# On Selecting the Right Agent*

Geoffroy de Clippel[†] Kfir Eliaz[‡] Daniel Fershtman[§] Kareen Rozen[†]

October 2019

**Abstract**

Each period, a principal must assign one of two agents to a new task. Profit is stochastically higher when the agent is qualified for the task, but the principal cannot observe qualification. Her only decision is which of the two agents to assign, if any, given the public history of selections and profits, but she cannot commit to any rule. While she maximizes expected discounted profits, each agent maximizes his expected discounted selection probabilities. We fully characterize when the principal's first-best payoff is attainable in equilibrium, and identify a simple, belief-free, strategy profile achieving this first-best whenever feasible. Addionally, 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, mechanism design without transfers, mechanism design without commitment

*Special thanks 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 no. 003624. Daniel Fershtman gratefully acknowledges financial support from the German Research Foundation through CRC TR 224.

[†]Department of Economics, Brown University.

[‡]Corresponding author: Eitan Berglas School of Economics, Tel-Aviv University and the David Eccles School of Business, University of Utah. Kfire@tauex.tau.ac.il

[§]Eitan Berglas School of Economics, Tel-Aviv University.

# 1 Introduction

Consider the following decision problem faced by a principal who repeatedly interacts with two agents. Each period, the principal faces a new task and needs to select one of the two agents to carry it out. At the start of a period, each agent privately learns if he is qualified for the current task (or, under an alternative interpretation, has enough time to perform it well), and decides whether to apply to do it. The principal can select only among the submitted applications. A completed task generates either a high or low profit for the principal, while a task that is unassigned generates zero profit. The agent assigned to a task, and the profit he generates, is publicly observed by all. A qualified agent is more likely to generate high profit than an unqualified one, but the principal is unable to observe qualification. The principal wants to maximize the expectation of her (average) discounted sum of profits, while each agent wants to maximize the (average) discounted number of times he is selected. The only action the principal can take each period is which applicant, if any, to select; there are no transfers. Moreover, she cannot commit in advance to any plan of action. What is the best outcome the principal can attain in equilibrium, and how?

This abstract problem shares stylized features with many economically relevant situations. Consider a manager who must decide which employee to assign to a new project or client; 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. More generally, our model can be viewed as the repetition of a stage game with the classic, persuasion payoff structure: the principal wants to choose the most qualified agent available, while each agent simply wants to be chosen.

In the benchmark model, each agent $i$ has a commonly known 'ability' parameter $\theta_i$, which is his probability of being qualified for each new task (each period agent $i$ makes a new, independent draw of his qualification). The realized profit from the project can be either high or low. A qualified agent who is selected generates high profit with probability $\gamma \in (0, 1)$. An unqualified agent generates high profit with a strictly smaller probability $\beta \in [0, \gamma)$, but expected profit is nonnegative. Both agents share the same discount factor $\delta$, and both receive a payoff of one whenever they are selected. The principal's *first-best* outcome is to pick the most qualified agent in *every* period. Our first main result in Section 2 concerns the principal's ability to attain her first-best in a perfect public equilibrium (PPE). We characterize the full set of parameter values $(\theta_1, \theta_2, \beta, \gamma, \delta)$ 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 that we characterized.

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 so long as all the principal's past profits were high. Otherwise, the agent of last resort is the most recent agent who generated a low profit.

The MLR profile has a number of 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 (who is the mechanism designer) to ensure she follows the strategy: MLR remains an equilibrium even when the principal's discount factor is zero. Third, it is robust to privately observed heterogeneity in the agents' abilities. To demonstrate this, we enrich our benchmark model by having each agent privately draw his ability from the interval $[\underline{\theta}, 1]$. We characterize the parameters $(\underline{\theta}, \beta, \gamma, \delta)$ for which the principal's first-best is attainable in a *belief-free* equilibrium, in the sense that it constitutes a PPE for *any* pair of realized abilities. Moreover, we show that whenever the parameters belong to this set, the MLR profile attains the first-best in a belief-free equilibrium.

In Section 3, we turn to analyze the challenging case of more than two agents. The MLR strategy profile easily generalizes to this case: the only modification needed is that whenever two or more discerning agents propose themselves, the principal chooses one of them at random. 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 the highest ability among the agents is below $1 - \sqrt[n-1]{\frac{1}{n}}$. We then characterize the sets of parameters for which the MLR profile constitutes a PPE and a belief-free equilibrium. We know that for any profile of abilities strictly superior to $1- \sqrt[n-1]{\frac{1}{n}}$, the MLR will be an equilibrium when agents are patient enough and realized profits are sufficiently informative of qualification. In that sense, we have characterized the widest range of *abilities* for which first-best is achievable, and shown that it is achievable by MLR. We do not know if 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, where we have such a characterization. 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, we do not know if it is feasible to bring more than one agent to the lowest PPE payoff.[1]

This leaves open the question of whether some other strategy profile attains the principal's first-best in PPE for a wider range for parameters than MLR. To at least partially address this question, we compare the performance of 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? Focusing on homogenous abilities, we show that (1) no hierarchical strategy profile 'dominates' MLR in the sense of attaining the first-best in PPE whenever MLR does, and (2) MLR dominates any hierarchical profile that sends a 'failing' agent to the bottom of the ranking.

There are numerous possible extensions of our benchmark model. In Section 4, we briefly discuss how our analysis can accommodate agents who also care about their performance; how to incorporate a continuum of profit levels; the implications

[1]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.

of incurring losses (in expectation) when selecting an unqualified agent; and the effect of imposing a refinement in the spirit of renegotiation-proofness (but tailored towards problems of mechanism design), in which case we fully characterize the set of PPE payoffs.

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, not just for the basic model but for several variants. 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.

## 1.1 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 of aligning incentives. Strategic use of favoritism also arises in static mechanism-design environments without monetary transfers. For example, in Ben-Porath, Dekel and Lipman (2014), a principal allocates a good or task among multiple agents, each of whom is privately informed about the principal's value from allocating it to him. In their static setting, the principal can pay a cost to learn a single agent's type *before* deciding who to select. They show all optimal mechanisms are essentially randomizations over 'favored-agent mechanisms', which consist of a favored agent and a threshold value. If all other agents report values below the threshold, the good or task is allocated to the favored agent. Otherwise, the agent who reports the highest value is checked, and receives the good if and only if his report is confirmed. In our setting, by observing the profit from a selected agent, the principal receives for *free* an *imperfect* signal about the truthfulness of the agent's *past* claim. An agent who is likely to have lied is then punished only in *future* allocations. It would be inefficient to never again pick an agent suspected of lying, and would also violate the principal's equilibrium incentive constraint. In fact, some form of redemption must occur with positive probability in our dynamic setting: the principal must treat a suspected liar less favorably by decreasing his discounted likelihood of being picked in the future, but others will be suspected of lying in the future since profits provide only an imperfect signal. Despite terminology, the last resort agent in our MLR strategy shares some similarity with the favored agent in Ben Porath et al. A novelty in our approach is to select who that agent will be based on past realizations. For a wide range of parameters, the first best allocation becomes achievable even if types are not verifiable. Should types be verifiable at a cost, as in Ben Porath et al., our paper suggests that the principal can oftentimes save on these costs when interacting repeatedly with the agents, by conditioning her future allocation rule.

The problem we study may be thought of as dynamic mechanism design without transfers when the planner is a player (and therefore, cannot commit). In our model, there is no institutional device that enables the principal to credibly commit to a policy, and the agents' payoffs cannot be made contingent on the payoff to the principal. Among other settings, these features arise in political environments where voters (or a median voter) elect one of multiple candidates to an office. A number of papers study infinitely repeated elections in which candidates have privately known types. 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. There are structural differences between our framework and this literature. Banks and Sundaram (1993,1998), for instance, include moral hazard and model private information as being persistent. By contrast, our model has persistence in the agent's underlying ability, and the agents' private qualification varies over time.

The recent literature on dynamic mechanism design with neither transfers nor commitment includes Lipnowski and Ramos (2016) and Li, Matouschek and Powell (2017). Both study an infinitely repeated game in which a principal decides whether to entrust a task to a *single* agent, who is better informed. Both papers predict different and interesting non-stationary dynamics in equilibrium. By contrast, the competition between agents in our model is a driving factor in the results: 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. Both papers allow the use of transfers. Board (2011) considers a hold-up problem, where the chosen contractor each period decides how much to repay the principal for her investment. Assuming the principal can commit to the selection rule, Board shows that it is optimal to be loyal to a subset of 'insider' contractors, because the rents the principal must promise to entice the contractor to repay act as an endogenous switching cost. This bias towards loyalty extends when the principal cannot commit, so long as she is sufficiently patient. Relaxing Board's assumption of commitment and introducing imperfect monitoring in the moral hazard problem, Andrews and Barron (2016) consider a firm who repeatedly faces multiple, ex-ante symmetric suppliers. The firm and suppliers use a common discount factor. A supplier's productivity level is redrawn each period but is observable to the principal. The principal approaches a supplier and, upon agreeing to the relationship, the supplier makes a hidden, binary effort choice yielding a stochastic profit for the principal. Each supplier observes only his own history with the principal. They suggest an allocation rule, the 'favored supplier' rule, and characterize the range of discount factors for which it is part of an equilibrium that attains first-best. They provide additional parameter restrictions which guarantee no equilibrium attains first-best for lower discount factors. The favored supplier rule has the feature that in every period, the principal chooses the supplier with the highest observed productivity level, breaking ties in favor of whoever most recently yielded high profit.

There are several interesting differences between these two papers and ours. First, in our environment the principal cannot use transfers as a means of aligning incentives. Second, we study a problem of adverse selection: the principal cannot observe the distinguishing characteristic – the agents' qualification for the task at hand. In our model, an aim of the principal's selection rule is to influence her set of proposers; thus the set of possible agents in each period is endogenous to the problem. Additional features distinguishing our analysis from Andrews and Barron (2016) is that we provide a full characterization of first-best for the two-agent case and allow for ex-ante asymmetric agents. Without any restriction on the parameters, we show that the MLR attains first-best if and only if first-best is attainable. Interestingly, in contrast to the MLR, Andrews and Barron's favored-supplier rule "favors past success and tolerates past failure." In their environment, monetary incentives allow the principal to punish by way of withholding compensation while rewarding through future promises. In our environment, where monetary incentives are absent, these dynamics are reversed - the principal does not tolerate past failure, and does not favor past success. Furthermore, Andrews and Barron point out that if they were to relax private monitoring, the agents could collectively punish the principal and the optimal allocation rule would become stationary. By contrast, our results rely on the history being at least partially public (the identity of the current agent of last resort must be known to all players), and the MLR does not rely on punishing the principal: whenever it is an equilibrium, it remains so for any discount factor of the principal, even if she is fully myopic.

Our first-best 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" in order 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. Since our focus is on features of the equilibrium strategy profiles rather than properties of the equilibrium payoff set, this condition plays a crucial role and allows us to show that the MLR strategy profile attains first best whenever it is attainable. Whereas the general approach of the collusion literature has been to model problems of private information on costs and imperfect monitoring of prices separately, in our model both agents' actions (whether or not to propose) and their performance (a signal of their qualification) are observable, and neither perfectly reveals a deviation. Lastly, our characterization of first best allows for heterogeneity across agents.

Finally, our work is also related to the literature on "trading favors" originating in Mobius (2001) and Hauser and Hopenhayn (2008), where players have private opportunities to do favors for one another. Among other differences, an important distinguishing feature is that in this literature the players benefit (in the stage game) at the expense of one another.

## 2 A Model

There is one principal and two agents, 1 and 2. Each period $t = 0, 1, 2, ...$ 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 project is either high ($H$) or low ($L$), where $H > L$, and depends stochastically on whether or not the agent assigned to carry it out is qualified to do so. A qualified agent has probability $\gamma \in (0, 1)$ of generating high profit for the principal; while a non-qualified agent generates high profit with a strictly smaller probability $\beta \in [0, \gamma)$. We assume $\beta H + (1 - \beta)L \geq 0$ , so that the principal prefers to hire a non-qualified agent over hiring no one. In each period $t$, the probability that agent $i$ is qualified for the $t$-th project is constant and equal to $\theta_i \in [\underline{\theta}, 1)$, where $\underline{\theta} > 0$. Thus, the parameter $\theta_i$ captures the ability of agent $i$. Each agent privately observes whether he is qualified for the specific project at hand, but the agents' general abilities (the probabilities $\theta_1$ and $\theta_2$) are 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.

Agent $i$ gets a positive payoff in period $t$ if the principal picks him in that period. We normalize this payoff to one (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 $\delta$ is each agent's discount factor, $1\{\cdot\}$ is the indicator function and $x_t \in \{1, 2\} \cup \{\emptyset\}$ 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.

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 $\delta_0$ is the principal's discount factor and $y_t \in \{0, L, H\}$ 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.[2] We define a *public history* at any period $t$ as the sequence $h^t = ((x_0, y_0, S_0), \ldots, (x_{t-1}, y_{t-1}, S_{t-1}))$, where $S_\tau \subseteq \{1, 2\} \cup \{\emptyset\}$ is the set of agents who made a proposal in each period $\tau < t$ and, as defined above, $x_\tau$ and

[2]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.

$y_\tau$ 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.

We view this game as a mechanism design problem without commitment. The principal wants to design a selection rule to maximize her payoff, but cannot commit to a rule. Instead, her rule must be justified endogenously, as an optimal response to that of the agents in equilibrium. The principal cannot influence nature (the probability that each agent is qualified, and the stochasticity of profit), but would ideally like to overcome the incentive problem of agents. The *first-best outcome* from the principal's point of view is to be able to select, in every period, the qualified agent whenever one exists, and any agent otherwise.

**Discussion.** A proposal in our model can be thought of as a packet of documents that lays out a detailed plan. Figuring out which agent is better qualified prior to making the selection is time consuming and costly for the principal. However, we will show that the principal may take advantage of the repeated nature of her interactions with the agents to reach her first-best, even when her time is very limited and she cannot review the agents' proposals before making a selection. There is thus no need to explicitly model a proposal-review stage or review-cost function to make this point. Furthermore, the assumption that agents simply want to be selected regardless of the end profit from the project may capture 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.[3]

There are several key features in our model. First, there is no institutional device that enables the principal to credibly commit to a selection policy. Second, the principal is better off selecting some agent than not selecting any. The idea is that the loss from not performing a task overweighs the loss from not doing it perfectly.[4]

[3]Our analysis would not change if each agent $i$ also received some fixed bonus $\lambda$ when profits are high (see Section 4), but would be considerably more tedious. The proof of Proposition 1 actually allows any $\lambda \geq 0$.

[4]In Section 4, we discuss the case where selecting an unqualified agent leads to expected losses.

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 explicitly lay out his plans. Finally, 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 contraints that preclude such contracts (as in most public organizations where subordinates, who receive a constant wage, may propose themselves to an executive decision maker).

## 2.1 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.[5] 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 (MLR)* 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 remains in that role so 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 condition is this profile a PPE?

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

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

$$\delta \geq \frac{1}{\beta + (\theta_1 + \theta_2)(\gamma - \beta)}. \tag{1}$$

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

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:

$$\theta_1 + \theta_2 > \frac{1}{\iota} \tag{2}$$

where $\iota = \frac{\gamma - \beta}{1 - \beta} = 1 - \frac{1-\gamma}{1-\beta} < 1$ measures how informative low profits are of qualification.

The key incentive constraint, which generates condition (1), is the one facing an unqualified discerning agent. This constraint explains why the incentives for player $i$ are determined both by $\theta_i$ *and* $\theta_{-i}$: When agent $i$ is discerning, the probability with which he is selected is $\theta_i$, but he must also consider the case in which he becomes last resort, and the likelihood of exiting that role is determined by $\theta_{-i}$.

Since abilities matter through their sum, the marginal contribution of having a second agent - even of low ability - can be significant, so long as (1) holds. Suppose agent 1 has rather high ability, say $\theta_1 = 3/4$. If he is the only agent, then the principal has no means to incentivize him: he will submit a proposal independently of his qualification, and in the first-best the principal achieves an expected payoff of $3/4H + 1/4L$. However, if there were an additional agent - even one with a much lower ability, say slightly higher than 1/4 - the expected payoff to the principal in the new first-best equilibrium increases to $(15/16)H + (1/16)L$, which is significantly higher than the principal's first-best equilibrium with only a single agent.

To understand why the MLR attains the first-best for the widest range of parameters, note that in any period, exactly one agent is discerning and one is last-resort. As shown in the proof, the last-resort agent is worse off than the discerning agent. 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. Note that the MLR treats agents with different abilities symmetrically. The reason is that although agents are asymmetric in expectation, they are identical conditional on being qualifed (or unqualified).

The proof, in the Appendix, proceeds as follows. We first observe that if the principal attains the first-best, then the following must be true. At each history $h$, there is an agent $i(h)$ who proposes regardless of his qualification, while the other agent (the 'discerning' agent) proposes if and only if he is qualified. The principal picks the discerning agent whenever he makes a proposal, and picks $i(h)$ otherwise. We thus refer to agent $i(h)$ as the agent of 'last resort.' This structure allows us to write the equilibrium payoff of an agent, both when he is last-resort and when he is discerning, as a function of the continuation payoffs following the selection of some agent and the profit level he generates (in the first-best path, an agent is picked in each period).

We then proceed in several steps, to identify a condition on the parameters necessary for the existence of a first-best equilibrium.[6] Assuming first-best equilibria exist, denote by $\underline{\sigma}_i, \overline{\sigma}_i$ the minimal and maximal payoffs each agent $i$ can obtain in a first-best equilibrium given the parameters.[7] We first find continuation payoffs that minimize agent $i$'s first-best equilibrium payoff subject to the incentive constraints that a discerning agent does not propose himself when he is unqualified (ignoring the constraint that an agent should propose when he is qualified, which can only raise the equilibrium payoff), the feasibility constraints on the continuation payoffs of both agents implied by $(\underline{\sigma}_1, \overline{\sigma}_1, \underline{\sigma}_2, \overline{\sigma}_2)$, as well as the observation that the agents' payoffs must sum up to one in the first-best. We guess that an agent's payoff in the principal's first-best equilibrium is minimized when he is the agent of last resort (we later verify this), and argue that when the last-resort agent's payoff is minimized, the incentive constraint of the discerning agent must be binding. Using this binding constraint, we solve for $\underline{\sigma}_i$ as a function of the parameters and the maximal equilibrium payoff $\overline{\sigma}_i$. This leads to two cases, corresponding to two possible solutions for $\underline{\sigma}_i$, depending on which of the feasibility constraints on the continuation payoffs bind.[8]

[6]Following Abreu, Pearce, and Stacchetti (1990; henceforth, APS), the equilibrium payoff set is the largest self-generating set. However, this property by itself does not guarantee an explicit characterization of the equilibrium payoff set as a function of the parameters. The crux of the proof is to use the property to derive a necessary condition for the existence of first-best equilibria, as a function of the model's parameters.

[7]Following APS, the set of PPE payoffs is compact. Hence, if the set of first-best PPE payoffs is non-empty, such minimum and maximum payoffs exist. Note that since agents are not symmetric, their maximal and minimal payoffs need not coincide.

[8]The minimization of agent $i$ first-best payoff involves increasing his continuation payoff when the other agent generates a high profit while decreasing his continuation payoff when the other agent generates a low profit. These continuation payoffs are both constrained by $\underline{\sigma}_i$ and $\overline{\sigma}_i$, and

The second step is analogous to the first, except that we find the continuation payoffs that *maximize* agent $i$'s equilibrium payoff subject to the incentive constraint that a discerning agent does not propose himself when he is unqualified, the feasibility constraints on the continuation payoffs of the agents, and the observation that the agents' payoff must sum up to one in the first-best. We derive an expression for the maximal equilibrium payoff $\overline{\sigma}_i$ as a function of the parameters and the minimal equilibrium payoff $\underline{\sigma}_i$ (again, two different potential solutions must be considered).

In the third step we consider the four possible solutions for $(\underline{\sigma}_i, \overline{\sigma}_i)$ and show that the inequality (1) is necessary for each of them, and hence necessary for the existence of a first-best equilbrium. Finally, we verify that it is indeed the case that $\overline{\sigma}_i$ ($\underline{\sigma}_i$) is attained when an agent is discerning (last resort).

To show that condition (1) is also *sufficient* for attaining the principal's first-best in a PPE, we argue that the MLR profile is indeed a PPE if and only if (1) holds.

The MLR profile, despite using very little information about the environment and history of past play, attains the first-best PPE payoff over the entire region of parameters for which a first-best PPE exists. It also has several desirable properties. First, 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. Furthermore, despite heterogeneity in agents' abilities, the principal's strategy does not bias selection decisions based on these differences.

Second, the principal's selection rule is optimal for her (thereby providing endogenous commitment) without relying on the agents to punish her if she deviates from it. While efficient equilibria in the literature often rely on any deviator to be punished by others, in our environment we would find it unnatural 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 $\delta_0$.

Third, the MLR addresses questions of equilibrium robustness. From the analysis in Proposition 1, it is clear that the MLR strategy profile is in fact an *ex-post PPE*

different solutions to the minimization problem may arise depending on which of the constraints on the continuation payoffs binds.

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 or not).[9] In light of Proposition 1, such robustness comes for free in our environment. In an ex-post equilibrium, stringent (simultaneous and private) communication protocols are not necessary.[10] Such robustness is particularly relevant for environments where it may be difficult or undesirable to restrict how agents share information with one another.

Finally, as we show below, the MLR strategy profile achieves the principal's first best in a belief-free way when there is uncertainty about the agents' abilities. Suppose the principal has little information about agents' abilities and would like to guarantee her first-best outcome in all realizations. The notion of *belief-free equilibrium* directly addresses the question of equilibrium robustness to such information.[11] A strategy profile is a belief-free equilibrium if it forms a PPE for *any* realized pair of abilities in $[\underline{\theta}, 1]^2$. From condition (1), it follows that the first best 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 is sum of the abilities of the agents. Hence, first-best is attainable in PPE for every possible realization of $(\theta_1, \theta_2)$ if and only if (1) holds for $\theta_1 = \theta_2 = \underline{\theta}$. Combined with Corollary 1, this implies the following result.

**Proposition 2.** *A belief-free equilibrium that attains the principal's first-best exists if and only if*

$$\delta \geq \frac{1}{\beta + 2\underline{\theta}(\gamma - \beta)}. \tag{3}$$

*The MLR strategy achieves the objective when that condition holds.*

It is worth noting the MLR strategy profile's simplicity in comparison to other conceivable strategies when information about abilities is incomplete. As in bandit problems, the principal could try to balance learning agents' abilities and exploiting the agent she currently believes has highest ability. The difficulty is that agents respond strategically to the principal's selection rule, which can impact her ability to learn.

[9]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 Valimaki (2010). The latter use the term "periodic ex-post." Miller (2012) considers ex-post PPE in a model of collusion with adverse selection.

[10]Such ex-post equilibria are also robust to the introduction of payoff-irrelevant signals and high-order beliefs; see Bergemann and Morris (2005).

[11]Equilbria here are "ex-post" with respect to the agents' abilities rather than the realization of the agents' private information.

Continuing the analogy with bandit problems, whether an arm is available to pull becomes endogenous, and may vary at equilibrium with the principal's strategy. The MLR strategy profile simplifies the problem by using this feature to the principal's advantage: there is no need to learn the agents' abilities, and the first-best is achieved, if one agent is provided incentives to submit proposals only when qualified.

# 3 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, with $\vec{\theta}$ denoting the vector of these agents' abilities. 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 3.** *If* $\max_{i \in \mathcal{A}} \theta_i < \theta^*$, *then there is no PPE (and even no Nash equilbirium) that attains the principal's first-best.*

When is the first-best is achievable, and how can the principal achieve it? To start answering these questions, we observe that the underlying principle from our earlier analysis generalizes to $n \geq 2$ agents: at each history $h$, there must be $n-1$ discerning agents each of whom proposes himself if and only if he is qualified, and one agent of last resort who proposes himself irrespective of his qualifications.

We will generalize the MLR strategy by treating all the $n-1$ discerning agents in a symmetric manner, with the principal randomizing uniformly when selecting among discerning agents who have proposed. We will show that in the many agents case, MLR constitutes a belief-free equilibrium for sufficiently patient agents if and only if *all* agents have ability strictly higher than the threshold ability $\theta^*$. Along the way, we find a necessary and sufficient condition for the MLR to form a PPE when $\theta_i > \theta^*$ for all agents $i$. Finally, we will consider the 'optimality' of this generalization of MLR in terms of whether another strategy profile is capable of sustaining the principal's first best in PPE for a wider range of parameters. In particular, we show that there is no domination relationship with some 'hierarchical' strategy profiles, in which the principal does not treat discerning agents symmetrically.

### 3.1 Characterizing when MLR is a PPE

Under the MLR strategy generalized to $n \geq 2$, the behavior prescribed for the principal and agent of last resort are clearly best responses to the discerning agents' strategies. The only question is whether a discerning agent is willing to propose himself when qualified, and refrain from proposing when not. The main difference between having two or many agents play the MLR is that a discerning agent's payoff *depends on the abilities of other discerning agents*, through how often the others propose. A discerning agent's payoff is thus impacted by which of the $n-1$ other agents is removed from the discerning pool to serve as the agent of last resort.

To understand incentives, we must thus understand the probability a given agent is selected under these different possible circumstances. We denote by $\rho_i(\vec{\theta})$ the probability $i$ is picked when he is the agent of last resort. When $\ell$ is the agent of last resort, we denote by $\sigma_i(\vec{\theta}, \ell)$ the probability that a discerning agent $i$ is picked, *conditional on him proposing*. When $\ell$ is the agent of last resort, we let $p_j(\vec{\theta}, i, \ell)$ denote the probability that a discerning agent $j$ is picked, *conditional on another discerning agent $i$ proposing but not being picked*. Finally, when $\ell$ is the agent of last resort, we let $q_j(\vec{\theta}, i, \ell)$ denote the probability a discerning agent $j$ is picked, *conditional on the discerning agent $i$ not proposing*. These probabilities are given by:

$$
\begin{aligned}
\rho_i(\vec{\theta}) &= \prod_{k \neq i}(1-\theta_k), \\
\sigma_i(\vec{\theta}, \ell) &= \frac{\sum_{S \subseteq \mathcal{A} \setminus \{\ell\}: i \in S} \frac{1}{|S|} \prod_{k \in S} \theta_k \prod_{k \notin S, k \neq \ell}(1-\theta_k)}{\theta_i}, \\
p_j(\vec{\theta}, i, \ell) &= \frac{\sum_{S \subseteq \mathcal{A} \setminus \{\ell\}: i,j \in S} \frac{1}{|S|} \prod_{k \in S} \theta_k \prod_{k \notin S, k \neq \ell}(1-\theta_k)}{\theta_i(1-\sigma_i(\vec{\theta}, \ell))}, \\
q_j(\vec{\theta}, i, \ell) &= \frac{\sum_{S \subseteq \mathcal{A} \setminus \{i,\ell\}: j \in S} \frac{1}{|S|} \prod_{k \in S} \theta_k \prod_{k \notin S, k \neq \ell}(1-\theta_k)}{1-\theta_i}.
\end{aligned}
$$

The expression for $\rho_i(\vec{\theta})$ follows because a last resort agent is selected under the MLR strategy profile if and only if all discerning agents are unqualified. To understand the expression for $\sigma_i(\vec{\theta}, \ell)$, observe that while agent $i$'s proposal is selected uniformly among any set of discerning agents' proposals, we must consider all different possible sets of proposers and their probabilities. The probabilities $\rho_i(\vec{\theta})$ and $\sigma_i(\vec{\theta}, \ell)$ are needed to characterize the equilibrium value functions of agents. The final two probabilities $p_j(\vec{\theta}, i, \ell)$ and $q_j(\vec{\theta}, i, \ell)$, whose expressions follow from similar reasoning, will be needed to capture incentive conditions.

With these probabilities in mind, we turn our attention to understanding agents' payoffs and their resulting incentives. We denote by $V_i^D(\vec{\theta}, \ell)$ agent $i$'s average discounted payoff under the MLR strategy profile when he is discerning and agent $\ell$ is the agent of last resort. We denote by $V_i^{LR}(\vec{\theta})$ agent $i$'s average discounted payoff under the MLR strategy profile when he is the agent of last resort himself. These are jointly determined by the following recursive system of equations for all possible agents $\ell \neq i$:

$$V_i^{LR}(\vec{\theta}) = \overbrace{\rho_i(\vec{\theta})}^{\substack{i \text{ is} \\ \text{chosen}}} \Big((1-\delta_i)u_i + \delta V_i^{LR}(\vec{\theta})\Big) + \sum_{j \neq i} \overbrace{\theta_j \sigma_j(\vec{\theta}, i)}^{\substack{j \text{ chosen when} \\ i \text{ is last resort}}} \Big(\gamma \delta V_i^{LR}(\vec{\theta}) + \overbrace{(1-\gamma)\delta V_i^D(\vec{\theta}, j)}^{\substack{\text{low profit,} \\ j \text{ becomes last resort}}}\Big),$$

$$\begin{aligned} V_i^D(\vec{\theta}, \ell) = & \overbrace{\theta_i \sigma_i(\vec{\theta}, \ell)}^{\substack{i \text{ chosen when} \\ \ell \text{ is last resort}}} \Big((1-\delta)u_i + \gamma \delta V_i^D(\vec{\theta}, \ell) + \overbrace{(1-\gamma)\delta V_i^{LR}(\vec{\theta})}^{\substack{\text{low profit,} \\ i \text{ becomes last resort}}}\Big) \\ & + \sum_{j \neq i, \ell} \underbrace{\theta_j \sigma_j(\vec{\theta}, \ell)}_{\substack{j \text{ chosen when} \\ \ell \text{ is last resort}}} \Big(\gamma \delta V_i^D(\vec{\theta}, \ell) + \underbrace{(1-\gamma)\delta V_i^D(\vec{\theta}, j)}_{\substack{\text{low profit,} \\ j \text{ becomes last resort}}}\Big) + \underbrace{\rho_\ell(\vec{\theta})}_{\substack{\ell \text{ is} \\ \text{chosen}}} \delta V_i^D(\vec{\theta}, \ell). \end{aligned} \tag{4}$$

Of course, following the MLR strategy requires certain incentive conditions to be satisfied. The incentive condition for a discerning agent $i$ who turns out to be unqualified *not to propose* in a period when $\ell$ is the agent of last resort, is given by:

$$\begin{aligned} & \overbrace{\frac{\rho_\ell(\vec{\theta})}{1-\theta_i}}^{\substack{\text{last resort} \\ \text{agent chosen}}} \delta V_i^D(\vec{\theta}, \ell) + \sum_{j \neq i, \ell} \overbrace{q_j(\vec{\theta}, i, \ell)}^{\substack{\text{discerning} \\ j \text{ chosen}}} \Big(\gamma \delta V_i^D(\vec{\theta}, \ell) + \overbrace{(1-\gamma)\delta V_i^D(\vec{\theta}, j)}^{\substack{\text{low profit, j} \\ \text{becomes last resort}}}\Big) \\ \geq & \overbrace{\sigma_i(\vec{\theta}, \ell)}^{i \text{ chosen}} \Big((1-\delta)u_i + \beta \delta V_i^D(\vec{\theta}, \ell) + \overbrace{(1-\beta)\delta V_i^{LR}(\vec{\theta})}^{\substack{\text{low profit, i} \\ \text{becomes last resort}}}\Big) \\ & + (1-\sigma_i(\vec{\theta}, \ell)) \sum_{j \neq i, \ell} \underbrace{p_j(\vec{\theta}, i, \ell)}_{\substack{\text{discerning j} \\ \text{chosen instead}}} \Big(\gamma \delta V_i^D(\vec{\theta}, \ell) + \underbrace{(1-\gamma)\delta V_i^D(\vec{\theta}, j)}_{\substack{\text{low profit, j} \\ \text{becomes last resort}}}\Big). \end{aligned} \tag{IC$_U$}$$

Similarly, the incentive condition for a qualified discerning agent $i$ to *propose* in a period when $\ell$ is the agent of last resort, is:

$$\begin{aligned}&\sigma_i(\vec{\theta},\ell)\Big((1-\delta)u_i+\gamma\delta V_i^D(\vec{\theta},\ell)+(1-\gamma)\delta V_i^{LR}(\vec{\theta})\Big)\\&\quad+(1-\sigma_i(\vec{\theta},\ell))\sum_{j\neq i,\ell}p_j(\vec{\theta},i,\ell)\Big(\gamma\delta V_i^D(\vec{\theta},\ell)+(1-\gamma)\delta V_i^D(\vec{\theta},j)\Big)\\&\geq\frac{\rho_\ell(\vec{\theta})}{1-\theta_i}\delta V_i^D(\vec{\theta},\ell)+\sum_{j\neq i,\ell}q_j(\vec{\theta},i,\ell)\Big(\gamma\delta V_i^D(\vec{\theta},\ell)+(1-\gamma)\delta V_i^D(\vec{\theta},j)\Big),\end{aligned}\tag{IC$_Q$}$$

which differs from Condition $\text{IC}_U$ both in the direction of the inequality and because the probability that agent $i$ generates low profit is $\gamma$ instead of $\beta$.

Incentive conditions $\text{IC}_U$ and $\text{IC}_Q$ are linear in the equilibrium payoffs. As will be seen below, $\text{IC}_U$ and $\text{IC}_Q$ depend on these payoffs *only* through the differences

$$\Delta V_i(\vec{\theta},\ell)=V_i^D(\vec{\theta},\ell)-V_i^{LR}(\vec{\theta})$$

in average discounted payoffs from being discerning instead of being the agent of last resort, which vary with the identity of the agent of last resort when abilities are heterogeneous. Furthermore, we will show that these payoff differences themselves depend on the vector of abilities $\vec{\theta}$ only through the *likelihood premiums* of being picked by the principal when discerning versus when the agent of last resort. Formally, agent $i$'s likelihood premium of being picked when discerning *while $\ell$ is the agent of last resort*, versus when $i$ himself is the agent of last resort, is:

$$\pi_{\ell i}(\vec{\theta})=\theta_i\sigma_i(\vec{\theta},\ell)-\rho_i(\vec{\theta})$$

For each $i$ and each $\vec{\theta}$, let $\vec{\sigma}_i(\vec{\theta})$, $\Delta\vec{V}_i(\vec{\theta})$ and $\vec{\pi}_i(\vec{\theta})$ be the $(n-1)$-column vectors whose $\ell$-component is $\sigma_i(\vec{\theta},\ell)$, $\Delta V_i(\vec{\theta},\ell)$ and $\pi_{\ell i}(\vec{\theta})$, respectively, for each $\ell\neq i$. These vectors thus list the selection probabilities, average payoff differences and likelihood premiums, respectively, that are relevant for $i$ as a function of the agent of last resort.

The claims above are established as intermediate steps in characterizing when the MLR strategy profiles constitutes a PPE. Stating the characterization requires defining three matrices: $M_i^Q(\vec{\theta})$, which collects terms from $\text{IC}_Q$; $M_i^U(\vec{\theta})$, which collects terms from $\text{IC}_U$; and $B_i(\vec{\theta})$, which collects terms from the recursive system (4). Given its phrasing in terms of matrix inequalities, the characterization may not seem insightful to the naked eye, but it is very useful in two respects. First, it provides straightforward inequalities to numerically check whether MLR constitutes a PPE.

Second, the characterization is a critical intermediate step to understanding when MLR constitutes a belief-free equilibrium (as studied in the next subsection), for which a far more transparent characterization emerges.

**Proposition 4.** *Suppose $\theta_i > \underline{\theta}^*$ for all $i \in \mathcal{A}$. The MLR strategy profile constitutes a PPE if and only if for all agents i:*

$$M_i^Q(\vec{\theta})B_i(\vec{\theta})^{-1}\vec{\pi}_i(\vec{\theta}) \leq \vec{\sigma}_i(\vec{\theta}) \leq M_i^U(\vec{\theta})B_i(\vec{\theta})^{-1}\vec{\pi}_i(\vec{\theta}), \tag{5}$$

*with the $(n-1)\times(n-1)$ matrices $M_i^Q(\vec{\theta})$, $M_i^U(\vec{\theta})$, $B_i(\vec{\theta})$ are defined in (6-8) below.*

The proof, in the Appendix, has three main steps. First, we manipulate incentive conditions $\text{IC}_U$ and $\text{IC}_Q$ to show that they depend on average discounted continuation payoffs *only through the payoff differences* $\Delta\vec{V}_i(\vec{\theta})$. In particular, we show that the MLR strategy profile constitutes a PPE if and only if

$$\delta(1-\gamma)M_i^Q(\vec{\theta})\Delta\vec{V}_i(\vec{\theta}) \leq (1-\delta)\vec{\sigma}_i(\vec{\theta}) \leq \delta(1-\gamma)M_i^U\Delta\vec{V}_i(\vec{\theta}),$$

where these matrices are defined by

$$[M_i^Q(\vec{\theta})]_{\ell\ell'} = \begin{cases} q_{\ell'}(\vec{\theta},i,\ell) - p_{\ell'}(\vec{\theta},i,\ell)(1-\sigma_i(\vec{\theta},\ell)) & \text{if } \ell \neq \ell', \\ \rho_\ell(\vec{\theta})/(1-\theta_i) & \text{if } \ell = \ell'; \end{cases} \tag{6}$$

and

$$[M_i^U(\vec{\theta})]_{\ell\ell'} = \begin{cases} [M_i^Q(\vec{\theta})]_{\ell\ell'} & \text{if } \ell \neq \ell', \\ \left[M_i^Q(\vec{\theta})\right]_{\ell\ell'} + \frac{\gamma-\beta}{1-\gamma}\sigma_i(\vec{\theta},\ell) & \text{if } \ell = \ell'. \end{cases} \tag{7}$$

This provides only a partial characterization of equilibrium conditions, since the payoff differences are not yet expressed in terms of exogenous parameters of the problem. Second, we manipulate the recursive system (4) defining payoffs themselves, to show that the differences in payoffs depend on the ability vector $\vec{\theta}$ *only through the likelihood premiums* (of which the matrix $B_i(\vec{\theta})$ is a function). Namely, we show that:

$$B_i(\vec{\theta})\Delta\vec{V}_i(\vec{\theta}) = \frac{1-\delta}{\delta(1-\gamma)}\vec{\pi}_i(\vec{\theta}),$$

where

$$[B_i(\vec{\theta})]_{\ell\ell'} = \begin{cases} \pi_{i\ell'}(\vec{\theta}) - \pi_{\ell\ell'}(\vec{\theta}) & \text{if } \ell \neq \ell', \\ 1 + \pi_{i\ell}(\vec{\theta}) + (1-\delta)/(\delta(1-\gamma)) & \text{if } \ell = \ell'. \end{cases} \tag{8}$$

The third and final step is establishing that the matrix $B_i(\vec{\theta})$ is invertible, which turns out to be nontrivial. We prove when $\theta_i > \theta^*$ for all agents $i$ that the matrix $B_i(\vec{\theta})$ has a special property ensuring invertibility: it is strictly diagonally dominant, which means that for every row, the absolute value of the diagonal element is strictly larger than the sum of the absolute values of the off-diagonal elements.

Generalizing one of our points from Section 2, note that the equilibrium conditions are independent of the principal's discount factor $\delta_0$, which means they hold even if the principal were fully myopic. The equilibrium thus doesn't require he principal's behavior to be enforced by the threat of punishments from agents, which we consider a natural property in a mechanism design context where the principal is the authority.

## 3.2 The MLR as a belief-free equilibrium

The principal may have little information about agents' abilities, and yet hope to guarantee her first-best outcome. The MLR strategy profile is a belief-free equilibrium if it forms a PPE for *any* realized vector of abilities $\vec{\theta}$ in the set $[\underline{\theta}, 1]^{\mathcal{A}}$ of possible abilities. The necessary and sufficient condition for this depends on the minimal probability premium $\pi_{\ell i}(\vec{\theta})$ for agent $i$ when considering *all possible* ability levels and last resort agents. As shown in the Appendix, the minimal probability premium is the following function of $\underline{\theta}$, which is the lower envelope of its two components:

$$\underline{\pi} = \begin{cases} \frac{\underline{\theta}}{n-1} & \text{if } \underline{\theta} \geq 1 - \sqrt[n-2]{\frac{1}{n}} \\ \frac{1-n(1-\underline{\theta})^{n-1}}{n-1} & \text{otherwise.} \end{cases} \tag{9}$$

This characterization allows us to derive the agents' minimal discount factor that sustains the MLR as a belief-free equilibrium.

**Proposition 5.** *The MLR forms a belief-free equilibrium if and only if for each agent,*

$$\delta \geq \frac{1}{\gamma + (\gamma - \beta)\underline{\pi}},$$

*where $\underline{\pi}$ is positive if and only if $\underline{\theta} > \theta^*$.*

Note that $\frac{1-n(1-\underline{\theta})^{n-1}}{n-1}$ is the probability premium in the homogenous case where all agents have an ability $\underline{\theta}$. Thus, by Proposition 5, the set of discount factors sustaining the MLR as a belief-free equilibrium for ability profiles in $[\underline{\theta}, 1]^{\mathcal{A}}$ is the *same* set that sustains it as an equilibrium with homogenous abilities known to be $\underline{\theta}$ when there are two agents or $\underline{\theta}$ falls below $1 - \sqrt[n-2]{\frac{1}{n}}$. Otherwise, the range of discount factors supporting the belief-free equilibrium is *smaller* than in the case where the agents are commonly known to be $\underline{\theta}$.

Why is this so? In view of Proposition 5, we need to understand at which profile of abilities the probability premium is minimized. Agent $i$'s probability premium $\pi_{\ell i}(\vec{\theta})$ is increasing in both $\theta_i$ and $\theta_\ell$, so it is minimized by setting both equal to $\underline{\theta}$. On the other hand, the abilities of discerning agents other than $i$ have two opposing effects on $\pi_{\ell i}(\vec{\theta})$. When these discerning agents have higher abilities, they reduce the probability $\sigma_i(\vec{\theta}, \ell)$ that $i$ is selected when he proposes (which lowers the premium), but they also reduce the probability $\rho_i(\vec{\theta})$ that $i$ is picked when he is the agent of last resort (which raises the premium). The effect associated to $\sigma_i(\vec{\theta}, \ell)$ becomes relatively more important as $\underline{\theta}$ grows because $\sigma_i(\vec{\theta}, \ell)$ is premultiplied by $\theta_i = \underline{\theta}$ in the definition of the probability premium, while $\rho_i(\vec{\theta})$ is independent of $\theta_i$. Thus the ability vector minimizing the probability premium has all agents with ability $\underline{\theta}$ when it is relatively low, but involves some high-ability opponents otherwise.

The main challenge in proving Proposition 5 stems from the fact that all possible combinations of abilities must be considered, and that inverting $B_i(\vec{\theta})$ is non-trivial with heterogenous abilities. Fortunately, Lemma 4 shows the equilibrium conditions depend directly on the vector $B_i(\vec{\theta})^{-1}\vec{\pi}_i(\vec{\theta})$. That vector can be shown to satisfy

$$B_i(\vec{\theta})^{-1}\vec{\pi}_i(\vec{\theta}) = [Id - \frac{1-\delta_i\gamma}{\delta_i(1-\gamma)}B_i(\vec{\theta})^{-1}]\vec{1},$$

because the sum over any row $\ell$ of the matrix $B_i(\vec{\theta})$ is equal to $1 + \frac{1-\delta_i}{\delta_i(1-\gamma)} + \pi_{\ell i}(\vec{\theta})$. This reduces the problem at hand to understanding the vector $B_i(\vec{\theta})^{-1}\vec{1}$, that is, the vector of row sums of $B_i(\vec{\theta})^{-1}$. Next, a power series development of $B_i(\vec{\theta})^{-1}$ establishes that $B_i(\vec{\theta})^{-1}\vec{1}$ is decreasing in $\theta_i$, or that $B_i(\vec{\theta})^{-1}\vec{\pi}_i(\vec{\theta})$ is increasing in $\theta_i$. Since $M_i^U$ is a positive matrix, the equilibrium constraint for discerning agents not to make a proposal when he is unqualified is most challenging when $\theta_i = \underline{\theta}$. After observing that the matrix $B_i(\vec{\theta})$ is an M-matrix[12] in that case, we can apply the Ahlberg-Nilson-Varah bound to provide a sharp upper-bound the row sums of $B_i(\vec{\theta})^{-1}$. Some algebra then establishes that a discerning agent does not want to make a proposal when he is unqualified when his discount factor is above the bound stated in Proposition 5.

[12] I.e., a strictly diagonally dominant matrix with positive diagonal entries and negative off-diagonal entries.

Similar techniques establish that discerning agents always want to submit proposals when they are qualified, independently of their discount factors. As for necessity in Proposition 5, we can just look at the equilibrium conditions stated in Lemma 4 for the ability vector that achieves $\underline{\pi}$. Although abilities are heterogenous when $\underline{\theta}$ is higher than $1 - \sqrt[n-2]{\frac{1}{n}}$, the matrix $B_i(\vec{\theta})$ remains easy to invert in that case because agents other than $i$ are all symmetric.

Propositions 3 and 5 together imply the following result.

**Corollary 1.** *Consider the ability threshold $\theta^*$ defined in Proposition 3. We have:*

*(i) If $\underline{\theta} < \theta^*$, then the principal's first best cannot be achieved in any belief-free equilibrium.*

*(ii) If $\underline{\theta} > \theta^*$, then for all $(\beta, \gamma)$ with $\frac{1-\beta}{1-\gamma} \geq \frac{1+\underline{\pi}}{\underline{\pi}}$, the MLR strategy profile attains the principal's first best in a belief-free equilibrium.*

The principal's ability to achieve her first best in a belief-free manner thus hinges on her *worst* possible agent, the organization's 'weakest link.' Only when she is certain that all agents have abilities greater than $\theta^*$ can she incentivize them to be discerning. A principal may or may not be able to screen agents to ensure this minimal standard. The threshold $\theta^*$ decreases in $n$, and is always smaller than 1/2, so it would suffice that agents are simply more likely to be qualified than not.

## 3.3 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 $n = 2$).[13]

[13] We are not aware of applications of APS to derive simple closed-form solutions in problems with more than two players and *no transfers*.

We thus propose to evaluate the performance of MLR against an intuitive class of alternative strategy profiles. To simplify algebra, we focus on the case of homogenous abilities ($\theta_1 = ... = \theta_n = \theta$). 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:

(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.

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

**Proposition 6.** *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 $(\beta, \gamma, \delta, \theta)$ at which $s'$ does. Then:*

*(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.*

One rough intuition for why a hierarchy-based profile may not dominate MLR is because the punishment MLR employs is "uniformly more severe," in the sense that there are only two tiers in the hierarchy, and a failing agent falls from the top of the hierarchy to the bottom. In contrast, in a multi-tier hierarchy, the decrease in the probability of being chosen for an agent in the second-to-last tier is not as severe.

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 6 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.

# 4 Extensions

We now consider some extensions in the original context of a principal and two agents.

## 4.1 When agents also enjoy success

Agents may also care about their reputation or enjoy positive psychological reinforcement from successfully carrying out a project. We extend the analysis to allow agents' payoffs to depend on their performance as well as participation in the project. This generalization allows for another possible interpretation, whereby an indivisible resource (e.g., server processing capacity, a common space, a piece of equipment, etc.) is allocated in each period to one of two agents, each of whom knows their probability distribution (using $\beta$ or $\gamma$) of getting a high or low payoff from using it that period. In this interpretation, the principal is a social planner who can, but need not, derive personal payoff. Following a utilitarian objective, in each period she would allocate the common good to the agent with highest expected payoff.

Formally, an agent receives an additional utility $\lambda \geq 0$ for generating high profit $H$, on top of the utility $u \geq 0$ he enjoys, irrespective of the outcome, from being selected to carry out the project. The case $\lambda = 0$, $u = 1$ corresponds to our original model, whereas the other extreme $u = 0$ corresponds to an environment in which the interests of the principal and the agents are most aligned (though not entirely: an agent only cares about his own performance). The following extension of our earlier results is proved in the Appendix.

**Proposition 7.** *Suppose agents obtain additional utility $\lambda$ for generating high profit in addition to the utility $u$ for being picked.*

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

$$\delta \geq \frac{1}{\beta + (\theta_1 + \theta_2)(\gamma - \beta) + \theta_1\theta_2 \left(\frac{\lambda(\gamma-\beta)(1-\beta)}{1+\lambda\beta}\right)}. \tag{10}$$

*(b) The MLR strategy profile is a PPE if and only if (10) is satisfied; that is, there exists a strategy profile attaining first-best if and only if the MLR attains it.*

Thus, the MLR strategy profile still attains the first-best whenever doing so is possible, independently of the agents' payoff structure $(u, \lambda)$. Notice that agents benefitting more from high outcomes introduces an additional positive component $\theta_1\theta_2 \left(\frac{\lambda(\gamma-\beta)(1-\beta)}{1+\lambda\beta}\right)$ into the denominator of (10), extending the region for which first-best is attainable. Intuitively, the bonus utility helps to align incentives. This has two implications. First, MLR is a belief-free equilibrium for a larger range of abilities $[\underline{\theta}, 1]^2$ than before.[14] Second, the first-best region depends on the *composition* of abilities, not their sum. When $\lambda = 0$, differences in agents' abilities had no implications for the possibility of attaining first-best. Such differences, however, play an important role when agents directly care about their performance. Holding fixed the sum of the agents' abilities, the first-best region is maximized when agents are identical, whereas heterogeneity reduces the positive effect from the alignment of incentives.

## 4.2 Principal-renegotiation-proofness

In standard problems of mechanism design, the notion of partial implementation only requires the mechanism to have some equilibrium that achieves the principal's desired outcome. In this view, the principal has the ability to make her desired equilibrium focal. How does one extend this idea to dynamic mechanism design without commitment? As a starting point, consider the refinement of renegotiation proofness (Farrell and Maskin, 1989). Loosely speaking, an equilibrium strategy profile for a repeated game is weakly renegotiation-proof if at every history, there is no other continuation equilibrium under that strategy profile which all players would unanimously agree to switch to. The ability to always switch to a more-preferred equilibrium has a similar flavor to the principal's ability in mechanism design, but

[14] The analog of Proposition 3 adds the term $\underline{\theta}^2 \left(\frac{\lambda(\gamma-\beta)(1-\beta)}{1+\lambda\beta}\right)$ in the threshold $\delta$'s denominator.

gives equal power to all players. That is, standard renegotiation-proofness treats the principal just like any other agent.

We propose instead a notion of *principal-renegotiation-proofness* that applies to dynamic-mechanism design without commitment. A PPE satisfies this refinement if at every history, there is no other continuation equilibrium under that PPE that achieves a higher expected payoff for the principal. Our notion is precisely that of renegotiation-proofness with power given to a single agent instead of the group as a whole (unanimity). More generally, one can define $S$-renegotiation-proof equilibrium, where $S$ represents a coalition that has the power to select equilibria (conventions, norms) within a game. The principal may not be able to commit in our dynamic mechanism design problem, and thus faces incentive constraints of her own; but she does not lose the power to select the equilibrium being played in her organization at any point. Hence the natural choice in our framework is $S = \{principal\}$.

Under this refinement, we identify a stark characterization of the payoffs that can be sustained, which also address cases in which the principal's first-best *cannot* be achieved in a PPE.

**Proposition 8.** *Consider a principal-renegotiation-proof PPE in which agents use pure strategies.*[15] *The principal's first-best is achieved (e.g, using the MLR strategy) if (1) holds, and otherwise the one-shot Nash equilibrium is played in every stage game.*

## 4.3 Losses

A key feature of our model is that the principal weakly prefers to choose a non-qualified agent to carry out a project, than to choose no agent at all. Suppose instead that a non-qualified agent who is hired for the task generates losses in expectation: $\beta H + (1 - \beta)L < 0$, which requires $L < 0$ and for $\beta, H$ to be relatively small. 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. Can this first-best payoff be attained in a PPE? We address this question in the case of equally able agents and focus on pure-strategy PPEs.

**Proposition 9.** *Assume $\theta_1 = \theta_2 = \theta$ and $\beta H + (1 - \beta)L < 0$. There is no pure-strategy PPE in which every period, every agent proposes if and only if he is qualified.*

[15] The principal is allowed to randomize in the event both agents make proposals.

Thus it is impossible for the principal to never pick the wrong agent, while simultaneously ensuring that the right agent always makes himself available. The proof proceeds along the same lines as our proof of the necessary condition for attaining the principal's first-best payoff in a PPE in our original model.

## 4.4 General profit distributions

Our model assumes there are only two profit levels, $H$ and $L$. Suppose instead that profit follows a more general distribution, conditional on an agent's qualification. Formally, suppose the principal's profit $y$ in any period is drawn from $[\underline{y}, \bar{y}]$ according to the CDF $Q$ ($U$) when the agent is qualified (unqualified). We allow for $\underline{y} < 0$; we only require that the expected profit from an unqualified agent is positive, and strictly lower than the expected profit from a qualified agent. This setting includes environments where qualified agents first-order stochastically dominate unqualified ones, or where qualified agents have higher variance.

For which profit levels should an agent be punished in this case? Notice that the MLR strategy profile can be adapted by *endogenizing* $\beta$ and $\gamma$. A discerning agent still proposes only when he is qualified, and the agent of last resort still proposes regardless of qualification. The principal selects an agent as before. The only difference is that a discerning agent becomes the new agent of last resort when generating a profit in some *punishment set* $Y \subset [\underline{y}, \bar{y}]$ which has positive measure according to $U$. Then $\gamma^* = 1 - \int_{y \in Y} dQ(y)$ is the probability a qualified agent generates a payoff outside the punishment set, and $\beta^* = 1 - \int_{y \in Y} dU(y)$ is the probability that an unqualified agent does so. If this adjusted-MLR strategy profile is an equilibrium, then the principal obtains her first best. In the Appendix, we examine how the punishment set $Y$ should be chosen to sustain the equilibrium, when possible. In particular, we may want to select the punishment set so that first-best is achievable for the largest range of discount factors. We show, for instance, that under the monotone likelihood ratio property, the punishment set comprises all profit levels below some threshold $y^*$.[16]

Analogous reasoning accommodates settings where outcomes are judged through lenses other than profit (an invention, a work of art, a research article) and may depend on the principal's perception. The principal may have gradations in her

[16] We are thus able to extend our necessary and sufficient conditions for MLR to be an equilibrium to the case of a continuum of profits. However, it remains an open question whether in this case, MLR attains the first-best in equilibrium for the widest range of parameters.

assessments of outcomes, but it only matters how she pools those into 'high' and 'low' categories to determine when to punish discerning agents. Her perception of outcomes need only be sufficiently astute to sustain equilibrium. In such settings, the distribution of the principal's possible assessments, conditional on an agent's qualification, must be common knowledge. The principal's assessment itself, however, need not be observed by agents. It suffices to allow her to publicly announce the next agent of last resort, as she has an incentive to speak truthfully.

# 5 Concluding Remarks

The literature on dynamic mechanism design has accumulated a rich set of results on what is the best outcome a principal can achieve in a variety of contexts, and what incentive schemes she should use for that purpose. That literature, however, requires the principal to credibly commit to her incentive schemes, and typically uses monetary transfers as a means for providing incentives. The repeated games literature, on the other hand, treats the principal as just another player (meaning it assumes away commitment), and has developed tools for characterizing the set of payoffs that can be sustained in equilibrium. However, most of the sharp results in that literature consider the limit case when the players are infinitely patient, or when transfers are allowed. There are no 'off-the-shelf' results that are applicable to an arbitrary game to obtain the best equilibrium payoffs a player can obtain for any combination of the game's parameters. Results tend to rely on complex strategy profiles, calibrated to the game's parameters, as a means for delineating the equilibrium payoff set.

This paper studies a simple, repeated interaction between a principal and a group of agents, which 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. Many of these examples can be seen as a 'pure persuasion' problem: the candidates or applicants simply want to be selected, while the decision-maker wants to select an individual satisfying some requirements (e.g., if he's qualified for the task). Oftentimes, the decision-maker 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 which 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. Furthermore, if we view the principal as a figure of authority who can steer the agents away from equilibria that are inferior (in her eyes), then either the repeated interaction leads to the best outcome for the principal, or it doesn't help the principal at all (she gets the same payoff as if the game was static), and regardless of parameter values, the players follow simple Markov strategies. Given rising interest in dynamic mechanism design, mechanism design without transfers, and mechanism design without commitment, we hope our notion of principal-renegotiation-proofness will prove useful in analyzing the intersection of these three areas.

# Appendix

## A1 Characterization of first-best with two agents

Proposition 1 corrresponds to $\lambda = 0$ in Proposition 7, so we prove the latter here.

**Proof of Proposition 7.** Suppose a first-best PPE exists, and denote the set of first-best equilibrium payoffs by $\mathcal{E}^{FB} \subset \mathbb{R}^3$. Given the reward scheme $(u, \lambda)$, the sum of the two agents' (average) continuation payoffs must equal

$$\sigma^* = u + \lambda \left[(1-\theta_1)(1-\theta_2)\beta + \left(1-(1-\theta_1)(1-\theta_2)\right)\gamma\right] \tag{11}$$

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 $\mathcal{E}^{FB}$. For convenience, we assume that after each period, firms can observe the realization of a public randomization device, and select continuation equilibria based on these realizations. This guarantees the convexity of the equilibrium payoff set, but is not needed for our results.

Denote by $[\underline{\sigma}_i, \overline{\sigma}_i]$ the set of average payoffs attainable in a first-best equilbrium for agent $i$.[17] The payoff sets may differ, since the agents may have different abilities. Let $p_i = \gamma\theta_i + \beta(1-\theta_i)$ be agent $i$'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 $\underline{\sigma}_1$.** Given the observations above, $\underline{\sigma}_1$ must be the minimal payoff of agent 1 that can be supported when promised continuation payoffs are restricted to $\mathcal{E}^{FB}$. Suppose $\underline{\sigma}_1$ is obtained when agent 1 is LR (we confirm this later). We assume $\underline{\sigma}_1$ actually solves the following weaker minimization problem, where some of the incentive constraints of the agents are ignored. Specifically, we assume $\underline{\sigma}_1$ minimizes

$$\begin{aligned}(1-\theta_2)\left[(1-\delta)(u+p_1\lambda)+p_1\delta\sigma_1(1S)+(1-p_1)\delta\sigma_1(1F)\right] \\ +\theta_2\delta\left[\gamma\sigma_1(2S)+(1-\gamma)\sigma_1(2F)\right]\end{aligned} \tag{12}$$

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

$$\delta\left[p_1\sigma_2(1S)+(1-p_1)\sigma_2(1F)\right] \geq (1-\delta)u+\beta\left((1-\delta)\lambda+\delta\sigma_2(2S)\right)+(1-\beta)\delta\sigma_2(2F),$$

as well as the feasibility constraints, i.e., the constraints on the continuation values, $\sigma_i \in [\underline{\sigma}_i, \overline{\sigma}_i]$, $i = 1, 2$. Adding the remaining IC constraints could make the minimum greater, for more stringent necessary conditions. However, this will be redundant since the necessary condition found will be sufficient.[18] Using the fact that agents' continuations sum to $\sigma^*$ for any realization, we can rewrite agent 2's IC constraint:

$$\delta\left(\beta\sigma_1(2S)+(1-\beta)\sigma_1(2F)\right) \geq (1-\delta)\left(u+\beta\lambda\right)+\delta\left[p_1\sigma_1(1S)+(1-p_1)\sigma_1(1F)\right].$$

Clearly, (12) is minimized only if $\sigma_1(1S) = \sigma_1(1F) = \underline{\sigma}_1$ (lowering these continuations decreases the objective and can only relax the constraint). Therefore, $\underline{\sigma}_1$ minimizes

$$(1-\theta_2)\left[(1-\delta)\left(u+p_1\lambda\right)+\delta\underline{\sigma}_1\right]+\theta_2\delta\left[\gamma\sigma_1(2S)+(1-\gamma)\sigma_1(2F)\right] \tag{13}$$

subject to the binding IC constraint $\delta\left(\beta\sigma_1(2S)+(1-\beta)\sigma_1(2F)\right) = (1-\delta)\left(u+\beta\lambda\right)+\delta\underline{\sigma}_1$ and the feaasibility constraints. Using the IC constraint, we see the coefficient

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

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

on $\sigma_1(2S)$ is $(\gamma-\beta)/(1-\beta)>0$, 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 cases to consider.

*Case 1: $\sigma_1(2S)=\underline{\sigma}_1$ does not violate the feasibility constraints.* Then $\sigma_1(2F)=\underline{\sigma}_1+\frac{(1-\delta)(u+\beta\lambda)}{\delta(1-\beta)}$ and feasibility requires $\sigma_1(2F)\leq\overline{\sigma}_1$. Setting $\underline{\sigma}_1$ equal to the objective in the minimization problem, we obtain $\underline{\sigma}_1=(1-\theta_2)(u+p_1\lambda)+\theta_2\frac{1-\gamma}{1-\beta}(u+\beta\lambda)$. To check whether the feasibility constraint $\underline{\sigma}_1+\frac{(1-\delta)(u+\beta\lambda)}{\delta(1-\beta)}\leq\overline{\sigma}_1$ is satisfied, we will consider later below the problem of *maximizing* 1's continuation payoff.

*Case 2: $\sigma_1(2F)=\overline{\sigma}_1$.* If $\sigma_1(2S)$ cannot be brought down further, then $\sigma_1(2F)$ must be at its maximum value, $\overline{\sigma}_1$. Then $\sigma_1(2S)=\frac{(1-\delta)(u+\beta\lambda)}{\delta\beta}+\frac{\underline{\sigma}_1}{\beta}-\frac{(1-\beta)\overline{\sigma}_1}{\beta}$ and, setting $\underline{\sigma}_1$ equal to the objective in the minimization problem,

$$\underline{\sigma}_1=\frac{(1-\delta)\left[(1-\theta_2)(u+p_1\lambda)+\theta_2\frac{\gamma(u+\beta\lambda)}{\beta}\right]-\delta\theta_2\overline{\sigma}_1\left[\frac{\gamma-\beta}{\beta}\right]}{1-\delta\left[(1-\theta_2)+\theta_2\frac{\gamma}{\beta}\right]}. \tag{14}$$

Feasibility requires that $\frac{1-\delta}{\delta\beta}+\frac{\underline{\sigma}_1}{\beta}-\frac{(1-\beta)\overline{\sigma}_1}{\beta}\in[\underline{\sigma}_1,\overline{\sigma}_1]$.

**Step 2. Solving for $\overline{\sigma}_1$.** Suppose that agent 1's first-best equlibrium payoff is maximized when 1 is discerning (this will later be confirmed). Analogously to step 1, we now solve for $\overline{\sigma}_1$ as a solution to the problem of maximizing 1's payoff

$$\begin{aligned}&\theta_1\left[(1-\delta)(u+\gamma\lambda)+\gamma\delta\sigma_1(1S)+(1-\gamma)\delta\sigma_1(1F)\right]\\&\quad+(1-\theta_1)\delta\left[p_2\sigma_1(2S)+(1-p_2)\sigma_1(2F)\right]\end{aligned}$$

subject to the IC constraint that agent 1 does not propose when he is unqualified,

$$\delta\left[p_2\sigma_1(2S)+(1-p_2)\sigma_1(2F)\right]\geq(1-\delta)u+\left(\beta\left((1-\delta)\lambda+\delta\sigma_1(1S)\right)+(1-\beta)\delta\sigma_1(1F)\right),$$

and feasibility constraints. As in step 1, ignoring remaining constraints is wlog. Setting $\sigma_1(2S),\sigma_1(2F)$ to $\overline{\sigma}_1$ (increases objective, relaxes IC), the objective becomes

$$\theta_1\left[(1-\delta)u+\left(\gamma\left((1-\delta)\lambda+\delta\sigma_1(1S)\right)+(1-\gamma)\delta\sigma_1(1F)\right)\right]+(1-\theta_1)\delta\overline{\sigma}_1$$

and the IC constraint, which must bind, becomes

$$\delta\overline{\sigma}_1=(1-\delta)(u+\beta\lambda)+\delta\left(\beta\sigma_1(1S)+(1-\beta)\sigma_1(1F)\right).$$

Solving this problem involves increasing $\sigma_1(1S)$ as much as possible (intuitively, increasing agent 1's payoff when he is discerning and succeeds). There are 2 cases:

*Case 3: $\sigma_1(1S) = \overline{\sigma}_1$ does not violate the feasibility constraints.* Then $\sigma_1(1F) = \overline{\sigma}_1 - \frac{(1-\delta)(u+\beta\lambda)}{\delta(1-\beta)}$ and feasibility requires $\sigma_1(1F) \geq \underline{\sigma}_1$. Setting $\overline{\sigma}_1$ equal to the objective in the maximization problem, we obtain $\overline{\sigma}_1 = \theta_1(\lambda + u)\left[\frac{\gamma-\beta}{1-\beta}\right]$.

*Case 4: $\sigma_1(1F) = \underline{\sigma}_1$.* Plugging $\sigma_1(1S) = \frac{\overline{\sigma}_1}{\beta} - \frac{(1-\beta)\underline{\sigma}_1}{\beta} - \frac{(1-\delta)(u+\beta\lambda)}{\delta\beta} \in [\underline{\sigma}_1, \overline{\sigma}_1]$ in the objective,

$$\overline{\sigma}_1 = \theta_1\left[\frac{\gamma}{\beta} - 1\right]\left(\frac{(1-\delta)u + \delta\underline{\sigma}_1}{\delta\theta_1\left[\frac{\gamma}{\beta} - 1\right] - (1-\delta)}\right). \tag{15}$$

In particular, note that it must be the case that $\delta\theta_1\left[\frac{\gamma}{\beta} - 1\right] - (1-\delta) > 0$.

**Step 3. Combining $\underline{\sigma}_1$ and $\overline{\sigma}_1$.** We now combine the possible cases.

*Cases 1 and 3.* Combining $\underline{\sigma}_1 = (1-\theta_2)(u + p_1\lambda) + \theta_2\frac{1-\gamma}{1-\beta}(u + \beta\lambda)$ and $\overline{\sigma}_1 = \theta_1(\lambda + u)\left[\frac{\gamma-\beta}{1-\beta}\right]$, together with the necessary conditions for these cases (which boil down to $\overline{\sigma}_1 - \underline{\sigma}_1 \geq \frac{(1-\delta)(u+\beta\lambda)}{\delta(1-\beta)}$), the following condition must hold:

$$\theta_1(\lambda+u)\left[\frac{\gamma-\beta}{1-\beta}\right] - \left((1-\theta_2)(u+p_1\lambda) + \theta_2\left[\frac{1-\gamma}{1-\beta}\right](u+\beta\lambda)\right) \geq \frac{(1-\delta)(u+\beta\lambda)}{\delta(1-\beta)}.$$

This condition simplifies to condition (10) in the statement of Proposition 7.

*Cases 2 and 4.* Combining (14) and (15), it can be shown that

$$\overline{\sigma}_1 - \underline{\sigma}_1 = (1-\delta)(u+\lambda\beta)\frac{1 + (\theta_1+\theta_2)\left[\frac{\gamma}{\beta} - 1\right] - \theta_1\theta_2\left[\frac{\gamma}{\beta} - 1\right]\frac{\lambda\beta}{u+\lambda\beta}}{\delta(\theta_1+\theta_2)\left[\frac{\gamma}{\beta} - 1\right] - (1-\delta)}.$$

Furthermore, the feasibility conditions for the two cases reduce to

$$\overline{\sigma}_1 - \underline{\sigma}_1 \in \left[\frac{1-\delta}{\delta}(u+\beta\lambda), \frac{1-\delta}{\delta(1-\beta)}(u+\beta\lambda)\right].$$

The requirement that $\overline{\sigma}_1 - \underline{\sigma}_1 \leq \frac{1-\delta}{\delta(1-\beta)}(u+\beta\lambda)$ is equivalent to the inequality,

$$\frac{\delta(1-\beta) + \delta(1-\beta)(\theta_1+\theta_2)\left[\frac{\gamma}{\beta} - 1\right] - \delta(1-\beta)\theta_1\theta_2\left[\frac{\gamma}{\beta} - 1\right]\frac{\lambda\beta}{u+\lambda\beta}}{\delta(\theta_1+\theta_2)\left[\frac{\gamma}{\beta} - 1\right] - (1-\delta)} \leq 1.$$

By the observation in case 4 that $\delta\theta_1\left[\frac{\gamma}{\beta}-1\right]-(1-\delta)>0$, the denominator is positive. The inequality can therefore be rewritten to again obtain (10). Finally, note that the conditions for cases 1 and 4 can be satisfied jointly only for a parameter set of measure zero, since case 1 requires $\overline{\sigma}_1-\underline{\sigma}_1 \geq \frac{(1-\delta)(u+\beta\lambda)}{\delta(1-\beta)}$, whereas case 4 requires $\overline{\sigma}_1-\underline{\sigma}_1 \leq \frac{(1-\delta)(u+\beta\lambda)}{\delta(1-\beta)}$. The same holds for the combination of cases 2 and 3.

We next verify our conjecture that agent 1's minimal (respectively, maximal) first-best equilibrium payoff is obtained when he is LR (respectively, discerning).

**Step 4. Verifying the postulated roles.**

**Claim 1.** *$\underline{\sigma}_1$ is attained when agent 1 is LR.*

*Proof.* If $\underline{\sigma}_1$ were met when agent 1 is discerning, his payoff would be

$$\begin{aligned}&\theta_1\left((1-\delta)\left(u+\gamma\lambda\right)+\gamma\delta\sigma_1(1S)+(1-\gamma)\delta\sigma_1(1F)\right)\\&+(1-\theta_1)\delta\left(p_2\sigma_1(2S)+(1-p_2)\sigma_1(2F)\right).\end{aligned}$$

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

$$\delta\left(p_2\sigma_1(2S)+(1-p_2)\sigma_1(2F)\right)\geq(1-\delta)\left(u+\lambda\beta\right)+\delta\left(\beta\sigma_1(1S)+(1-\beta)\sigma_1(1F)\right).$$

Therefore,

$$\begin{aligned}\underline{\sigma}_1 &\geq \theta_1\left((1-\delta)\left(u+\gamma\lambda\right)+\delta\left(\gamma\sigma_1(1S)+(1-\gamma)\sigma_1(1F)\right)\right)\\&\quad+(1-\theta_1)\delta\left(p_2\sigma_1(2S)+(1-p_2)\sigma_1(2F)\right)\\&\geq \theta_1\left((1-\delta)\left(u+\gamma\lambda\right)+\gamma\delta\sigma_1(1S)+(1-\gamma)\delta\sigma_1(1F)\right)\\&\quad+(1-\theta_1)\left((1-\delta)\left(u+\lambda\beta\right)+\delta(\beta\sigma_1(1S)+(1-\beta)\sigma_1(1F))\right)\\&\geq (1-\delta)\left(u+\lambda p_1\right)+\delta\underline{\sigma}_1,\end{aligned}$$

which implies that $\underline{\sigma}_1 \geq (1-\delta)(u+\lambda p_1)+\delta\underline{\sigma}_1$, or that $\underline{\sigma}_1 \geq u+\lambda p_1$. But $u+\lambda p_1$ is agent 1's average payoff when he is selected in all periods, a contradiction. □

**Claim 2.** *$\overline{\sigma}_1$ is attained when agent 1 is discerning.*

*Proof.* By contradiction. If $\overline{\sigma}_1$ is attained when agent 1 is LR, his average payoff is

$$\begin{aligned}&(1-\theta_2)\left((1-\delta)\left(u+p_1\lambda\right)+\delta\left(p_1\sigma_1(1S)+(1-p_1)\sigma_1(1F)\right)\right)\\&+\theta_2\delta\left(\gamma\sigma_1(2S)+(1-\gamma)\sigma_1(2F)\right).\end{aligned}$$

The IC constraint of the discerning agent 2 for not proposing when unqualified is:

$$(1-\delta)(u+\beta\lambda)+\delta(\beta\sigma_2(2S)+(1-\beta)\sigma_2(2F))\leq\delta(p_1\sigma_2(1S)+(1-p_1)\sigma_2(1F)).$$

Recalling that each outcome $x\in\{1S,1F,2S,2F\}$, $\sigma_1(x)+\sigma_2(x)=\sigma^*$, we have:

$$(1-\delta)(u+\beta\lambda)+\delta(p_1\sigma_1(1S)+(1-p_1)\sigma_1(1F))\leq\delta(\beta\sigma_1(2S)+(1-\beta)\sigma_1(2F)).$$

Therefore

$$\begin{aligned}\sigma_2&\leq(1-\theta_2)((1-\delta)(u+p_1\lambda)+\delta(p_1\sigma_1(1S)+(1-p_1)\sigma_1(1F)))\\&\quad+\theta_2\delta(\gamma\sigma_1(2S)+(1-\gamma)\sigma_1(2F))\\&\leq(1-\theta_2)\delta(\beta\sigma_1(2S)+(1-\beta)\sigma_1(2F))+\theta_2\delta(\gamma\sigma_1(2S)+(1-\gamma)\sigma_1(2F))\\&\quad+(1-\theta_2)(1-\delta)((u+p_1\lambda)-(u+\beta\lambda))\\&\leq\delta\overline{\sigma}_1+(1-\theta_2)(1-\delta)\lambda(p_1-\beta),\end{aligned}$$

which means $\overline{\sigma}_1\leq(1-\theta_2)\lambda(p_1-\beta)<(1-\theta_2)(u+\lambda p_1)$. But $(1-\theta_2)(u+\lambda p_1)$ is 1's average payoff when he is last resort in all periods, a contradiction. □

We conclude that we have in (10) a necessary condition for the existence of a first-best PPE. In fact, since (10) is also sufficient for cases 1 and 3 to hold jointly, this immediately implies (10) is also sufficient for the existence of a first-best PPE.[19] We next show directly that the MLR forms a (first-best) PPE whenever (10) holds.

**Step 5: Sufficient conditions for MLR.** 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. Then the IC constraint for an unqualified discerning agent not to propose is given by:

$$\delta V_1^D\geq(1-\delta)(u+\beta\lambda)+\beta\delta V_1^D+(1-\beta)\delta V_1^{LR}.$$

Subtracting $\delta V_1^{LR}$ from both sides of the inequality yields:

$$V_1^D-V_1^{LR}\geq\frac{(1-\delta)(u+\beta\lambda)}{\delta(1-\beta)}. \tag{16}$$

[19] More precisely, following APS, (10) guarantees that a non-empty, bounded, self-generating set of first-best payoffs (payoff vectors in which the principal obtains her first best) exists.

To express the LHS in terms of the parameters we solve for $V_1^D$ and $V_1^{LR}$ :

$$V_1^D = \theta_1[(1-\delta)(u+\lambda\gamma) + \gamma\delta V_1^D + (1-\gamma)\delta V_1^{LR}] + (1-\theta_1)\delta V_1^D,$$
$$V_1^{LR} = (1-\theta_2)[(1-\delta)(u+\lambda p_1) + \delta V_1^{LR}] + \theta_2[\gamma\delta V_1^{LR} + (1-\gamma)\delta V_1^D].$$

Rearranging, we have

$$V_1^D = \frac{\theta_1(1-\delta)(u+\lambda\gamma) + \theta_1(1-\gamma)\delta V_1^{LR}}{(1-\delta) + \theta_1(1-\gamma)\delta}, \tag{17}$$
$$V_1^{LR} = \frac{(1-\theta_2)(1-\delta)(u+\lambda p_1) + \theta_2(1-\gamma)\delta V_1^D}{(1-\delta) + \delta\theta_2(1-\gamma)}.$$

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

$$\frac{(1-\theta_2)(1-\delta)(u+\lambda p_1) + \theta_1(1-\theta_2)(1-\gamma)\delta(u+\lambda p_1) + \theta_1\theta_2\delta(1-\gamma)(u+\lambda\gamma)}{(1-\delta) + \delta(1-\gamma)(\theta_1+\theta_2)},$$

and from (17) it follows that

$$V_1^D - V_1^{LR} = \frac{\theta_1(1-\delta)(u+\lambda\gamma) - (1-\delta)V_1^{LR}}{(1-\delta) + \theta_1(1-\gamma)\delta}.$$

Plugging in the expression for $V_1^{LR}$ yields:

$$V_1^D - V_1^{LR} = (1-\delta)\frac{(u+\lambda\beta)(\theta_1+\theta_2-1) + \theta_1\theta_2\lambda(\gamma-\beta)}{(1-\delta) + \delta(1-\gamma)(\theta_1+\theta_2)},$$

which combined with the IC constraint (16) yields the condition (10). ■

## A2 Proofs for the many agents case

**Proof of Proposition 3.** Suppose some PPE achieves the principal's first best. Thus at each history $h$, there is $i(h) \in \mathcal{A}$ such that agents other than $i(h)$ propose themselves iff they are qualified, $i(h)$ always proposes himself, the principal picks $i(h)$ only when he is the sole proposer, and otherwise picks an agent other than $i(h)$.

An agent $j$ could follow the strategy of proposing himself in each round, whatever its quality. By doing this, the agent gets picked with probability at least $(1 - \max_{i\in\mathcal{A}} \theta_i)^{n-1}$ at any history $h$ with $j = i(h)$, and he gets picked with probability at least $(1 - \max_{i\in\mathcal{A}} \theta_i)^{n-2}$ at any history $h$ with $j \neq i(h)$. Each agent can thus secure himself a discounted likelihood of being picked which is larger than or equal to $(1-\max_{i\in\mathcal{A}}\theta_i)^{n-1}/(1-\delta)$.

To achieve her first best in equilibrium, the principal picks exactly one agent in each round. So, in total, the aggregate discounted likelihood of being picked is $1/(1-\delta)$. The equilibrium could not exist if $1/(1-\delta)$ were strictly smaller than the aggregate discounted likelihood of being picked that agents can minimally guarantee, that is, $n(1-\max_{i\in\mathcal{A}}\theta_i)^{n-1}/(1-\delta)$. That relationship holds iff $\max_{i\in\mathcal{A}}\theta_i < \theta^*$. ■

**Remark 1.** *Observe that* $\sum_{j\neq\ell}\theta_j\sigma_j(\vec{\theta},\ell)+\rho_\ell(\vec{\theta}) = 1$*, since the principal always selects some agent, resorting to the last resort agent if no discerning agent proposes. Moreover, note that* $\sum_{j\neq i,\ell}p_j(\vec{\theta},i,\ell) = 1$*, since the fact that player* $i$ *has proposed means that the selected agent will come from the discerning pool. On the other hand,* $\sum_{j\neq i,\ell}q_j(\vec{\theta},i,\ell)+\frac{\rho_\ell(\vec{\theta})}{1-\theta_i} = 1$*, since it is possible that no discerning agent will propose.*

**Proof of Proposition 4** The proof follows from Lemmas 1 and 2, combined with the invertibility of $B_i(\vec{\theta})$ proved in Lemma 4(d) further below.

**Lemma 1.** *The MLR strategy profile constitutes a PPE if and only if*

$$\frac{\delta(1-\gamma)}{(1-\delta)}M_i^Q(\vec{\theta})\Delta\vec{V}_i(\vec{\theta}) \leq \vec{\sigma}_i(\vec{\theta}) \leq \frac{\delta(1-\gamma)}{(1-\delta)}M_i^U\Delta\vec{V}_i(\vec{\theta}).$$

*Proof.* First note that the MLR strategy of the principal is first best for him, regardless of his discount factor and agents' types, so long as agents follow their strategies. Moreover, given that the principal follows this strategy, a last resort agent cannot change his probability of going back into the discerning pool of agents by his own actions. The last resort agent thus finds it optimal to propose himself with probability one, regardless of his discount factor and agents' types. It remains to check the incentive conditions for discerning agents.

Subtracting $\delta V_i^{LR}(\vec{\theta})$ from both sides of the condition ($\text{IC}_U$) for $i$ to refrain from proposing when unqualified and when $\ell$ is the last resort agent, we find that

$$\begin{aligned}
&\frac{\rho_\ell(\vec{\theta})}{1-\theta_i}\delta\Delta V_i^D(\vec{\theta},\ell)+\sum_{j\neq i,\ell}q_j(\vec{\theta},i,\ell)\left(\gamma\delta\Delta V_i^D(\vec{\theta},\ell)+(1-\gamma)\delta\Delta V_i^D(\vec{\theta},j)\right)\\
&\qquad\geq \sigma_i(\vec{\theta},\ell)\left(1-\delta+\beta\delta_i\Delta V_i^D(\vec{\theta},\ell)\right)\\
&\qquad\quad+(1-\sigma_i(\vec{\theta},\ell))\sum_{j\neq i,\ell}p_j(\vec{\theta},i,\ell)\left(\gamma\delta\Delta V_i^D(\vec{\theta},\ell)+(1-\gamma)\delta\Delta V_i^D(\vec{\theta},j)\right).
\end{aligned}$$

Collect all $\Delta V_i^D$ terms on the left-hand side, and multiply the inequality through by $\frac{1}{1-\delta}$. Then, for each $j \neq \ell$, the coefficient multiplying $\frac{(1-\gamma)\delta}{1-\delta}\Delta V_i^D(\vec{\theta}, j)$ is easily seen to be $[M_i^U(\vec{\theta})]_{\ell j}$. Using Remark 1, the coefficient multiplying $\frac{(1-\gamma)\delta}{1-\delta}\Delta V_i^D(\vec{\theta}, \ell)$ is

$$\frac{1}{1-\gamma}\left(\frac{\rho_\ell(\vec{\theta})}{1-\theta_i} + \gamma\sum_{j\neq i,\ell} q_j(\vec{\theta}, i, \ell) - \beta\sigma_i(\vec{\theta}, \ell) - \gamma(1-\sigma_i(\vec{\theta}, \ell))\sum_{j\neq i,\ell} p_j(\vec{\theta}, i, \ell)\right)$$
$$= \frac{1}{1-\gamma}\left(\frac{\rho_\ell(\vec{\theta})}{1-\theta_i} + \gamma(1 - \frac{\rho_\ell(\vec{\theta})}{1-\theta_i}) - \beta\sigma_i(\vec{\theta}, \ell) - \gamma(1-\sigma_i(\vec{\theta}, \ell))\right) = [M_i^U(\vec{\theta})]_{\ell\ell}.$$

Stacking the inequalities for $\ell \neq i$ yields the matrix inequality with $M_i^U(\vec{\theta})$.

Next, subtracting $\delta V_i^{LR}(\vec{\theta})$ from both sides of the condition ($\text{IC}_Q$) for agent $i$ to propose himself when qualified and when $\ell$ is the last resort agent, we find that

$$\sigma_i(\vec{\theta}, \ell)\left(1 - \delta + \gamma\delta\Delta V_i^D(\vec{\theta}, \ell)\right)$$
$$+ (1-\sigma_i(\vec{\theta}, \ell))\sum_{j\neq i,\ell} p_j(\vec{\theta}, i, \ell)\left(\gamma\delta\Delta V_i^D(\vec{\theta}, \ell) + (1-\gamma)\delta\Delta V_i^D(\vec{\theta}, j)\right)$$
$$\geq \sum_{j\neq i,\ell} q_j(\vec{\theta}, i, \ell)\left(\gamma\delta\Delta V_i^D(\vec{\theta}, \ell) + (1-\gamma)\delta\Delta V_i^D(\vec{\theta}, j)\right) + \frac{\rho_\ell(\vec{\theta})}{1-\theta_i}\delta\Delta V_i^D(\vec{\theta}, \ell).$$

Collect all $\Delta V_i^D$-terms on the right-hand side, and multiply the inequality through by $\frac{1}{1-\delta}$. Then the coefficient multiplying $\frac{(1-\gamma)\delta}{1-\delta}\Delta V_i^D(\vec{\theta}, j)$ is easily seen to be $[M_i^Q(\vec{\theta})]_{\ell j}$. Given Remark 1, the coefficient multiplying $\frac{(1-\gamma)\delta}{1-\delta}\Delta V_i^D(\vec{\theta}, \ell)$ reduces to

$$\frac{1}{1-\gamma}\left(\gamma\sum_{j\neq i,\ell} q_j(\vec{\theta}, i, \ell) + \frac{\rho_\ell(\vec{\theta})}{1-\theta_i} - \gamma\right) = [M_i^Q(\vec{\theta})]_{\ell\ell}.$$

Stacking the inequalities for $\ell \neq i$ yields the matrix inequality with $M_i^Q(\vec{\theta})$. □

**Lemma 2.** *For all $i$ and $\vec{\theta}$, the average discounted payoff differences $\Delta\vec{V}_i(\vec{\theta})$ satisfy:*

$$B_i(\vec{\theta})\Delta\vec{V}_i(\vec{\theta}) = \frac{u_i(1-\delta_i)}{\delta_i(1-\gamma)}\vec{\pi}_i(\vec{\theta}),$$

*where $B_i(\vec{\theta})$ is the $(n-1)$-square matrix whose $\ell\ell'$-entry, for any $\ell, \ell'$ in $\mathcal{A}\setminus\{i\}$, is:*

$$[B_i(\vec{\theta})]_{\ell\ell'} = \begin{cases} \pi_{i\ell'}(\vec{\theta}) - \pi_{\ell\ell'}(\vec{\theta}) & \text{if } \ell \neq \ell', \\ 1 + \pi_{i\ell}(\vec{\theta}) + (1-\delta_i)/(\delta_i(1-\gamma)) & \text{if } \ell = \ell'. \end{cases}$$

*Proof.* The value function $V_i^D$ is defined by the equation

$$\begin{aligned} V_i^D(\vec{\theta},\ell) &= \theta_i\sigma_i(\vec{\theta},\ell)\left(1-\delta+\gamma\delta V_i^D(\vec{\theta},\ell)+(1-\gamma)\delta V_i^{LR}(\vec{\theta})\right) \\ &+ \sum_{j\neq i,\ell}\theta_j\sigma_j(\vec{\theta},\ell)\left(\gamma\delta V_i^D(\vec{\theta},\ell)+(1-\gamma)\delta V_i^D(\vec{\theta},j)\right)+\rho_\ell(\vec{\theta})\delta V_i^D(\vec{\theta},\ell), \end{aligned} \tag{18}$$

while the value function $V_i^{LR}$ is defined by

$$V_i^{LR}(\vec{\theta}) = \rho_i(\vec{\theta})\left(1-\delta+\delta V_i^{LR}(\vec{\theta})\right)+\sum_{j\neq i}\theta_j\sigma_j(\vec{\theta},i)\left(\gamma\delta V_i^{LR}(\vec{\theta})+(1-\gamma)\delta V_i^D(\vec{\theta},j)\right). \tag{19}$$

Subtracting $\delta V_i^{LR}(\vec{\theta})$ from both sides of Equation (18), we find that

$$\begin{aligned} V_i^D(\vec{\theta},\ell)-\delta V_i^{LR}(\vec{\theta}) &= \theta_i\sigma_i(\vec{\theta},\ell)\left(1-\delta+\gamma\delta\Delta V_i^D(\vec{\theta},\ell)\right) \\ &+ \sum_{j\neq i,\ell}\theta_j\sigma_j(\vec{\theta},\ell)\left(\gamma\delta\Delta V_i^D(\vec{\theta},\ell)+(1-\gamma)\delta\Delta V_i^D(\vec{\theta},j)\right)+\rho_\ell(\vec{\theta})\delta\Delta V_i^D(\vec{\theta},\ell), \end{aligned} \tag{20}$$

In view of Remark 1, Equation (18) simplifies to

$$\begin{aligned} V_i^D(\vec{\theta},\ell)-\delta V_i^{LR}(\vec{\theta}) &= \theta_i\sigma_i(\vec{\theta},\ell)(1-\delta)+(1-\gamma)\delta\sum_{j\neq i,\ell}\theta_j\sigma_j(\vec{\theta},\ell)\Delta V_i^D(\vec{\theta},j) \\ &+ \delta\Delta V_i^D(\vec{\theta},\ell)\left(\gamma+(1-\gamma)\rho_\ell(\vec{\theta})\right). \end{aligned} \tag{21}$$

Similarly, subtracting $\delta V_i^{LR}(\vec{\theta})$ from both sides of Equation (19), we find that

$$V_i^{LR}(\vec{\theta})-\delta V_i^{LR}(\vec{\theta}) = \rho_i(\vec{\theta})(1-\delta)+(1-\gamma)\delta\sum_{j\neq i}\theta_j\sigma_j(\vec{\theta},i)\Delta V_i^D(\vec{\theta},j). \tag{22}$$

Subtracting Equation (22) from Equation (21), and using the definition of $\pi_{\ell\ell'}(\vec{\theta})$,

$$\begin{aligned} \Delta V_i^D(\vec{\theta},\ell) &= \pi_{\ell i}(\vec{\theta})(1-\delta)+\delta\Delta V_i^D(\vec{\theta},\ell)\left(\gamma-(1-\gamma)\pi_{i\ell}(\vec{\theta})\right) \\ &+ (1-\gamma)\delta\sum_{j\neq i,\ell}\left(\theta_j\sigma_j(\vec{\theta},\ell)-\theta_j\sigma_j(\vec{\theta},i)\right)\Delta V_i^D(\vec{\theta},j). \end{aligned} \tag{23}$$

Note that $\theta_j\sigma_j(\vec{\theta},\ell)-\theta_j\sigma_j(\vec{\theta},i)=\pi_{\ell j}(\vec{\theta})-\pi_{ij}(\vec{\theta})$. We can thus rearrange Equation (23) and divide through by $(1-\gamma)\delta$ to find that $B_i(\vec{\theta})\Delta\vec{V}_i(\vec{\theta})=\frac{1-\delta}{(1-\gamma)\delta}\vec{\pi}_i(\vec{\theta})$, as claimed. □

**Proof of Proposition 5.** The result follows from Lemmas 3-7. We let $\sigma^* = \sigma_i(\theta^*, \ldots, \theta^*, \ell)$ for any $i \neq \ell$ (the selection probability does not vary on $i$ and $\ell$ when agents are identical).

**Lemma 3.** *(a) For each agent $\ell \neq \ell'$, $\frac{\rho_{\ell'}(\vec{\theta})}{\sigma_{\ell'}(\vec{\theta},\ell)}$ is decreasing in $\theta_k$, for all $k \in \mathcal{A}$.*

*(b) $\pi_{\ell\ell'}(\vec{\theta}) > 0$ for all $\vec{\theta} \in [\underline{\theta}, 1]^n$ and any $\ell \neq \ell'$ in $\mathcal{A}$, if and only if $\underline{\theta} > \theta^*$.*

*(c) $(1-\theta^*)\sigma^* \leq 1/2$.*

*(d) Suppose $\underline{\theta} > \theta^*$, and $\ell \neq i$ is such that $\theta_\ell \leq \theta_i$. Then $\pi_{\ell i}(\vec{\theta}) - \pi_{i\ell}(\vec{\theta}) \leq 1/2$.*

*(e) The minimal probability premium $\underline{\pi} := \min_{\ell \in \mathcal{A}\setminus\{i\}} \min_{\vec{\theta} \in [\underline{\theta},1]^n} \pi_{\ell i}(\vec{\theta})$ is given by*

$$\underline{\pi} = \begin{cases} \frac{\underline{\theta}}{n-1} & \text{if } n \geq 3 \text{ and } \underline{\theta} \geq 1 - \sqrt[n-2]{\frac{1}{n}} \\ \frac{1-n(1-\underline{\theta})^{n-1}}{n-1} & \text{otherwise.} \end{cases}$$

*Proof. (a)* This is true since the following function is decreasing in $\theta_k$, for all $k \in \mathcal{A}$:

$$\begin{aligned} \frac{\rho_{\ell'}(\vec{\theta})}{\sigma_{\ell'}(\vec{\theta},\ell)} &= \frac{\prod_{j\neq\ell'}(1-\theta_j)}{\sum_{k=0}^{n-2}\frac{1}{k+1}\sum_{S\subseteq A\setminus\{\ell,\ell'\},|S|=k}\prod_{j\in S}\theta_j \prod_{j\in A\setminus S, j\neq \ell,\ell'} 1-\theta_j} \\ &= \frac{1-\theta_\ell}{\sum_{k=0}^{n-2}\frac{1}{k+1}\sum_{S\subseteq A\setminus\{\ell,\ell'\},|S|=k}\prod_{j\in S}\frac{\theta_j}{1-\theta_j}\prod_{j\in A\setminus S, j\neq\ell,\ell'}\frac{1-\theta_j}{1-\theta_j}} \\ &= \frac{1-\theta_\ell}{\sum_{k=0}^{n-2}\frac{1}{k+1}\sum_{S\subseteq A\setminus\{\ell,\ell'\},|S|=k}\prod_{j\in S}\frac{\theta_j}{1-\theta_j}}. \end{aligned}$$

*(b)* Notice that $\pi_{\ell\ell'}(\vec{\theta}) > 0$ iff $\theta_{\ell'} > \frac{\rho_{\ell'}(\vec{\theta})}{\sigma_{\ell'}(\vec{\theta},\ell)}$. From *(a)*, the RHS takes its highest value at $\vec{\underline{\theta}} = (\underline{\theta}, \ldots, \underline{\theta})$. Using this, notice $\pi_{\ell\ell'}(\vec{\theta}) > 0$ for all $\vec{\theta} \in [\underline{\theta}, 1]^n$ and any two distinct $\ell, \ell'$ in $\mathcal{A}$, if and only if $\underline{\theta} > \underline{\theta}(n-1)\frac{(1-\underline{\theta})^{n-1}}{1-(1-\underline{\theta})^{n-1}}$, or equivalently, $\underline{\theta} > \theta^* = \sqrt[n-1]{\frac{1}{n}}$.

*(c)* First note that the definition of $\sigma^*$ is independent of the choice of $i, \ell$ since $\sigma$ is evaluated when all abilities are equal to $\theta^*$. Then observe $(1-\theta^*)\sigma^* \leq 1/2$ if and only if $\frac{2}{n}\sqrt[n-1]{\frac{1}{n}} \leq 1 - \sqrt[n-1]{\frac{1}{n}}$, since, by construction, $\theta^*\sigma^* = \rho^* := \rho_i(\theta^*, \ldots, \theta^*)$ and $\theta^* = 1 - \sqrt[n-1]{\frac{1}{n}}$. The desired inequality is thus equivalent to $1 \leq \frac{n^n}{(n+2)^{n-1}}$ Taking ln of both sides, and adding/subtracting $\ln(n+2)$, $1 \leq \frac{n^n}{(n+2)^{n-1}}$ is equivalent to

$$n\left(\ln n - \ln(n+2)\right) + \ln(n+2) \geq 0. \tag{24}$$

The inequality $1 \leq \frac{n^n}{(n+2)^{n-1}}$, and thus (24), is satisfied for $n \in \{2,3,4\}$ (i.e., $1 \geq 1$, $27/25 \geq 1$ and $256/216 \geq 1$, resp.). We now show it holds for all larger $n$ by proving the derivative of the LHS of (24) is positive for $n \geq 4$. Indeed, that derivative is

$$\frac{3}{n+2} + \ln n - \ln(n+2) > \frac{3}{n+2} - \frac{2}{n} = \frac{n-4}{n(n+2)},$$

where the inequality follows by strict concavity of $\ln n$, so $\frac{\ln(n+2)-\ln n}{2} < \frac{d}{dn}\ln n = \frac{1}{n}$.

*(d)* Note that $\theta_i \geq \theta_\ell$ implies that

$$\pi_{\ell i}(\vec{\theta}) - \pi_{i\ell}(\vec{\theta}) = \theta_i \sigma_i(\vec{\theta}, \ell) - \rho_i(\vec{\theta}) - \theta_\ell \sigma_\ell(\vec{\theta}, i) + \rho_\ell(\vec{\theta})$$

$$= (\theta_i - \theta_\ell)\left(\sigma_i(\vec{\theta}, \ell) - \prod_{j \neq i,\ell}(1-\theta_j)\right) \leq (\theta_i - \theta_\ell)\sigma_i(\vec{\theta}, \ell) \leq (1-\theta^*)\sigma^*.$$

The proof concludes by applying the inequality from *(c)*.

*(e)* Notice that $\pi_{\ell i}(\vec{\theta})$ is increasing in $\theta_i$, so that one should take $\theta_i = \underline{\theta}$ to find the minimum. If $n = 2$, then the minimum is reached by taking $\theta_{-i} = \underline{\theta}$ as well. Suppose $n \geq 3$. The expression $\pi_{\ell i}(\vec{\theta})$ is linear in $\theta_k$ for all $k \neq i, \ell$. Thus one need only consider the cases $\theta_k \in \{\underline{\theta}, 1\}$ for all $k$. Notice, however, that $\rho_i(\vec{\theta}) = 0$ as soon as one such $\theta_k = 1$, in which case $\pi_{\ell i}(\vec{\theta})$ is decreasing in $\theta_j$ for $j \neq i, \ell, k$, and independent of $\theta_\ell$. In addition, if $\theta_k = \underline{\theta}$ for all $k \neq i, \ell$, then $\pi_{\ell i}(\vec{\theta})$ is strictly increasing in $\theta_\ell$ and the minimum will be reached at $\theta_\ell = \underline{\theta}$. To summarize, the minimal $\pi_{\ell i}(\vec{\theta})$ is reached at a profile $\vec{\theta}$ where $\theta_i = \underline{\theta}$, and other agents' abilities are either all $\underline{\theta}$ or all 1. The probability premium is[20] $\frac{1-n(1-\underline{\theta})^{n-1}}{n-1}$ in the former case, and $\frac{\underline{\theta}}{n-1}$ in the latter case. It is then easy to check that the former expression is smaller than the latter if and only if $\underline{\theta} \leq 1 - \sqrt[n-2]{\frac{1}{n}}$ (which is larger than $\theta^*$). □

**Lemma 4.** *The matrix $B_i(\vec{\theta})$ satisfies the following properties.*

*(a)* $B_i(\vec{\theta})\vec{1} = \frac{1-\delta\gamma}{\delta(1-\gamma)}\vec{1} + \vec{\pi}_i(\vec{\theta})$.

*(b) Diagonal entries of $B_i(\vec{\theta})$ are positive. Off-diagonal entries are positive on rows $\ell$ with $\theta_i > \theta_\ell$, negative on rows $\ell$ with $\theta_i < \theta_\ell$, and zero on rows $\ell$ with $\theta_i = \theta_\ell$.*

*(c) For each $\ell \neq i$, let $z_\ell$ be the difference between row $\ell$'s diagonal entry and the sum of the absolute value of its off-diagonal entries: $z_\ell = [B_i(\vec{\theta})]_{\ell\ell} - \sum_{\ell' \neq \ell} |[B_i(\vec{\theta})]_{\ell\ell'}|$. If $\theta_i \leq \theta_\ell$, then $z_\ell = \frac{1-\delta\gamma}{\delta(1-\gamma)} + \pi_{\ell i}$. If $\theta_i \geq \theta_\ell$, then $z_\ell = \frac{1-\delta\gamma}{\delta(1-\gamma)} + 2\pi_{i\ell} - \pi_{\ell i}$.*

[20]Indeed, agents other than $i$ are symmetric and the fact that one must be chosen implies $(n-1)\underline{\theta}\sigma_i(\vec{\underline{\theta}}, \ell) + \rho_\ell(\vec{\underline{\theta}}) = 1$, or $\underline{\theta}\sigma_i(\vec{\underline{\theta}}, \ell) = \frac{1-(1-\underline{\theta})^{n-1}}{n-1}$.

*(d)* $B_i(\vec{\theta})$ *is (row) strictly diagonally dominant, and thus invertible.*

*(e)* $||B_i(\vec{\theta})^{-1}||_\infty \leq \frac{1}{\min_{\ell\neq i} z_\ell}$.

*(f)* $B_i(\vec{\theta})^{-1}\vec{\pi}_i(\vec{\theta}) = [Id - \frac{1-\delta\gamma}{\delta(1-\gamma)}B_i(\vec{\theta})^{-1}]\vec{1}$.

*(g)* $B_i(\vec{\theta})^{-1} = \sum_{k=0}^{\infty}(-1)^k(\theta_i - \theta^*)^k(X_i^{-1}Y_i)^kX_i^{-1}$, *where* $X_i$ *is the matrix* $B_i(\vec{\theta})$ *evaluated at* $\theta_i = \theta^*$, *and* $Y_i$ *is the positive matrix whose* $\ell\ell'$*-entry is* $\frac{\rho_\ell(\vec{\theta})}{1-\theta_i}$ *if* $\ell = \ell'$, *and* $-\theta_{\ell'}\frac{d\sigma_{\ell'}}{d\theta_i}(\vec{\theta},\ell)$ *if* $\ell \neq \ell'$.

*(h) Each component of the vector* $B_i(\vec{\theta})^{-1}\vec{\pi}(\vec{\theta})$ *is increasing in* $\theta_i$, *and each component of the vector* $B_i(\vec{\theta})^{-1}\vec{1}$ *is decreasing in* $\theta_i$, *for* $\theta_i \in [\theta^*, 1]$.

*Proof.* *(a)* Notice that

$$\sum_{\ell'\neq i,\ell}(\pi_{i\ell'}(\vec{\theta})-\pi_{\ell\ell'}(\vec{\theta})) = \sum_{\ell'\neq i,\ell}\theta_{\ell'}(\sigma_{\ell'}(\vec{\theta},i)-\sigma_{\ell'}(\vec{\theta},\ell)) = \rho_\ell(\vec{\theta})-\rho_i(\vec{\theta})+\theta_i\sigma_i(\vec{\theta},\ell)-\theta_\ell\sigma_\ell(\vec{\theta},i).$$

Thus the sum over the columns of the entries of $B_i(\vec{\theta})$ appearing on row $\ell$ is equal to $1 + \frac{1-\delta}{\delta(1-\gamma)} + \pi_{\ell i}(\vec{\theta})$. Thus $B_i(\vec{\theta})\vec{1} = \frac{1-\delta\gamma}{\delta(1-\gamma)}\vec{1} + \vec{\pi}_i(\vec{\theta})$, as desired.

*(b)* The fact that diagonal entries are positive is obvious. Off-diagonal entries on row $\ell$ are of the form $\pi_{i\ell'}(\vec{\theta}) - \pi_{\ell\ell'}(\vec{\theta})$, which is equal to $\theta_{\ell'}(\sigma_{\ell'}(\vec{\theta},i) - \sigma_{\ell'}(\vec{\theta},\ell))$. The result about the sign of off-diagonal entries then follows as the likelihood for a discerning $\ell'$ to be picked diminishes when part of a better pool of discerning agents.

*(c)* By *(b)*, off-diagonal entries on a row $\ell$ are non-positive when $\theta_i \leq \theta_\ell$, in which case $z_\ell$ is simply the sum of the elements appearing on row $\ell$, whose value is given in *(a)*. Suppose now $\theta_i \geq \theta_\ell$. The first computation in the proof of *(a)* shows that the sum of the off-diagonal elements on row $\ell$ (which are all positive, by *(b)*) is equal to $\pi_{\ell i}(\vec{\theta}) - \pi_{i\ell}(\vec{\theta})$. Thus $z_\ell = \frac{1-\delta\gamma}{\delta(1-\gamma)} + \pi_{i\ell} - (\pi_{\ell i}(\vec{\theta}) - \pi_{i\ell}(\vec{\theta}))$, and the result follows.

*(d)* We need to check $z_\ell > 0$ for all $\ell$. Since $\frac{1-\delta\gamma}{\delta(1-\gamma)} > 1$, the result follows from the fact that $\pi_{\ell i} \geq 0$ for the case $\theta_\ell \geq \theta_i$, and from $\pi_{i\ell} \geq 0$ and $\pi_{\ell i} < 1$ for the case $\theta_\ell \leq \theta_i$.

*(e)* This follows from the Ahlberg-Nilson-Varah bound (see e.g. Varah (1975)) since $B_i(\vec{\theta})$ is strictly diagonally dominant.

*(f)* Since $B_i(\vec{\theta})$ is invertible by *(e)*, multiply both sides of *(a)* by $B_i(\vec{\theta})^{-1}$.

*(g)* Notice that the entries of $B_i(\vec{\theta})$ are affine functions of $\theta_i$. Indeed, the matrix $Y_i$ is obtained by taking the derivative with respect to $\theta_i$ of the entries of $B_i(\vec{\theta})$, and

is independent of $\theta_i$. Thus $B_i(\vec{\theta}) = X_i + (\theta_i - \theta^*)Y_i$. The result then follows from the power series expansion of matrix inverses, after showing that $||X_i^{-1}Y_i||_\infty < 1$. To check this, first notice that $||X_i^{-1}||_\infty < 1$ by *(e)* given that $\theta_\ell \geq \theta^*$ for all $\ell \neq i$. Consider $Y_i$ next. It is a positive matrix, so its infinite norm is obtained by computing for each row the sum of its entries, and then taking the maximum of these sums over the rows. Observed that $Y_i$ is the derivative with respect to $\theta_i$ of the matrix $B_i(\vec{\theta})$. Using the computations from *(a)*, the sum of the elements on row $\ell$ of $Y_i$ is simply the derivative with respect to $\theta_i$ of $\pi_{\ell i}(\vec{\theta})$, which is equal to $\sigma_i(\vec{\theta}, \ell)$. This expression is decreasing in $\vec{\theta}$ for each $\ell$, and thus lower or equal to $\sigma^*$, which is less than 1. Then $||X_i^{-1}Y_i||_\infty \leq ||X_i^{-1}||_\infty ||Y_i||_\infty < \sigma^* < 1$, as desired.

*(h)* By *(a)*, the derivative of $B_i(\vec{\theta})^{-1}\vec{\pi}(\vec{\theta})$ with respect to $\theta_i$ is equal to the opposite of the derivative of $B_i(\vec{\theta})^{-1}\vec{1}$, which by *(g)* is equal to $\sum_{k=1}^{\infty}(-1)^{k+1}k(\theta_i - \theta^*)^{k-1}(X_i^{-1}Y_i)^k X_i^{-1}\vec{1}$.

Notice that $2(\theta_i - \theta^*)Y_i X_i^{-1}\vec{1} \leq 2(1-\theta^*)Y_i X_i^{-1}\vec{1} < \vec{1}$. The first inequality follows from the facts that $Y_i$ and $X_i^{-1}$ (inverse of an $M$-matrix) are positive, and $\theta_i \leq 1$. The strict inequality follows from *(c)* in Lemma 3, since each component of the vector $Y_i X_i^{-1}\vec{1}$ is lower of equal to $||Y_i X_i^{-1}||_\infty$, which is strictly less than $\sigma^*$ (see *(g)* above).

Being a product of positive matrices, the matrix $X_i^{-1}Y_i X_i^{-1}$ is positive. Hence we know $X_i^{-1}Y_i X_i^{-1}\vec{1} - 2(\theta_i - \theta^*)(X_i^{-1}Y_i)^2 X_i^{-1}\vec{1}$ is a strictly positive vector. This corresponds to the first two terms in the above expression for the derivative of $B_i(\vec{\theta})^{-1}\vec{\pi}(\vec{\theta})$ with respect to $\theta_i$. A fortiori, $3X_i^{-1}Y_i X_i^{-1}\vec{1} - 4(\theta_i - \theta^*)(X_i^{-1}Y_i)^2 X_i^{-1}\vec{1}$ is a strictly positive vector, and hence $3(\theta_i - \theta^*)^2(X_i^{-1}Y_i)^3 X_i^{-1}\vec{1} - 4(\theta_i - \theta^*)^3(X_i^{-1}Y_i)^4 X_i^{-1}\vec{1}$ is a strictly positive vector as well (since $(\theta_i - \theta^*)^2(X_i^{-1}Y_i)^2$ is a positive matrix). This corresponds to the next two terms in the above expression for the derivative of $B_i(\vec{\theta})^{-1}\vec{\pi}(\vec{\theta})$ with respect to $\theta_i$. Iterating the argument shows this derivative is strictly positive. □

**Lemma 5.** *Discerning agents are always willing to propose themselves when qualified.*

*Proof.* Remember that discerning agents propose themselves when qualified if and only if $M_i^Q(\vec{\theta})B_i(\vec{\theta})^{-1}\vec{\pi}_i(\vec{\theta}) \leq \vec{\sigma}_i(\vec{\theta})$. To establish this inequality, it is sufficient to show that $||B_i(\vec{\theta})^{-1}\vec{\pi}_i(\vec{\theta})||_\infty \leq 1$, since $M_i^Q(\vec{\theta})$ is a positive matrix with the sum of entries on any row $\ell$ equal to $\sigma_i(\vec{\theta}, \ell)$. It is sufficient to establish the upper-bound on the infinite norm under the assumption that $\theta_i = 1$, because of *(h)* in Lemma 4. Using $||B_i(\vec{\theta})^{-1}\vec{\pi}_i(\vec{\theta})||_\infty \leq ||B_i(\vec{\theta})^{-1}||_\infty ||\vec{\pi}_i(\vec{\theta})||_\infty$, combined with *(c)* and *(e)* from

Lemma 4, it is sufficient to check that

$$\pi_{ki}(\vec{\theta}) < \frac{1-\delta\gamma}{\delta(1-\gamma)} - \pi_{\ell i}(\vec{\theta}) + 2\pi_{i\ell}(\vec{\theta}), \tag{25}$$

where $k$ is an agent $j \neq i$ that maximizes $\pi_{ji}(\vec{\theta})$ and $\ell$ is an agent $j \neq i$ that minimizes $2\pi_{ij}(\vec{\theta}) - \pi_{ji}(\vec{\theta})$. Inequality (25) holds when $k = \ell$, since $\pi_{\ell i}(\vec{\theta}) - \pi_{i\ell}(\vec{\theta}) \leq 1/2$ by *(c)* in Lemma 3, and $\frac{1-\delta\gamma}{\delta(1-\gamma)} > 1$. Suppose then $k \neq \ell$. Inequality (25) becomes (as $\theta_i = 1$)

$$\sigma_i(\vec{\theta}, k) - \theta_\ell \sigma_i(\vec{\theta}, \ell) - 2\rho_i(\vec{\theta}) + (1-\theta_\ell)\sigma_i(\vec{\theta}, \ell) < \frac{1-\delta\gamma}{\delta(1-\gamma)}.$$

It is sufficient to check that $\sigma_i(\vec{\theta}, k) - \theta_\ell \sigma_i(\vec{\theta}, \ell) + (1-\theta_\ell)\sigma_i(\vec{\theta}, \ell) \leq 1$. Notice that the expression on the LHS is linear in $\theta_\ell$, and it is thus maximized by taking $\theta_\ell = 1$ or $\theta^*$. The inequality is obvious if $\theta_\ell = 1$, so let's assume that $\theta_\ell = \theta^*$. Thus it is sufficient to prove that $\sigma_i((\theta^*, \vec{\theta}_{-\ell}), k) - \theta^* \sigma_i(\vec{\theta}, \ell) + (1-\theta^*)\sigma_i(\vec{\theta}, \ell) \leq 1$. Remember $\theta^* \leq 1/2$ when $n \geq 2$, so the total weight on $\sigma_i(\vec{\theta}, \ell)$ is positive. The expression on the LHS is thus $\leq (2 - 2\theta^*)\sigma^*$. The desired inequality follows from *(c)* in Lemma 3. □

**Lemma 6.** *Unqualified discerning agents do not propose if* $\delta \geq \frac{1}{\gamma + (\gamma-\beta)\underline{\pi}}$.

*Proof.* Remember that discerning agents do not propose themselves when unqualified if and only if $\vec{\sigma}_i(\vec{\theta}) \leq M_i^U(\vec{\theta}) B_i(\vec{\theta})^{-1} \vec{\pi}_i(\vec{\theta})$. By *(f)* from Lemma 4, this is equivalent to

$$\vec{\sigma}_i(\vec{\theta}) + \frac{1-\delta\gamma}{\delta(1-\gamma)} M_i^U(\vec{\theta}) B_i(\vec{\theta})^{-1} \vec{1} \leq M_i^U(\vec{\theta}) \vec{1} = \vec{\sigma}_i(\vec{\theta}) + \frac{\gamma-\beta}{1-\gamma} \vec{\sigma}_i(\vec{\theta}),$$

or

$$\frac{1-\delta\gamma}{\delta(1-\gamma)} M_i^U(\vec{\theta}) B_i(\vec{\theta})^{-1} \vec{1} \leq \frac{\gamma-\beta}{1-\gamma} \vec{\sigma}_i(\vec{\theta}). \tag{26}$$

The RHS is independent of $\theta_i$, while all the components of the LHS vector are decreasing in $\theta_i$ (by *(h)* from Lemma 4, and since $M_i^U$ is a positive matrix). It is thus sufficient to prove this inequality for $\theta_i = \underline{\theta}$, which we assume from now on. Since $M_i^U$ is positive, the LHS vector is smaller or equal to $\frac{1-\delta\gamma}{\delta(1-\gamma)} ||B_i(\vec{\theta})^{-1}||_\infty M_i^U(\vec{\theta}) \vec{1}$. Using *(e)* from Lemma 4 and the fact that $M_i^U(\vec{\theta}) \vec{1} = \frac{1-\beta}{1-\gamma} \vec{\sigma}_i(\vec{\theta})$, it is sufficient to check that

$$\frac{1-\delta\gamma}{\delta(1-\gamma)} \frac{1}{\frac{1-\delta\gamma}{\delta(1-\gamma)} + \min_{\ell \neq i} \pi_{\ell i}(\vec{\theta})} \frac{1-\beta}{1-\gamma} \leq \frac{\gamma-\beta}{1-\gamma},$$

or $\delta \geq \frac{1}{\gamma + (\gamma-\beta)\min_{\ell \neq i} \pi_{\ell i}(\vec{\theta})}$. Observe $\underline{\pi} \leq \min_{\ell \neq i} \pi_{\ell i}(\vec{\theta})$, for all $\vec{\theta} \in [\underline{\theta}, 1]^n$ s.t. $\theta_i = \underline{\theta}$. □

**Lemma 7.** *If the MLR is a belief-free equilibrium then* $\delta \geq \frac{1}{\gamma+(\gamma-\beta)\underline{\pi}}$ *for all* $i$.

*Proof.* The proof of Lemma 6 shows that condition (26) is necessary and sufficient for discerning agents to refrain from proposing when unqualified. Given $\underline{\theta}$, consider the ability vector $\vec{\theta}$ for which the minimal probability premium $\underline{\pi}$ is achieved. For the MLR to be a belief-free equilibrium, it is necessary that it is an ex-post equilibrium for this $\vec{\theta}$. By Lemma 3*(e)*, this ability vector either has all agent abilities equal to $\underline{\theta}$, or there is some agent $i$ with ability $\underline{\theta}$ and all others have ability 1. In both cases, the value of $\pi_{\ell i}(\vec{\theta})$ is constant in $\ell$. By the characterization in Lemma 4*(a)*, for this $\vec{\theta}$ we have that $B_i(\vec{\theta})\vec{1} = \left(\frac{1-\delta\gamma}{\delta(1-\gamma)} + \underline{\pi}\right)\vec{1}$. If a matrix has constant row sums equal to $s$, then the inverse has constant row sums equal to $1/s$. Thus $B_i^{-1}(\vec{\theta})\vec{1} = \frac{1}{\frac{1-\delta\gamma}{\delta(1-\gamma)}+\underline{\pi}}\vec{1}$. Applying this expression as well as the fact that $M_i^b(\vec{\theta})\vec{1} = \frac{1-\beta}{1-\gamma}\vec{\sigma}_i(\vec{\theta})$ in the necessary condition (26), we immediately obtain the desired condition on $\delta$. □

**Proof of Proposition 6.** 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 $n-1$:

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

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.[21] 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 RHS is strictly positive. In particular, $V^k \geq V^n + \alpha(k)\frac{1-\delta}{\delta(1-\beta)}$, for all $k$, where $\alpha(k)$ is the number of times $j(\cdot)$ must be iterated to reach $n$. We have:

$$1 \geq \sum_{k=1}^{n} V^k \geq nV^n + \sum_{k=1}^{n-1}\alpha(k)\frac{1-\delta}{\delta(1-\beta)}. \tag{27}$$

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

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

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

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 $V^n \geq (1-\theta)^{n-1} + \frac{P(1-\gamma)}{(1-\beta)}$, 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 (27), for the hierarchical strategy profile to be an equilibrium requires:

$$1 \geq n\Big((1-\theta)^{n-1} + \frac{P(1-\gamma)}{(1-\beta)}\Big) + \sum_{k=1}^{n-1} \alpha(k)\frac{(1-\delta)}{\delta(1-\beta)}. \tag{28}$$

On the other hand, MLR forms an equilibrium *if and only if*[22]

$$1 \geq n\Big((1-\theta)^{n-1} + \frac{(1-(1-\theta)^{n-1})(1-\gamma)}{(1-\beta)}\Big) + (n-1)\frac{(1-\delta)}{\delta(1-\beta)}. \tag{29}$$

Consider the necessary condition (28) for the case of hierarchical strategy profiles that send failing agents to the bottom. Here, $P = 1-(1-\theta)^{n-1}$ and $\alpha(k) = 1$ 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 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 $\sum_{k=1}^{n-1} \alpha(k) \geq n$. Thus in this case, (28) implies the following necessary condition for the hierarchy to form an equilibrium:

$$1 \geq n\Big((1-\theta)^{n-1} + \frac{\theta(1-\theta)^{n-2}(1-\gamma)}{(1-\beta)} + \frac{(1-\delta)}{\delta(1-\beta)}\Big).$$

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 $\theta$'s; 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 $\gamma$ 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

[22] One can check directly that the same condition on $\delta$ as in Proposition 5 but with $\underline{\pi}$ 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 $V^D - V^{LR} = \frac{1-\delta}{\delta(1-\beta)}$, in which case we can derive the exact values for $V^{LR}$ and $V^D$, and the equation $V^{LR} + (n-1)V^D = 1$ gives the largest range of parameters.

after him in the ranking. The recursive equations that give the agents' payoffs are:

$$V^1 = \theta(1-\delta) + p_1\delta V^2 + (1-p_1)\delta V^1$$
$$V^2 = (1-\theta)\theta(1-\delta) + p_1\delta V^1 + p_2\delta V^3 + (1-p_1-p_2)\delta V^2$$
$$V^3 = (1-\theta)^2(1-\delta) + p_2\delta V^2 + (1-p_2)\delta V^3,$$

where $p_1 = \theta(1-\gamma)$ 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 $\beta = 0$, $\gamma = 4/5$, $\delta = 5/6$ and $\theta = 1$. The RHS of inequality (29) is $3/5 + 2/5 = 1$. Thus, MLR is a PPE for these parameters, but it ceases to be one for any lower $\theta$. 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 $V^1 - V^2 \geq \frac{1-\delta}{\delta(1-\beta)}$ 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 $\theta$ a bit will only change those values a bit, and the ICs will still hold. ■

## A3 Characterization of principal-renegotiation-proof PPE

**Proof of Proposition 8.** Our first observation is that principal-renegotiation-proof PPEs have a very simple payoff structure.

**Lemma 8.** *Consider a PPE where the principal gets the same (maximum) discounted expected payoff at the start of period, no matter the history. Then the principal gets the same expected equilibrium payoff within each stage game.*

*Proof.* Let $X$ be the principal's discounted payoff at the start of any period. Let $x$ be the principal's expected equilibrium payoff within that period. Then $X = (1-\delta)x + \delta X$, and hence $x = X$, which is independent of the history. □

This observation reduces the type of strategies that the principal employs in equilibrium. For example, it cannot be that there is a history after which the principal switches from not selecting the last agent who generated a low profit, to always selecting that agent, because these can give different expected payoffs to the principal within a stage game. Hence, characterizing the principal-renegotiation-proof PPEs reduces to finding which stage game behaviors lead to the same outcome and whether those stage games in the same 'equivalence class' can be sequenced in a way that forms a repeated-game equilibrium. It turns out that any PPE that satisfies our refinement gives the principal either his first-best payoff or his one-shot Nash payoff.

Note that each agent has four strategies in the stage game: propose regardless of qualification, don't propose regardless of qualification, propose only when qualified, and propose only when unqualified. There are thus sixteen combinations to consider for the agents. As for the principal, renegotiation-proofness implies that she gets the same discounted payoff at the beginning of any new round in the game, independently of what happened in the past. Hence it must be that she selects agents optimally in each repetition of the stage game taken individually (e.g. picking the discerning agent instead of the last resort in the MLR strategy profile when both agents make proposals). Otherwise, she has a profitable unilateral deviation by picking the one that has a higher likelihood (given equilibrium reporting strategies) of being qualified.

We already analyzed the following cases: (*i*) both agent propose regardless of their qualification, (*ii*) the most able agent proposes regardless of his qualification while the other agent does not propose regardless of his qualification, and (*iii*) one agent proposes only when qualified and the remaining agent proposes regardless of his qualification (note that these are two cases since the identity of the constant proposer can change). Cases (*i*) and (*ii*) correspond to the one-shot Nash equilibrium outcome, while case (*iii*) corresponds to the MLR strategy profile.

A fourth possible case is when every period both agents propose only when they are qualified. This generates the payoff $\big(1-(1-\theta_1)(1-\theta_2)\big)\big(\gamma H+(1-\gamma)L\big)$ to the principal, which is higher than the one-shot Nash payoff if $\min\{\theta_1,\theta_2\} > \big(\beta H+(1-\beta)L\big)/\big(\gamma H+(1-\gamma)L\big)$. We establish the following observation:

**Lemma 9.** *Suppose there exists a PPE in which the agents propose if and only if they are qualified and the principal picks one of the proposing agents. Then the MLR strategy profile is also a PPE.*

**Proof of Lemma 9.** We follow the same methodology as in the proof of Proposition 7. Using the same notation as in that proof, we let $\sigma_i(\emptyset)$ denote agent $i$'s promised continuation payoff when no agent is selected.

**Step 1. Deriving $\underline{\sigma}_1$.** If $\underline{\sigma}_1$ is obtained when agent 1 is LR, then to find $\underline{\sigma}_1$ minimize

$$(1-\theta_2)\theta_1\left[(1-\delta)+\delta\left(\gamma\sigma_1(1S)+(1-\gamma)\sigma_1(1F)\right)\right]+$$
$$(1-\theta_2)(1-\theta_1)\delta\sigma_1(\emptyset)+\theta_2\delta\left[\gamma\sigma_1(2S)+(1-\gamma)\sigma_1(2F)\right]$$

subject to the IC constraints that *both* agents do not propose when unqualified:[23]

$$\begin{aligned}&\delta\left[\theta_1(\gamma\sigma_2(1S)+(1-\gamma)\sigma_2(1F))+(1-\theta_1)\sigma_2(\emptyset)\right]\\&\geq(1-\delta)+\delta\left(\beta\sigma_2(2S)+(1-\beta)\sigma_2(2F)\right),\end{aligned}$$

for agent 2, and for agent 1: $\delta\sigma_1(\emptyset)\geq(1-\delta)+\delta\left[\beta\sigma_1(1S)+(1-\beta)\sigma_1(1F)\right]$. As the sum of continuation payoffs is always $1-(1-\theta_1)(1-\theta_2)$, we rewrite agent 2's IC as

$$\delta\beta\sigma_1(2S)+\delta(1-\beta)\sigma_1(2F)\geq(1-\delta)+\delta\theta_1\gamma\sigma_1(1S)+\delta(1-\gamma)\theta_1\sigma_1(1F)+\delta(1-\theta_1)\sigma_1(\emptyset).$$

Hence, we can decrease $\sigma_1(1S),\sigma_1(1F)$ all the way to $\underline{\sigma}_1$ (reduces the continuation payoff and can only relax the IC). We then have the following problem:

$$\min\ (1-\theta_2)\theta_1\left[(1-\delta)+\delta\underline{\sigma}_1\right]+(1-\theta_2)(1-\theta_1)\delta\sigma_1(\emptyset)+\theta_2\delta\left[\gamma\sigma_1(2S)+(1-\gamma)\sigma_1(2F)\right]$$

s.t. the feasibility constraint that continuation payoffs lie in $[\underline{\sigma}_1,\overline{\sigma}_1]$, and the IC's

$$\delta\left(\beta\sigma_1(2S)+(1-\beta)\sigma_1(2F)\right)\ \ \geq\ \ (1-\delta)+\delta\theta_1\underline{\sigma}_1+\delta(1-\theta_1)\sigma_1(\emptyset)$$

and $\delta\sigma_1(\emptyset)\geq(1-\delta)+\delta\underline{\sigma}_1$. Substituting the latter IC (which must clearly bind) into the former, and also into the objective function, we wish to minimize

$$(1-\theta_2)\left[(1-\delta)+\delta\underline{\sigma}_1\right]+\theta_2\delta\left[\gamma\sigma_1(2S)+(1-\gamma)\sigma_1(2F)\right]$$

subject to feasibility and $\sigma_1(2F)=\frac{1-\delta}{\delta(1-\beta)}(2-\theta_1)+\frac{\underline{\sigma}_1}{1-\beta}-\frac{\beta}{1-\beta}\sigma_1(2S)$. Plugging this back into the objective function we obtain that the coefficient on $\sigma_1(2S)$ is $\frac{\gamma-\beta}{1-\beta}>0$. We therefore wish to reduce $\sigma_1(2S)$ as much as possible, noting that a decrease in $\sigma_1(2S)$ yields an increase in $\sigma_1(2F)$. There are therefore two cases to consider:

*Case 1.* $\sigma_1(2S)=\underline{\sigma}_1$ *and* $\sigma_1(2F)=\frac{1-\delta}{\delta(1-\beta)}(2-\theta_1)+\underline{\sigma}_1\leq\overline{\sigma}_1$ . In this case, $\overline{\sigma}_1-\underline{\sigma}_1\geq\frac{1-\delta}{\delta(1-\beta)}(2-\theta_1)$ must hold. Setting $\underline{\sigma}_1$ equal to the objective in the minimization, we obtain $\underline{\sigma}_1=(1-\theta_2)+\theta_2\frac{1-\gamma}{1-\beta}(2-\theta_1)$. The necessary condition for Case 1 is therefore:

[23] As in the symmetric case, we ignore the remaining constraints, which will turn out to be without loss.

$$(1-\theta_2)+\theta_2\frac{1-\gamma}{1-\beta}(2-\theta_1)+\frac{1-\delta}{\delta(1-\beta)}(2-\theta_1)\leq\overline{\sigma}_1.$$

To check when it is satisfied, we later examine maximizing 1's continuation payoff.

*Case 2.* $\sigma_1(2F)=\overline{\sigma}_1$ *and* $\sigma_1(2S)=\frac{1-\delta}{\delta\beta}(2-\theta_1)+\frac{\underline{\sigma}_1}{\beta}-\frac{1-\beta}{\beta}\overline{\sigma}_1\in[\underline{\sigma}_1,\overline{\sigma}_1]$. Setting $\underline{\sigma}_1$ equal to the objective in the minimization problem, we obtain that

$$\underline{\sigma}_1=\frac{(1-\delta)\left[(1-\theta_2)+\theta_2\frac{\gamma}{\beta}(2-\theta_1)\right]-\delta\theta_2\overline{\sigma}_1\left[\frac{\gamma-\beta}{\beta}\right]}{1-\delta\left[(1-\theta_2)+\theta_2\frac{\gamma}{\beta}\right]}. \tag{30}$$

**Step 2. Deriving $\overline{\sigma}_1$.** We now maximize 1's continuation payoff. Suppose first that this occurs when 1 is discerning. We maximize

$$\begin{aligned}&\theta_1\left[(1-\delta)+\delta\left(\gamma\sigma_1(1S)+(1-\gamma)\sigma_1(1F)\right)\right]\\&+(1-\theta_1)\delta\left[\theta_2(\gamma\sigma_1(2S)+(1-\gamma)\sigma_1(2F))+(1-\theta_2)\sigma_1(\emptyset)\right]\end{aligned}$$

subject to the IC that neither agent wants to propose when unqualified:

$$\begin{aligned}&\delta\left[\theta_2(\gamma\sigma_1(2S)+(1-\gamma)\sigma_1(2F))+(1-\theta_2)\sigma_1(\emptyset)\right]\geq(1-\delta)+\delta\left(\beta\sigma_1(1S)+(1-\beta)\sigma_1(1F)\right)\\&\delta\sigma_2(\emptyset)\geq(1-\delta)+\delta\left(\beta\sigma_2(2S)+(1-\beta)\sigma_2(2F)\right)\end{aligned}$$

and $\sigma_1\in[\underline{\sigma}_1,\overline{\sigma}_1]$. As continuation payoffs following each event sum to $1-(1-\theta_1)(1-\theta_2)$, we rewrite agent 2's IC as $\delta\left(\beta\sigma_1(2S)+(1-\beta)\sigma_1(2F)\right)\geq(1-\delta)+\delta\sigma_1(\emptyset)$. Setting $\sigma_1(2S),\sigma_1(2F)=\overline{\sigma}_1$ (increases objective and only relaxes IC), we wish to maximize

$$\theta_1\left[(1-\delta)+\delta\left(\gamma\sigma_1(1S)+(1-\gamma)\sigma_1(1F)\right)\right]+(1-\theta_1)\theta_2\delta\overline{\sigma}_1+(1-\theta_1)(1-\theta_2)\delta\sigma_1(\emptyset)$$

subject to feasibility,

$$\delta\theta_2\overline{\sigma}_1+(1-\theta_2)\delta\sigma_1(\emptyset)\quad\geq\quad(1-\delta)+\delta\left(\beta\sigma_1(1S)+(1-\beta)\sigma_1(1F)\right),$$

and $\delta\overline{\sigma}_1\geq(1-\delta)+\delta\sigma_1(\emptyset)$. Since the latter must bind, plugging into the other the first IC, we obtain $\delta\overline{\sigma}_1\geq(1-\delta)(2-\theta_2)+\delta\left(\beta\sigma_1(1S)+(1-\beta)\sigma_1(1F)\right)$, which clearly must bind. Therefore the objective of maximization becomes:

$$(1-\delta)\left(\theta_1-(1-\theta_1)(1-\theta_2)\right)+\theta_1\delta\left(\gamma\sigma_1(1S)+(1-\gamma)\sigma_1(1F)\right)+(1-\theta_1)\delta\overline{\sigma}_1.$$

To solve this problem, we must increase $\sigma_1(1S)$ as much as possible (intuitively, increase agent 1's payoff when he is discerning and succeeds), and have two cases:

*Case 3.* $\sigma_1(1S) = \overline{\sigma}_1$ *and* $\sigma_1(1F) = \overline{\sigma}_1 - \frac{(1-\delta)}{\delta(1-\beta)}(2-\theta_2) \geq \underline{\sigma}_1$. Setting $\overline{\sigma}_1$ equal to the objective, $\overline{\sigma}_1 = \theta_1(2-\theta_2)\frac{\gamma-\beta}{1-\beta} - (1-\theta_2)$. So the necessary condition is

$$\theta_1(2-\theta_2)\frac{\gamma-\beta}{1-\beta} - (1-\theta_2) - \frac{(1-\delta)}{\delta(1-\beta)}(2-\theta_2) \geq \underline{\sigma}_1.$$

*Case 4.* $\sigma_1(1F) = \underline{\sigma}_1$ *and* $\sigma_1(1S) = \frac{\overline{\sigma}_1}{\beta} - \frac{(1-\beta)\underline{\sigma}_1}{\beta} - \frac{1-\delta}{\delta\beta}(2-\theta_2) \in [\underline{\sigma}_1, \overline{\sigma}_1]$. Plugging into the objective yields,

$$\overline{\sigma}_1 = \frac{(1-\delta)\left(\theta_1 - (1-\theta_1)(1-\theta_2) - \theta_1(2-\theta_2)\frac{\gamma}{\beta}\right) - \delta\underline{\sigma}_1\theta_1\left(\frac{\gamma}{\beta}-1\right)}{1-\delta\left(\theta_1\frac{\gamma}{\beta} + (1-\theta_1)\right)}. \tag{31}$$

Since $\theta_1 - (1-\theta_1)(1-\theta_2) - \theta_1(2-\theta_2)\frac{\gamma}{\beta} < 0$, it must be that $1-\delta\left(\theta_1\frac{\gamma}{\beta} + (1-\theta_1)\right) < 0$.

**Step 3. Combining cases 1 and 3.** From Case 3, $\overline{\sigma}_1 = (\theta_1 - (1-\theta_1)(1-\theta_2)) - \theta_1(2-\theta_2)\frac{1-\gamma}{1-\beta}$ and $\overline{\sigma}_1 - \underline{\sigma}_1 \geq \frac{1-\delta}{\delta(1-\beta)}(2-\theta_2)$, and from Case 1 we have $\underline{\sigma}_1 = (1-\theta_2) + \theta_2\frac{1-\gamma}{1-\beta}(2-\theta_1)$ and $\overline{\sigma}_1 - \underline{\sigma}_1 \geq \frac{1-\delta}{\delta(1-\beta)}(2-\theta_1)$. Therefore, we have

$$\overline{\sigma}_1 - \underline{\sigma}_1 = 2\left(\theta_1 + \theta_2 - \theta_1\theta_2\right)\left(\frac{\gamma-\beta}{1-\beta}\right) - 2 + \theta_1\theta_2.$$

And the combined necessary condition for the two cases is

$$\delta \;\geq\; \max_{i\in\{1,2\}}\left\{\frac{1}{\frac{g(\theta_1,\theta_2,\gamma,\beta)}{2-\theta_i}+1}\right\}, \tag{32}$$

where $g(\theta_1,\theta_2,\gamma,\beta) = 2\left(\theta_1+\theta_2-\theta_1\theta_2\right)(\gamma-\beta) - (1-\beta)(2-\theta_1\theta_2)$. Note that the effective constraint is the one with the smaller $\theta_1$.

First assume $\theta_1 \leq \theta_2$. We want to verify the necessary constraint for the candidate equilibrium (proposing only when qualified) is more restrictive than that for MLR:

$$\frac{2-\theta_1}{g(\theta_1,\theta_2,\gamma,\beta)+2-\theta_1} > \frac{1}{\beta + (\theta_1+\theta_2)(\gamma-\beta)}.$$

This inequality holds if and only if $(\theta_2-\theta_1)(\gamma-\beta) > -(1-\beta)(1-\theta_2)$, which holds since $\gamma > \beta$. The analogous argument holds for $\theta_1 > \theta_2$. Hence the necessary conditions of cases 1 and 3 are more stringent than the condition assuring MLR is a PPE.

**Step 4. Combining cases 2 and 4.** From Case 2 we have (30) and from Case 4 we have (31). Combining the two yields:

$$\overline{\sigma}_1 - \underline{\sigma}_1 = \frac{(1-\delta)\left[2\left(\theta_1+\theta_2-\theta_1\theta_2\right)\left(\frac{\gamma}{\beta}-1\right)+2-\theta_1\theta_2\right]}{\delta(\theta_1+\theta_2)\left(\frac{\gamma}{\beta}-1\right)-(1-\delta)},$$

and the necessary conditions for these two cases reduce to

$$\max_{i\in\{1,2\}}\left\{\frac{(1-\delta)(2-\theta_i)}{\delta}\right\} \leq \overline{\sigma}_1 - \overline{\sigma}_1 \leq \min_{i\in\{1,2\}}\left\{\frac{(1-\delta)(2-\theta_i)}{\delta(1-\beta)}\right\}.$$

Suppose first that $\theta_2 \geq \theta_1$. Then it suffices to check the upper bound $\frac{(1-\delta)(2-\theta_2)}{\delta(1-\beta)}$ and the lower bound $\frac{(1-\delta)(2-\theta_1)}{\delta}$. Starting with the upper bound:

$$\overline{\sigma}_1 - \underline{\sigma}_1 = \frac{(1-\delta)\left[2\left(\theta_1+\theta_2-\theta_1\theta_2\right)\left(\frac{\gamma}{\beta}-1\right)+2-\theta_1\theta_2\right]}{\delta(\theta_1+\theta_2)\left(\frac{\gamma}{\beta}-1\right)-(1-\delta)} \leq \frac{(1-\delta)(2-\theta_2)}{\delta(1-\beta)}$$

which can be rewritten as

$$\delta \geq \frac{2-\theta_2}{\left(\frac{\gamma}{\beta}-1\right)\left[\left(2-\theta_2\right)\left(\theta_1+\theta_2\right)-(1-\beta)2\left(\theta_1+\theta_2-\theta_1\theta_2\right)\right]+\left(2-\theta_2\right)-\left(1-\beta\right)\left(2-\theta_1\theta_2\right)}.$$

We want to show that that this constraint is more restrictive than the one for MLR. That is, that the RHS of the last inequality is greater than $\frac{1}{\beta+(\theta_1+\theta_2)(\gamma-\beta)}$. After some algebra, it can be shown that this is equivalent to $(\theta_2-\theta_1)\frac{\gamma}{\beta}\left(\beta-1\right) < (1-\theta_2)(1-\beta)$, which is clearly satisfied since the LHS is negative. So for $\theta_2 \geq \theta_1$ it must be that the combination of cases 2 and 4 hold only under conditions more restrictive than the equilibrium condition for MLR; equivalently, the condition for the existence of a first-best equilibrium (there is no need to check the lower bound).

Next suppose $\theta_2 < \theta_1$. Then it suffices to check the upper bound $\frac{(1-\delta)(2-\theta_1)}{\delta(1-\beta)}$ and the lower bound $\frac{(1-\delta)(2-\theta_2)}{\delta}$. As before, we start with the upper bound:

$$\overline{\sigma}_1 - \underline{\sigma}_1 = \frac{(1-\delta)\left[2\left(\theta_1+\theta_2-\theta_1\theta_2\right)\left(\frac{\gamma}{\beta}-1\right)+2-\theta_1\theta_2\right]}{\delta(\theta_1+\theta_2)\left(\frac{\gamma}{\beta}-1\right)-(1-\delta)} \leq \frac{(1-\delta)(2-\theta_1)}{\delta(1-\beta)},$$

or equivalently

$$\delta \geq \frac{2-\theta_1}{(2-\theta_1)(\theta_1+\theta_2)\left(\frac{\gamma}{\beta}-1\right)+(2-\theta_1)-(1-\beta)\left[2\left(\theta_1+\theta_2-\theta_1\theta_2\right)\left(\frac{\gamma}{\beta}-1\right)+2-\theta_1\theta_2\right]}. \tag{33}$$

We therefore want to show that the RHS of this last inequality is greater than $\frac{1}{\beta+(\theta_1+\theta_2)(\gamma-\beta)}$. This is equivalent to the inequality $(1-\beta)(1-\theta_1) > \frac{\gamma}{\beta}\left(\theta_1-\theta_2\right)\left(\beta-1\right)$, which holds since the RHS is negative. It follows that the conditions for cases 2 and 4 are more stringent than the condition for attaining the first-best in PPE.

**Step 5. Combining cases 1 and 4.** From Case 1 we have $\underline{\sigma}_1 = (1-\theta_2) + \theta_2\frac{1-\gamma}{1-\beta}(2-\theta_1)$, and $\overline{\sigma}_1 - \underline{\sigma}_1 \geq \frac{1-\delta}{\delta(1-\beta)}(2-\theta_1)$, and from Case 4 we have

$$\overline{\sigma}_1 = \frac{(1-\delta)\left((1-\theta_1)(1-\theta_2)-\theta_1+\theta_1(2-\theta_2)\frac{\gamma}{\beta}\right)+\delta\underline{\sigma}_1\theta_1\left(\frac{\gamma}{\beta}-1\right)}{\delta\left(\theta_1\frac{\gamma}{\beta}+(1-\theta_1)\right)-1} \tag{34}$$

and $\overline{\sigma}_1 - \underline{\sigma}_1 \in \left[\frac{(1-\delta)(2-\theta_2)}{\delta}, \frac{(1-\delta)(2-\theta_2)}{\delta(1-\beta)}\right]$. Combining these, a necessary condition is:

$$\frac{(2-\theta_1)[1-\theta_2(\frac{\gamma-\beta}{1-\beta})]-\theta_1+\theta_1(2-\theta_2)\frac{\gamma}{\beta}}{\delta\left(\theta_1\frac{\gamma}{\beta}+(1-\theta_1)\right)-1} \leq \frac{(2-\theta_2)}{\delta(1-\beta)}.$$

An implicit requirement for Case 4 is $\delta\left(\theta_1\frac{\gamma}{\beta}+(1-\theta_1)\right)-1>0$, since the numerator in the expression (34) is positive and hence the denominator must also be positive to guarantee $\overline{\sigma}_1 > 0$. Therefore, rearranging the necessary condition above yields:

$$\delta \geq \frac{2-\theta_2}{\theta_1(2-\theta_2)\gamma+(1-\theta_1)(2-\theta_2)+\theta_2(2-\theta_1)(\gamma-\beta)-2(1-\beta)(1-\theta_1)}. \tag{35}$$

We want to show the RHS of (35) is greater than $\frac{1}{\beta+(\theta_1+\theta_2)(\gamma-\beta)}$. Using algebra, the last inequality is equivalent to $(1-\beta)-\theta_1(1-\gamma)-(\gamma-\beta)\theta_2 > 0$, which clearly holds.

**Step 6. Combining cases 2 and 3.** From Case 3 we have

$$\overline{\sigma}_1 = (\theta_1-(1-\theta_1)(1-\theta_2))-\theta_1(2-\theta_2)\frac{1-\gamma}{1-\beta}$$

and $\overline{\sigma}_1 - \underline{\sigma}_1 \geq \frac{1-\delta}{\delta(1-\beta)}(2-\theta_2)$, and from Case 2 we have (30) and $\overline{\sigma}_1 - \underline{\sigma}_1 \in \left[\frac{(1-\delta)(2-\theta_1)}{\delta}, \frac{(1-\delta)(2-\theta_1)}{\delta(1-\beta)}\right]$.

Solving for $\overline{\sigma}_1 - \underline{\sigma}_1$, we get

$$\overline{\sigma}_1 - \underline{\sigma}_1 = (1-\delta)\frac{2(1-\theta_2) + \frac{\gamma}{\beta}\theta_2(2-\theta_1) - \theta_1(2-\theta_2)\frac{\gamma-\beta}{1-\beta}}{\delta\left[(1-\theta_2) + \theta_2\frac{\gamma}{\beta}\right] - 1}.$$

One necessary condition is therefore:

$$\frac{2(1-\theta_2) + \frac{\gamma}{\beta}\theta_2(2-\theta_1) - \theta_1(2-\theta_2)\frac{\gamma-\beta}{1-\beta}}{\delta\left[(1-\theta_2) + \theta_2\frac{\gamma}{\beta}\right] - 1} \leq \frac{(2-\theta_1)}{\delta(1-\beta)}.$$

Suppose first that $\delta\left[(1-\theta_2) + \theta_2\frac{\gamma}{\beta}\right] - 1 > 0$, i.e., $\delta > \frac{1}{(1-\theta_2)+\theta_2\frac{\gamma}{\beta}}$. Then after some algebra we obtain that the necessary condition can be rewritten as

$$2-\theta_1 \leq \delta\Big((2-\theta_1)(1-\theta_2) + (2-\theta_1)\theta_2\frac{\gamma}{\beta} + \theta_1(2-\theta_2)(\gamma-\beta) - (1-\beta)2(1-\theta_2) - \frac{\gamma}{\beta}\theta_2(2-\theta_1) + \gamma\theta_2(2-\theta_1)\Big).$$

Note that if the RHS is negative, we are done, since a necessary condition for cases 2 and 3 cannot be satisfied. We can therefore divide to obtain

$$\delta \geq \frac{2-\theta_1}{(2-\theta_1)(1-\theta_2) + \theta_1(2-\theta_2)(\gamma-\beta) - (1-\beta)2(1-\theta_2) + \gamma\theta_2(2-\theta_1)}. \quad (36)$$

We want to show that the RHS of (36) is greater than $\frac{1}{\beta+(\theta_1+\theta_2)(\gamma-\beta)}$. This inequality reduces to $(1-\beta) - (\gamma-\beta)\theta_1 - \theta_2(1-\gamma) > 0$, which holds since $(1-\beta) - (\gamma-\beta)\theta_1 - \theta_2(1-\gamma) > (1-\beta) - (\gamma-\beta) - (1-\gamma) = 0$. It remains to consider the case $\delta\left[(1-\theta_2) + \theta_2\frac{\gamma}{\beta}\right] - 1 < 0$, i.e., $\delta < \frac{1}{(1-\theta_2)+\theta_2\frac{\gamma}{\beta}}$. Recall that another necessary condition for cases 2 and 3 is that

$$\overline{\sigma}_1 - \underline{\sigma}_1 = \frac{2(1-\theta_2) + \frac{\gamma}{\beta}\theta_2(2-\theta_1) - \theta_1(2-\theta_2)\frac{\gamma-\beta}{1-\beta}}{\delta\left[(1-\theta_2) + \theta_2\frac{\gamma}{\beta}\right] - 1} \geq \frac{(2-\theta_1)}{\delta}.$$

Rearranging, since $\delta\left[(1-\theta_2) + \theta_2\frac{\gamma}{\beta}\right] - 1 < 0$, we get

$$2-\theta_1 \leq \delta\theta_1\left((2-\theta_2)\frac{\gamma-\beta}{1-\beta} - (1-\theta_2)\right).$$

If $(2-\theta_2)\frac{\gamma-\beta}{1-\beta}-(1-\theta_2)<0$, we are done. If $(2-\theta_2)\frac{\gamma-\beta}{1-\beta}-(1-\theta_2)>0$, we get

$$\frac{2-\theta_1}{\theta_1\left[(2-\theta_2)\frac{\gamma-\beta}{1-\beta}-(1-\theta_2)\right]}\leq\delta. \tag{37}$$

We therefore get a contradiction if we show that

$$\frac{2-\theta_1}{\theta_1\left[(2-\theta_2)\frac{\gamma-\beta}{1-\beta}-(1-\theta_2)\right]}>\frac{1}{(1-\theta_2)+\theta_2\frac{\gamma}{\beta}} \tag{38}$$

since this, together with (37), contradicts $\delta<\frac{1}{(1-\theta_2)+\theta_2\frac{\gamma}{\beta}}$. Indeed, (38) simplifies to

$$(2-\theta_2)\left(1-\theta_1\frac{\gamma-\beta}{1-\beta}\right)+\theta_2\left((2-\theta_1)\frac{\gamma}{\beta}-1\right)>0,$$

which holds since $1-\theta_1\frac{\gamma-\beta}{1-\beta}>0$ and $(2-\theta_1)\frac{\gamma}{\beta}-1>0$.

**Step 7. Verifying the postulated configuration of roles**.

**Claim 1.** *$\underline{\sigma}_1$ is attained when agent 1 is last-resort.*

**Proof.** Assume, by contradiction, that $\underline{\sigma}_1$ is attained when agent 1 is discerning,

$$\begin{aligned}\underline{\sigma}_1\geq\min\theta_1[(1-\delta)+\delta\left(\gamma\sigma_1^D(1S)+(1-\gamma)\sigma_1^D(1F)\right)\\ +(1-\theta_1)\delta\left(\theta_2\left(\gamma\sigma_1^D(2S)+(1-\gamma)\sigma_1^D(2F)\right)+(1-\theta_2)\sigma_1^D(\emptyset)\right).\end{aligned}$$

The IC constraint of agent 1 for not proposing when unqualified is:

$$\begin{aligned}&\delta\left(\theta_2\left(\gamma\sigma_1^D(2S)+(1-\gamma)\sigma_1^D(2F)\right)+(1-\theta_2)\sigma_1^D(\emptyset)\right)\\ &\geq(1-\delta)+\delta\left(\beta\sigma_1^D(1S)+(1-\beta)\sigma_1^D(1F)\right).\end{aligned}$$

Hence, we must have:

$$\begin{aligned}&\theta_1\left((1-\delta)+\delta\left(\gamma\sigma_1^D(1S)+(1-\gamma)\sigma_1^D(1F)\right)\right)\\ &+(1-\theta_1)\delta\left(\theta_2\left(\gamma\sigma_1^D(2S)+(1-\gamma)\sigma_1^D(2F)\right)+(1-\theta_2)\sigma_1^D(\emptyset)\right)\\ &\geq\theta_1\left((1-\delta)+\delta\left(\gamma\sigma_1^D(1S)+(1-\gamma)\sigma_1^D(1F)\right)\right)\\ &+(1-\theta_1)\left((1-\delta)+\delta\left(\beta\sigma_1^D(1S)+(1-\beta)\sigma_1^D(1F)\right)\right)\geq(1-\delta)+\delta\underline{\sigma}_1.\end{aligned}$$

But this implies that $\underline{\sigma}_1\geq(1-\delta)+\delta\underline{\sigma}_1$ or that $\underline{\sigma}_1\geq 1$, a contradiction.

**Claim 2.** *$\overline{\sigma}_1$ is attained when agent 1 is discerning.*

**Proof.** By contradiction, assume $\overline{\sigma}_1$ is attained when agent 1 is last-resort. Then

$$\begin{aligned}\overline{\sigma}_1 \leq \max(1-\theta_2)\left(\theta_1\left((1-\delta)+\delta\left(\gamma\sigma_1^{LR}(1S)+(1-\gamma)\sigma_1^{LR}(1F)\right)\right)+\delta(1-\theta_1)\sigma_1^{LR}(\emptyset)\right)\\ +\delta\theta_2\left(\gamma\sigma_1^{LR}(2S)+(1-\gamma)\sigma_1^{LR}(2F)\right)\end{aligned}$$

The IC constraint of the discerning agent 2 for not proposing when unqualified is:

$$(1-\delta)+\delta\left(\beta\sigma_2^{D}(2S)+(1-\beta)\sigma_2^{D}(2F)\right)\leq\delta\theta_1\left(\gamma\sigma_2^{D}(1S)+(1-\gamma)\sigma_2^{D}(1F)\right)+\delta(1-\theta_1)\sigma_2^{D}(\emptyset).$$

Since $\sigma_1^{LR}(x)+\sigma_2^{D}(x)=1$ for $x\in\{1S,1F,2S,2F\}$, we can rewrite the constraint:

$$(1-\delta)+\delta\theta_1\left(\gamma\sigma_1^{LR}(1S)+(1-\gamma)\sigma_1^{LR}(1F)\right)+\delta(1-\theta_1)\sigma_1^{LR}(\emptyset)\leq\delta\left(\beta\sigma_1^{LR}(2S)+(1-\beta)\sigma_1^{LR}(2F)\right).$$

Therefore:

$$\begin{aligned}&(1-\theta_2)\left(\theta_1\left((1-\delta)+\gamma\delta\sigma_1^{LR}(1S)+(1-\gamma)\delta\sigma_1^{LR}(1F)\right)+(1-\theta_1)\delta\sigma_1^{LR}(\emptyset)\right)\\ &\quad+\theta_2\delta\left(\gamma\sigma_1^{LR}(2S)+(1-\gamma)\sigma_1^{LR}(2F)\right)\\ &\leq-(1-\theta_2)(1-\theta_1)(1-\delta)+(1-\theta_2)\delta\left(\beta\sigma_1^{LR}(2S)+(1-\beta)\sigma_1^{LR}(2F)\right)\\ &\quad+\theta_2\delta\left(\gamma\sigma_1^{LR}(2S)+(1-\gamma)\sigma_1^{LR}(2F)\right)\\ &\leq\delta\overline{\sigma}_1.\end{aligned}$$

But this implies that $\overline{\sigma}_1\leq\delta\overline{\sigma}_1$ or $1\leq\delta$. □

This implies that if the principal's first-best cannot be attained in a PPE, then there cannot be a PPE where the agents propose if and only if they are qualified. It is straightforward to verify that none of the remaining cases lead to an expected stage-game payoff for the principal that is higher than that of the one-shot Nash. ■

## A4 Losses

**Proof of Proposition 9.** Recall the necessary conditions derived for the four cases in the proof of Lemma 9 (inequalities (32)-(33)). When $\theta_1=\theta_2=\theta$ the lower bound for cases (1+3), (2+3) and (1+4) are exactly the same, and lower than that of (2+4). Hence, a necessary condition for this case is:

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

This requires $1 - \frac{1-\gamma}{1-\beta} > \frac{2-\theta^2}{2\theta(2-\theta)}$, which is impossible given $\frac{2-\theta^2}{2\theta(2-\theta)} > 1$. ■

## A5 General profit distributions

Recall the definitions of $Q$, $U$, $\underline{y}$, $\bar{y}$, $\gamma^*$, $\beta^*$, and the adjusted-MLR strategy from the main text. We now examine how the punishment set $Y$ should be chosen to sustain the equilibrium, when possible. Consider, for instance, the model with uncertain abilities in $[\underline{\theta}, 1]^2$ that we characterized in Proposition 2. As seen from that result, the first-best is achievable in a belief-free equilibrium if and only if for all agents $i$,

$$\delta \geq \frac{1}{\beta^* + 2\underline{\theta}(\gamma^* - \beta^*)} = \frac{1}{1 - P_U(Y) + 2\underline{\theta}(P_U(Y) - P_Q(Y))}. \tag{39}$$

First, it is clear from (39) that the punishment set $Y$ must be more likely for an unqualified agent than a qualified one (i.e., $\gamma^* > \beta^*$). Intuitively, the ICs would be impossible to satisfy if this were not the case. Moreover, the punishment set must be chosen so that the denominator in (39) is strictly larger than one, or equivalently:

$$2\underline{\theta} > \frac{P_U(Y)}{P_U(Y) - P_Q(Y)} = \frac{P_U(Y)/P_Q(Y)}{P_U(Y)/P_Q(Y) - 1}. \tag{40}$$

The smallest $\underline{\theta}$ for which this is possible is obtained by picking $Y$ to maximize the likelihood ratio $P_U(Y)/P_Q(Y)$ that the punishment set comes from an unqualified agent versus a qualified one. If there is a profit $y$ in the support of $B$ but not $G$, then this ratio is made arbitrarily large by setting $Y = [y - \varepsilon, y + \varepsilon]$ for small enough $\varepsilon$.

What happens when unqualified agents can't be identified with certainty (i.e., the support of $U$ is contained in the support of $Q$)? Suppose, for instance, that $U$ and $Q$ have continuous densities $u$ and $q$ satisfying the monotone likelihood ratio property, with $u(y)/q(y)$ decreasing in $y$.[24] Assuming $\underline{y} \in \text{supp } u$, the maximum of $P_U(Y)/P_Q(Y)$ can be shown to be $\lim_{y \to \underline{y}} u(y)/q(y)$.[25] Hence a belief-free equilibrium achieves first-best if $2\underline{\theta} > \frac{\lim_{y \to \underline{y}} u(y)/q(y)}{\lim_{y \to \underline{y}} u(y)/q(y) - 1}$. If the likelihood ratio goes to infinity as $y$

[24]The case of probability mass functions $u, q$ satisfying MLRP is similar.

[25]For $Y = [\underline{y}, y]$, we have $\lim_{y \to \underline{y}} P_U(Y)/P_Q(Y) = \lim_{y \to \underline{y}} U(y)/Q(y) = \lim_{y \to \underline{y}} u(y)/Q(y)$ by l'Hôpital's rule. For any other $Y$ with positive measure under $U$ (and thus $Q$, by assumption),

$$\frac{P_U(Y)}{P_Q(Y)} = \frac{\int_{y \in Y} u(y) dy}{\int_{y \in Y} q(y) dy} = \frac{\int_{y \in Y} \frac{u(y)}{q(y)} q(y) dy}{\int_{y \in Y} q(y) dy} \leq \lim_{y \to \underline{y}} \frac{u(y)}{q(y)}.$$

decreases to $\underline{y}$, then for any $\underline{\theta} > 1/2$, one can find $y^*$ low enough to ensure first-best can be achieved in a belief-free way with $Y = [\underline{y}, y^*]$ for sufficiently patient agents.

We may want to select the punishment set so that first-best is achievable for the largest range of $\delta$'s. Given (39), choose $Y$ to maximize the objective:

$$-P_U(Y) + 2\underline{\theta}[P_U(Y) - P_Q(Y)] = \int_{y \in Y} \Big((2\underline{\theta} - 1)u(y) - 2\underline{\theta}q(y)\Big)dy.$$

To that end, a profit level $y$ should be included in the punishment set if and only if

$$\frac{u(y)}{q(y)} \geq \frac{2\underline{\theta}}{2\underline{\theta} - 1}. \tag{41}$$

Under the monotone likelihood ratio property, the optimal punishment set will be an interval $Y = [\underline{y}, y^*]$, where $y^*$ satisfies condition (41) with equality.

## References


**Abreu, D., D. Pearce, and E. Stacchetti**, 1990. "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." *Econometrica* 1041-1063.

**Andrews, I. and D. Barron**, 2016. "The Allocation of Future Business." *American Economic Review* **106**, 2742–2759.

**Athey, S. and K. Bagwell**, 2001. "Optimal Collusion with Private Information." *RAND Journal of Economics* 32:428-465.

**Athey, S. and D. A. Miller**, 2007. "Efficiency in Repeated Trade with Hidden Valuations." *Theoretical Economics* 2:299-354.

**Banks, J. and R. Sundaram**, 1993. "Adverse Selection and Moral Hazard in a Repeated Elections Model," in: William Barnett et al. (Ed.), *Political Economy: Institutions, Competition and Representation.* Cambridge University Press.

**Banks, J. and R. Sundaram**, 1998. "Optimal Retention in Agency Problems." *Journal of Economic Theory* **82**, 293–323.

**Ben-Porath, E., E. Dekel, and B. L. Lipman**, 2014. "Optimal Allocation with Costly Verification." *American Economic Review* 104:3779-3813.

**Bergemann, D. and S. Morris**, 2005. "Robust Mechanism Design." *Econometrica* 73:1771-1813.

**Bergemann, D. and J. Välimäki**, 2010. "The Dynamic Pivot Mechanism." *Econometrica* 78:771-789.

**Board, S.**, 2011. "Relational Contracts and the Value of Loyalty." *American Economic Review* **101**, 3349–3367.

**Carnicer, J.M., T.N.T. Goodman and J.M. Peña**, 1999. "Linear Conditions for Positive Determinants." *Linear Algebra and Its Applications* **292**, 39-59.

**Duggan, J. and C. Martinelli**, 2017. "The Political Economy of Dynamic Elections: Accountability, Commitment and Responsiveness." *Journal of Economic Literature* 55(3), 916-84.

**Farrell, J., and E. Maskin**, 1989. "Renegotiation in Repeated Games." *Games and Economic Behavior*, 1(4), 327-360.

**Hauser, C. and H. Hopenhayn**, 2008. "Trading Favors: Optimal Exchange and Forgiveness." *Working paper.*

**Li, J., N. Matouscheck and M. Powell**, 2017. "Power Dynamics in Organizations." *American Economic Journal: Microeconomics*, **9**, 217–241.

**Lipnowski, E. and J. Ramos**, 2016. "Repeated Delegation." *Working Paper.*

**Miller, D. A.**, 2012. "Robust Collusion with Private Information." *The Review of Economic Studies* 79:778– 811.

**Möbius, M.**, 2001. "Trading favors." *Working paper.*

**Varah, J.M.**, 1975. "A Lower Bound for the Smallest Singular Value of a Matrix." *Linear Algebra and Its Applications* **11**, 3-5.