1^{\mathrm{LR}} = \frac{(1-\delta)(2\theta-1)}{1-\delta+2\theta\delta(1-\alpha)}. \tag{4}$$

Plugging this expression into the left-hand side of (3) yields condition (1).

The intuition for why the MLR attains the first-best for the widest range of parameters is more subtle. First, note that attaining the first-best restricts the type of punishments that can be levied on agents. In particular, at no history can there be an agent who is selected with probability 0 *regardless* of the agents' choices in that period. The reason is that during those periods that agent may be the only one qualified and would, therefore, need to be chosen to attain the first-best. Second, note that in any period, exactly one agent is discerning and one is last resort. Since the last-resort agent proposes himself regardless of his qualifications, he cannot be incentivized. As evident from (4), the last-resort agent is worse off than the discerning agent if $\theta > \frac{1}{2}$, which must be true for (2) to hold. Hence, the harshest possible punishment is to keep an agent as last resort for as long as possible, conditional on motivating the other agent. The best possible reward is to make an agent discerning for as long as possible, conditional on motivating him. MLR does both.

The MLR profile has several desirable properties. First, it is robust to heterogeneous abilities: Proposition 1 extends to the case in which agent i's probability of being qualified is θ_i with $\theta_1 \neq \theta_2$. In this case, the term 2θ in (1) is simply replaced with the sum of abilities $\theta_1+\theta_2$ (the proof is analogous to that of Proposition 1, only slightly more tedious; for details, see our working paper, de Clippel et al. (2019)).[7] Thus, the first-best

[7]Allowing for heterogeneity in success probability, MLR is a PPE as soon as i's discount factor is larger than or equal to $\frac{1}{\beta_i+2\theta(\bar{\alpha}-\beta_i)}$, where $\bar{\alpha}$ is the average of α_1 and α_2. Notice, however, that the principal's first-best is unattainable when these probabilities are common knowledge. Indeed, it requires choosing the high

becomes harder to attain in PPE (in the sense of having a smaller range of parameter values for which the first-best is attainable) the lower are the agents' abilities.

This last observation implies that the MLR strategy profile also achieves the principal's first-best in a "belief-free" way when there is unobservable heterogeneity in the agents' ability. Specifically, suppose each agent privately draws his ability from some distribution over the interval $[\underline{\theta}, 1)$, and the principal wants to guarantee her first-best outcome for *all* realizations of (θ_1, θ_2). To accommodate this form of robustness, define a *belief-free equilibrium* to be a strategy profile that constitutes a PPE for *any* realized vector of abilities. It follows that the first-best is attainable in a belief-free equilibrium if and only if (1) holds for $\theta = \underline{\theta}$. Furthermore, the MLR profile achieves the first-best in a belief-free equilibrium whenever this is feasible.

Second, the principal and the agents need not observe, nor remember, much information about past behavior. At any history, the principal's selection decision is based only on the identity of the current last-resort agent—which changes if and only if a discerning agent fails—and the set of agents who propose. In particular, past proposals play no direct role and high profit realizations do not trigger changes in the identity of the last-resort agent.

Third, the principal's selection rule is optimal for her without relying on the agents to punish her if she deviates from it, thereby providing endogenous commitment. While efficient equilibria in the literature often rely on any deviator to be punished by others, we would find it unnatural in our environment if the principal were to follow her part of an equilibrium that achieves her first-best only because she fears the agents will punish her otherwise. Indeed, the MLR strategy profile remains a PPE independently of the principal's discount factor.

Relative to the discussion, in our model there is no institutional device that enables the principal to credibly commit to a selection policy. However, since we focus on attaining the principal's first-best, allowing the principal to commit would not enlarge the set of parameters for which the first-best is attainable. What restricts this set of parameters are the agents' incentive constraints.

Fourth, the MLR strategy profile addresses questions of equilibrium robustness. From the proof of Proposition 1, it is clear that the MLR strategy profile is in fact an ex post PPE whenever (1) holds: taking expectations over the future path of play, each agent's proposal decision remains optimal irrespective of his belief about the other agent's current private information (i.e., whether the other agent is qualified).[8] In an ex post equilibrium, stringent (simultaneous and private) communication protocols are not necessary. This robustness, which is particularly relevant for environments where it is difficult to restrict how agents share information, comes for free in our environment.

probability agent whenever he is qualified, which is possible only if he is discerning at all rounds. But this cannot be done, as it eliminates the possibility of punishment. Whether the MLR achieves a second-best for the widest range of parameters remains an open question.

[8]Such notions of equilibrium, imposing ex post incentive compatibility in each period taking expectations over the future path of play, were introduced separately by Athey and Miller (2007) and Bergemann and Välimäki (2010). The latter use the term "periodic ex post." Miller (2012) considers ex post PPE in a model of collusion with adverse selection. In addition, ex post equilibria are robust to the introduction of payoff-irrelevant signals and high-order beliefs; see Bergemann and Morris (2005).

If we focus on ex post PPE, then MLR also achieves the first-best for the widest range of parameters when the principal can choose an agent even if he did not propose himself. Details can be found in de Clippel et al. (2019).

Finally, our results imply that even if the principal could incur a cost to figure out which agent is better qualified prior to making the selection, under condition (1) the repeated nature of the interactions with the agents allows her to reach the first-best without having to resort to this costly verification.

4. Many agents

In the previous section we established that when the principal faces two agents, there is a simple and intuitive strategy profile—the MLR—that attains the principal's first-best in PPE whenever the first-best is attainable in PPE.

In this section, we examine how some of our results generalize when there is a set $\mathcal{A} = \{1, 2, \ldots, n\}$ of $n \geq 2$ agents, each with ability θ.[9] Our first observation identifies a necessary condition for the existence of any PPE that attains the principal's first-best. To present this result, define the threshold ability level $\theta^* = 1 - \sqrt[n-1]{\frac{1}{n}}$, which decreases in n (starting from $1/2$ for $n = 2$) and tends to 0 as n tends to infinity.

Proposition 2. *If $\theta < \theta^*$, then there is no PPE (and even no Nash equilibrium) that attains the principal's first-best.*

Proof. First, observe that the underlying principle from our earlier analysis generalizes to $n \geq 2$ agents. To achieve the principal's first-best in PPE, at each history h, there must be $n-1$ discerning agents, each of whom proposes himself if and only if he is qualified, one agent $i(h)$ of last resort who proposes himself irrespective of his qualifications, and the principal who must pick some $i \neq i(h)$ whenever possible, using $i(h)$ only as a last resort. Notice that $i(h)$ is picked with probability $(1-\theta)^{n-1}$, corresponding to no discerning agent being qualified. If a discerning agent were to deviate and propose himself regardless of qualification, he could guarantee a probability at least $(1-\theta)^{n-2}$ of getting picked, which exceeds $(1-\theta)^{n-1}$. By proposing oneself in all periods, each agent can thus secure a discounted probability of being chosen that is at least $(1-\theta)^{n-1}/(1-\delta)$. The principal will pick exactly one agent in each round in her first-best PPE, so the aggregate discounted probability of being picked is $1/(1-\delta)$. The equilibrium could not exist if $1/(1-\delta)$ were strictly smaller than the sum of aggregate discounted probabilities that each agent can guarantee. The cutoff θ^* above is the smallest θ satisfying this constraint. □

[9] We refer the reader to our working paper for extensions with heterogenous abilities $\vec{\theta}$, where θ_i is the ability of agent i. Proposition 2 below extends verbatim. Proposition 3, which characterizes the set of parameters for which the MLR (generalized to n agents) forms a PPE, extends after taking care of the following complexities. An agent i's continuation payoff from being discerning depends both on $\vec{\theta}$ and the identity of the agent currently removed from the discerning pool, and similarly the continuation payoff from being last resort depends on all of $\vec{\theta}$. The ICs are thus potentially asymmetric, but they can be summarized in matrix form, and we use linear algebra to characterize when the MLR strategy profile forms a PPE. We also provide the inequality on parameters that characterizes when MLR forms a belief-free PPE, that is, a PPE whatever the agents' abilities are in the set $[\underline{\theta}, 1]^n$.

4.1 *Characterizing when MLR is a PPE*

We now generalize the MLR strategy profile by treating all $n-1$ discerning agents in a symmetric manner, with the principal randomizing uniformly when selecting among discerning agents who have proposed. At the very beginning of a period—before agents learn their qualifications—we have the following scenarios:

- The ex ante probability that the last resort agent is chosen is $\rho = (1-\theta)^{n-1}$.
- The ex ante probability of being selected as a discerning agent is $\frac{1-\rho}{n-1}$,
- The premium (in terms of the increased ex ante probability of selection) from being a discerning agent, instead of the agent of last resort, is $\pi = \frac{1-\rho}{n-1} - \rho = \frac{1-n\rho}{n-1}$.

We may now characterize when the MLR strategy profile is a PPE.

PROPOSITION 3. *The MLR strategy profile is a PPE if and only if*

$$\delta \geq \frac{1}{\alpha + (\alpha-\beta)\pi}. \tag{5}$$

PROOF. Assume that players follow the MLR strategy profile. Clearly, neither the last-resort agent nor the principal have profitable unilateral deviations. We need to check that a discerning agent proposes himself if and only if he is qualified. Let σ be the probability that a discerning agent is picked conditional on proposing himself, that is, $\sigma = \frac{1-\rho}{(n-1)\theta}$. A discerning agent i who is unqualified refrains from proposing himself if

$$\delta V_i^{\mathrm{D}} \geq \underbrace{\sigma}_{i \text{ selected}} \big((1-\delta) + \underbrace{\beta\delta V_i^{\mathrm{D}}}_{\text{high profit}} + \underbrace{(1-\beta)\delta V_i^{\mathrm{LR}}}_{\substack{\text{low profit, } i \text{ becomes} \\ \text{last-resort agent}}}\big) + \underbrace{(1-\sigma)\delta V_i^{\mathrm{D}}}_{i \text{ not selected}}, \tag{6}$$

where V_i^{D} and V_i^{LR} represent i's average discounted payoff (before learning his qualification) under the MLR strategy profile when discerning and last resort, respectively. Similarly, a discerning agent i who is qualified proposes himself if

$$\underbrace{\sigma}_{i \text{ selected}} \big((1-\delta) + \underbrace{\alpha\delta V_i^{\mathrm{D}}}_{\text{high profit}} + \underbrace{(1-\alpha)\delta V_i^{\mathrm{LR}}}_{\substack{\text{low profit, } i \text{ becomes} \\ \text{last-resort agent}}}\big) + \underbrace{(1-\sigma)\delta V_i^{\mathrm{D}}}_{i \text{ not selected}} \geq \delta V_i^{\mathrm{D}}. \tag{7}$$

Let us first examine incentive condition (6). We subtract δV_i^{LR} from both sides of the inequality (6) and let Δ_i represent $V_i^{\mathrm{D}} - V_i^{\mathrm{LR}}$. Then incentive condition (6) is equivalent to

$$\delta\Delta_i \geq \sigma(1-\delta) + \sigma\beta\delta\Delta_i + (1-\sigma)\delta\Delta_i,$$

which can be rearranged to obtain the inequality

$$\Delta_i \geq \frac{(1-\delta)}{(1-\beta)\delta}. \tag{8}$$

Similar computations show that inequality (7) is equivalent to

$$\Delta_i \leq \frac{(1-\delta)}{(1-\alpha)\delta}. \tag{9}$$

The payoff difference Δ_i from being a discerning agent instead of the last-resort agent can be computed through the recursive equations defining V_i^{D} and V_i^{LR}:

$$\begin{aligned} V_i^{\mathrm{D}} &= \overbrace{\theta\sigma}^{\substack{\text{qualified,}\\ \text{selected}}} \left((1-\delta) + \alpha\delta V_i^{\mathrm{D}} + (1-\alpha)\delta V_i^{\mathrm{LR}}\right) + \overbrace{(1-\theta\sigma)}^{\substack{\text{unqualified or}\\ \text{not selected}}} \delta V_i^{\mathrm{D}}, \\ V_i^{\mathrm{LR}} &= \underbrace{\rho}_{\text{selected}} \left((1-\delta) + \delta V_i^{\mathrm{LR}}\right) + \underbrace{(1-\rho)}_{\text{not selected}} \left(\alpha\delta V_i^{\mathrm{LR}} + \underbrace{(1-\alpha)\delta V_i^{\mathrm{D}}}_{\substack{\text{low profit, switch}\\ \text{to discerning}}}\right). \end{aligned} \tag{10}$$

Replacing V_i^{D} by $V_i^{\mathrm{LR}} + \Delta_i$, notice that the expression for V_i^{LR} can be rewritten as

$$V_i^{\mathrm{LR}} = \rho(1-\delta) + \delta V_i^{\mathrm{LR}} + (1-\rho)(1-\alpha)\delta\Delta_i.$$

Subtracting this new expression for V_i^{LR} from that for V_i^{D} in (10), we get

$$\Delta_i = \pi(1-\delta) + \theta\sigma\alpha\delta\Delta_i + (1-\theta\sigma)\delta\Delta_i - (1-\rho)(1-\alpha)\delta\Delta_i$$

or

$$\Delta_i = \frac{\pi(1-\delta)}{1-\delta+\delta(1-\alpha)(1+\pi)}.$$

Using this expression for Δ_i, we conclude that the incentive condition (9) (proposing when qualified) is always satisfied, and that the incentive condition (8) (not proposing when unqualified) is satisfied if and only if (5) holds, as claimed. □

From Proposition 3, we see that the MLR strategy profile forms a PPE when agents are patient enough if and only if $\alpha + (\alpha-\beta)\pi > 1$ or

$$\pi > \frac{1-\alpha}{\alpha-\beta}. \tag{11}$$

For (11) to hold, low profits need to be sufficiently informative of the agent's lack of qualification. This follows from observing that the inverse of the right-hand side, $\frac{\alpha-\beta}{1-\alpha} = \frac{1-\beta}{1-\alpha} - 1$, increases with the likelihood ratio $\frac{1-\beta}{1-\alpha}$, which measures the extent to which it is more likely that low profits originated from a nonqualified agent. Thus, when π is strictly positive, the principal's first-best is achievable in equilibrium, provided that agents are patient enough and profits are sufficiently informative of qualification. Moreover, π is strictly positive if and only if $\theta > \theta^*$ (as can be seen using $\pi = \frac{1-n\rho}{n-1}$ and $\rho = (1-\theta)^{n-1}$). Combined with Proposition 2, we conclude that we have identified the widest range of abilities (θ) for which the principal can achieve her first-best in a PPE. Contrary to the two-agent case, we have not been able to identify the widest range of all parameters (α, β, δ, and θ) for which this is feasible. The next subsection elaborates on this.

4.2 *Hierarchies*

A natural question is whether a strategy profile other than MLR achieves the principal's first-best in PPE for a wider range of parameters. A complete characterization of the necessary and sufficient conditions for attaining the first-best in PPE is a challenging task with three or more agents. It is not immediately clear how the proof technique used for the $n = 2$ case extends to $n \geq 3$. First, solving the minimization problem to find the lowest discounted probability with which an agent is picked in equilibrium is very challenging to solve. Second, and more importantly, it is not clear that finding this minimum would allow to characterize the range of parameters for which the principal's first-best is achievable. This is because we do not know the shape of the convex set of equilibrium payoffs (which must be an interval for $n = 2$).[10]

We thus propose to evaluate the performance of MLR against an intuitive class of alternative strategy profiles. A strategy profile is *hierarchical* if following each history h, the principal uses a ranking (i.e., strict ordering) R_h of all the agents such that the following scenarios exist:

(i) In the period following history h, the principal picks the proposing agent ranked highest according to R_h.

(ii) If high profit is generated in the period following h or if the lowest-ranked agent under R_h was picked, then the ranking in the next period remains R_h.

(iii) If low profit is generated, then a deterministic rule is applied to generate the next period's ranking as a function of the current rank k of the failing agent. Under this rule, agents ranked above agent k keep their positions.

(iv) The top $(n-1)$-ranked agents under R_h propose if and only if they are qualified (i.e., are discerning), while the bottom-ranked agent always proposes himself.

Some examples of rules that determine how the agents' rankings change when a discerning agent generates low profit are (a) the failing agent drops to the bottom of the ranking and every agent ranked below i moves up one rank, (b) the failing agent switches ranks with the bottom-ranked agent, and (c) the failingagent switches ranks with the agent right below him. There are many possibilities, but none clearly dominates MLR.

PROPOSITION 4. *For two strategy profiles s and s' achieving the principal's first-best, say s dominates s' if s forms a PPE for all values $(\beta, \alpha, \delta, \theta)$ at which s' does. Then there are two cases:*

(a) No hierarchical strategy profile dominates MLR.

(b) MLR does not dominate all hierarchical strategy profiles, but it does dominate any such profile that sends a failing agent to the bottom of the ranking.

[10]We are not aware of applications of Abreu et al. (1990) to derive simple closed-form solutions in problems with more than two players and *no transfers*.

To get some rough intuition for (b), suppose there are three agents, with 1 ranked first, 2 ranked second, and 3 the last resort. The tough constraint is to get agent 2 to report honestly, since his continuation value is lower than 1's, meaning that 1's IC is slack when 2's binds. If we treat 1 and 2 equally, we can introduce slack into 2's IC and enlarge the set of parameter values for which first-best is achievable.

It remains an open question whether there exists some strategy profile, which is not MLR and lies outside the class of hierarchical strategy profiles, that achieves the principal's first-best in PPE for the widest range of parameters. If no such profile exists, then Proposition 4 suggests a more complex picture, where different strategy profiles have to be used for different values of parameters to maximize the range of parameters where first-best is achievable in PPE. In the proof (in the Appendix), we show that MLR works for some parameter values, while switching a failing agent with the next in the hierarchy works for others.

5. Related literature

Our paper relates to several strands of literature. In our problem, the principal uses a form of dynamic favoritism, the promise (threat) of future (dis)advantage, as a means to align incentives. Strategic use of favoritism also arises in static mechanism-design environments without monetary transfers. For example, Ben-Porath et al. (2014) characterize the optimal mechanism for allocating a task to one of several agents when each agent's profitability is private information and the principal can pay a cost to learn a single agent's type *before* deciding who to select. They show that an optimal mechanism is characterized by a favored agent and a threshold value, such that when all other agents report values below the threshold, the favored agent gets the task, while in all other cases, it is given to the agent with the highest reported value if and only if his report is verified. Our MLR strategy is similar in spirit to this mechanism in the sense that the discerning agent gets the task whenever he proposed himself and in all other cases, the last resort is chosen.

Our model is also related to a small literature that analyzes infinitely repeated elections in which a median voter needs to elect one of multiple candidates with private types. Two notable examples are Banks and Sundaram (1993, 1998), where private types are persistent and chosen candidates take a hidden action. They restrict attention to a particular class of equilibria where the median voter retains the current incumbent as long as his payoff is above some threshold and lower cost incumbents take higher actions. According to a recent survey by Duggan and Martinelli (2017), this literature has remained small due to the "difficult theoretical issues related to updating of voter beliefs" and has examined various restrictions to simplify this difficulty. We circumvent this difficulty since the MLR strategy profile achieves the first-best in a belief-free equilibrium for the widest range of parameters.

Lipnowski and Ramos (2020) and Li et al. (2017) study an infinitely repeated game in which a principal decides whether to entrust a task to a *single* privately informed agent. The presence of multiple competing agents is crucial to our analysis: If there were only one agent, the principal could achieve no better than having him propose regardless of qualification.

Our paper relates to a small literature on relational contracts with multiple agents. Board (2011) and Andrews and Barron (2016) study how a principal (firm) chooses each period among multiple contractors or suppliers whose characteristics are perfectly observed by the principal, but whose post-selection action is subject to moral hazard. In contrast to these papers, our framework has no transfers, our MLR strategy profile attains the first-best whenever the first-best is attainable, and it does not rely on threats by agents to punish the principal.

Board (2011) shows that under commitment, an optimal contract favors agents with whom the principal has traded in the past, and when the principal is sufficiently patient, such a contract is self-enforcing. Andrews and Barron (2016) show that under certain conditions, the first-best can be attained by a *favored producer allocation* (FPA) rule that picks in each period the agent who generated high output most recently among the set of agents with the highest type in that period (if none of the agents in this set has produced high output, an agent from the set is randomly chosen). Although our environment is very different, the MLR shares the feature of favoring an agent: if the agent chosen most recently has generated high output, he is chosen again. However, the MLR rule differs from FPA in that the latter tilts future allocations to reward success, while the former tilts future allocations to punish failure. To understand why the two mechanisms differ, suppose there are two agents, with 1 being favored. In our case, we do not need to provide agent 2 any incentives, so that agent 1 remains favored even if agent 2 succeeds. In the model of Andrews and Barron (2016), agent 2 must be incentivized not to hold up the principal, which requires making him the favored agent if he succeeds.

By focusing on the first-best, our analysis relates to Athey and Bagwell (2001), where two colluding, ex ante symmetric firms play a repeated Bertrand game and are privately informed about their respective costs. In a binary-types model, they show that the firms can use future "market-share favors" so as to achieve first-best payoffs. Besides differences in the game structure, a key feature distinguishing our analysis is our derivation of a condition (on *all* parameters) that is not only sufficient for first-best, but also necessary. This condition allows us to identify an intuitive strategy profile that attains first-best whenever it is attainable. In addition, our characterization of first-best allows for heterogeneity across agents.

6. Concluding remarks

This paper studies a simple, repeated interaction between a principal and a group of agents that naturally arises in many contexts: deciding which worker is best for a new project, which team member's idea has the most potential, which candidate to hire. In many of these examples, the candidates or applicants simply want to be selected, while the decision-maker (the principal) wants to select an individual who satisfies some requirements (e.g., if he is qualified for the task). Oftentimes, the principal in these scenarios cannot make contingent transfers and has no credible means of committing to a decision rule.

Intuition suggests that the principal should contemplate selecting someone else after an agent generates a disappointing outcome if she hopes to incentivize at least some

of the agents to be discerning. It is not obvious. however, whether the principal should act after a single failure, whether her decision rule should depend on the number of past successes or failures, or whether the best outcome is attained by a rule that is sensitive to the parameters of the environment. It is, therefore, interesting to learn that whenever the principal's first-best outcome is achievable in equilibrium, it is achievable by a simple Markov strategy, which is independent of the environment's parameters.

There are numerous interesting ways to extend our model. Some are easy to accommodate, while others are more challenging. One natural extension is to the case that an unqualified agent generates losses in expectations (by letting low profits be negative). In this case, the principal attains his first-best payoff if in every period, he chooses a qualified agent whenever one is available and chooses no one otherwise. However, it can be shown that there is no PPE that achieves this (for details, see de Clippel et al. 2019).

A second natural extension is when the principal's profit follows a more general distribution on some some interval $[0, y]$, conditional on an agent's qualification, such that the expected profit from an unqualified agent is positive and strictly lower than the expected profit from a qualified agent. The MLR strategy profile can be adapted to this setting by *endogenizing* α and β: A discerning agent becomes the new agent of last resort when generating a profit in some *punishment set* $X \subset [0, y]$. We can then select X so that first-best is achievable for the largest range of discount factors. For instance, under the monotone likelihood ratio property, the punishment set comprises all profit levels below some threshold y^*. However, it remains an open question whether MLR achieves the first-best for the widest range of parameters.

A more challenging extension is to allow for multiple, or possibly a continuum of, qualifications levels such that a higher-qualified agent is more likely to generate high profits. Addressing this extension would require us to consider a framework with cheap-talk announcements where agents report their qualification levels. The first-best may not be attainable in equilibrium, in which case it is unclear what is the best payoff the principal can achieve. Addressing these questions would require different techniques than those employed in this paper.

Appendix

A.1 *Characterization of first-best with two agents*

Proof of Proposition 1. Suppose a first-best PPE exists, and denote the set of first-best equilibrium payoffs by $\mathcal{E}^{\text{FB}} \subset \mathbb{R}^3$. The sum of the two agents' (average) continuation payoffs must equal 1 at any history. Furthermore, in each stage game, it must be that one of the agents, say agent i, is discerning (D) and proposes if and only if he is qualified; the other, last-resort, agent (LR), $-i$, proposes regardless of his qualification, and the principal selects i if he proposes and $-i$ otherwise. Following APS, each pair of first-best equilibrium payoffs for the players can be supported by such a stage-game action profile and a rule specifying *promised (average) continuation payoff vectors*, one for each outcome of the stage game, each of which belongs to $\mathcal{E}^{\text{FB}}$. For convenience, we assume after each period, firms can observe the realization of a public randomization device,

based on which they select continuation equilibria. This guarantees convexity of the equilibrium payoff set, but is not needed for our results.

Denote by $[\underline{\sigma}, \overline{\sigma}]$ the set of average payoffs attainable in a first-best equilibrium for each of the agents.[11] Note that $\underline{\sigma} = 1 - \overline{\sigma}$. Let $p = \alpha\theta + \beta(1-\theta)$ be an agent's ex ante probability of carrying out a project successfully, and let $\sigma_i(jS)$ (respectively, $\sigma_i(jF)$) denote i's continuation payoff when j is picked and succeeds (respectively, fails). We proceed in several steps to derive necessary conditions on the parameters for existence of a first-best equilibrium.

Step 1: Solving for $\underline{\sigma}$. Given the observations above, $\underline{\sigma}$ must be the minimal payoff of agent 1 that can be supported when promised continuation payoffs are restricted to $\mathcal{E}^{\mathrm{FB}}$. Suppose $\underline{\sigma}$ is obtained when agent 1 is LR (we confirm this later). We assume $\underline{\sigma}$ actually solves the following weaker minimization problem, where some incentive constraints of the agents are ignored. Specifically, we assume $\underline{\sigma}$ minimizes

$$(1-\theta)\big(1-\delta+p\delta\sigma_1(1S)+(1-p)\delta\sigma_1(1F)\big)+\theta\delta\big(\alpha\sigma_1(2S)+(1-\alpha)\sigma_1(2F)\big) \tag{12}$$

subject to the IC constraint that agent 2 does not propose when unqualified,

$$\delta\big(p_1\sigma_2(1S)+(1-p_1)\sigma_2(1F)\big)\geq 1-\delta+\beta\delta\sigma_2(2S)+(1-\beta)\delta\sigma_2(2F),$$

as well as the feasibility constraints, i.e., the constraints on the continuation values, $\sigma_i \in [\underline{\sigma}, \overline{\sigma}]$, $i = 1, 2$. Adding the remaining IC constraints could make the minimum greater for more stringent necessary conditions. However, this is redundant since the necessary condition found are sufficient.[12] Using the fact that agents' continuations sum to 1 for any realization, we can rewrite agent 2's IC constraint as

$$\delta\big(\beta\sigma_1(2S)+(1-\beta)\sigma_1(2F)\big)\geq 1-\delta+\delta\big(p_1\sigma_1(1S)+(1-p_1)\sigma_1(1F)\big).$$

Clearly, (12) is minimized only if $\sigma_1(1S) = \sigma_1(1F) = \underline{\sigma}$ (lowering these continuations reduces the objective and can only relax the constraint). Therefore, $\underline{\sigma}$ minimizes

$$(1-\theta)(1-\delta+\delta\underline{\sigma})+\theta\delta\big(\alpha\sigma_1(2S)+(1-\alpha)\sigma_1(2F)\big) \tag{13}$$

subject to the binding IC constraint

$$\delta\big(\beta\sigma_1(2S)+(1-\beta)\sigma_1(2F)\big)=1-\delta+\delta\underline{\sigma}$$

and the feasibility constraints. Using the IC constraint, we see the coefficient on $\sigma_1(2S)$ is $(\alpha-\beta)/(1-\beta) > 0$ and, hence, (13) is increasing in $\sigma_1(2S)$. Since a decrease in $\sigma_1(2S)$ yields an increase in $\sigma_1(2F)$, there are two possible cases to consider.

Case 1: $\sigma_1(2S) = \underline{\sigma}$ does not violate the feasibility constraints. Then $\sigma_1(2F) = \underline{\sigma} + \frac{1-\delta}{\delta(1-\beta)}$ and feasibility requires $\sigma_1(2F) \leq \overline{\sigma}$. Setting $\underline{\sigma}$ equal to the objective in the minimization problem, we obtain $\underline{\sigma} = 1-\theta+\theta\frac{1-\alpha}{1-\beta}$. The feasibility constraint $\underline{\sigma} + \frac{1-\delta}{\delta(1-\beta)} \leq$

[11] Compactness of the PPE payoff set follows from standard arguments.

[12] Alternatively, once obtained, it can be verified that the solution to the relaxed minimization problem also solves the original one.

$\overline{\sigma} = 1 - \underline{\sigma}$ is, therefore, satisfied if and only if

$$\left(1 - \theta + \theta\frac{1-\alpha}{1-\beta}\right) + \frac{1-\delta}{\delta(1-\beta)} \leq 1 - \left(1 - \theta + \theta\frac{1-\alpha}{1-\beta}\right),$$

which can be rearranged to obtain

$$\delta \geq \frac{1}{\beta + 2\theta(\alpha - \beta)}. \tag{14}$$

This condition is therefore necessary for Case 1.

Case 2: $\sigma_1(2F) = \overline{\sigma}$. If $\sigma_1(2S)$ cannot be brought down further, then $\sigma_1(2F)$ must be at its maximal feasible continuation, $\overline{\sigma}$. From the IC,

$$\sigma_1(2S) = \frac{1-\delta}{\delta\beta} + \frac{\underline{\sigma}}{\beta} - \frac{(1-\beta)\overline{\sigma}}{\beta}.$$

Setting $\underline{\sigma}$ equal to the objective in the minimization problem and solving for it gives

$$\underline{\sigma} = 1 + \frac{\theta\left(\dfrac{\alpha-\beta}{\beta}\right)}{1 - \delta - 2\delta\theta\left(\dfrac{\alpha-\beta}{\beta}\right)}. \tag{15}$$

For Case 2 to be possible, it must be that $\sigma_1(2S) \in [\underline{\sigma}, \overline{\sigma}]$. Rearranging, and using $\underline{\sigma} = 1 - \overline{\sigma}$ and (15), this is equivalent to

$$\frac{1}{\delta} \leq \frac{2\theta\left(\dfrac{\alpha-\beta}{\beta}\right) + 1}{2\delta\theta\left(\dfrac{\alpha-\beta}{\beta}\right) - 1 + \delta} \leq \frac{1}{\delta(1-\beta)}. \tag{16}$$

It is easy to verify that the right inequality of (16) is equivalent to (14).

Therefore, (14) is necessary for both Cases 1 and 2. To show it is necessary for the existence of a first-best equilibrium, it remains to verify our conjecture that agent 1's minimal first-best equilibrium payoff is indeed obtained when he is LR.

Step 2: $\underline{\sigma}$ *is attained when agent* 1 *is LR.* If $\underline{\sigma}$ were attained when agent 1 is discerning, his payoff would be

$$\theta\big(1 - \delta + \alpha\delta\sigma_1(1S) + (1-\alpha)\delta\sigma_1(1F)\big) + (1-\theta)\delta\big(p_2\sigma_1(2S) + (1-p_2)\sigma_1(2F)\big).$$

The IC constraint for agent 1 not proposing when he is unqualified is

$$\delta\big(p_2\sigma_1(2S) + (1-p_2)\sigma_1(2F)\big) \geq 1 - \delta + \delta\big(\beta\sigma_1(1S) + (1-\beta)\sigma_1(1F)\big).$$

Therefore,

$$\begin{aligned}
\underline{\sigma} &\geq \theta\big(1 - \delta + \delta\big(\alpha\sigma_1(1S) + (1-\alpha)\sigma_1(1F)\big)\big) + (1-\theta)\delta\big(p_2\sigma_1(2S) + (1-p_2)\sigma_1(2F)\big) \\
&\geq 1 - \delta + \theta\big(\alpha\delta\sigma_1(1S) + (1-\alpha)\delta\sigma_1(1F)\big) + (1-\theta)\big(\delta\big(\beta\sigma_1(1S) + (1-\beta)\sigma_1(1F)\big)\big) \\
&\geq 1 - \delta + \delta\underline{\sigma},
\end{aligned}$$

which implies $\underline{\sigma} \geq 1$, a contradiction.

From Steps 1 and 2 we conclude that we have in (14) a necessary condition for the existence of a first-best PPE. In fact, since (14) is also sufficient for Case 1 to hold, this immediately implies (14) is also sufficient for the existence of a first-best PPE.[13]

We next show directly that the MLR forms a (first-best) PPE whenever (14) holds.

Step 3*: Sufficient conditions for MLR.* Recall from Section 3 that V_1^{D} and V_1^{LR} represent agent 1's average discounted payoff (prior to learning his qualification status) under the MLR strategy profile when he is discerning and when he is last resort, respectively. In Section 3, we showed that the IC constraint for an unqualified discerning agent not to propose is given by

$$V_1^{\mathrm{D}} - V_1^{\mathrm{LR}} \geq \frac{1-\delta}{\delta(1-\beta)}, \tag{17}$$

where V_1^{D} and V_1^{LR} are defined as

$$\begin{aligned} V_1^{\mathrm{D}} &= \theta\big(1-\delta+\alpha\delta V_1^{\mathrm{D}} + (1-\alpha)\delta V_1^{\mathrm{LR}}\big) + (1-\theta)\delta V_1^{\mathrm{D}}, \\ V_1^{\mathrm{LR}} &= (1-\theta)\big(1-\delta+\delta V_1^{\mathrm{LR}}\big) + \theta\big(\alpha\delta V_1^{\mathrm{LR}} + (1-\alpha)\delta V_1^{\mathrm{D}}\big). \end{aligned}$$

Rearranging, we have

$$\begin{aligned} V_1^{\mathrm{D}} &= \frac{\theta(1-\delta) + \theta(1-\alpha)\delta V_1^{\mathrm{LR}}}{1-\delta+\delta\theta(1-\alpha)}, \\ V_1^{\mathrm{LR}} &= \frac{(1-\theta)(1-\delta) + \theta(1-\alpha)\delta V_1^{\mathrm{D}}}{1-\delta+\delta\theta(1-\alpha)}. \end{aligned} \tag{18}$$

Solving explicitly for V_1^{LR}, we find it equals

$$\frac{(1-\theta)(1-\delta) + \theta(1-\theta)(1-\alpha)\delta + \theta^2\delta(1-\alpha)}{1-\delta+2\theta\delta(1-\alpha)},$$

and from (18) it follows that

$$V_1^{\mathrm{D}} - V_1^{\mathrm{LR}} = \frac{\theta(1-\delta) - (1-\delta)V_1^{\mathrm{LR}}}{1-\delta+\theta(1-\alpha)\delta}.$$

Plugging in the expression for V_1^{LR} yields

$$V_1^{\mathrm{D}} - V_1^{\mathrm{LR}} = \frac{(1-\delta)(2\theta-1)}{1-\delta+2\theta\delta(1-\alpha)},$$

which combined with the IC constraint (17) yields the condition (14). □

[13] More precisely, following APS, (14) guarantees that a nonempty, bounded, self-generating set of first-best payoffs (payoff vectors in which the principal obtains her first-best) exists.

A.2 *Hierarchies in the many-agents case*

PROOF OF PROPOSITION 4. Let V^k denote the normalized discounted expected utility of an agent in position k of the ranking. Consider the incentive constraint of not proposing for an unqualified agent whose rank is between 1 and $n-1$,

$$X + p\delta V^k \geq X + p\big[1 - \delta + \beta\delta V^k + (1-\beta)\delta V^{j(k)}\big],$$

where $j(k)$ is the rank ($\geq k$) where the agent of rank k is sent after low profit, p is the probability all agents ranked above are unqualified, and X is the expected continuation value for an agent at rank k when the principal selects a higher-priority (lower ranked) agent.[14] The inequality is written more concisely as

$$V^k - V^{j(k)} \geq \frac{1-\delta}{\delta(1-\beta)}.$$

In particular, we see that $j(k)$ must be strictly larger than k as the right-hand side is strictly positive. In particular,

$$V^k \geq V^n + \phi(k)\frac{1-\delta}{\delta(1-\beta)}$$

for all k, where $\phi(k)$ is the number of times $j(\cdot)$ must be iterated to reach n. We have

$$1 \geq \sum_{k=1}^{n} V^k \geq nV^n + \sum_{k=1}^{n-1} \phi(k)\frac{1-\delta}{\delta(1-\beta)}. \tag{19}$$

We can also determine a lower bound for V^n. Notice that

$$V^n = (1-\theta)^{n-1}(1-\delta) + \delta V^n + \sum_{k=1}^{n-1} p(k)(1-\alpha)\delta\big(V^{j'(k)} - V^n\big),$$

where $j'(k)$ is the rank where n is sent if the agent at rank k gets low profit and $p(k) = (1-\theta)^{k-1}\theta$ is the probability the agent of rank k is chosen. Thus,

$$V^n \geq (1-\theta)^{n-1} + \frac{P(1-\alpha)}{(1-\beta)},$$

where P is the probability an agent of rank k with $j'(k) \neq n$ is picked (the sum of those $p(k)$s).

Given (19), for the hierarchical strategy profile to be an equilibrium requires

$$1 \geq n\left((1-\theta)^{n-1} + \frac{P(1-\alpha)}{(1-\beta)}\right) + \sum_{k=1}^{n-1} \phi(k)\frac{(1-\delta)}{\delta(1-\beta)}. \tag{20}$$

[14] It is notationally heavy to develop X in terms of the Vs, as k may reshuffle position even if others follow equilibrium strategies since $\alpha < 1$, but it does not matter since the term appears on both sides.

Alternatively, MLR forms an equilibrium *if and only if*[15]

$$1 \geq n\left((1-\theta)^{n-1} + \frac{\left(1-(1-\theta)^{n-1}\right)(1-\alpha)}{(1-\beta)}\right) + (n-1)\frac{(1-\delta)}{\delta(1-\beta)}. \tag{21}$$

Consider the necessary condition (20) for the case of hierarchical strategy profiles that send failing agents to the bottom. Here, $P = 1-(1-\theta)^{n-1}$ and $\phi(k) = 1$ for all k, which proves the second half of the result in (b).

Consider next any hierarchical strategy profile. Observe that $P \geq \theta(1-\theta)^{n-2}$ since $j(k) = n$ for at least one agent of rank $k \leq n-1$, with $k = n-1$ in the worst-case scenario. If the strategy profile does not send all failing agents to the bottom (the case we have already treated), then $\sum_{k=1}^{n-1}\phi(k) \geq n$. Thus, in this case, (20) implies the following necessary condition for the hierarchy to form an equilibrium:

$$1 \geq n\left((1-\theta)^{n-1} + \frac{\theta(1-\theta)^{n-2}(1-\alpha)}{(1-\beta)} + \frac{(1-\delta)}{\delta(1-\beta)}\right).$$

The second term is smaller than the corresponding term for MLR because $\theta(1-\theta)^{n-2} < 1-(1-\theta)^{n-1}$ over the relevant range of θs, but the last term is larger as there is at least an extra $\frac{1-\delta}{\delta(1-\beta)}$. It is easy to find (e.g., taking α near 1) parameter combinations for which the MLR inequality is verified, but the above inequality is violated. This proves (a).

Finally, we prove the first part of (b) by example. We let $n = 3$ and consider the hierarchical strategy profile where the failing agent trades his spot with the one right after him in the ranking. The recursive equations that give the agents' payoffs are

$$\begin{aligned}
V^1 &= \theta(1-\delta) + p_1\delta V^2 + (1-p_1)\delta V^1,\\
V^2 &= (1-\theta)\theta(1-\delta) + p_1\delta V^1 + p_2\delta V^3 + (1-p_1-p_2)\delta V^2,\\
V^3 &= (1-\theta)^2(1-\delta) + p_2\delta V^2 + (1-p_2)\delta V^3,
\end{aligned}$$

where $p_1 = \theta(1-\alpha)$ is the ex ante probability the top player drops to second and $p_2 = (1-\theta)p_1$ is the ex ante probability the player in the second spot drops to third.

Now consider the case of $\beta = 0$, $\alpha = 4/5$, $\delta = 5/6$, and $\theta = 1$. The right-hand side of inequality (21) is $3/5 + 2/5 = 1$. Thus, MLR is a PPE for these parameters, but it ceases to be one for any lower θ. Let us now look back at the recursive equations for the hierarchical equilibrium. They become $V^1/3 - V^2/6 = 1/6$, $V^2/3 - V^1/6 = 0$, and $V^3 = 0$ or $V^1 = 2/3$, $V^2 = 1/3$, and $V^3 = 0$. The IC constraints (as derived earlier in the proof, using $j(k) = k+1$) are $V^1 - V^2 \geq \frac{1-\delta}{\delta(1-\beta)}$

[15]One can check directly that this is the same condition on δ as in Proposition 3 but with π replaced with $\frac{1-n(1-\theta)^{n-1}}{n-1}$. However, there is also an intuition why this must be true: For MLR, P is just the probability that a discerning agent is picked, or $1-(1-\theta)^{n-1}$, and each of the IC constraints (only one common IC constraint really, because of symmetry of the MLR) must be binding to get the widest range of parameters, or $V^{\mathrm{D}} - V^{\mathrm{LR}} = \frac{1-\delta}{\delta(1-\beta)}$, in which case we can derive the exact values for V^{LR} and V^{D}, and the equation $V^{\mathrm{LR}} + (n-1)V^{\mathrm{D}} = 1$ gives the largest range of parameters.

and $V^2 - V^3 \geq \frac{1-\delta}{\delta(1-\beta)}$, both of which hold strictly since $\frac{1-\delta}{\delta(1-\beta)} = 1/5$. The determinant of the matrix defining continuation values is strictly positive at these parameters, so diminishing θ a bit will only change those values a bit and the ICs will still hold. □

References

Abreu, Dilip, David Pearce, and Ennio Stacchetti (1990), "Toward a theory of discounted repeated games with imperfect monitoring." *Econometrica*, 58, 1041–1063. [392]

Andrews, Isaiah and Daniel Barron (2016), "The allocation of future business: Dynamic relational contracts with multiple agents." *American Economic Review*, 106, 2742–2759. [394]

Athey, Susan and Kyle Bagwell (2001), "Optimal collusion with private information." *Rand Journal of Economics*, 32, 428–465. [394]

Athey, Susan and David A. Miller (2007), "Efficiency in repeated trade with hidden valuations." *Theoretical Economics*, 2, 299–354. [388]

Banks Jeffrey, S. and Rangarajan K. Sundaram (1998), "Optimal retention in agency problems." *Journal of Economic Theory*, 82, 293–323. [393]

Banks, Jeffrey S. and Rangarajan K. Sundaram (1993), "Adverse selection and moral hazard in a repeated elections model." In *International Symposia in Economic Theory and Econometrics* (William A. Barnett, Norman Schofield, and Melvin Hinich, eds.), 295–311, Cambridge University Press, New York, New York. [393]

Ben-Porath, Elchanan, Eddie Dekel, and Barton L. Lipman (2014), "Optimal allocation with costly verification." *American Economic Review*, 104, 3779–3813. [393]

Bergemann, Dirk and Stephen Morris (2005), "Robust mechanism design." *Econometrica*, 73, 1771–1813. [388]

Bergemann, Dirk and Juuso Välimäki (2010), "The dynamic pivot mechanism." *Econometrica*, 78, 771–789. [388]

Board, Simon (2011), "Relational contracts and the value of loyalty." *American Economic Review*, 101, 3349–3367. [394]

de Clippel, Geoffroy, Kfir Eliaz, Daniel Fershtman, and Kareen Rozen (2019), *On Selecting the Right Agent*. Unpublished paper, CEPR Discussion Paper No. DP13891. [387, 389, 395]

Duggan, John and César Martinelli (2017), "The political economy of dynamic elections: Accountability, commitment and responsiveness." *Journal of Economic Literature*, 55, 916–984. [393]

Li, Jin, Niko Matouschek, and Michael Powell (2017), "Power dynamics in organizations." *American Economic Journal: Microeconomics*, 9, 217–241. [393]

Lipnowski, Elliot and João Ramos (2020), "Repeated delegation." *Journal of Economic Theory*, 188, 105040. [393]

Miller, David A. (2012), "Robust collusion with private information." *Review of Economic Studies*, 79, 778–811. [388]

Co-editor Simon Board handled this manuscript.

Manuscript received 18 November, 2019; final version accepted 12 July, 2020; available online 20 July, 2020.