

Taming Discrimination in Hiring: A Dynamic Delegation Approach*

PAOLO MARTELLINI

NYU and NBER

GUIDO MENZIO

NYU and NBER

October 2026

Abstract

We study the optimal delegation of hiring decisions to an agent who has private information about the quality of different applicants, but discriminates against workers from some demographic group. The optimal delegation is dynamic. The hiring decision is initially delegated to the agent. The agent is rewarded for hiring and punished for rejecting applicants from the discriminated group. The reward takes the form of weaker incentives, which allow the agent to lean into its bias, and more time before the hiring decision is given to the principal, which allows the agent to exercise its bias for longer. The punishment takes the form of stronger incentives, and a reduction in the time before the hiring decision is given to the principal. No matter how many discriminated applicants are hired by the agent, the decision power is never bestowed upon the agent on a permanent basis. After sufficiently many discriminated applicants are consecutively rejected by the agent, the decision power is permanently given to the principal.

JEL Codes: D82, J71.

Keywords: Mechanism Design, Delegation, Labor Discrimination.

*Martellini: Department of Economics, NYU Stern School of Business, Henry Kaufman Management Center, 44 West 4th Street, NY 10012 (email: pm1987@nyu.edu). Menzio: Department of Economics, New York University, 19 West 4th Street, New York, NY 10012 (email: gm1310@nyu.edu). We previously circulated the manuscript as “The Dean and the Chair.” We are grateful to Manuel Amador, Mark Aguiar, Jarda Borovicka, Borys Grochulski, Marina Halac, Oleg Itskhoki, Andreas Kleiner, Natalia Kovrijnykh, Jennifer La’O, Elliot Lipnowski, Lee Ohanian, Russell Wong, and Pierre Yared for helpful suggestions. We are also thankful to our audiences at the NBER Summer Institute, the SITE Workshop on Dynamic Contracts, Columbia, ASU, EIEF, CMU, FRB Chicago and FRB Richmond.

1 Introduction

How should the Dean of the College organize the hiring of new professors when the Chair of the Department has superior information about the quality and fit of different scholars, but disagrees on the importance of making the faculty more diverse? How should the owner of a firm set up the hiring process given that the manager is better suited to identifying which applicants are a good fit and which are not, but holds biased views against a particular demographic group? How should a labor market authority resolve the trade-off between letting firms hire whomever they see fit, and taming gender or race discrimination?

We formalize these questions as a mechanism design problem without transfers (i.e., a delegation problem) between a principal (e.g., the Dean of the College, the owner of the firm, a labor market authority) and an agent (e.g., the Chair of the Department, the manager of the firm, a private firm) who has private information about the quality of applicants, but who discriminates against applicants from some demographic group. In general, the optimal mechanism neither assigns the hiring decision to the agent permanently, nor it permanently assigns it to the principal. Instead, the optimal mechanism is dynamic. Depending on the realized history of hiring, the mechanism may assign the hiring decision to the agent (delegation state) or to the principal (control state). In the delegation state, the mechanism rewards the agent for hiring an applicant from the discriminated demographic group, and it punishes the agent for rejecting such an applicant. The reward and the punishment induce the agent to follow hiring criteria that are less biased. The reward is delivered as a weaker incentive to hire discriminated applicants in the future, and as an extension of the time until the hiring decision is taken away from the agent. The punishment is delivered as a stronger incentive to hire discriminated applicants in the future, and as reduction of the time until the hiring decision is taken away from the agent. In the control state, the principal hires only applicants from the discriminated group. No matter how many discriminated applicants are hired by the agent, the agent does not get permanent control over hiring. After sufficiently many discriminated applicants are consecutively rejected by the agent, the principal gets permanent control over hiring.

In Section 2, we study a simple environment, which turns out to be the basis for richer and more realistic scenarios. In the simple environment, a new applicant from the discriminated group is available for hire in every period. If the applicant is hired, the payoff to the principal is given by x and the payoff to the agent is given by $x - \eta$, where x is a continuous random variable that describes the quality of the applicant whose realization is privately observed by the agent, and $\eta > 0$ is a parameter that describes the agent's bias against the applicant's demographic group. If the applicant is not hired, the payoff to the principal and the payoff to the agent are normalized to 0.

The principal commits to a direct-revelation mechanism in order to maximize its ex-

pected value. In each period, the direct-revelation mechanism solicits a report from the agent about the quality of the applicant. Depending on the realized history of play, the mechanism may make use of the agent’s report to decide whether to hire the applicants, or it may ignore the agent’s report. In the first case, the mechanism produces the same outcomes as if the agent was making the hiring decision directly, while taking into account the consequences of its decision on the future prescriptions of the mechanism. We refer to the first case as “delegation”. In the second case, the mechanism produces the same outcomes as if the principal was making the hiring decision. We refer to the second case as “control” (or, in the academic jargon, receivership) Since the upper envelope of the principal’s value under delegation and control is not generally concave in the agent’s value, the optimal mechanism includes a public lottery between delegation and control at the beginning of every period. We restrict attention to mechanisms with the property that: (i) once the mechanism assigns control of the hiring decision to the principal, it does so forever; (ii) when in control of the hiring decision, the principal hires every applicant. In Section 3, we prove that these restrictions are made without any loss in generality.

We formulate the mechanism design problem recursively, using the agent’s promised value as an auxiliary state variable. We first characterize the optimal lottery between delegation and control as a function of the agent’s promised value V . If V is lower than some threshold V_C , the optimal lottery is non-degenerate, in the sense that it assigns positive probability to both delegation and control. If the outcome of the lottery is control, the principal hires every future applicant, and the agent’s value falls to its minimum V_P . If the outcome of the lottery is delegation, the agent keeps making the hiring decisions for the time being and its continuation value $\hat{V}$ moves up to the threshold V_C . If the agent’s promised value V is higher than V_C , the optimal lottery is degenerate. The hiring decision is delegated to the agent, and the agent’s continuation value $\hat{V}$ is equal to its promised value V .

We then characterize the optimal incentives in delegation, as a function of the agent’s promised value $\hat{V}$. We show that, if the agent hires the applicant in the current period, it gets rewarded with a continuation value V_1 that is strictly greater than $\hat{V}$. If the agent does not hire the applicant, it gets punished with a continuation value V_0 that is strictly smaller than $\hat{V}$. The gap between V_1 and V_0 induces the agent to adopt a reservation quality for hiring, R , that is tilted away from η , the agent’s preferred quality cutoff, and towards 0, the principal’s preferred quality cutoff. The mechanism delivers the reward $V_1 - \hat{V}$ by increasing the expected time until the hiring decision is taken away from the agent, and by weakening the incentives faced by the agent in the future, which allows the agent to adopt a reservation quality R that is closer to its preferred one. Conversely, the mechanism delivers the punishment $\hat{V} - V_0$ by decreasing the expected time until the hiring decision is taken away from the agent, and by strengthening the incentives faced by the agent, which forces the agent to adopt a reservation quality R that is further away from its preferred one. If the agent keeps hiring applicants, its value approaches

(but never reaches) V_F , which denotes the value that the agent could obtain if it had permanent discretion over hiring. If the agent keeps rejecting applicants, the mechanism eventually gives the control of the hiring process to the principal, and the agent's value reaches V_P .

In Section 4, we exploit the fact that the quality x of the applicant is a continuous random variable to effortlessly extend the characterization of the optimal mechanism in the basic environment to richer and more realistic scenarios. The optimal mechanism in the basic environment is exactly the same as in the case where the agent can choose between an applicant from a demographic group against which it is biased (a contentious applicant), and an applicant from a demographic group against which it holds no bias (an uncontentious applicant). The analysis of the optimal mechanism in the basic environment can be readily extended to the case in which the agent can choose between n contentious applicants and m uncontentious applicants. The analysis of the optimal mechanism in the basic environment can also be adapted to the case in which the agent's bias is positive rather than negative, and to the case in which the agent privately observes whether a contentious applicant is available for hire or not.

The paper relates to the literature on delegation. The delegation literature focuses on the optimal design of the choice set for an agent who has private information about the value of different actions, but whose preferences are imperfectly aligned with those of the principal. Much of the delegation literature is static (see, e.g., Holmstrom 1977, Aghion and Tirole 1997, Alonso and Matouschek 2008, Amador and Bagwell 2013). The delegation problem is interesting because the set of actions is rich. In our paper, the set of actions only includes hiring or not hiring an applicant and, hence, the agent's possible choice sets boils down to either having the right to hire an applicant or not. The problem, however, is still interesting because the model is dynamic. Indeed, we show that by making the agent's choice set a function of its past decisions, the mechanism can use some of the agent's private information in favor of the principal.

There is a smaller literature on dynamic delegation problems. Some of the work is concerned with delegation in environments where the decision-maker has a time-inconsistency problem (Athey, Atkeson, and Kehoe 2005, and Angeletos, Werning, and Amador 2006, Halac and Yared 2018). These papers find that, as long as the state privately observed by the agent is independent over time, the agent's optimal choice set does not depend on the history of play. In contrast, we find that the optimal choice set does depend on the history of play. Intuitively, our findings are different because the principal and the agent disagree both on current and future actions. Like us, Guo and Horner (2020) study an optimal delegation problem between an agent and a principal that persistently disagree over a binary action. Like us, they find that the optimal delegation is dynamic. Unlike us, they find that, depending on the history of play, the decision may be permanently delegated to the agent, or permanently assigned to the principal (as in the repeated moral-hazard models studied by Clementi and Hopenhayn 2006, Sannikov 2008, and Padro i Miquel

and Yared 2012). The difference is due to the fact that the state can only take two values in Guo and Horner (2020), and it can take a continuum of values in our paper. Intuitively, when the state can take a continuum of values (i.e., the quality of the applicant is a continuous random variable), the marginal rate of substitution between the value of the mechanism to the agent and the value of the mechanism to the principal becomes infinite when the mechanism allows the agent to make decisions that are closer and closer to its preferred ones (i.e., follow a cutoff quality R closer and closer to η). Hence, rewarding the agent with additional value becomes infinitely costly in terms of the value of the principal (a similar argument can be found in Carrasco, Fuchs, and Fukuda 2019, while a different argument for the same result can be found in Li, Matouscheck, and Powell 2017, and Lipnowski and Ramos 2021). Importantly, by moving from a binary to a continuous state, our analysis can be immediately generalized to the most natural environment, one in which there are many applicants, some contentious and some not. This generalization is not possible if the quality of applicants is a discrete random variable.

Our paper relates to the literature on quotas, which are a simple type of mechanism that does not involve monetary transfers. In the labor context of our paper, quotas are mechanisms that force the agent to hire a particular fraction of applicants from the discriminated group. Jackson and Sonnenschein (2007) show that, when there are sufficiently many hiring decisions to be taken in a given interval of time, it is possible to approach first-best outcomes by putting a quota on the number of times in which the agent can make a given report about the quality of the candidate (see also Escobar and Toikka 2013). Frankel (2016) identifies a class of principal-agent problems in which a generalized notion of quotas is optimal, even when the hiring decisions are infrequent. The key property of this class of problems is that the agent’s payoff is not allowed to depend on the quality of the applicant and, hence, this class of problems does not include ours. Frankel (2021) studies a static environment and characterizes the optimal hiring quotas as a function of the hard information that is available to both the principal and the agent about the quality of the applicants. In contrast to our model, Frankel (2021) assumes that the firm has a large number of positions to fill, and faces a large population of applicants.

If one is willing to interpret the agent as a private firm and the principal as a labor market regulator, our paper can be seen as a contribution to the debate about hiring discrimination and its remedies. Friedman (1962) proposed to let hiring discrimination be reflected in wages, so that unbiased firms could drive prejudiced ones out of the market. Search frictions in the labor market, however, limit the strength of this solution. If explicit wage discrimination is forbidden, tackling hiring discrimination requires intervening on the firms’ employment decisions. Some of the literature has advocated for hiring quotas (see, e.g., Frankel 2023). As discussed above, however, hiring quotas need not be optimal when firms hire infrequently. Some of the literature has advocated for blind hiring, a process by which applicants are reviewed while their demographics are hidden (see, e.g., Goldin and Rouse 2000). While ideal, blind hiring may be technologically unfeasible in

many situations. The analysis in this paper provides a characterization of the optimal non-monetary mechanism to curb hiring discrimination when quotas are too restrictive, and blind hiring is unfeasible.

2 Baseline model

In this section, we consider a simple environment. In every period, the principal has the option to hire a new applicant from some particular demographic group. The principal does not have the expertise to assess the quality of the applicant, but the agent does. The agent, however, is biased against applicants from that demographic group. The principal commits to a direct-revelation mechanism in order to elicit the information of the agent while minimizing the consequences of the agent's bias. The mechanism does not allow for monetary transfers between the principal and the agent and, in this sense, the mechanism design problem is a delegation problem. In Section 2.1, we describe and interpret the environment. In Section 2.2, we formulate the mechanism design problem recursively as a two-stage problem, using the agent's value as an auxiliary state variable. In Section 2.3, we characterize the solution of the first-stage problem—a lottery between delegation and control, where delegation means that, effectively, the agent directly chooses whether to hire the applicant, and control means that, effectively, the principal directly chooses whether to hire the applicant. In Section 2.4, we characterize the solution of the second-stage problem—the design of the mechanism when the hiring decision is delegated to the agent.

2.1 Environment

In every period $t = 0, 1, 2, \dots$, a new applicant is available for hire. The quality x of the applicant is drawn from a continuously differentiable cumulative distribution function $F(x)$ with mean 0 and support $X = [\underline{x}, \bar{x}]$.¹ The agent observes the quality of the applicant, but the principal does not. If the applicant is hired, the agent obtains a flow payoff of $(1 - \beta)(x - \eta)$ and the principal obtains a flow payoff of $(1 - \beta)x$, where $\beta \in (0, 1)$ is the factor at which the principal and the agent discount future payoffs, and $\eta \in (0, \bar{x})$ is the agent's bias against hiring the applicant. If the applicant is not hired, both the agent and the principal obtain a flow payoff of 0. No monetary transfers between the principal and the agent are allowed.

In period $t = 0$, the principal commits to a direct-revelation mechanism. In every period t , the mechanism solicits a report $\hat{x} \in X$ from the agent about the quality $x \in X$ of the applicant and decides whether or not to hire the applicant $a \in \{0, 1\}$. Depending on the history h_t of play observed up to period t , the mechanism may act on the report

¹The assumption $\mathbb{E}[x] = 0$ is not critical for any of our results. The case in which $\mathbb{E}[x] < \eta$ only requires minor modifications in the proofs. The case in which $\mathbb{E}[x] \geq \eta$ is more involved. Numerically, however, we find that our results also hold when $\mathbb{E}[x] \geq \eta$.

from the agent, in the sense that it does or does not hire the applicant depending on the report, or it may ignore the report from the agent, in the sense that it may or may not hire the applicant irrespective of the report. In the first case, we say that the mechanism is delegating the hiring decision to the agent. In the second case, we say that the mechanism is letting the principal control the hiring decision.²

Consider the case in which hiring is delegated to the agent. Let X_0 denote the set of reports such that the mechanism does not hire the applicant, i.e. $a(\hat{x}, h_t) = 0$ for all $\hat{x} \in X_0$. Similarly, let X_1 denote the set of reports such that the mechanism does hire the applicant, i.e. $a(\hat{x}, h_t) = 1$ for all $\hat{x} \in X_1$. If the agent reports $\hat{x}$, its flow payoff is $a(\hat{x}, h_t)(1 - \beta)(x - \eta)$ and its continuation value is some $V(\hat{x}, h_t)$. Since a direct-revelation mechanism must induce the agent to report the applicant's quality truthfully, it follows that $V(\hat{x}, h_t) = V_0(h_t)$ for all $\hat{x} \in X_0$ and $V(\hat{x}, h_t) = V_1(h_t)$ for all $\hat{x} \in X_1$. In words, the mechanism must give the same continuation value $V_0(h_t)$ to the agent for all the reports that make the mechanism pass on the applicant, and the same continuation value $V_1(h_t)$ for all the reports that make the mechanism hire the applicant. Moreover, since the mechanism must induce the agent to report the applicant's quality truthfully, it has to be the case that $X_0 = [x_\ell, R(h_t))$ and $X_1 = [R(h_t), x_h]$, where $R(h_t)$ is the quality of the applicant that makes the agent indifferent between $(1 - \beta)(x - \eta) + \beta V_1(h_t)$ and $\beta V_0(h_t)$. Overall, when the mechanism uses the agent's report, it is as if the agent was directly choosing whether or not to hire the applicant, taking into account that its continuation value is $V_0(h_t)$ if it does not hire the applicant and $V_1(h_t)$ if it does.

Consider the case in which hiring is controlled by the principal. In this case, the mechanism either does not hire the applicant irrespective of the agent's report, i.e. $a(\hat{x}, h_t) = 0$ for all $\hat{x} \in X$, or it hires the applicant irrespective of the agent's report, i.e. $a(\hat{x}, h_t) = 1$ for all $\hat{x} \in X$. In either case, since the mechanism must induce the agent to report truthfully the quality of the applicant, the agent's continuation value $V(\hat{x}, h_t)$ must be independent of the agent's report $\hat{x}$. Overall, when the mechanism is ignoring the agent's report, it is as if the principal was controlling the hiring process directly.

Clearly, the combinations of agent's and principal's values that the mechanism can implement under delegation and under control are different and, in turn, the upper envelope of the agent's and principal's values under delegation and under control need not be concave. For this reason, it is natural to let the mechanism specify a public lottery between delegation and control at the beginning of every period t . Specifically, at the beginning of every period, the mechanism specifies some probability $p(h_t)$ with which hiring is controlled by the principal, and some probability $1 - p(h_t)$ with which hiring is

²We restrict attention to mechanisms that specify a hiring probability $a(x)$ that is either 0 or 1. The restriction is not without loss in generality. The restriction, however, is quite natural. If the mechanism specifies a hiring probability $a(x) \in \{0, 1\}$, a direct-revelation mechanism can be implemented as a delegation mechanism—a mechanism that assigns the hiring decision to either the agent or the principal, depending on the history of play. If the mechanism is allowed to specify a hiring probability $a(x) \in [0, 1]$, the direct-revelation mechanism cannot be implemented as a delegation mechanism. Moreover, the restriction allows us to deal with a quality x that is a continuous random variable.

delegated to the agent.

In this section, we restrict attention to mechanisms such that: (i) if the mechanism hands the control of the hiring process to the principal, it does so forever; (ii) if the mechanism hands control of the hiring process to the principal, the mechanism instructs the principal to hire every applicant. In the next section, we show that these restrictions are without loss in generality. Even though starting with a restriction on the space of mechanisms and then showing that it is without loss in generality may appear awkward to some of our readers, it does lead to a neater derivation of the optimal mechanism.

The environment described above is standard in the delegation literature (see, e.g., Holmstrom 1977, Aghion and Tirole 1997, Alonso and Matouschek 2008, Amador and Bagwell 2013). The principal (the Dean of the College) is faced with a decision (hiring or not hiring an applicant from some demographic group), but some of the relevant information for making the decision (the quality of the applicant) is privately held by the agent (the Chair of the Department), whose preferences are not perfectly aligned with those of the principal (as the agent does not fully internalize the value of a diverse faculty). The principal commits to a direct-revelation mechanism in order to make use of some of the agent's information. As in the delegation literature, we assume that the mechanism does not allow for monetary transfers between the principal and the agent. The assumption is motivated by the idea that arranging explicit monetary transfers between members of the same organization is too costly (i.e., it would be too costly to set up a payment system between the Dean and the Chair). As in the delegation literature, we assume that the agent's private information does not become public. The assumption is motivated by the view that the principal cannot pinpoint the contribution of particular decisions on its payoff (i.e., the Dean cannot isolate the contribution of an individual hire on the quality of the College).

In contrast to the dynamic delegation literature (see, e.g., Li, Matoushek and Powell 2017, Guo and Horner 2020, Lipnowski and Ramos 2021), we assume that the state observed by the agent (the quality of the applicant) is a continuous random variable, rather than a discrete random variable. In the context of hiring applicants, the assumption seems quite natural. Moreover, the assumption leads to new insights (e.g., the hiring standards used by the agent evolve over time depending on the realized history of play), and it allows us to effortlessly generalize the analysis to richer and more realistic environments (e.g., an environment in which the decision involves hiring one among several applicants some of whom are from a discriminated group and some of whom are not).

2.2 Recursive formulation and preliminaries

The principal's mechanism design problem can be formulated recursively, using the value of the mechanism to the agent as an auxiliary state variable. In the first-stage of the recursive problem, the mechanism chooses a lottery between delegation and control, subject to delivering a particular value to the agent. In the second-stage, which is associated

with the delegation branch of the lottery, the mechanism chooses the agent's continuation value conditional on the agent hiring the applicant, and the agent's continuation value conditional on the agent not hiring the applicant, subject to delivering a particular expected value to the agent.

Formally, the first-stage problem is

$$J(V) = \max_{p \in [0,1], \hat{V} \in \hat{\mathcal{V}}} pJ_P + (1-p)\hat{J}(\hat{V}), \quad (2.1)$$

subject to the promise-keeping constraint

$$V = pV_P + (1-p)\hat{V}. \quad (2.2)$$

The first-stage problem is easy to understand. The mechanism chooses the probability p with which it permanently assigns the hiring decision to the principal (control), the probability $1-p$ with which this period's hiring decision is made by the agent (delegation), as well as the value $\hat{V}$ to the agent in case the outcome of the lottery is delegation. The mechanism makes these choices to maximize the principal's expected value, subject to delivering some expected value V to the agent. The value to the principal is J_P if hiring is controlled by the principal, and $\hat{J}(\hat{V})$ if hiring is delegated to the agent. The value to the agent is V_P if hiring is controlled by the principal, and $\hat{V}$ if hiring is delegated to the agent. The values J_P and V_P , with P standing for "punishment", are

$$J_P \equiv \int_{\underline{x}}^{\bar{x}} x dF(x) = 0, \quad V_P \equiv \int_{\underline{x}}^{\bar{x}} (x - \eta) dF(x) = -\eta. \quad (2.3)$$

The second-stage problem is

$$\hat{J}(\hat{V}) = \max_{V_0, V_1 \in \mathcal{V}} (1-\beta) \int_R x dF(x) + \beta [F(R)J(V_0) + (1-F(R))J(V_1)], \quad (2.4)$$

subject to the promise-keeping constraint

$$\hat{V} = (1-\beta) \int_R (x - \eta) dF(x) + \beta [F(R)V_0 + (1-F(R))V_1], \quad (2.5)$$

and the incentive-compatibility constraint

$$R = \eta - \frac{\beta}{1-\beta} (V_1 - V_0). \quad (2.6)$$

The second-stage problem is also easy to understand. The mechanism chooses the agent's continuation value V_0 conditional on the agent not hiring the applicant, and the agent's continuation value V_1 conditional on the the agent hiring the applicant. Given V_0 and V_1 , the agent finds it optimal to hire the applicant if and only if the applicant's quality x exceeds the reservation threshold R in (2.6). The mechanism maximizes the principal's value, subject to delivering the expected value $\hat{V}$ to the agent. In the current period, the

payoff to the principal is $(1 - \beta) \int_R x dF(x)$ and the payoff to the agent is $(1 - \beta) \int_R (x - \eta) dF(x)$. If the agent does not hire the applicant, an event that occurs with probability $F(R)$, the continuation value to the principal is $J(V_0)$, and the continuation value to the agent is V_0 . If the agent hires the applicant, an event that occurs with probability $1 - F(R)$, the continuation value to the principal is $J(V_1)$, and the continuation value to the agent is V_1 .

The formulation of the first and second-stage problems (2.1) and (2.4) is still incomplete, as it does not specify the choice sets $\mathcal{V}$ and $\hat{\mathcal{V}}$ for the agent's continuation values. The agent's continuation value $\hat{V}$ in (2.1) must be implementable in the second stage, in the sense that there exists a mechanism that delivers $\hat{V}$ to the agent in the second stage. The agent's continuation values V_0 and V_1 in (2.4) must be implementable in the first stage, in the sense that there exist mechanisms that deliver V_0 and V_1 to the agent in the first stage. The set $\mathcal{V}$ denotes the agent's values that can be implemented in the first-stage. The set $\hat{\mathcal{V}}$ denotes the agent's values that can be implemented in the second-stage.

Lemma 1 below characterizes the implementable sets $\mathcal{V}$ and $\hat{\mathcal{V}}$. The lemma shows that, in the first stage, the implementable set is the interval $[V_P, V_F]$, where V_P is the value to the agent of a mechanism in which the principal has permanent control over hiring, and V_F is the value to the agent of a mechanism in which hiring is permanently delegated to the agent, with F standing for "freedom." The values V_P and V_F are implementable. Values between V_P and V_F can be implemented by choosing the appropriate lottery between control and delegation. Values lower than V_P and values greater than V_F cannot be implemented in the first stage. In the second stage, the implementable set is the interval $[V_L, V_F]$, where V_L is the value to the agent of a mechanism in which hiring is in the hands of the agent in the current period and in the hands of the principal from the next period onwards. Values between V_L and V_F can be implemented by appropriately choosing the agents' continuation values. Values lower than V_L and values greater than V_F cannot be implemented in the second stage.

Lemma 1: (Implementability) *The sets $\mathcal{V}$ and $\hat{\mathcal{V}}$ are, respectively, given by*

$$\mathcal{V} = [V_P, V_F], \quad \hat{\mathcal{V}} = [V_L, V_F], \quad (2.7)$$

where V_L and V_F are defined as

$$V_L \equiv (1 - \beta)V_F + \beta V_P, \quad (2.8)$$

$$V_F \equiv \int_{\underline{x}}^{\bar{x}} \max\{x - \eta, 0\} dF(x) = \int_{\eta}^{\bar{x}} (x - \eta) dF(x). \quad (2.9)$$

Proof: Consider the first-stage problem (2.1). The choice $p = 1$ is feasible. For $p = 1$, the first-stage value to the agent is V_P . Hence, $V_P \in \mathcal{V}$. Consider the second-stage problem (2.4). The choices $V_0 = V_P$ and $V_1 = V_P$ are feasible. Given $(V_0, V_1) = (V_P, V_P)$, the

second-stage value to the agent is

$$\begin{aligned}\bar{V}_1 &= (1 - \beta) \int_{\eta} (x - \eta) dF(x) + \beta V_P \\ &= (1 - \beta) V_F + \beta V_P,\end{aligned}\tag{2.10}$$

where the first line in (2.10) makes use of the fact that $V_0 = V_1$ implies $R = \eta$, and the second line in (2.10) makes use of the definition of V_F . Hence, $\bar{V}_1 \in \hat{\mathcal{V}}$. Notice that, since $V_F > V_P$, $\bar{V}_1 > V_P$ and $\bar{V}_1 < V_F$.

Return to the first-stage problem (2.1). The choices $p = (\bar{V}_1 - V)/(\bar{V}_1 - V_P)$ and $\hat{V} = \bar{V}_1$ are feasible for any $V \in [V_P, \bar{V}_1]$, since $p \in [0, 1]$ and $\bar{V}_1 \in \hat{\mathcal{V}}$. Given the choices $p = (\bar{V}_1 - V)/(\bar{V}_1 - V_P)$ and $\hat{V} = \bar{V}_1$, the first-stage value to the agent is V . Hence, any $V \in [V_P, \bar{V}_1]$ belongs to $\mathcal{V}$. Now, consider the second-stage problem (2.4). The choices $V_0 = V$ and $V_1 = V$ are feasible for any $V \in [V_P, \bar{V}_1]$. Given $(V_0, V_1) = (V, V)$, the second-stage value to the agent is $\hat{V} = (1 - \beta)V_F + \beta V$. Hence, any value $\hat{V}$ in the interval $[\bar{V}_1, \bar{V}_2]$ belongs to $\hat{\mathcal{V}}$, where

$$\bar{V}_2 = (1 - \beta)V_F + \beta \bar{V}_1.\tag{2.11}$$

Since $\bar{V}_1 < V_F$, it follows that $\bar{V}_2 > \bar{V}_1$ and $\bar{V}_2 < V_F$.

Repeating the above argument k times yields that any value $V \in [V_P, \bar{V}_k]$ belongs to $\mathcal{V}$, and any value $\hat{V} \in [\bar{V}_1, \bar{V}_{k+1}]$ belongs to $\hat{\mathcal{V}}$, where

$$\bar{V}_{k+1} = (1 - \beta)V_F + \beta \bar{V}_k.\tag{2.12}$$

The sequence $\{\bar{V}_i\}_{i=1}^{k+1}$ is strictly increasing and converges to V_F . Since a mechanism that permanently delegates hiring to the agent is worth V_F to the agent, V_F belongs to both $\mathcal{V}$ and $\hat{\mathcal{V}}$. These observations imply that $[V_P, V_F] \subseteq \mathcal{V}$ and $[\bar{V}_1, V_F] \subseteq \hat{\mathcal{V}}$.

Now, notice that the periodical payoff to the agent is such that

$$v \geq \min \left\{ \min_R (1 - \beta) \int_R (x - \eta) dF(x), (1 - \beta) \int_{\underline{x}}^{\bar{x}} (x - \eta) dF(x) \right\}.\tag{2.13}$$

The right-hand side of (2.13) is the minimum between two terms. The first term is the lowest periodical payoff that an agent can attain when hiring is delegated to him. The second term is the periodical payoff that an agent attains when hiring is controlled by the principal. The minimum between the two terms is $-(1 - \beta)\eta = (1 - \beta)V_P$. This observation implies that a mechanism cannot implement any agent's value V smaller than V_P . That is, $V \notin \mathcal{V}, \hat{\mathcal{V}}$ for any $V < V_P$.

The periodical payoff v to the agent is such that

$$v \leq \max \left\{ \max_R (1 - \beta) \int_R (x - \eta) dF(x), (1 - \beta) \int_{\underline{x}}^{\bar{x}} (x - \eta) dF(x) \right\}.\tag{2.14}$$

The right-hand side of (2.14) is the maximum between two terms. The first term is the

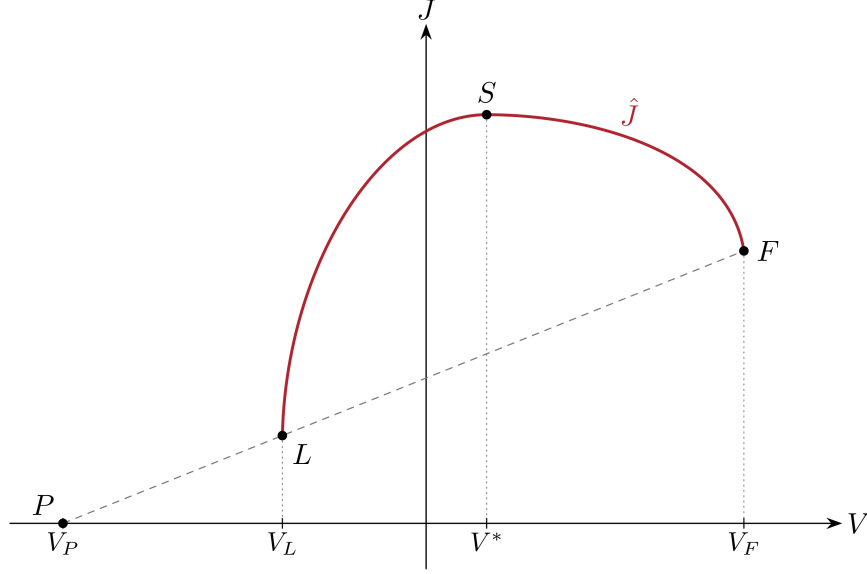


Figure 1: Points P , L , F and S , and frontier $(\hat{V}, \hat{J}(\hat{V}))$.

highest periodical payoff that the agent can attain when hiring is delegated to him. The second term is the periodical payoff that the agent attains when the hiring is controlled by the principal. The maximum of the two terms is $(1 - \beta)V_F$. Therefore, a mechanism cannot implement any value V strictly greater than V_F . That is, $V \notin \mathcal{V}$, $\hat{\mathcal{V}}$ for any $V > V_F$.

Finally, notice that $V \notin \hat{\mathcal{V}}$ for any $V < \bar{V}_1 = V_L$. To see why this is the case notice that, in the second stage, the agent's value is such that

$$\begin{aligned}
 \hat{V} &= \max_R (1 - \beta) \int_R (x - \eta) dF(x) + \beta F(R)V_0 + \beta(1 - F(R))V_1 \\
 &\geq \max_R (1 - \beta) \int_R (x - \eta) dF(x) + \beta F(R)V_P + \beta(1 - F(R))V_P \\
 &= (1 - \beta)V_F + \beta V_P = V_L,
 \end{aligned} \tag{2.15}$$

where the second line makes use of the fact that V_0 and V_1 must be greater than V_P , and the third line makes use of the fact that the agent finds it optimal to set $R = \eta$ when V_0 and V_1 are both equal to V_P . ■

Before turning to the characterization of the optimal mechanism, it is useful to examine some points along the (V, J) and $(\hat{V}, \hat{J})$ frontiers. First, consider the agent's value V_P . Since $V_P \in \mathcal{V}$ and $V_P \notin \hat{\mathcal{V}}$, a mechanism can deliver the value V_P to the agent only in the first stage. It follows immediately from (2.1) that a mechanism can only deliver V_P to the agent through a lottery such that the hiring decision is delegated to the agent with probability 0 and permanently controlled by the principal with probability 1. Hence, $J(V_P) = J_P$. We denote as P the point (V_P, J_P) . The point P belongs to the (V, J) frontier.

Second, consider the agent's value V_F . Since $V_F \in \mathcal{V}$ and $V_F \in \hat{\mathcal{V}}$, a mechanism can deliver the value V_F to the agent in both the first and the second stage. In the first stage, the mechanism can deliver the value V_F to the agent only through a lottery where the hiring decision is delegated to the agent with probability 1, and the agent's second-stage value is V_F . Hence, $J(V_F)$ is equal to $\hat{J}(V_F)$. In the second stage, the mechanism can only deliver the value V_F to the agent by setting both V_0 and V_1 equal to V_F . Since $V_0 = V_1$ implies $R = \eta$, it follows that $\hat{J}(V_F)$ is equal to $(1 - \beta) \int_{\eta} x dF(x) + \beta J(V_F)$. Since $J(V_F) = \hat{J}(V_F)$, it follows that $\hat{J}(V_F)$ is equal to $\int_{\eta} x dF(x)$, which we denote as J_F . We denote as F the point (V_F, J_F) , which belongs to both the (V, J) and $(\hat{V}, \hat{J})$ frontiers. Note that, from the above characterization of the optimal mechanism, it follows that F is an absorbing state, in the sense that, once the mechanism reaches F , it remains there forever and, hence, the hiring decision is permanently delegated to the agent.

Third, consider the agent's value V_L , which is defined as $(1 - \beta)V_F + \beta V_P$. Since V_L belongs to $\hat{\mathcal{V}}$, a mechanism can deliver the value V_L to the agent in the second stage. In the second stage, the mechanism can only deliver the value V_L to the agent by setting both V_0 and V_1 to V_P . Since $V_0 = V_1$ implies $R = \eta$, $\hat{J}(V_L)$ is equal to $(1 - \beta) \int_{\eta} x dF(x) + \beta J(V_P)$. Since $J(V_P) = J_P$ and $\int_{\eta} x dF(x) = J_F$, it follows that $\hat{J}(V_L)$ is equal to $(1 - \beta)J_F + \beta J_P$, which we denote as J_L . We denote as L the point (V_L, J_L) . The point L belongs to the $(\hat{V}, \hat{J})$ frontier but, as we will show soon, it does not belong to the (V, J) frontier.

Note that the points P , L and F are on a line, as illustrated in Figure 1. Formally, P , L and F are on the line

$$H(V) = J_P + \frac{J_F - J_P}{V_F - V_P}(V - V_P). \quad (2.16)$$

The slope of $H(V)$ is

$$H' = \frac{J_F - J_P}{V_F - V_P}. \quad (2.17)$$

The slope is strictly positive because $J_F = \int_{\eta} x dF(x) > 0$, $J_P = \int_{\underline{x}} x dF(x) = 0$, $V_F = \int_{\eta} (x - \eta) dF(x) > 0$ and $V_P = \int_{\underline{x}} (x - \eta) dF(x) = -\eta$. The slope is smaller than 1 because $V_F - V_P$ is equal to $J_F - J_P + \eta F(\eta)$.

Lastly, we denote as S the point (V^*, J^*) , where V^* is the agent's value such that the principal's value $J(V)$ is maximized. That is, V^* and J^* are, respectively, the arg-maximum and the maximum of $J(V)$ with respect to $V \in \mathcal{V}$. The point S is the initial position of the optimal mechanism, since the initial value of the mechanism to the agent is unconstrained by prior promises. For some parameter values, S is equal to F . In these cases, the optimal mechanism is such that the hiring decision is permanently delegated to the agent. For other parameter values, S is different from F and, hence, $J^* > J_F$ and $V^* < V_F$. In these cases, the optimal mechanism must be such that, after some history, the principal takes over the hiring process. In any case, V^* and J^* are also the arg-maximum and the maximum of $\hat{J}(\hat{V})$ with respect to $\hat{V} \in \hat{\mathcal{V}}$. This is so because $J(V)$

is a convex combination between J_P and some $\hat{J}(\hat{V})$, and $\hat{J}(V_F) > J_P$. Therefore, the optimal mechanism is always such that the hiring decision is initially delegated to the agent.

The following lemma identifies a sufficient condition on the parameters of the model such that $S \neq F$ and, hence, the optimal mechanism is dynamic. In the remainder of the paper, we assume that the sufficient condition holds.

Lemma 2: (Sufficient condition for a dynamic mechanism) *The starting position $S = (V^*, J^*)$ of the optimal mechanism is such that $J^* > J_F$ as long as η and F are such that*

$$\frac{\eta F'(\eta)}{F(\eta)} > \frac{\int_{\eta} x dF(x)}{\int_{\eta} x dF(x) + \eta F(\eta)}. \quad (2.18)$$

Proof: Consider the following mechanism. In the current period, the hiring decision is delegated to the agent. If the agent hires the applicant, the agent's continuation value V_1 is V_F . That is, if the agent hires the applicant, the hiring decision is permanently delegated to the agent. If the agent does not hire the applicant, the agent's continuation value V_0 is $\epsilon V_P + (1 - \epsilon)V_F$ for some $\epsilon > 0$. The continuation value V_0 is delivered as a lottery that assigns probability ϵ to V_P and probability $1 - \epsilon$ to V_F . That is, if the agent does not hire the applicant, the hiring decision is permanently delegated to the agent with probability $1 - \epsilon$, and it is permanently controlled by the principal with probability ϵ . The mechanism need not be optimal, but it is feasible.

The value $\Gamma(\epsilon)$ of the mechanism to the principal is

$$\Gamma(\epsilon) = (1 - \beta) \int_R x F'(x) dx + \beta F(R) [\epsilon J_P + (1 - \epsilon) J_F] + \beta (1 - F(R)) J_F, \quad (2.19)$$

where the reservation quality R is

$$R = \eta - \frac{\beta}{1 - \beta} [V_F - \epsilon V_P - (1 - \epsilon) V_F]. \quad (2.20)$$

The derivative of $\Gamma(\epsilon)$ with respect to ϵ is

$$\Gamma'(\epsilon) = -\beta F(R) (J_F - J_P) - \frac{dR}{d\epsilon} \{ (1 - \beta) R F'(R) + \beta F'(R) \epsilon [J_F - J_P] \}, \quad (2.21)$$

where

$$\frac{dR}{d\epsilon} = -\frac{\beta}{1 - \beta} (V_F - V_P). \quad (2.22)$$

When evaluated at $\epsilon = 0$, the derivative of $\Gamma(\epsilon)$ with respect to ϵ becomes

$$\Gamma'(0) = -\beta F(\eta) (J_F - J_P) + \beta \eta F'(\eta) (V_F - V_P). \quad (2.23)$$

The above expression is strictly positive as long as

$$\frac{\eta F'(\eta)}{F(\eta)} > \frac{J_F - J_P}{V_F - V_P} = \frac{\int_{\eta} x dF(x)}{\int_{\eta} x dF(x) + \eta F(\eta)}. \quad (2.24)$$

Note that the value $\Gamma(\epsilon)$ is such that $\Gamma(0) = J_F$. Hence, if condition (2.24) is satisfied, there exists some $\epsilon^* > 0$ such that $\Gamma(\epsilon^*) > J_F$. Next, note that the value $\Gamma(\epsilon^*)$ is such that $\Gamma(\epsilon^*) \leq J^*$, since the mechanism is feasible but need not be optimal. Hence, if condition (2.24) is satisfied, $J^* > J_F$. ■

Intuitively, condition (2.18) is a lower bound on the agent's bias η given the distribution F of applicant's quality. For instance, if the distribution of applicant's quality is uniform over some interval $[-\delta, \delta]$, condition (2.18) boils down to $\eta > \delta/2$. That is, as long as the agent's bias is large relative to the dispersion of the applicant's quality, the optimal mechanism does not permanently delegate the hiring decision to the agent.³

2.3 Optimal lottery

In this subsection, we characterize the solution to the first-stage problem in (2.1), which is the design of the lottery between delegation and control. In order to simplify the analysis, we assume that the second-stage value to the principal, $\hat{J}(\hat{V})$, is strictly concave and differentiable.⁴

Using the promise-keeping constraint (2.2) to substitute out p , we can rewrite (2.1) as

$$\begin{aligned} J(V) = \max_{\hat{V} \in \hat{\mathcal{V}}} & \frac{\hat{V} - V}{\hat{V} - V_P} J_P + \frac{V - V_P}{\hat{V} - V_P} \hat{J}(\hat{V}), \\ \text{s.t. } & \hat{V} \geq V. \end{aligned} \quad (2.25)$$

The necessary condition for the optimality of $\hat{V}$ is

$$\begin{aligned} \frac{V - V_P}{\hat{V} - V_P} \left[\hat{J}'(\hat{V}) - \frac{\hat{J}(\hat{V}) - J_P}{\hat{V} - V_P} \right] &= 0 \quad \text{if } \hat{V} \in (\max\{V, V_L\}, V_F), \\ &\leq 0, \quad \text{if } \hat{V} = \max\{V, V_L\}, \\ &\geq 0, \quad \text{if } \hat{V} = V_F. \end{aligned} \quad (2.26)$$

Let us examine the term in square brackets on the left-hand side of (2.26). The function $\hat{J}'(\hat{V})$ is the derivative of the second-stage problem value function. Since $\hat{J}(\hat{V})$ is strictly

³Even in the case of other standard distributions, such as a Normal or an Exponential, condition (2.24) implies a lower bound on η that is increasing in the dispersion of the applicant's quality.

⁴We could guarantee the weak concavity of $\hat{J}(\hat{V})$, and in turn differentiability almost everywhere, by introducing lotteries in the second-stage problem. Weak concavity is enough to derive our key results, but it makes the analysis much more cumbersome. It is harder to guarantee the differentiability of $\hat{J}(\hat{V})$. Using an envelope argument, we can show that $\hat{J}(\hat{V})$ is differentiable if, starting from $\hat{V}$, there is a history that is reached with positive probability and in which the first-stage lottery is non-degenerate. In our numerical examples, we always find this to be the case.

concave and attains its maximum at some $V^* \in (V_L, V_F)$, it follows that $\hat{J}'(\hat{V})$ is strictly decreasing and such that $\hat{J}'(V_L) > 0$, $\hat{J}'(V^*) = 0$, and $\hat{J}'(V_F) < 0$. Since $\hat{J}(\hat{V})$ is strictly concave and such that $\hat{J}(V_L) = H(V_L)$, $\hat{J}(V_F) = H(V_F)$, and $\hat{J}(V^*) > H(V_F) > H(V^*)$, it follows that $\hat{J}'(V_L) > H'$ and $\hat{J}'(V_F) < 0 < H'$. For the same reason, $\hat{J}(\hat{V}) > H(\hat{V})$ for all $\hat{V} \in (V_L, V_F)$.

The function $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ is the slope of the line connecting the points (V_P, J_P) and $(\hat{V}, \hat{J}(\hat{V}))$. Since $\hat{J}(V_L) = H(V_L)$ and $\hat{J}(V_F) = H(V_F)$, it follows that $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ is equal to H' for $\hat{V} = V_L$ and $\hat{V} = V_F$. Since $\hat{J}(\hat{V}) > H(\hat{V})$ for all $\hat{V} \in (V_L, V_F)$, $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ is strictly greater than H' for all $\hat{V} \in (V_L, V_F)$. These observations imply that $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ is strictly smaller than $\hat{J}'(\hat{V})$ at $\hat{V} = V_L$ and strictly greater than $\hat{J}'(\hat{V})$ at $\hat{V} = V_F$. Therefore, there must exist at least one $V_C \in (V_L, V_F)$ such that $(\hat{J}(V_C) - J_P)/(V_C - V_P)$ is equal to $\hat{J}'(V_C)$. Clearly, any V_C is strictly greater than V_L . Similarly, any V_C is strictly smaller than V^* , since $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ is strictly positive and $\hat{J}'(\hat{V})$ is non-positive for any $\hat{V} \in [V^*, V_F]$.

It is easy to verify that the derivative of the function $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ with respect to $\hat{V}$ equals zero if and only if $\hat{V} = V_C$, and that the function attains a local maximum at V_C (i.e., the second derivative is strictly negative). Therefore, V_C must be unique and the function $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ must attain its global maximum at V_C . Moreover, since $\hat{J}'(\hat{V})$ is strictly decreasing and the function $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ is equal to $\hat{J}'(\hat{V})$ and attains its maximum at V_C , it follows that $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ must be strictly smaller than $\hat{J}'(\hat{V})$ for all $\hat{V} \in [V_L, V_C]$. Similarly, $(\hat{J}(\hat{V}) - J_P)/(\hat{V} - V_P)$ must be strictly greater than $\hat{J}'(\hat{V})$ for all $\hat{V} \in (V_C, V_F]$.

The above observations imply that, for any $V \in [V_P, V_C]$, the optimality condition (2.26) and the promise-keeping constraint (2.2) are satisfied if and only if

$$\hat{V} = V_C, \quad p = \frac{V_C - V}{V_C - V_P} > 0. \quad (2.27)$$

Plugging the optimal choice for $\hat{V}$ in (2.25) yields

$$J(V) = J_P + \frac{J_C - J_P}{V_C - V_P}(V - V_P), \quad (2.28)$$

where $J_C \equiv \hat{J}(V_C)$. Differentiating (2.28) with respect to V yields

$$J'(V) = \frac{J_C - J_P}{V_C - V_P} < \hat{J}'(V), \quad (2.29)$$

where the last inequality follows from the fact that $(J_C - J_P)/(V_C - V_P) = \hat{J}'(V_C)$ and $\hat{J}'(V) > \hat{J}'(V_C)$.

For any $V \in [V_C, V_F]$, the optimality condition (2.26) and the promise-keeping con-

straint (2.2) are satisfied if and only if

$$\hat{V} = V, \quad p = 0. \quad (2.30)$$

Plugging the optimal choice for $\hat{V}$ in (2.25) yields

$$J(V) = \hat{J}(V). \quad (2.31)$$

Differentiating (2.31) with respect to V yields

$$J'(V) = \hat{J}'(V). \quad (2.32)$$

The proposition below summarizes the characterization of the first-stage problem.

Proposition 1: (Optimal lottery) *The solution to the first-stage problem (2.1) is:*

1. *For $V \in [V_P, V_C)$, the hiring process is permanently controlled by the principal with probability p , and it is delegated to the agent with probability $1 - p$, where $p = (V_C - V)/(V_C - V_P)$. Conditional on delegation, the agent's value is V_C .*
2. *For $V \in [V_C, V_F]$, the hiring process is permanently controlled by the principal with probability 0, and it is delegated to the agent with probability 1. Conditional on delegation, the agent's value is V .*
3. *The value to the principal is*

$$J(V) = \begin{cases} J_P + \frac{J_C - J_P}{V_C - V_P}(V - V_P), & \text{if } V \in [V_P, V_C), \\ \hat{J}(V), & \text{if } V \in [V_C, V_F]. \end{cases} \quad (2.33)$$

4. *The point $C = (V_C, J_C)$ is such that $J_C = \hat{J}(V_C)$ and $V_C \in (V_L, V^*)$ is the unique solution to*

$$\hat{J}'(V_C) = \frac{J_C - J_P}{V_C - V_P}. \quad (2.34)$$

The properties of the optimal lottery between delegation and control are illustrated in Figures 2 and 3. The properties are intuitive. The lottery allows the mechanism to achieve any combinations of the agent's and the principal's values that lie on a line connecting the control values, given by the point P , and one of the delegation values, given by a point along the $(\hat{V}, \hat{J})$ frontier. The optimal lottery is the highest of such lines, which is the line that connects P and C , the point where the lottery line is tangent to the $(\hat{V}, \hat{J})$ frontier. It follows that, for any $V \in (V_P, V_C)$, the optimal lottery randomizes with between P and C with non-degenerate probabilities. For any $V \in [V_C, V_F]$, the lottery is useless, in the sense that the lottery assigns the hiring decision to the agent with probability 1.

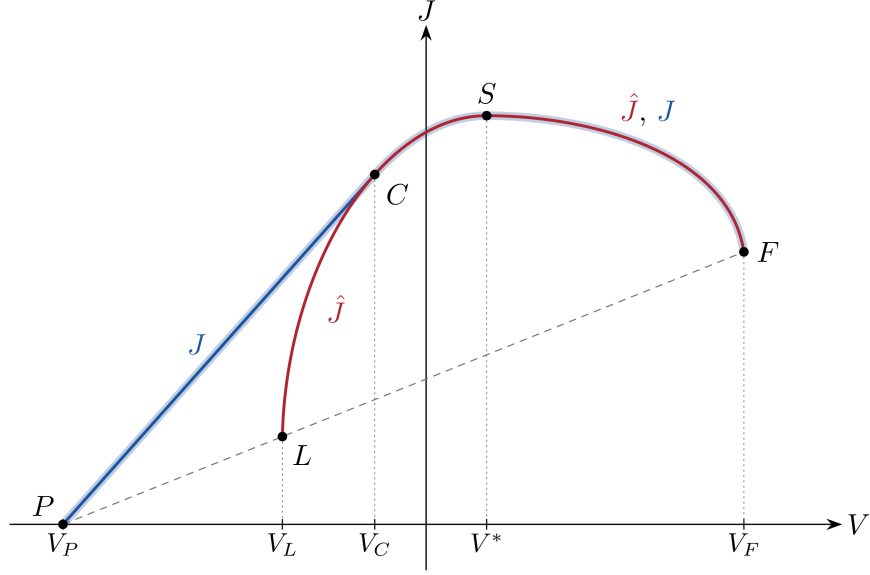


Figure 2: Value functions $J(V)$ and $\hat{J}(\hat{V})$.

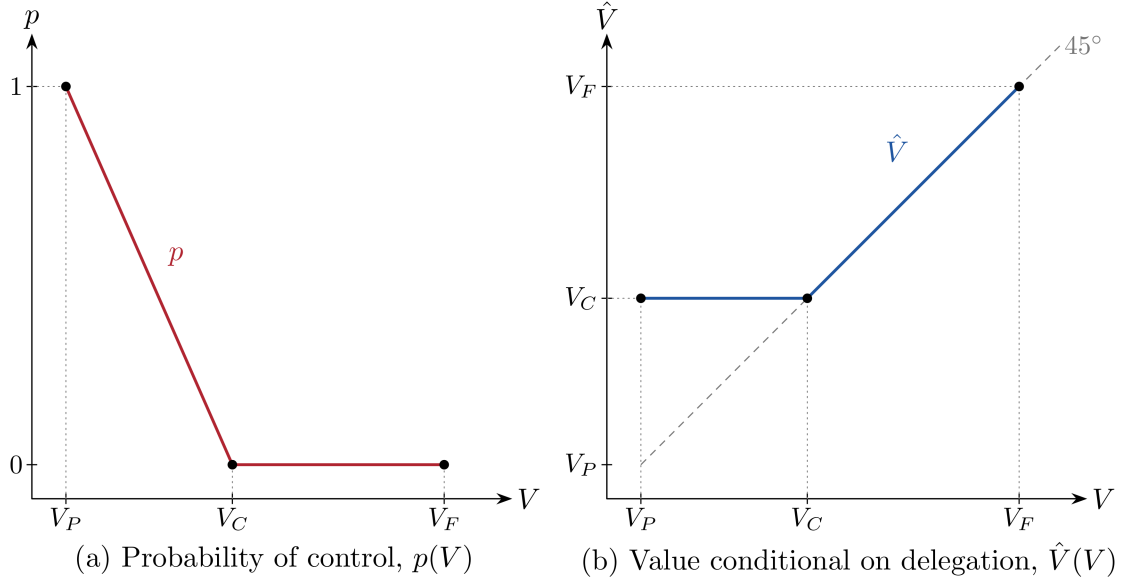


Figure 3: Optimal lottery p and $\hat{V}$ as function of the promised value V .

2.4 Optimal incentives in delegation

In this subsection, we characterize the solution to the second-stage problem in (2.4), which is the design of the mechanism conditional on the hiring decision being delegated to the agent in the current period. We focus the analysis on agent's promised values $\hat{V}$ in the interval $[V_C, V_F]$, since Proposition 1 implies that values outside of this interval are never reached.

In Proposition 2 we show that the solution to the second-stage problem is such that the agent's reservation quality R is in the interior of the support of the quality distribution F . Proposition 2 is useful, as it guarantees that there is a strictly positive probability that the agent hires the applicant, in which case the agent's continuation value is V_1 , and a strictly positive probability that the agent does not hire the applicant, in which case the agent's continuation value is V_0 .

The intuition behind Proposition 2 is simple. If R was smaller than $\underline{x}$, the agent would hire the applicant for sure. The agent's and the principal's values would then be a non-degenerate convex combination between the point P and some point along the $(\hat{V}, \hat{J})$ frontier. This convex combination is achievable in the first-stage problem (2.1) but, as shown in Proposition 1, a non-degenerate convex combination is not optimal. Therefore, $\hat{J}(\hat{V})$ would be strictly smaller than $J(\hat{V})$. Proposition 1, however, guarantees that $J(\hat{V})$ is equal to $\hat{J}(\hat{V})$ for all $\hat{V} \in [V_C, V_F]$. Similarly, if R was greater than $\bar{x}$, the agent would not hire the applicant. Therefore, the agent's and principal's values would be a convex combination between the point $(0, 0)$ and some point on the frontier $(\hat{V}, \hat{J})$. We show that such a convex combination also implies that $\hat{J}(\hat{V})$ would be strictly smaller than $J(\hat{V})$, which would contradict Proposition 1.

Proposition 2: (Optimal reservation quality) *For all $\hat{V} \in [V_C, V_F]$, the solution to the second-stage problem (2.4) is such that the reservation quality R belongs to $(\underline{x}, \bar{x})$.*

Proof: On the way to a contradiction, suppose that, for some $\hat{V}^a \in [V_C, V_F]$, the solution to the second-stage problem (2.4) specifies continuation values $V_0^a, V_1^a \in \mathcal{V}$ such that

$$R^a = \eta - \frac{\beta}{1 - \beta}(V_1^a - V_0^a) \leq \underline{x}. \quad (2.35)$$

In this case, the value to the agent is

$$\begin{aligned} \hat{V}^a &= (1 - \beta) \int_{\underline{x}}^{\bar{x}} (x - \eta) dF(x) + \beta V_1^a \\ &= (1 - \beta) V_P + \beta V_1^a. \end{aligned} \quad (2.36)$$

The value to the principal is

$$\begin{aligned} \hat{J}(\hat{V}^a) &= (1 - \beta) \int_{\underline{x}}^{\bar{x}} x dF(x) + \beta J(V_1^a) \\ &= (1 - \beta) J_P + \beta J(V_1^a) \\ &= (1 - \beta) J_P + \beta \hat{J}(V_1^a) \end{aligned} \quad (2.37)$$

The second line in (2.36) makes use of the definition of V_P . The second line in (2.37) makes use of the definition of J_P . The third line in (2.37) makes use of the fact that $\hat{V}^a > V_P$ implies $V_1^a > \hat{V}^a \geq V_C$ and, in light of Proposition 1, $J(V_1^a) = \hat{J}(V_1^a)$.

Now, consider the first-stage problem (2.1) for $\hat{V}^a$. The problem is

$$\begin{aligned} J(\hat{V}^a) &= \max_{p \in [0,1], \hat{V} \in \hat{\mathcal{V}}} pJ_P + (1-p)\hat{J}(\hat{V}), \\ \text{s.t. } \hat{V}^a &= pV_P + (1-p)\hat{V}. \end{aligned} \quad (2.38)$$

The lottery $(p, \hat{V}) = (\beta, V_1^a)$ is a feasible choice for (2.38), since $\beta \in (0, 1)$, $V_1^a \in \hat{\mathcal{V}}$, and $\beta V_P + (1-\beta)V_1^a$ is equal to $\hat{V}^a$. The lottery (β, V_1^a) is not the optimal choice for (2.38), since $\hat{V}^a \geq V_C$ and Proposition 1 states that the optimal lottery for any $V \geq V_C$ is $(p, \hat{V}) = (0, V)$. From (2.37), it follows that the value to the principal of the feasible and suboptimal lottery (β, V_1^a) is $\hat{J}(\hat{V}^a)$. Since the lottery (β, V_1^a) is suboptimal, it follows that $\hat{J}(\hat{V}^a)$ is strictly smaller than $J(\hat{V}^a)$. However, Proposition 1 states that $J(V) = \hat{J}(V)$ for all $V \geq V_C$, including for $V = \hat{V}^a$. We have thus reached the desired contradiction.

On the way to a second contradiction, suppose that, for some $\hat{V}^b \in [V_C, V_F]$, the solution to the second-stage problem (2.4) specifies continuation values $V_0^b, V_1^b \in \mathcal{V}$ such that

$$R = \eta - \frac{\beta}{1-\beta}(V_1^b - V_0^b) \geq \bar{x}. \quad (2.39)$$

In this case, the value to the agent is

$$\hat{V}^b = \beta V_0^b. \quad (2.40)$$

The value to the principal is

$$\begin{aligned} \hat{J}(\hat{V}^b) &= 0 + \beta J(V_0^b) \\ &< (1-\beta)J(0) + \beta J(V_0^b) \\ &\leq J((1-\beta)0 + \beta V_0^b) \\ &= J(\hat{V}^b) \end{aligned} \quad (2.41)$$

The second line in (2.41) makes use of the fact that $J(0) > 0$ because $J(V) \geq H(V) > 0$ for all $V > V_P$ and $V_P < 0$. The third line makes use of the fact that $J(V)$ is weakly concave. The fourth line makes use of the fact that $V_0^b = \hat{V}^b/\beta$. Comparing the first and the last line yields $\hat{J}(\hat{V}^b) < J(\hat{V}^b)$. However, $\hat{J}(\hat{V}^b) = J(\hat{V}^b)$, since Proposition 1 implies that $J(V) = \hat{J}(V)$ for all $V \geq V_C$, including for $V = \hat{V}^b$. We have thus reached the desired contradiction. ■

We now turn to the characterization of the optimal agent's continuation values in the

second-stage problem. The necessary condition for the optimality of an interior V_0 is

$$0 = \beta F(R)J'(V_0) + \beta F'(R)[J(V_0) - J(V_1)] \frac{dR}{dV_0} - (1 - \beta)RF'(R) \frac{dR}{dV_0} + \lambda \left\{ \beta F(R) + \beta F'(R)(V_0 - V_1) \frac{dR}{dV_0} - (1 - \beta)(R - \eta)F'(R) \frac{dR}{dV_0} \right\}, \quad (2.42)$$

where λ denotes the Lagrange multiplier on the promise-keeping constraint (2.5) and dR/dV_0 denotes the derivative of the agent's reservation quality (2.6) with respect to V_0 , i.e.

$$\frac{dR}{dV_0} = \frac{\beta}{1 - \beta}. \quad (2.43)$$

The first line on the right-hand side of (2.42) is the derivative of the principal's value with respect to V_0 . This derivative is given by the change in the principal's continuation value due to the change in V_0 and the change in the principal's flow payoff and continuation value due to change in the agent's reservation quality R . The second line on the right-hand side of (2.42) is the derivative of the agent's value with respect to V_0 multiplied by λ . This derivative is given by the change in the agent's continuation value due to the change in V_0 and the change in the agent's flow payoff and continuation value due to change in the agent's reservation quality R .

Using the fact that the agent's reservation quality R is given by (2.6) and that the derivative dR/dV_0 is given by (2.43), we can rewrite (2.42) as

$$J'(V_0) + \lambda = \frac{F'(R)}{F(R)} \left\{ R + \frac{\beta}{1 - \beta} [J(V_1) - J(V_0)] \right\}. \quad (2.44)$$

The necessary condition for optimality for an interior V_1 is

$$0 = \beta(1 - F(R))J'(V_1) + \beta F'(R)[J(V_0) - J(V_1)] \frac{dR}{dV_1} - (1 - \beta)RF'(R) \frac{dR}{dV_1} + \lambda \left\{ \beta(1 - F(R)) + \beta F'(R)(V_0 - V_1) \frac{dR}{dV_1} - (1 - \beta)(R - \eta)F'(R) \frac{dR}{dV_1} \right\}, \quad (2.45)$$

where dR/dV_1 denotes the derivative of the agent's reservation quality (2.6) with respect to V_1 , i.e.

$$\frac{dR}{dV_1} = -\frac{\beta}{1 - \beta}. \quad (2.46)$$

Using the fact that the agent's reservation quality R is given by (2.6) and that the derivative dR/dV_1 is given by (2.46), we can rewrite (2.45) as

$$J'(V_1) + \lambda = -\frac{F'(R)}{1 - F(R)} \left\{ R + \frac{\beta}{1 - \beta} [J(V_1) - J(V_0)] \right\}. \quad (2.47)$$

Lastly, the derivative $\hat{J}'(\hat{V})$ of the principal's value with respect to the agent's promised value $\hat{V}$ is equal to the negative of the Lagrange multiplier λ on the promise-keeping

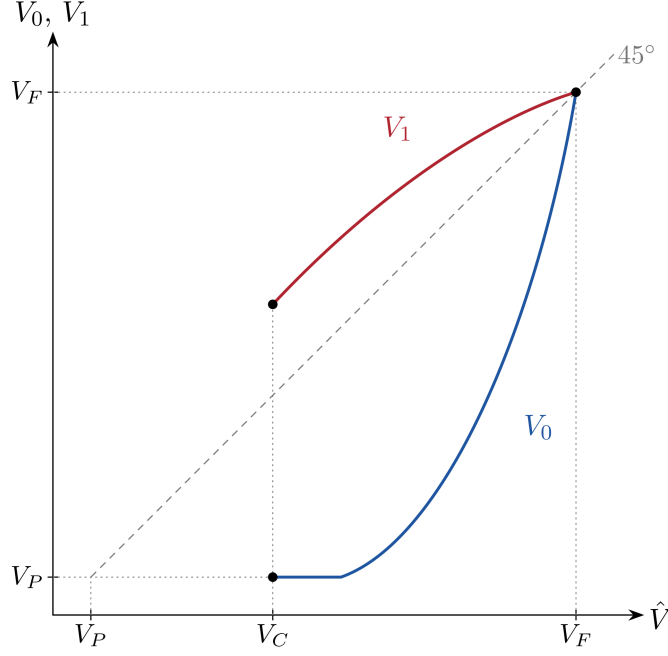


Figure 4: Optimal continuation values V_0 and V_1 as functions of the promised value $\hat{V}$.

constraint (2.5). That is,

$$\hat{J}'(\hat{V}) = -\lambda. \quad (2.48)$$

In the next proposition, we further characterize the solution to the second-stage problem (2.4). In order to keep the analysis simple, we assume that the agent's continuation values V_0 and V_1 are continuous functions of the value $\hat{V}$ promised to the agent. Using the first-order conditions (2.44) and (2.47) and the envelope condition (2.48), we show that, for any $\hat{V} \in (V_C, V_F)$, the agent's continuation value V_0 conditional on not hiring the applicant is strictly smaller than $\hat{V}$. Conversely, the agent's continuation value V_1 conditional on hiring the applicant is strictly greater than $\hat{V}$. Since $V_0 < V_1$, the agent's reservation quality R is strictly smaller than the agent's preferred quality cutoff η . These findings are intuitive. The mechanism rewards the agent for hiring and punishes the agent for not hiring so as to induce the agent to lower its reservation quality. We then show that, for $\hat{V} = V_C$, $V_0 = V_P$ and $V_1 > V_C$ and, hence, $R < \eta$. These findings are also intuitive. The mechanism uses the worst punishment $V_0 = V_P$ (certain and permanent loss of control over the hiring process) when the agent reaches its lowest possible value $\hat{V} = V_C$ in delegation. Lastly, we show that V_1 is strictly smaller than V_F for all $\hat{V} \in [V_C, V_F)$. It is easy to understand why the mechanism never rewards the agent with $V_1 = V_F$ (permanent control over the hiring process). If $V_1 = V_F$, the agents will use its preferred reservation quality $R = \eta$. If $V_1 = V_F - \epsilon$, the agent will use a reservation quality slightly below its preferred one. This has second-order effects on the value of the agent, but first-order effects on the value of the principal. Hence, the marginal cost to the principal of increasing the reward V_1 from $V_F - \epsilon$ to V_F is infinitely large. These characterization results are

illustrated in Figure 4.

Proposition 3: (Optimal incentives in delegation). *The solution to the second-stage problem (2.4) is such that:*

1. For all $\hat{V} \in (V_C, V_F)$, the agent's continuation value conditional on not hiring is $V_0 < \hat{V}$, the agent's continuation value conditional on hiring the applicant is $V_1 > \hat{V}$, and the reservation quality is $R < \eta$.
2. For all $\hat{V} \in [V_C, V_F)$, the agent's continuation value conditional on hiring the applicant is $V_1 < V_F$.
3. For $\hat{V} = V_C$, the agent's continuation value conditional on not hiring the applicant is $V_0 = V_P$, the agent's continuation value conditional on hiring the applicant is $V_1 > V_C$, and the reservation quality is $R < \eta$. For $\hat{V} = V_F$, $V_0 = V_F$, $V_1 = V_F$, and $R = \eta$.

Proof of part (1). We break down the proof in three claims.

Claim 1: For any value $\hat{V}$ promised to the agent such that $\hat{V} \in [V_C, V_F)$, the continuation values V_0 and V_1 are different.

On the way to a contradiction, suppose that the continuation values V_0 and V_1 are both equal to some V , with $V \in (V_P, V_F)$. The necessary conditions (2.44) and (2.47) become

$$\begin{aligned} J'(V_0) + \lambda &= \frac{F'(R)}{F(R)} \left\{ R + \frac{\beta}{1-\beta} [J(V_1) - J(V_0)] \right\} = \frac{F'(\eta)}{F(\eta)} \eta, \\ J'(V_1) + \lambda &= -\frac{F'(R)}{1-F(R)} \left\{ R + \frac{\beta}{1-\beta} [J(V_1) - J(V_0)] \right\} = -\frac{F'(\eta)}{1-F(\eta)} \eta, \end{aligned} \tag{2.49}$$

where the second equality in both lines makes use of the fact that $V_0 = V_1$ implies $J(V_1) = J(V_0)$ and $R = \eta$. The necessary conditions (2.49) imply that $J'(V_0) + \lambda > 0 > J'(V_1) + \lambda$. However, $V_0 = V_1$ implies that $J'(V_0) + \lambda = J'(V_1) + \lambda$. A contradiction.

On the way to another contradiction, suppose that the continuation values V_0 and V_1 are both equal to V_P . The promise-keeping constraint (2.5) becomes

$$\begin{aligned} \hat{V} &= (1-\beta) \int_R (x-\eta) dF(x) + \beta V_P \\ &= (1-\beta) \int_\eta (x-\eta) dF(x) + \beta V_P \\ &= (1-\beta) V_F + \beta V_P = V_L, \end{aligned} \tag{2.50}$$

where the second line makes use of the fact that $V_0 = V_1$ implies $R = \eta$, and the third line makes use of the definitions of V_F and V_L . Since $\hat{V} \geq V_C$ and $V_C > V_L$, we have reached a contradiction.

Lastly, suppose that the continuation values V_0 and V_1 are both equal to V_F . In this

case, the promise-keeping constraint (2.5) becomes

$$\begin{aligned}
\hat{V} &= (1 - \beta) \int_R (x - \eta) dF(x) + \beta V_F \\
&= (1 - \beta) \int_\eta (x - \eta) dF(x) + \beta V_F \\
&= (1 - \beta) V_F + \beta V_F = V_F.
\end{aligned} \tag{2.51}$$

where the second line makes use of the fact that $V_0 = V_1$ implies $R = \eta$, and the third lines makes use of the definition of V_F . Since $\hat{V} < V_F$, we have reached a contradiction.

Claim 2: For any value $\hat{V}$ promised to the agent such that $\hat{V} \in [V_C, V_F)$, the continuation values V_0 and V_1 are such that $V_0 < V_1$.

Since $V_0 \neq V_1$ for all $\hat{V} \in [V_C, V_F)$ and V_0 and V_1 are continuous functions of $\hat{V}$, it follows that either $V_0 > V_1$ for all $\hat{V} \in [V_C, V_F)$, or $V_0 < V_1$ for all $\hat{V} \in [V_C, V_F)$. On the way to a contradiction, suppose that $V_0 > V_1$ for all $\hat{V} \in [V_C, V_F)$. If that is the case, the agent's reservation quality R is strictly greater than η for all $\hat{V} \in [V_C, V_F)$. For $\hat{V} = V_F$, the only feasible continuation values are $V_0 = V_1 = V_F$ and, hence, the agent's reservation quality R is equal to η . Therefore, whenever hiring is delegated to the agent, the principal's flow payoff is

$$\begin{aligned}
u &\leq \max_{R \geq \eta} \int_R x dF(x) \\
&= \int_\eta x dF(x) = (1 - \beta) J_F,
\end{aligned} \tag{2.52}$$

where the second line uses the fact that the integral in the first line is strictly decreasing in R for all $R > 0$. Whenever hiring is controlled by the principal, the principal's flow payoff is

$$u = \int_{\underline{x}}^{\bar{x}} x dF(x) = (1 - \beta) J_P. \tag{2.53}$$

Since the value of the mechanism to the principal is the discounted sum of flow payoffs, it follows that

$$J(\hat{V}) \leq \frac{(1 - \beta) \max \{J_F, J_P\}}{1 - \beta} = J_F. \tag{2.54}$$

The above inequality holds for any $\hat{V} \in [V_L, V_F]$. However, $V^* \in [V_C, V_F)$ and $J(V^*) > J_F$. A contradiction.

Claim 3: For any value $\hat{V}$ promised to the agent such that $\hat{V} \in (V_C, V_F)$, the continuation values V_0 and V_1 are such that $V_0 < \hat{V} < V_1$.

Consider some value $\hat{V} \in (V_C, V_F)$ promised to the agent. The continuation values V_0 and V_1 are such that $V_0 < V_1$. Suppose that V_0 and V_1 are both interior. Using the envelope condition (2.48) to substitute out the Lagrange multiplier λ , we can rewrite the necessary

conditions (2.44) and (2.47) as

$$\begin{aligned} J'(V_0) - \hat{J}'(\hat{V}) &= \frac{F'(R)}{F(R)} \left\{ R + \frac{\beta}{1-\beta} [J(V_1) - J(V_0)] \right\}, \\ J'(V_1) - \hat{J}'(\hat{V}) &= -\frac{F'(R)}{1-F(R)} \left\{ R + \frac{\beta}{1-\beta} [J(V_1) - J(V_0)] \right\}. \end{aligned} \quad (2.55)$$

Subtracting the second equation in (2.55) from the first yields

$$J'(V_0) - J'(V_1) = \left[\frac{F'(R)}{F(R)} + \frac{F'(R)}{1-F(R)} \right] \left\{ R + \frac{\beta}{1-\beta} [J(V_1) - J(V_0)] \right\}. \quad (2.56)$$

Suppose that $V_0 \geq V_C$. Since $V_0 \geq V_C$ and $V_1 > V_0 \geq V_C$, Proposition 1 guarantees that $J'(V_0)$ is equal to $\hat{J}'(V_0)$ and $J'(V_1)$ is equal to $\hat{J}'(V_1)$. Since $V_1 > V_0$ and $\hat{J}$ is strictly concave, $\hat{J}'(V_0)$ is strictly greater than $\hat{J}'(V_1)$. Putting these observations together yields $J'(V_0) > J'(V_1)$. Since $J'(V_0) > J'(V_1)$, (2.56) implies

$$R + \frac{\beta}{1-\beta} [J(V_1) - J(V_0)] > 0. \quad (2.57)$$

Using (2.57) in the first necessary condition in (2.55) yields $J'(V_0) > \hat{J}'(\hat{V})$. Since $J'(V_0)$ is equal to $\hat{J}'(V_0)$ and $\hat{J}$ is strictly concave, $V_0 < \hat{V}$. Using (2.57) into the second necessary condition in (2.55) yields $J'(V_1) < \hat{J}'(\hat{V})$. Since $J'(V_1)$ is equal to $\hat{J}'(V_1)$ and $\hat{J}$ is strictly concave, $V_1 > \hat{V}$.

Now, suppose that $V_0 < V_C$. Since $V_0 < V_C$, Proposition 1 guarantees that $J'(V_0)$ is equal to $\hat{J}'(V_C)$. Since $\hat{V} > V_C$ and $\hat{J}$ is strictly concave, it follows that $\hat{J}'(\hat{V}) < \hat{J}'(V_C)$. Putting these observations together yields $\hat{J}'(\hat{V}) < J'(V_0)$. Since $\hat{J}'(\hat{V}) < J'(V_0)$, the first necessary condition in (2.55) implies

$$R + \frac{\beta}{1-\beta} [J(V_1) - J(V_0)] > 0. \quad (2.58)$$

Using (2.58) in the second necessary condition in (2.55) yields $J'(V_1) < \hat{J}'(\hat{V})$. Since $\hat{J}'(\hat{V})$ is equal to $J'(\hat{V})$ and J is weakly concave, $V_1 > \hat{V}$. Moreover, since $V_0 < V_C$ and $\hat{V} > V_C$, we also have $V_0 < \hat{V}$.

Similar arguments can be used to prove that $V_0 < \hat{V} < V_1$ even when the agent's continuation values are not interior. ■

Proof of part (2). We first establish that $\hat{J}'(\hat{V})$ diverges to $-\infty$ for $\hat{V} \rightarrow V_F$. Recall that $V_0 < \hat{V}$ for all $\hat{V} \in (V_C, V_F)$ and $V_0 = V_F$ for $\hat{V} = V_F$. By continuity of V_0 , it follows that there exists a $\delta > 0$ such that $V_0 \in (V_C, V_F)$ for all $\hat{V} \in (V_F - \delta, V_F)$. Hence, for all $\hat{V} \in (V_F - \delta, V_F)$, V_0 is interior and its optimality condition is

$$J'(V_0) - \hat{J}'(\hat{V}) = \frac{F'(R)}{F(R)} \left\{ R + \frac{\beta}{1-\beta} [J(V_1) - J(V_0)] \right\}, \quad (2.59)$$

On the way to a contradiction, let us suppose that the limit for $\hat{V} \rightarrow V_F$ of $\hat{J}'(\hat{V})$ exists and is equal to some κ . In this case, the limit for $\hat{V} \rightarrow V_F$ of the left-hand side of (2.59) is 0, since $J'(V_0) = \hat{J}'(V_0)$ for all $V_0 \geq V_C$, and V_0 converges to V_F . The limit for $\hat{V} \rightarrow V_F$ of the right-hand side of (2.59) is $F'(\eta)\eta/F(\eta) > 0$, since V_0 and V_1 converge to V_F and R converges to η . Since V_0 converges to V_F , $\hat{J}'(\hat{V}_0)$ converges to κ for $\hat{V} \rightarrow V_F$, it follows that the left-hand side of (2.59) converges to 0. We have thus reached a contradiction, and $\hat{J}'(\hat{V})$ must diverge to $-\infty$ for $\hat{V} \rightarrow V_F$.

Next, we establish that $V_1 < V_F$ for all $\hat{V} \in [V_C, V_F)$. For any $\hat{V} \in [V_C, V_F)$, $V_1 = V_F$ is optimal only if

$$J'(V_F) - \hat{J}'(\hat{V}) \geq -\frac{F'(R)}{1 - F(R)} \left\{ R + \frac{\beta}{1 - \beta} [J_F - J(V_0)] \right\}. \quad (2.60)$$

The condition above cannot be satisfied. The left-hand side of (2.60) is equal to $-\infty$ since $J'(V_F) = \hat{J}'(V_F)$ and $\hat{J}'(V_F) = -\infty$. The right-hand side of (2.60) is bounded. Indeed, R and $F'(R)/(1 - F(R))$ are bounded because $R \in (\underline{x}, \bar{x})$, and $J_F - J(V_0)$ is bounded between $J_F - J^*$ and $J_F - J_P$. ■

Proof of part (3). Consider the second-stage problem for $\hat{V} = V_C$. From part (1) and the continuity of V_0 and V_1 with respect to $\hat{V}$, it follows that $V_0 < V_1$ and $V_0 \leq V_C \leq V_1$. From part (2), it follows that $V_1 < V_F$.

We want to establish that $V_0 = V_P$. On the way to a contradiction, let us suppose that $V_0 > V_P$. If this is the case, $V_0 \in (V_P, V_C]$ and the necessary condition for the optimality of V_0 is

$$J'(V_0) - \hat{J}'(V_C) = \frac{F'(R)}{F(R)} \left\{ R + \frac{\beta}{1 - \beta} [J(V_1) - J(V_0)] \right\}. \quad (2.61)$$

Since $V_0 \in [V_P, V_C]$, it follows from Proposition 1 that $J'(V_0) = \kappa \equiv (J_C - J_P)/(V_C - V_P)$. It also follows from Proposition 1 that $\hat{J}'(V_C) = J'(V_C)$, which, in turn, is equal to κ . Hence, the left-hand side of (2.61) is equal to zero and

$$R + \frac{\beta}{1 - \beta} [J(V_1) - J(V_0)] = 0. \quad (2.62)$$

Since $V_1 \in [V_C, V_F)$, the necessary condition for the optimality of V_1 is

$$J'(V_1) - \hat{J}'(V_C) = -\frac{F'(R)}{1 - F(R)} \left\{ R + \frac{\beta}{1 - \beta} [J(V_1) - J(V_0)] \right\}. \quad (2.63)$$

The right-hand side of (2.63) is equal to zero and, hence, $J'(V_1) = \hat{J}'(V_C)$. Since $V_1 \geq V_C$, $J'(V_1) = \hat{J}'(V_1)$ and, hence, $\hat{J}'(V_1) = \hat{J}'(V_C)$. Since $\hat{J}$ is strictly concave, $\hat{J}'(V_1) = \hat{J}'(V_C)$ implies $V_1 = V_C$.

The incentive-compatibility constraint (2.6) implies

$$R = \eta - \frac{\beta}{1 - \beta} (V_1 - V_0). \quad (2.64)$$

The optimality condition (2.62) for V_0 implies

$$R = -\frac{\beta}{1-\beta} [J(V_1) - J(V_0)]. \quad (2.65)$$

Combing the expressions above and using the fact that $V_1 = V_C$ yields

$$J(V_C) - J(V_0) = V_C - V_0 - \frac{1-\beta}{\beta} \eta. \quad (2.66)$$

Since $J(V_C) - J(V_0) = \kappa(V_C - V_0)$, (2.66) gives

$$V_C - V_0 = \frac{1-\beta}{\beta} \frac{\eta}{1-\kappa}, \quad J(V_C) - J(V_0) = \frac{1-\beta}{\beta} \frac{\kappa\eta}{1-\kappa}. \quad (2.67)$$

Since $V_0 < V_1$, it follows that $V_C - V_0 > 0$ and, hence, κ must be strictly smaller than 1.

Since the continuation values V_0 and V_1 are optimal, they must deliver the value $\hat{J}(V_C)$ to the principal:

$$\hat{J}(V_C) = (1-\beta) \int_R x dF(x) + \beta [F(R)J(V_0) + (1-F(R))J(V_1)]. \quad (2.68)$$

Since $\hat{J}(V_C) = J(V_C)$ and $V_1 = V_C$, we can rewrite (2.68) as

$$\begin{aligned} J(V_C) &= \int_R x dF(x) + \frac{\beta}{1-\beta} F(R)(J(V_0) - J(V_C)) \\ &= \int_R x dF(x) - F(R) \frac{\kappa\eta}{1-\kappa}, \end{aligned} \quad (2.69)$$

where the second line makes use of (2.67). Since the continuation values V_0 and V_1 are optimal, they are feasible and they must deliver the promised value V_C to the agent:

$$V_C = (1-\beta) \int_R (x - \eta) dF(x) + \beta [F(R)V_0 + (1-F(R))V_1]. \quad (2.70)$$

Since $V_1 = V_C$, we can rewrite (2.70) as

$$\begin{aligned} V_C &= \int_R x dF(x) - \eta(1-F(R)) + \frac{\beta}{1-\beta} F(R)(V_0 - V_C) \\ &= \int_R x dF(x) - \eta(1-F(R)) - F(R) \frac{\eta}{1-\kappa}, \end{aligned} \quad (2.71)$$

where the second line makes use of (2.67). Combining (2.69) and (2.71) yields

$$J(V_C) = V_C + \eta.$$

Since $J(V_C) = V_C + \eta$, $J_P = 0$, and $V_P = -\eta$, it follows that

$$\kappa \equiv \frac{J_C - J_P}{V_C - V_P} = \frac{V_C + \eta}{V_C + \eta} = 1.$$

Since (2.67) implies $\kappa < 1$, we have reached the desired contradiction. Therefore, $V_0 = V_P$. The same argument can be used to reach a contradiction when $V_0 = V_P$ and $V_1 = V_C$. Therefore, $V_1 > V_C$. ■

Propositions 1, 2, and 3 provide a complete characterization of the optimal mechanism. The initial value of the mechanism to the principal is J^* , and the initial value of the mechanism to the agent is V^* . The initial value J^* of the mechanism to the principal is greater than the value J_P that the uninformed principal could obtain by making all the hiring decisions without any input from the agent. The initial value J^* of the mechanism to the principal is also greater than the value J_F that the principal could obtain by delegating all the hiring decisions to the informed but biased agent. The mechanism uses the threat of giving control of the hiring process to the uninformed principal to induce the agent to select applicants in a way that partly reflects the preferences of the principal. In doing so, the mechanism succeeds in providing the principal with a value J^* greater than both J_P and J_F .

The initial value V^* of the mechanism to the agent is greater than V_C , which means that the mechanism starts by delegating the hiring decision to the agent. The agent hires applicants whose quality is above a cutoff R , where the cutoff R is tilted away from the value η that would maximize the flow payoff of the biased agent, and towards the value 0 that would maximize the flow payoff of the unbiased principal. The quality cutoff R chosen by the agent is lower than η because the mechanism rewards the agent with an increase in value when it does hire an applicant, and it punishes the agent with a decrease in value when it does not. The increase in value pushes the agent away from the region $[V_P, V_C)$, where the mechanism may assign the hiring decision to the principal. The increase in value pushes the agent towards V_F , where the mechanism lets the agent choose applicants according to its preferences ($R = \eta$). Heuristically, the mechanism increases the agent's value by letting the agent control the hiring process for a longer period of time, and by letting the agent hire applicants that are closer to its preferences. Conversely, the mechanism lowers the agent's value by letting the agent control the hiring process for a shorter period of time, and by inducing the agent to hire applicants that are further from its preferences.

As long as the agent keeps hiring applicants, the value of the mechanism to the agent increases monotonically and converges to V_F , which is the value that the agent would obtain if it had permanent control over the hiring process. Yet, no matter how many applicants the agent hires, the value of the mechanism to the agent never reaches V_F . In other words, no matter how many applicants the agent hires, the mechanism never gives it permanent control over the hiring process. Indeed, as the agent's value approaches

V_F , the hiring cutoff R moves closer and closer to the one that maximizes the agent's flow payoff, and further and further away from the one that maximizes the principal's flow payoff. Hence, as the agent's value approaches V_F , the marginal cost (in terms of principal's value) of increasing the agent's value becomes infinitely large.

As long as the agent keeps rejecting applicants, the value of the mechanism to the agent falls monotonically and, after sufficiently many rejections, the value of the mechanism to the agent enters the region $[V_P, V_C)$. In this region, the mechanism runs a lottery to assign the control of the hiring process to either the principal or the agent. With some strictly positive probability, the mechanism permanently assigns the control of the hiring process to the principal. In this case, the principal hires every applicant, the value of the mechanism to the principal is J_P , and the value of the mechanism to the agent is V_P . With complementary probability, the mechanism keeps the agent in control of the hiring process. In this case, the agent gets to decide whether to hire the next applicant or not. If the agent hires the applicant, the value of the mechanism to the agent exits the region $[V_P, V_C)$. If the agent rejects the next applicant, the mechanism permanently assigns the control of the hiring process to the principal with probability 1.

The mechanism features ex-post inefficiencies. For instance, when the mechanism permanently assigns the control of the hiring process to the principal, the value to the agent is V_P , and the value to the principal is J_P . If the mechanism permanently assigned control of the hiring process to the agent, the value to the agent would be $V_F > V_P$, and the value to the principal would be $J_F > J_P$. More generally, when the mechanism reaches any point on the upward-sloping part of the (V, J) frontier, the mechanism could be modified so as to make both the agent and the principal better off. These ex-post inefficiencies, however, are necessary to maximize the ex-ante value of the mechanism to the principal. Indeed, if the mechanism never assigned the control of the hiring process to the principal, the value of the mechanism to the principal would be $J_F < J^*$, because the agent would select candidates according to its biased preferences without any fear of retribution.

Figure 5 shows properties of the optimal mechanism for $\beta = 0.9$, $\eta = 0.5$, and $F(x)$ uniform on the interval $[-1, 1]$. Panel (a) plots the principal's values, J and $\hat{J}$, in the first and in the second stage as a function of the agent's values, V and $\hat{V}$, in the first and in the second stage. Panel (b) plots the agent's continuation values, V_0 and V_1 , conditional on rejecting or hiring the applicant as a function of the agent's value $\hat{V}$. These two panels show that, consistently with our simplifying assumption, $\hat{J}$ is strictly concave in $\hat{V}$, and V_0 and V_1 are continuous in $\hat{V}$. Moreover, in this example, the sufficient condition for the differentiability of $\hat{J}$ is satisfied.

Panel (c) plots the reservation quality R as a function of the agent's value $\hat{V}$. Panel (d) plots the expected time until the mechanism assigns the control of the hiring process to the principal as a function of the agent's value $\hat{V}$. The reservation quality R is strictly increasing in $\hat{V}$, except for a region near V_C where the solution of the delegation problem

is at a corner. The expected time until the mechanism assigns the control of the hiring process to the principal is also strictly increasing in $\hat{V}$. Overall, these two panels confirm that the mechanism delivers higher value to the agent by moving the reservation quality closer to the agent's preferred value of η , and by letting the agent control the hiring process for a longer period of time.

Panel (e) plots the dynamics of the probability distribution of the value of the mechanism to the agent. On average, the value of the mechanism to the agent declines over time. The 25th percentile of the probability distribution of the agent's value also decreases over time, and it reaches the minimum V_P after about 50 periods. The 75th percentile of the probability distribution of the agent's value is at first increasing and then decreasing over time, and it reaches the minimum V_P after about 175 periods. Any percentile of the probability distribution of the agent's value eventually reaches the minimum V_P . These findings show that the agent's expected value falls over time, and that the agent's value converges towards V_P almost surely. These findings are reminiscent of the "immiseration in the limit" results in Thomas and Worrall (1990) and Atkeson and Lucas (1992).

3 Optimal control

In Section 2, we restricted attention to mechanisms such that: (i) if the mechanism hands the control of the hiring process to the principal, it does so forever; (ii) if the mechanism hands control of the hiring process to the principal, the mechanism instructs the principal to hire every applicant. In this section, we show that these seemingly arbitrary restrictions to the mechanism space are without loss in generality.

Let us start by relaxing restriction (i) and consider mechanisms that allow for the hiring decision to be either controlled by the principal or delegated to the agent in every period and after every history. We maintain the restriction to mechanisms that require the principal to hire the applicant whenever the principal controls the hiring decision. As in Section 2, the principal's mechanism design problem can be written recursively as a two-stage problem. The first-stage problem is

$$\begin{aligned} \Gamma(V) = & \max_{p, \tilde{V}, \hat{V}} p\tilde{J}(\tilde{V}) + (1-p)\hat{J}(\hat{V}), \\ \text{s.t. } & V = p\tilde{V} + (1-p)\hat{V}, \\ & p \in [0, 1], \quad \tilde{V} \in \mathcal{V}, \quad \hat{V} \in \hat{\mathcal{V}}, \end{aligned} \tag{3.1}$$

where $\tilde{V}$ and $\tilde{J}(\tilde{V})$ denote the second-stage values to the agent and the principal conditional on the hiring decision being controlled by the principal in the current period, and $\hat{V}$ and $\hat{J}(\hat{V})$ denote the second-stage values to the agent and the principal conditional on the hiring decision being delegated to the agent in the current period. The value $\hat{J}(\hat{V})$ is

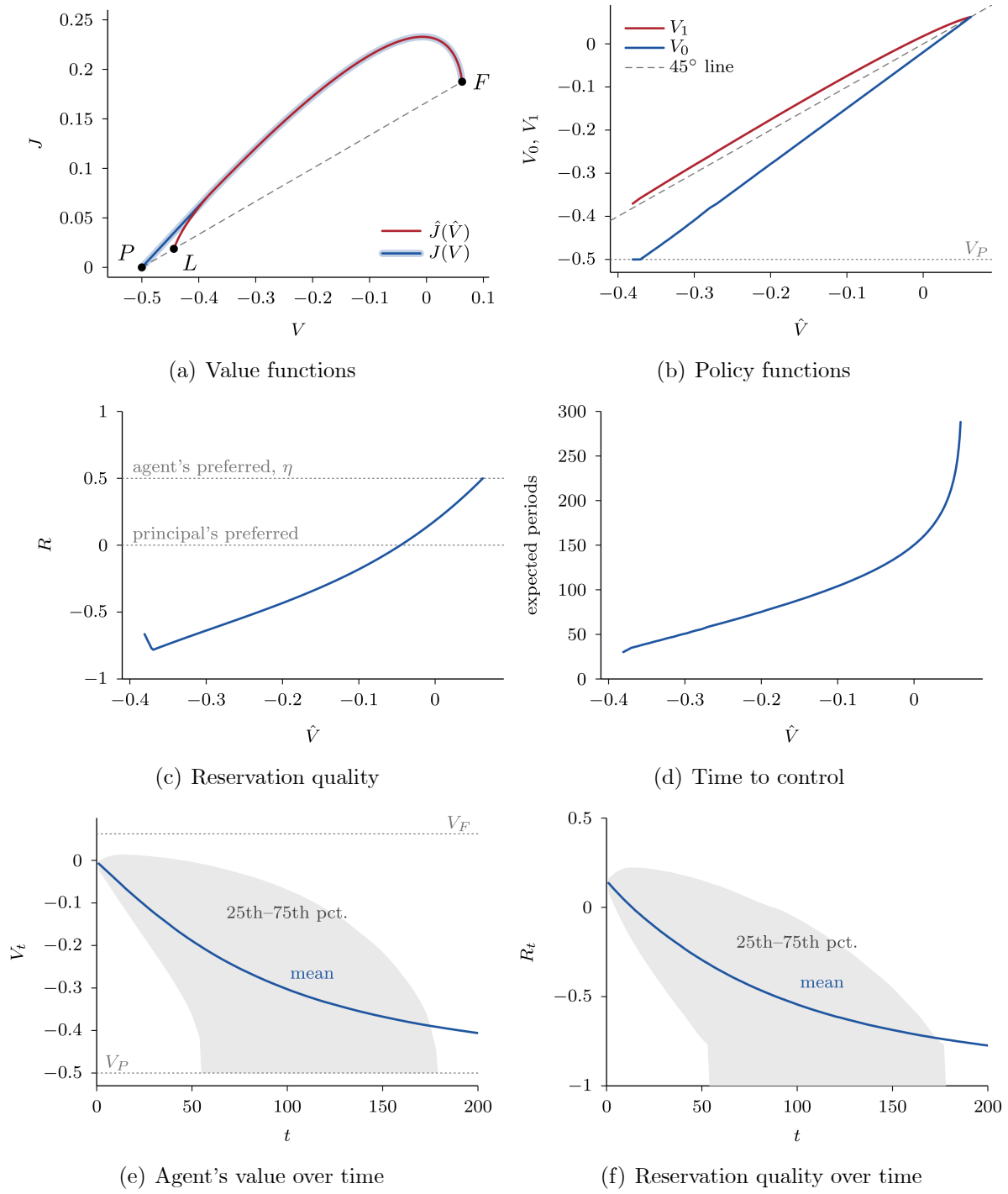


Figure 5: Properties of the mechanism for $\beta = 0.9$, $\eta = 0.5$, and $F(x)$ uniform on $[-1, 1]$.

given by (2.4). The value $\tilde{J}(\tilde{V})$ is given by

$$\begin{aligned}\tilde{J}(\tilde{V}) &= (1 - \beta) \int_{\underline{x}}^{\bar{x}} x dF(x) + \beta \Gamma(V_+) \\ &= (1 - \beta) J_P + \beta \Gamma(V_+),\end{aligned}\tag{3.2}$$

where

$$\begin{aligned}\tilde{V} &= (1 - \beta) \int_{\underline{x}}^{\bar{x}} (x - \eta) dF(x) + \beta V_+ \\ &= (1 - \beta) V_P + \beta V_+.\end{aligned}\tag{3.3}$$

Consider some arbitrary value V_0 of the first-stage problem to the agent. Using (3.2) to substitute $\tilde{J}(\tilde{V})$ on the right-hand side of (3.1), we can write $\Gamma(V_0)$ as

$$\Gamma(V_0) = p_0(1 - \beta) J_P + (1 - p_0) \hat{J}(\hat{V}_0) + p_0 \beta \Gamma(V_1),\tag{3.4}$$

where p_0 , $\tilde{V}_0$ and $\hat{V}_0$ are the solution to (3.1) for $V = V_0$, and V_1 is obtained from (3.3) and hence it is such that $\tilde{V}_0 = (1 - \beta) V_P + \beta V_1$. Note that the coefficients in front of J_P , $\hat{J}(\hat{V}_0)$ and $\Gamma(V_1)$ are positive and they sum to 1.

Using (3.1) to substitute $\Gamma(V_1)$ on the right-hand side of (3.4) and (3.2) to substitute $\tilde{J}(\tilde{V})$ in (3.1), we can write $\Gamma(V_0)$ as

$$\begin{aligned}\Gamma(V_0) &= p_0(1 - \beta) J_P + \beta p_0 p_1 (1 - \beta) J_P \\ &\quad + (1 - p_0) \hat{J}(\hat{V}_0) + \beta p_0 (1 - p_1) \hat{J}(\hat{V}_1) + \beta^2 p_0 p_1 \Gamma(V_2),\end{aligned}\tag{3.5}$$

where p_1 , $\tilde{V}_1$ and $\hat{V}_1$ are the solution to (3.1) for $V = V_1$, and $\tilde{V}_1$ is obtained from (3.3) and hence it is such that $\tilde{V}_1 = (1 - \beta) V_P + \beta V_2$. Again, the coefficients in front of J_P , $\hat{J}(\hat{V}_0)$, $\hat{J}(\hat{V}_1)$ and $\Gamma(V_2)$ are positive and they sum to 1. This is because, when we replace $\Gamma(V_1)$ on the right-hand side of (3.4) with $p_1 [(1 - \beta) J_P + \beta \Gamma(V_2)] + (1 - p_1) \hat{J}(\hat{V}_1)$, the coefficients on J_P , $\Gamma(V_2)$ and $\hat{J}(\hat{V}_1)$ are positive and they sum to 1.

Repeating the same steps as above T times, we can write $\Gamma(V_0)$ as

$$\begin{aligned}\Gamma(V_0) &= \left[(1 - \beta) \sum_{t=0}^T \beta^t \left(\prod_{i=0}^t p_i \right) \right] J_P \\ &\quad + \sum_{t=0}^T \left[\beta^t (1 - p_t) \left(\prod_{i=0}^{t-1} p_i \right) \hat{J}(\hat{V}_t) \right] + \left(\beta^{T+1} \prod_{i=0}^T p_i \right) \Gamma(V_{T+1}).\end{aligned}\tag{3.6}$$

For $T \rightarrow \infty$, the expression above becomes

$$\begin{aligned}\Gamma(V_0) &= \left[(1 - \beta) \sum_{t=0}^{\infty} \beta^t \left(\prod_{i=0}^t p_i \right) \right] J_P \\ &\quad + \sum_{t=0}^{\infty} \left[\beta^t (1 - p_t) \left(\prod_{i=0}^{t-1} p_i \right) \hat{J}(\hat{V}_t) \right].\end{aligned}\tag{3.7}$$

Again, the coefficients in front of J_P and $\hat{J}(\hat{V}_t)$ are positive and sum up to 1.

Note that we can write (3.7) as

$$\begin{aligned}
\Gamma(V_0) &= \bar{p}J_P + (1 - \bar{p}) \sum_{t=0}^{\infty} \left[\frac{\beta^t(1 - p_t) (\prod_{i=0}^{t-1} p_i)}{1 - \bar{p}} \hat{J}(\hat{V}_t) \right] \\
&\leq \bar{p}J_P + (1 - \bar{p}) \hat{J}(\bar{V}), \\
&\leq J(V_0),
\end{aligned} \tag{3.8}$$

where $\bar{p}$ and $\bar{V}$ are given by

$$\bar{p} = (1 - \beta) \sum_{t=0}^{\infty} \beta^t \left(\prod_{i=0}^t p_i \right), \text{ and } \bar{V} = \frac{\beta^t(1 - p_t) (\prod_{i=0}^{t-1} p_i)}{1 - \bar{p}} \hat{V}_t. \tag{3.9}$$

The second line in (3.8) makes use of the fact that $\hat{J}$ is concave. The third line in (3.8) makes use of the fact that the lottery $(p, \hat{V}) = (\bar{p}, \bar{V})$ is a feasible choice in the first-stage problem (2.1) for $V = V_0$.

The inequality in the third line of (3.8), $\Gamma(V_0) \leq J(V_0)$, shows that a mechanism that allows the hiring decision to be assigned to either the agent or the principal in every period and after every history cannot give a higher value to the principal than a mechanism that, when it assigns the hiring decision to the principal, it does so forever. Since a mechanism that assigns the hiring decision to the principal once and for all is a special case of a mechanism that allows the hiring decision to be assigned to either party in every period, it follows that $\Gamma(V_0) \geq J(V_0)$. Combining these observations yields $\Gamma(V_0) = J(V_0)$. That is, the optimal mechanism that allows the hiring decision to be given to either party in every period is payoff-equivalent to the optimal mechanism in which control by the principal is permanent. The intuition for this finding is simple and can be seen from the analysis above. Namely, a lottery between temporary control by the principal and delegation to the agent is equivalent to a lottery between permanent control by the principal and delegation to the agent that attaches a lower probability on control.

Next, we relax restriction (ii) by considering mechanisms that, when the principal is in control of the hiring decision, are allowed to instruct the principal to either hire or not hire the applicants. We maintain the restriction to mechanisms that hand the hiring decision to the principal on a permanent basis. As in Section 2, the principal's mechanism design problem can be written recursively as a two-stage problem. The first-stage problem is

$$\begin{aligned}
\Gamma(V) &= \max_{p, q, \hat{V}} pJ_P + qJ_Q + (1 - p - q)\hat{J}(\hat{V}), \\
\text{s.t. } &V = pV_P + qV_Q + (1 - p - q)\hat{V}, \\
&p, q \in [0, 1], \quad p + q \leq 1, \quad \hat{V} \in \hat{\mathcal{V}},
\end{aligned} \tag{3.10}$$

where V_Q and J_Q are the values to the agent and the principal if the principal has control over hiring and he is prescribed not hire the applicants, i.e. $(V_Q, J_Q) = (0, 0)$, V_P and J_P are the values to the agent and the principal if the principal has control over hiring

and is prescribed to hire the applicants, i.e. $(V_P, J_P) = (-\eta, 0)$, and $\hat{V}$ and $\hat{J}(\hat{V})$ are the values to the agent and the principal if the hiring decision is delegated to the agent in the current period.

Let p_0, q_0 and $\hat{V}_0$ denote a solution to (3.10) for V equal to some arbitrary V_0 . Suppose that $V_Q \geq V_L$. Then, the value $\Gamma(V_0)$ to the principal is such that

$$\begin{aligned}\Gamma(V_0) &= p_0 J_P + q_0 J_Q + (1 - p_0 - q_0) \hat{J}(\hat{V}_0) \\ &< p_0 J_P + (1 - p_0) \left[\frac{q_0}{1 - p_0} \hat{J}(V_Q) + \frac{1 - p_0 - q_0}{1 - p_0} \hat{J}(\hat{V}_0) \right] \\ &\leq p_0 J_P + (1 - p_0) \hat{J}(\bar{V}),\end{aligned}\tag{3.11}$$

where

$$\bar{V} = \frac{q_0}{1 - p_0} V_Q + \frac{1 - p_0 - q_0}{1 - p_0} \hat{V}_0.\tag{3.12}$$

The second line in (3.11) uses the fact that $\hat{J}(V_Q) > 0 = J_Q$. The third line in (3.11) makes use of the concavity of $\hat{J}$. Similarly, notice that the value V_0 to the agent is such that

$$V_0 = p_0 V_P + (1 - p_0) \bar{V}.\tag{3.13}$$

Note that the lottery $(p, \hat{V}) = (p_0, \bar{V})$ is a feasible choice in the first-stage problem (2.1). In turn, this implies that $J(V_0) \geq \Gamma(V_0)$. Since any mechanism that involves the principal hiring applicants when in control is a feasible choice for (3.11), it follows that $\Gamma(V_0) \geq J(V_0)$. Combining the two inequalities yields $\Gamma(V_0) = J(V_0)$.

Suppose that $V_Q < V_L$ and, hence, $\hat{V}_0 > V_Q$. Notice that the value $\Gamma(V_0)$ to the principal is such that

$$\begin{aligned}\Gamma(V_0) &= p_0 J_P + q_0 J_Q + (1 - p_0 - q_0) \hat{J}(\hat{V}_0) \\ &\leq \bar{p} J_P + (1 - \bar{p}) \hat{J}(\hat{V}_0),\end{aligned}\tag{3.14}$$

where $\bar{p}$ is defined as

$$\bar{p} = p_0 + q_0 \frac{\hat{V}_0 - V_Q}{\hat{V}_0 - V_P} \in [p_0, p_0 + q_0].\tag{3.15}$$

The inequality in (3.14) makes use of the fact that $J_P = J_Q < \hat{J}(\hat{V}_0)$ and $1 - \bar{p} \geq 1 - p_0 - q_0$. Similarly, notice that the value V_0 to the agent is such that

$$\begin{aligned}V_0 &= p_0 V_P + q_0 V_Q + (1 - p_0 - q_0) \hat{V}_0 \\ &= \bar{p} V_P + (1 - \bar{p}) \hat{V}_0.\end{aligned}\tag{3.16}$$

Therefore, the lottery $(p, \hat{V}) = (\bar{p}, \hat{V}_0)$ is a feasible choice in the first-stage problem (2.1). Hence, $J(V_0) \geq \Gamma(V_0)$. Since any mechanism that involves the principal hiring applicants when in control is a feasible choice for (3.11), it follows that $\Gamma(V_0) \geq J(V_0)$. Combining the two inequalities yields $\Gamma(V_0) = J(V_0)$.

Whether $V_Q \geq V_L$ or $V_Q < V_L$, $\Gamma(V_0)$ is equal to $J(V_0)$. That is, any optimal mechanism that, when the principal is in control of the hiring decision, is allowed to instruct the

principal to either hire or not hire the applicants is payoff-equivalent to the optimal mechanism that, when the principal is in control of the hiring decision, instructs the principal to hire the applicants. This finding is also intuitive. If the mechanism gives control over hiring to the principal and instructs the principal to hire, the first-stage values to the agent and the principal, V and $J(V)$, are a convex combination between the point $P = (-\eta, 0)$ and the point $C = (V_C, J_C)$. Since $V_C > -\eta$ and $J_C > 0$, the convex combination generates a $J(V)$ such that $J(0) > 0$. If the mechanism is allowed to instruct the principal not to hire, the first-stage values to the agent and the principal, V and $J(V)$, can be further convexified with the point $Q = (0, 0)$. This convexification, however, is useless because Q lies below the (V, J) frontier (see Figure 2).

It is straightforward to combine the arguments developed in this section and obtain the following proposition.

Proposition 4 (Optimal control). *Consider the class of mechanisms that can assign the hiring decision to either the principal or the agent in any period and after any history, and that can prescribe that the principal hires or does not hire the applicant in any period in which the hiring decision is assigned to the principal. Let Γ denote the first-stage value function for this class of mechanisms and let J be the first-stage value function in the restricted class of mechanisms considered in Section 2. Then, $\Gamma = J$.*

4 Extensions

In this section, we use the analysis of the baseline model to characterize the optimal mechanism in richer and more realistic environments. In Section 4.1, we consider an environment in which the agent has the option to hire either a contentious applicant, an applicant from a demographic group against which the agent is biased, or a non-contentious applicant, an applicant from a demographic group against which the agent holds no bias. In Section 4.2, we consider an environment in which the agent has the option to hire one among n contentious applicants, and an environment in which the agent has the option to hire one among n contentious applicants and m uncontentious applicants. In Section 4.3, we consider a version of the baseline model in which the agent is positively biased. In Section 4.4, we consider an environment in which the principal cannot observe whether an applicant is available or not. The analysis of these extensions follows directly from the analysis of the baseline model.

4.1 Contentious and uncontentious applicants

In the baseline model of Section 2 we considered an environment in which there is only one applicant that can be hired in each period and the agent is biased against the applicant. This environment describes well a situation in which the principal operates a technology with constant returns to scale in labor and, hence, applicants do not compete against

each other but only against the option of not hiring. The assumption of constant returns to scale in labor is made by most search-theoretic models of the labor market, such as Mortensen and Pissarides (1994), Burdett and Mortensen (1998), Postel-Vinay and Robin (2002), Menzio and Shi (2011). In other situations, however, it may be more realistic to assume that applicants compete against each other because the principal operates a technology with decreasing returns to scale. A simple formulation of this situation is an environment in which multiple applicants compete for a single vacancy. This is the assumption in the search-theoretic models by Montgomery (1991) and Burdett, Shi and Wright (2001). In this case, it is also natural to assume that the agent is biased against some applicants, but not against others. In this subsection, we characterize the optimal mechanism in this alternative environment.

We consider a version of the model in which there are multiple applicants for each vacancy. Suppose that applicants come from two different groups. Some applicants come from group X , and some applicants come from group Y . The quality x of an X -applicant is drawn from a distribution $F_x(x)$, with mean 0 and support $[\underline{x}, \bar{x}]$. The quality y of a Y -applicant is drawn from a distribution $F_y(y)$, with mean 0 and support $[\underline{y}, \bar{y}]$. The agent is negatively biased towards X -applicants, with a bias given by $\eta \in (0, \bar{x})$. The agent is unbiased towards Y -applicants. In this sense, X -applicants are contentious, and Y -applicants are not. The group of a particular applicant is known to both the principal and the agent, as it may reflect some readily-observable demographic characteristics. The quality of a particular applicant is known only to the agent.

In order to keep the analysis simple, let us assume that there is exactly one X -applicant and one Y -applicant for each vacancy⁵ and that the vacancy needs to be filled.⁶ It is useful to denote as z the difference between the quality x of an X -applicant and the quality y of a Y -applicant. That is, z equals $x - y$. The joint distribution of the random variables y and z is determined by the quality distribution F_x and F_y . For our purposes, it is useful to describe the joint distribution of y and z with the marginal distribution $F_z(z)$ of the random variable z and the distribution $G(y|z)$ of the random variable y conditional on z . Note that the support of z is $[\underline{z}, \bar{z}]$, with $\underline{z} = \underline{x} - \bar{y}$ and $\bar{z} = \bar{x} - \underline{y}$.

The mechanism design problem can be written recursively as a two-stage problem.

⁵In the next subsection, we extend the analysis to the case of multiple contentious and non-contentious applicants.

⁶If the vacancy does not need to be filled, the structure of the mechanism design problem is qualitatively different than in the baseline. In particular, in the second-stage problem, the mechanism can specify different continuation values for the agent depending on whether the agent hired the contentious applicant, the agent hired the uncontentious applicant, or the agent did not hire anyone. We leave the characterization of this mechanism design problem to future work.

The first-stage problem is

$$\begin{aligned}
J(V) = & \max_{p, \hat{V}} pJ_P + (1-p)\hat{J}(\hat{V}) \\
\text{s.t. } & V = pV_P + (1-p)\hat{V}, \\
& p \in [0, 1], \quad \hat{V} \in \hat{\mathcal{V}},
\end{aligned} \tag{4.1}$$

where V_P and J_P are respectively given by

$$V_P = -\eta, \quad J_P = 0. \tag{4.2}$$

For the same reasons as in Section 3, it is without loss in generality to restrict attention to mechanisms such that: (i) when the mechanism gives control to the principal, it does so forever; (ii) when the mechanism gives control to the principal, the mechanism instructs the principal to hire the contentious applicant. For this reason, V_P is equal to $-\eta$ and J_P is equal to 0.

At the second stage, the value of the mechanism to the agent is

$$\hat{V} = \int_z \left[\int_y \max \{ (1-\beta)(y+z-\eta) + \beta V_1, (1-\beta)y + \beta V_0 \} dG_y(y|z) \right] dF_z(z). \tag{4.3}$$

The above expression is easy to understand. Consider a particular realization of the random variables y and z . If the agent hires the X -applicant, the agent's flow payoff is $(1-\beta)(y+z-\eta)$, where $y+z$ is the quality x of the X -applicant, and the agent's continuation value is V_1 . If the agent does not hire the X -applicant, the agent's flow payoff is $(1-\beta)y$, where y is the quality of the Y -applicant, and the agent's continuation value is V_0 . The agent chooses whether to hire or not hire the X -applicant so as to maximize the sum of its flow and continuation payoffs. Hence, the agent hires the X -applicant if and only if $z \geq R$, where

$$R = \eta - \frac{\beta}{1-\beta}(V_1 - V_0). \tag{4.4}$$

Using the definition of the reservation quality R in (4.4), we can rewrite (4.3) as

$$\begin{aligned}
\hat{V} &= \int^R \left[\int_y [(1-\beta)y + \beta V_0] dG_y(y|z) \right] dF_z(z) \\
&\quad + \int_R \left[\int_y [(1-\beta)(y+z-\eta) + \beta V_1] dG_y(y|z) \right] dF_z(z) \\
&= (1-\beta) \int_R (z-\eta) dF_z(z) + \beta F_z(R)V_0 + \beta (1-F_z(R))V_1,
\end{aligned} \tag{4.5}$$

where the last line makes use of the fact that the mean of the random variable y is 0.

At the second stage, the value of the mechanism to the principal is

$$\begin{aligned}
\hat{J} &= \int^R \left[\int_y [(1-\beta)y + \beta J(V_0)] dG_y(y|z) \right] dF_z(z) \\
&\quad + \int_R \left[\int_y [(1-\beta)(y+z) + \beta J(V_1)] dG_y(y|z) \right] dF_z(z) \\
&= (1-\beta) \int_R z dF_z(z) + \beta F_z(R) J(V_0) + \beta (1 - F_z(R)) J(V_1).
\end{aligned} \tag{4.6}$$

The above expression is also easy to understand. Consider a particular realization of the random variables y and z . If $z \geq R$, the agent hires the X -applicant. The principal's flow payoff is $(1-\beta)(y+z)$, where $y+z$ is the quality x of the X -applicant, and the principal's continuation value is $J(V_1)$. If $z < R$, the agent hires the Y -applicant. The principal's flow payoff is $(1-\beta)y$, where y is the quality of the Y -applicant, and the principal's continuation value is $J(V_0)$.

Using (4.5) and (4.6), we can write the second-stage problem as

$$\begin{aligned}
\hat{J}(\hat{V}) &= \max_{V_0, V_1 \in \mathcal{V}} (1-\beta) \int_R z dF_z(z) + \beta [F_z(R) J(V_0) + (1 - F_z(R)) J(V_1)], \\
\text{s.t. } \hat{V} &= (1-\beta) \int_R (z - \eta) dF_z(z) + \beta F_z(R) V_0 + \beta (1 - F_z(R)) V_1 \\
R &= \eta - \frac{\beta}{1-\beta} (V_1 - V_0).
\end{aligned} \tag{4.7}$$

Following the same argument as in Lemma 1, it is easy to show that the implementable set $\mathcal{V}$ is given by the interval $[V_P, V_F]$, and the implementable set $\hat{\mathcal{V}}$ is given by the interval $[V_L, V_F]$, where

$$V_L \equiv (1-\beta)V_F + \beta V_P, \quad V_F \equiv \int_{\eta} (z - \eta) dF_z(z). \tag{4.8}$$

Following the same argument as in Lemma 2, it is east to show that

$$J^* \equiv \max_{V \in \mathcal{V}} J(V) > \int_{\eta} z dF_z(z) \equiv J_F \tag{4.9}$$

if the marginal distribution F_z and the bias η are such that

$$\frac{\eta F'_z(\eta)}{F_z(\eta)} > \frac{\int_{\eta} z dF_z(z)}{\int_{\eta} z dF_z(z) + \eta F(\eta)}. \tag{4.10}$$

Note that the mechanism design problem with contentious and uncontentious applicants is identical to the mechanism design problem with a single contentions applicant, except that the distribution F_z of the difference in quality between the contentious and the uncontentious applicants takes the place of the distribution F of the quality of the contentious applicant. Since the distribution F_z has a mean of 0, just like the distribution

F , the mechanism design problem with contentious and uncontentious applicants is a special case of the mechanism design problem with a single contentious applicant. Therefore, under condition (4.10), we can directly apply Propositions 1, 2 and 3. This would not be possible if quality had been a binary random variable, as the difference z between x and y would no longer be a binary random variable.

Proposition 5: (Contentious and uncontentious applicants) *The optimal mechanism has the following properties:*

1. *The solution to the first-stage problem in (4.1) is a lottery $(p, \hat{V})$ such that: $p > 0$ and $\hat{V} = V_C$ for all $V \in [V_P, V_C]$; $p = 0$ and $\hat{V} = V$ for all $V \in [V_C, V_F]$, and V_C is given by (2.34).*
2. *The solution to the second-stage problem in (4.5) is a pair of agent's continuation values (V_0, V_1) such that: $V_0 = V_P$ and $V_1 \in (V_C, V_F)$ for $\hat{V} = V_C$; $V_0 < \hat{V}$ and $V_1 \in (\hat{V}, V_F)$ for all $\hat{V} \in (V_C, V_F)$; $V_0 = V_F$ and $V_1 = V_F$ for $\hat{V} = V_F$. For all $\hat{V} \in [V_C, V_F]$, the agent's cutoff on the difference between the quality of the contentious and the uncontentious applicant is an $R \in (\underline{z}, \eta)$.*

The version of the model in which a contentious and an uncontentious applicant are available to fill each vacancy and each vacancy needs to be filled is isomorphic to the baseline model, except that the distribution F needs to be interpreted as the distribution of the quality gap between the contentious and the uncontentious applicants and not hiring needs to be interpreted as hiring the uncontentious applicant. It is useful to read Figure 5 in light of this alternative interpretation. The figure shows that the reservation quality gap is smaller than η , it is strictly positive when the agent's promised value is sufficiently high, and it is strictly negative when the agent's promised value is sufficiently low. This implies that, when the agent's promised value is sufficiently high, the agent's hiring decision is still biased against the contentious applicant and in favor of the uncontentious applicant, although the extent of the bias is less than under unfettered delegation. When the agent's promised value is sufficiently low, the agent's hiring decision becomes biased in favor of the contentious applicant and against the uncontentious applicant, since the agent's reservation quality gap becomes negative. Therefore, the optimal mechanism is such that, after some histories, the agent is induced to switch the direction of its bias. Even though this is ex-post inefficient, it is ex-ante optimal.

4.2 Multiple applicants

In this subsection, we characterize the optimal mechanism in a version of the baseline model in which multiple contentious applicants and multiple uncontentious applicants compete for each of vacancy. It is useful to start the analysis from the case in which there are $n \in \mathbb{N}$ contentious applicants and no uncontentious applicants. Each contentious applicant has a quality x that is independently drawn from the distribution $F(x)$. Let x_n denote the maximum of the quality of the n contentious applicants. Let $F_n(x_n)$ denote

the distribution of x_n , i.e., $F_n(x) = F(x)^n$.

The principal's optimal mechanism design problem can be written recursively as a two-stage problem. The first-stage of the recursive problem is

$$\begin{aligned} J(V) = & \max_{p, \hat{V}} pJ_P + (1-p)\hat{J}(\hat{V}) \\ \text{s.t. } & V = pV_P + (1-p)\hat{V}, \\ & p \in [0, 1], \quad \hat{V} \in \hat{\mathcal{V}}, \end{aligned} \quad (4.11)$$

where the values V_P and J_P are given by

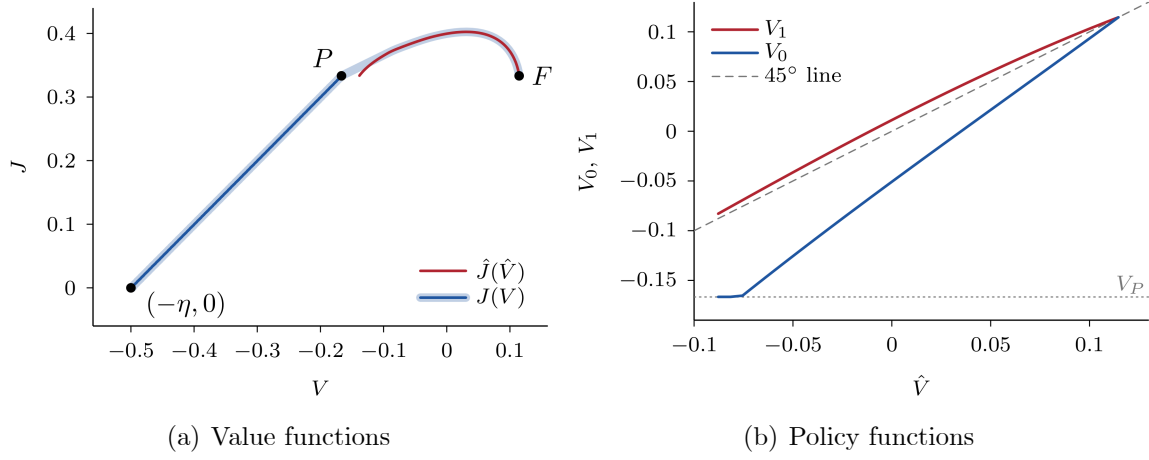
$$V_P = \int_{\underline{x}}^{\bar{x}} (x_n - \eta) dF_n(x_n), \quad J_P = \int_{\underline{x}}^{\bar{x}} x_n dF_n(x_n). \quad (4.12)$$

The second-stage of the recursive problem is

$$\begin{aligned} \hat{J}(\hat{V}) = & \max_{V_0, V_1 \in \mathcal{V}} (1-\beta) \int_R x_n dF_n(x_n) + \beta [F_n(R)J(V_0) + (1-F_n(R))J(V_1)], \\ \text{s.t. } & \hat{V} = (1-\beta) \int_R (x_n - \eta) dF_n(x) + \beta [F_n(R)V_0 + (1-F_n(R))V_1] \\ & R = \eta - \frac{\beta}{1-\beta} (V_1 - V_0). \end{aligned} \quad (4.13)$$

The formulation of (4.11) subsumes a restriction to mechanisms such that: (i) if the mechanism gives control of the hiring process to the principal, it does so permanently; (ii) if the principal has control over the hiring process, the mechanism instructs the agent to report the ranking of the applicants, which the agent has an incentive to disclose truthfully, and then it instructs the principal to hire the highest-ranked applicant. The first restriction on the mechanisms is without loss in generality. The second restriction may not be without loss in generality. If the principal has control over hiring and hires a randomly-selected applicant, the values to the agent and the principal are $-\eta$ and 0. If the principal has control over hiring and hires the best applicant, the value to the agent is $V_P = E[x_n] - \eta$ and the value to the principal is $J_P = E[x_n]$. The optimal lottery between delegation and control may be a convex combination between a delegation point (V_C, J_C) and the control point $(-\eta, 0)$. Alternatively, the optimal lottery may be a convex combination between a delegation point (V_C, J_C) and the control point (V_P, J_P) for $V \in [V_P, V_C)$ and a convex combination between the control points $(-\eta, 0)$ and (V_P, J_P) for $V \in [-\eta, V_P)$. If the optimal mechanism in delegation is such that $V_0 \geq V_P$ for all $\hat{V} \geq V_C$, the only relevant control is the point (V_P, J_P) and the restriction (ii) is innocuous. This is the case in the numerical example illustrated in Figure 6.

Taking as given the restrictions on the space of mechanisms implicit in (4.11), it is immediate to show that the implementable set $\mathcal{V}$ is given by the interval $[V_P, V_F]$, and the



Notes: Value functions and policy functions given parameters $n = 2$, $\beta = 0.9$, $\eta = 0.5$, and $F(x)$ uniform on the interval $[-1, 1]$.

Figure 6: Numerical example: Value and policy functions with n applicants

implementable set $\hat{\mathcal{V}}$ is given by the interval $[V_L, V_F]$, where

$$V_L \equiv (1 - \beta)V_F + \beta V_P, \quad V_F \equiv \int_{\eta} (x_n - \eta) dF_n(x_n). \quad (4.14)$$

Similarly, it is straightforward to show that J^* , which is defined as the maximum of $J(V)$ with respect to V , is strictly greater than J_F , which is defined as $\int_{\eta} x_n dF_n(x_n)$, if the quality distribution F_n and the bias η are such that

$$\frac{\eta F'_n(\eta)}{F_n(\eta)} > \frac{\int_{\eta} x_n dF_n(x_n) - \int_{\underline{x}} x_n dF_n(x_n)}{\int_{\eta} (x_n - \eta) dF_n(x_n) - \int_{\underline{x}} (x_n - \eta) dF_n(x_n)}. \quad (4.15)$$

Note that the mechanism design with multiple contentious applicants is identical to the mechanism design problem with a single applicant, except that the distribution F_n of the highest quality x_n among n applicants replaces the distribution F of the quality of a single applicant. The only substantive difference between the distribution F_n and the distribution F is that the former has a strictly positive mean and the latter has a mean equal to 0. Yet, under condition (4.15), it is easy to verify that the characterization results in Propositions 1, 2 and 3 also apply to the case of a distribution with a strictly positive mean.

Proposition 6: (Multiple contentious applicants) *The optimal mechanism has the following properties:*

1. *The solution to the first-stage problem in (4.11) is a lottery $(p, \hat{V})$ such that: $p > 0$ and $\hat{V} = V_C$ for all $V \in [V_P, V_C]$; $p = 0$ and $\hat{V} = V$ for all $V \in [V_C, V_F]$, and V_C is given by (2.34).*
2. *The solution to the second-stage problem in (4.13) is a pair of continuation values*

(V_0, V_1) such that: $V_0 = V_P$ and $V_1 \in (V_C, V_F)$ for $\hat{V} = V_C$; $V_0 < \hat{V}$ and $V_1 \in (\hat{V}, V_F)$ for all $\hat{V} \in (V_C, V_F)$; $V_0 = V_F$ and $V_1 = V_F$ for $\hat{V} = V_F$. For $\hat{V} \in [V_C, V_F]$, the agent's reservation quality is an $R \in (\underline{x}, \eta)$.

In light of the results of the previous subsection, it is easy to extend Proposition 6 to the case in which $n \in \mathbb{N}$ contentious applicants and $m \in \mathbb{N}$ non-contentious applicants compete for each vacancy, and each vacancy needs to be filled. Let x_n denote the maximum quality of the contentious applicants, and let $F_{x,n}(x_n)$ denote the distribution of x_n . Let y_m denote the maximum quality of uncontentious applicants, and let $F_{y,m}(y_m)$ denote the distribution of y_m . Let z denote the difference between the maximum quality x_n of the contentious applicants and the maximum quality y_m of the uncontentious applicants, and let $F_z(z)$ denote the distribution of the random variable z .

The mechanism design problem can be written recursively as a two-stage problem.⁷ The first-stage problem is (4.11), where V_P and J_P are respectively given by

$$V_P = \mu + \int_{\underline{z}}^{\bar{z}} (z - \eta) dF_z(z), \quad J_P = \mu + \int_{\underline{z}}^{\bar{z}} z dF_z(z). \quad (4.16)$$

and μ denotes the unconditional mean of the random variable y_m and is given by

$$\mu = \int y_m dF_{y,m}(y_m). \quad (4.17)$$

The second-stage problem is

$$\begin{aligned} \hat{J}(\hat{V}) &= \max_{V_0, V_1 \in \mathcal{V}} (1 - \beta) [\mu + \int_R z dF_z(z)] + \beta [F_z(R)J(V_0) + (1 - F_z(R))J(V_1)], \\ \text{s.t. } \hat{V} &= (1 - \beta) [\mu + \int_R (z - \eta) dF_z(z)] + \beta F_z(R)V_0 + \beta (1 - F_z(R))V_1 \\ R &= \eta - \frac{\beta}{1 - \beta}(V_1 - V_0). \end{aligned} \quad (4.18)$$

The implementable set $\mathcal{V}$ is given by the interval $[V_P, V_F]$, and the implementable set $\hat{\mathcal{V}}$ is given by the interval $[V_L, V_F]$, where

$$V_L \equiv (1 - \beta)V_F + \beta V_P, \quad V_F \equiv \mu + \int_{\eta}^{\bar{z}} (z - \eta) dF_z(z). \quad (4.19)$$

The optimal mechanism is not unfettered delegation, in the sense that

$$J^* \equiv \max_{V \in \mathcal{V}} J(V) > \mu + \int_{\eta}^{\bar{z}} z dF_z(z) \equiv J_F \quad (4.20)$$

⁷We maintain the restriction to mechanisms such that: (i) if the mechanism gives control of the hiring process to the principal, it does so permanently; (ii) if the principal has control over the hiring process, the mechanism instructs the agent to report the ranking of the contentious applicants, and it instructs the principal to hire the highest-ranked contentious applicant.

if the marginal distribution F_z and the bias η are such that

$$\frac{\eta F'_z(\eta)}{1 - F_z(\eta)} > \frac{\int_{\eta} z dF_z(z) - \int_{\underline{z}}^{\bar{z}} z dF_z(z)}{\int_{\eta} (z - \eta) dF_z(z) - \int_{\underline{z}}^{\bar{z}} (z - \eta) dF_z(z)}. \quad (4.21)$$

The mechanism design with multiple contentious and multiple uncontentious applicants is identical to the mechanism design problem with a multiple contentious applicant, except that the distribution F_z of the gap between the highest quality x_n among the n contentious applicants and the highest quality y_m among the m uncontentious applicants replaces F_n , and the payoffs to the agent and the principal are shifted by a constant μ . Propositions 1, 2 and 3 generalize to the case in which the distribution has a non-zero mean and to the case in which the payoffs are shifted by a constant. Therefore, under condition (4.21), Propositions 1, 2 and 3 and, in turn, Proposition 6 apply to the version of the model with multiple contentious and uncontentious applicants.

4.3 Positive bias

In the baseline model of Section 2, we characterize the optimal mechanism when the agent is biased against the applicants. In this subsection, we want to characterize the optimal mechanism in an environment where the agent is biased in favor of the applicants. In particular, we consider a version of the baseline model in which the difference between the payoff to the agent and the payoff to the principal if an applicant is hired is a strictly positive constant ϕ .

The principal's mechanism design problem can be written recursively as a two-stage problem. The first-stage of the recursive problem is

$$\begin{aligned} J_+(V) = & \max_{p_+, \hat{V}_+} p_+ J_P^+ + (1 - p_+) \hat{J}_+(\hat{V}_+) \\ \text{s.t. } & V = p_+ V_P^+ + (1 - p_+) \hat{V}_+, \\ & p_+ \in [0, 1], \quad \hat{V}_+ \in \hat{\mathcal{V}}_+, \end{aligned} \quad (4.22)$$

where the values V_P^+ and J_P^+ are given by

$$V_P^+ = 0, \quad J_P^+ = 0. \quad (4.23)$$

The second-stage of the recursive problem is

$$\begin{aligned}
\hat{J}_+(\hat{V}) &= \max_{V_0^+, V_1^+ \in \mathcal{V}_+} (1 - \beta) \int_{R_+} x dF(x) + \beta [F(R_+)J(V_0^+) + (1 - F(R_+))J(V_1^+)], \\
\text{s.t. } \hat{V} &= (1 - \beta) \int_{R_+} (x + \phi) dF(x) + \beta [F(R_+)V_0^+ + (1 - F(R_+))V_1^+], \\
R_+ &= -\phi - \frac{\beta}{1 - \beta} (V_1^+ - V_0^+).
\end{aligned} \tag{4.24}$$

The formulation of (4.22) assumes a restriction to mechanisms such that: (i) if the mechanism gives control of the hiring process to the principal, it does so permanently; (ii) if the principal has control over the hiring process, the mechanism prescribes that the principal hires none of the applicants. The first restriction is without loss in generality, for the same reasons as in Section 3. The second restriction is also without loss in generality. This is intuitive. If the principal has control over hiring and hires all the applicants, the values to the agent and the principal are ϕ and 0. If the principal has control over hiring and does not hire any applicants, the value to both the agent and the principal is 0. Clearly, the optimal lottery between delegation and control is a convex combination between a point on the delegation frontier $(\hat{V}, \hat{J}_+)$ and the control point $(0, 0)$.

As in Lemma 1, we can prove that the implementable set $\mathcal{V}_+$ is given by the interval $[V_P^+, V_F^+]$, and the implementable set $\hat{\mathcal{V}}_+$ is given by the interval $[V_L^+, V_F^+]$, where

$$V_L^+ \equiv (1 - \beta)V_F^+ + \beta V_P^+, \quad V_F^+ \equiv \int_{-\phi} (x + \phi) dF(x). \tag{4.25}$$

As in Lemma 2, we can prove that J_+^* , which is defined as the maximum of $J_+(V)$ with respect to V , is strictly greater than J_F^+ , which is defined as $\int_{-\phi} x dF(x)$, if the quality distribution F and the bias ϕ are such that

$$\frac{\phi F'(-\phi)}{1 - F(-\phi)} > \frac{\int_{-\phi} x dF(x)}{\int_{-\phi} (x + \phi) dF(x)}. \tag{4.26}$$

Under condition (4.26), we can use the same arguments as in Propositions 1, 2 and 3 to show that the optimal mechanism has the same qualitative features as in the baseline model. There are only two substantive differences between the optimal mechanism when the agent is positively biased and the optimal mechanism when the agent is negatively biased. First, as we have already mentioned, if the agent is positively biased, the optimal mechanism is such that, when the hiring decision is controlled by the principal, the principal is required to hire none of the applicants. If the agent is negatively biased, the optimal mechanism is such that, when the hiring decision is controlled by the principal, the principal is required to hire all of the applicants. Second, if the agent is positively biased, the optimal mechanism is such that, when the hiring decision is delegated to the agent, the agent is punished for hiring an applicant with a lower continuation values, and

it is rewarded for not hiring an applicant with a higher continuation value. If the agent is negatively biased, the optimal mechanism is such that the agent is rewarded for hiring an applicant and punished for not hiring an applicant. This difference is easy to understand. If the agent is positively biased, the agent wants to hire more applicants than the principal. The optimal mechanism reduces the gap between the preferences of the agent and the principal by rewarding the agent for not hiring applicants and by punishing him for hiring applicants.

The proposition below contains the characterization of the optimal mechanism when the agent is positively biased.

Proposition 7: (Positive bias) *The optimal mechanism has the following properties:*

1. *The solution to the first-stage problem in (4.22) is a lottery $(p_+, \hat{V}_+)$ such that: $p_+ > 0$ and $\hat{V}_+ = V_C$ for all $V \in [V_P^+, V_C^+]$; $p_+ = 0$ and $\hat{V}_+ = V$ for all $V \in [V_C^+, V_F^+]$, and V_C^+ is given by (2.34).*
2. *The solution to the second-stage problem in (4.24) is a pair of continuation values (V_0^+, V_1^+) such that: $V_1^+ = V_P^+$ and $V_0^+ \in (V_C^+, V_F^+)$ for $\hat{V} = V_C^+$; $V_0^+ \in (\hat{V}, V_F^+)$ and $V_1^+ < \hat{V}$ for all $\hat{V} \in (V_C^+, V_F^+)$; $V_0^+ = V_F^+$ and $V_1^+ = V_F^+$ for $\hat{V} = V_F^+$. For all $\hat{V} \in [V_C^+, V_F^+]$, the agent's reservation quality is an $R \in (-\phi, \bar{x})$.*

Under some conditions on ϕ and F , the solution to the mechanism design problem under positive bias can be expressed as a simple transformation of the solution to the mechanism design problem under negative bias. Specifically, if the quality distribution F is symmetric around 0, in the sense that $F(x) = 1 - F(-x)$ for all $x \in [\underline{x}, \bar{x}]$, and the positive bias has the same magnitude as the negative bias, in the sense that $\phi = \eta$, the value functions that solve (4.22) and (4.26) are an horizontal translation of the value functions that solve (2.1) and (2.4), and the policy functions associated with (4.22) and (4.24) are an horizontal translation of the policy functions that solve (2.1) and (2.4), with V_0^+ taking the place of V_1 and V_1^+ taking the place of V_0 .

The next proposition makes the statements above formal. The proof is in the Appendix

Proposition 8: (Positive and negative bias under symmetry). *Assume $F(x) = 1 - F(-x)$ for all $x \in [\underline{x}, \bar{x}]$ and $\phi = \eta$. Then:*

1. *The first-stage and second-stage value functions J_+ and $\hat{J}_+$ are such that*

$$J_+(V + \eta) = J(V), \quad \hat{J}_+(\hat{V} + \eta) = \hat{J}(\hat{V}). \quad (4.27)$$

2. *The solution to the first-stage problem in (4.22) given the promised value $V_+ = V + \eta$ is the lottery $(p_+, \hat{V}_+)$, where $p_+ = p$, $\hat{V}_+ = \hat{V} + \eta$, and $(p, \hat{V})$ is the solution to the first-stage problem in (2.1) given the promised value V .*
3. *The solution to the second-stage problem in (4.24) given the promised value $\hat{V}_+ = \hat{V} + \eta$ is a pair of continuation values (V_0^+, V_1^+) , where $V_0^+ = V_1 + \eta$, and $V_1^+ = V_0 + \eta$,*

and (V_0, V_1) is the solution to the second-stage problem in (2.4) given the promised value $\hat{V}$.

4. The agent's reservation quality associated with the solution to (4.24) given the promised value $\hat{V}_+ = \hat{V} + \eta$ is $R_+ = -R$, where R is the agent's reservation quality associated with the solution to (2.4) given the promised value $\hat{V}$.

4.4 Privately observed arrival of applicants

In the baseline model, we assume that the principal knows when there is an applicant that could be hired. In some situations, though, it may be more realistic to assume that the principal does not know whether an applicant is available or not. In this subsection, we characterize the optimal mechanism when the agent privately observes whether an applicant is available or not and, in the former case, the agent privately observes the quality of the applicant. Specifically, suppose that an applicant is available with probability $\rho \in (0, 1)$ and no applicant is available with probability $1 - \rho$. As in the baseline model, we assume that the agent is biased against the applicant.

As usual, the mechanism design problem can be written recursively as a two-stage problem. The first-stage problem is

$$\begin{aligned} J(V) = & \max_{p, \hat{V}} pJ_P + (1 - p)\hat{J}(\hat{V}) \\ \text{s.t. } & V = pV_P + (1 - p)\hat{V}, \\ & p \in [0, 1], \quad \hat{V} \in \hat{\mathcal{V}}, \end{aligned} \tag{4.28}$$

where V_p and J_P are respectively given by

$$V_P = 0, \quad J_P = 0. \tag{4.29}$$

The above formulation of the first-stage problem assumes that the optimal mechanism is such that: (i) if the mechanism gives control of hiring to the principal, it does so forever; (ii) when the principal has control over hiring, the mechanism instructs the principal not hire. These assumptions are without loss in generality. The first assumption is without loss in generality for the same reasons discussed in Section 3. The second assumption is without loss in generality because not hiring is the only feasible mechanism when the principal has control over hiring. In fact, as the principal does not know whether an applicant is available or not, it needs to rely on a report by the agent. Since the agent prefers not hiring an applicant than hiring an applicant irrespective of quality, the only mechanism that induces the agent to truthfully report whether an applicant is available or not is such that, irrespective of the agent's report, the principal does not hire.

The second-stage problem is

$$\begin{aligned}
\hat{J}(\hat{V}) &= \max_{V_0, V_1 \in \mathcal{V}} (1 - \beta) \rho \int_R x dF(x) + \beta [(1 - \rho(1 - F(R))) J(V_0) + \rho(1 - F(R)) J(V_1)], \\
\text{s.t. } \hat{V} &= (1 - \beta) \rho \int_R (x - \eta) dF(x) + \beta [(1 - \rho(1 - F(R))) V_0 + \rho(1 - F(R)) V_1] \\
R &= \eta - \frac{\beta}{1 - \beta} (V_1 - V_0).
\end{aligned} \tag{4.30}$$

The problem above is easy to understand. In the current period, an applicant is available with probability ρ . If the applicant's quality x is higher than R , the agent hires the applicant. In this case, the principal's flow payoff is x , the agent's flow payoff is $x - \eta$, the principal's continuation value is $J(V_1)$, and the agent's continuation value is V_1 . If an applicant is not available or if an applicant is available and their quality x is lower than R , the agent does not hire. In this case, the principal's and the agent's flow payoffs are 0, the principal's continuation value is $J(V_0)$, and the agent's continuation value is V_0 . Clearly, since the principal does not observe whether an applicant is or is not available, the agent's continuation payoff can only be contingent on whether an applicant is hired or not.

Following the same argument as in Lemma 1, it is easy to show that the implementable set $\mathcal{V}$ is given by the interval $[V_P, V_F]$, and the implementable set $\hat{\mathcal{V}}$ is given by the interval $[V_L, V_F]$, where

$$V_L \equiv (1 - \beta)V_F + \beta V_P, \quad V_F \equiv \rho \int_{\eta} (x - \eta) dF(x). \tag{4.31}$$

Following the same arguments as in Lemma 2, it is easy to show that J^* , which is defined as the maximum of $J(V)$ with respect to V , is strictly greater than J_F , which is given by $\rho \int_{\eta} x dF(x)$, if the quality distribution F , the bias η , and the arrival probability ρ are such that

$$\frac{\rho \eta F'(\eta)}{1 - \rho(1 - F(\eta))} > \frac{\rho \int_{\eta} x dF(x)}{\rho \int_{\eta} (x - \eta) dF(x)}. \tag{4.32}$$

The mechanism design problem where the arrival of applicants is privately observed by the agent is very similar to the baseline mechanism design problem in Section 2. There are three minor differences. First, the value to the agent and the principal in control are both equal to 0, rather than $-\eta$ and 0. Second, the agent and principal's flow payoffs in delegation are scaled by the factor ρ . Lastly, the probability that the agent hires an applicant in delegation is $\rho(1 - F(R))$, rather than $1 - F(R)$. Notwithstanding these differences, it is easy to generalize Propositions 1 and 3. Formally, under condition (4.32), we can establish the following result.

Proposition 9: (Privately observed arrival of applicants). *The optimal mechanism has the following properties:*

1. The solution to the first-stage problem in (4.28) is a lottery $(p, \hat{V})$ such that: $p > 0$

and $\hat{V} = V_C$ for all $V \in [V_P, V_C)$; $p = 0$ and $\hat{V} = V$ for all $V \in [V_C, V_F]$, and V_C is given by (2.34).

2. The solution to the second-stage problem in (4.30) is a pair of continuation values (V_0, V_1) such that: $V_0 = V_P$, and $V_1 \in (V_C, V_F)$ for $\hat{V} = V_C$; $V_0 < \hat{V}$ and $V_1 \in (\hat{V}, V_F)$ for all $\hat{V} \in (V_C, V_F)$; $V_0 = V_F$ and $V_1 = V_F$ for $\hat{V} = V_F$.

5 Conclusions

We studied the optimal delegation of hiring decisions to an agent who has private information about the quality of applicants, but discriminates against workers from some demographic group. We showed that the optimal delegation evolves over time according to the realized record of hires. Initially, the hiring decision is delegated to the agent. The agent is rewarded for hiring workers from the discriminated group. The agent is punished for rejecting workers from the discriminated group. The reward is delivered through weaker incentives in the future, which allow the agent to lean into its bias, and through an increase in the expected time until the hiring decision is given to the principal, which allows the agent to exert its bias for longer. The punishment is delivered through stronger incentives in the future, and through a decrease in the expected time until the hiring decision is given to the principal. Hiring decisions are never permanently delegated to the agent, no matter how many applicants are hired from the discriminated group. After sufficiently many applicants from the discriminated group are rejected, the hiring decision is permanently given to the principal.

The paper contributes to the still sparse literature on dynamic delegation. In contrast to previous work, we assume that the state that the agent privately observes is a continuous rather than a binary random variable. The assumption is not only natural in the context of hiring applicants. The assumption also leads us to uncover novel insights on the dynamic evolution of the strength of incentives, and on delegation in the long run. Moreover, the assumption allows us to effortlessly generalize the basic environment to the most relevant scenario, one in which the hiring decision involves hiring one of multiple applicants, some of whom are from the discriminated group, and some of whom are not. As long as one is willing to interpret the agent as a private firm and the principal as a labor market authority, the paper contributes to the debate on hiring discrimination and its remedies. In order to make progress, we made several simplifying assumptions that may be worth relaxing in the future. We assumed that the bias of the agent is known to the principal, while it may be more realistic to assume that it is not. We assumed that hiring is the only decision that might be delegated to the agent, while it would be more realistic to include a firing decision as well.

References

- [1] Atkeson, A., and R. Lucas. “On Efficient Distribution with Private Information.” *Review of Economic Studies*, 59: 427-453.
- [2] Aghion, P., and J. Tirole. 1997. “Formal and Real Authority in Organizations.” *Journal of Political Economy* 105: 1-29.
- [3] Alonso, R., and N. Matouschek. 2008. “Optimal Delegation.” *Review of Economic Studies*, 75: 259-293.
- [4] Amador, M., and K. Bagwell. 2013. “The Theory of Optimal Delegation with an Application to Tariff Caps.” *Econometrica*, 81: 1541-1599.
- [5] Angeletos, M., I. Werning, and M. Amador. 2006. “Commitment vs. Flexibility.” *Econometrica*, 74: 365-396.
- [6] Athey, S., A. Atkeson, and P. Kehoe. 2005. “The Optimal Degree of Discretion in Monetary Policy.” *Econometrica*, 73: 1431-1475.
- [7] Burdett, K., and D. Mortensen. 1998. “Wage Differentials, Employer Size, and Unemployment.” *International Economic Review*, 39: 955-970.
- [8] Burdett, K., S. Shi, and R. Wright. 2001. Pricing and Matching with Frictions. *Journal of Political Economy* 109: 1060–85.
- [9] Carrasco, V., W. Fuchs, and S. Fukuda. 2019. “From Equals to Despots: The Dynamics of Repeated Decision Making in Partnerships with Private Information.” *Journal of Economic Theory*, 182: 402-432.
- [10] Clementi, G., and H. Hopenhayn. 2006. “A Theory of Financing Constraints and Firm Dynamics.” *Quarterly Journal of Economics*, 121: 229-265.
- [11] Escobar, J., and J. Toikka. 2013. "Efficiency in Games with Markovian Private Information." *Econometrica*, 81: 1887-1934.
- [12] Frankel, A. 2016. “Discounted Quotas.” *Journal of Economic Theory*, 166: 396-444.
- [13] Frankel, A. 2021. “Selecting Applicants.” *Econometrica*, 89: 615-645.
- [14] Friedman, M. 1962. *Capitalism and Freedom*, University of Chicago Press.
- [15] Goldin, C., and C. Rouse. 2000. “Orchestrating Impartiality: The Impact of “Blind” Auditions on Female Musicians.” *American Economic Review*, 90: 715-741.
- [16] Guo, Y., and J. Horner. 2020. “Dynamic Allocation without Money.” Manuscript, Yale University.

- [17] Halac, M., and P. Yared. 2018. “Fiscal Rules and Discretion in a World Economy.” *American Economic Review*, 108: 2305-2334.
- [18] Holmstrom, B. 1977. “On Incentives and Control in Organizations.” Manuscript, Stanford University.
- [19] Jackson, M., and H. Sonnenschein. 2007. “Overcoming Incentive Constraints by Linking Decisions.” *Econometrica*, 75: 241-257.
- [20] Li, J., N. Matouschek, and M. Powell. 2017. “Power Dynamics in Organizations.” *American Economic Journal: Microeconomics*, 9: 217-241.
- [21] Lipnowski, E., and J. Ramos. 2021. “Repeated Delegation.” *Journal of Economic Theory*, 188: 1-41.
- [22] Menzio, G., and S. Shi. 2011. “Efficient Search on the Job and the Business Cycle.” *Journal of Political Economy*, 119: 468-510.
- [23] Montgomery, J. 1991. Equilibrium Wage Dispersion and Interindustry Wage Differentials. *Quarterly Journal of Economics* 106: 163–79.
- [24] Mortensen, D. and C. Pissarides. 1994. “Job Creation and Job Destruction in the Theory of Unemployment.” *Review of Economic Studies*, 61: 397-415.
- [25] Padró i Miquel, G., and P. Yared. “The Political Economy of Indirect Control.” *Quarterly Journal of Economics*, 127: 947-1015.
- [26] Postel Vinay, F., and J. Robin. 2002. “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity.” *Econometrica*, 70: 2295-2350.
- [27] Sannikov, Y. 2008. “A Continuous-Time Version of the Principal-Agent Problem.” *Review of Economic Studies*, 75: 957-984.
- [28] Thomas, J., and T. Worrall. 1990. “Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem.” *Journal of Economic Theory*, 51: 367-390.

Appendix

A Proof of Proposition 8

For $\hat{V}_+ = \hat{V} + \eta$, the functional equation (4.24) can be written

$$\begin{aligned}
& \hat{J}_+(\hat{V} + \eta) \\
&= \max_{V_0^+, V_1^+} (1 - \beta) \int_{R_+} x dF(x) + \beta [F(R_+)J_+(V_0^+) + (1 - F(R_+))J_+(V_1^+)] , \\
&\text{s.t. } \hat{V} + \eta = (1 - \beta) \int_{R_+} (x + \phi) dF(x) + \beta [F(R_+)V_0^+ + (1 - F(R_+))V_1^+] \quad (\text{A.1}) \\
& R_+ = -\phi - \frac{\beta}{1 - \beta} (V_1^+ - V_0^+), \quad V_0^+, V_1^+ \in [V_P^+, V_F^+].
\end{aligned}$$

Let us define V_0 and V_1 as, respectively, $V_1^+ - \eta$ and $V_0^+ - \eta$. Since the choice variables V_0^+ and V_1^+ must belong to the interval $[V_P^+, V_F^+]$, the alternative choice variables V_0 and V_1 must belong to the interval

$$\begin{aligned}
[V_P^+ - \eta, V_F^+ - \eta] &= \left[-\eta, \int_{-\phi} (x + \phi) dF(x) - \eta \right] \\
&= \left[-\eta, \int_{-\phi} x dF(x) + (1 - F(-\phi))\phi - \eta \right] \\
&= \left[-\eta, \int_{\eta} x dF(x) + F(\eta)\eta - \eta \right] \\
&= \left[-\eta, \int_{\eta} (x - \eta) dF(x) \right] = [V_P, V_F] \quad (\text{A.2})
\end{aligned}$$

The first line makes use of the definitions of V_P^+ and V_F^+ . The third line makes use of the fact that $1 - F(-\phi)$ equals $F(\phi)$ and $\phi = \eta$, as well as of the fact that, by symmetry of the quality distribution, $\int_{-\phi} x dF(x)$ equals $\int_{\phi} x dF(x)$. The fourth line makes use of the definition of V_P and V_F .

Using the above definitions and observations, we can rewrite (A.1) as

$$\begin{aligned}
& \hat{J}_+(\hat{V} + \eta) \\
&= \max_{V_0, V_1} (1 - \beta) \int_{R_+} x dF(x) + \beta [F(R_+)J_+(V_1 + \eta) + (1 - F(R_+))J_+(V_0 + \eta)] , \\
&\text{s.t. } \hat{V} + \eta = (1 - \beta) \int_{R_+} (x + \phi) dF(x) + \beta [F(R_+)V_1 + (1 - F(R_+))V_0 + \eta] \quad (\text{A.3}) \\
& R_+ = -\eta - \frac{\beta}{1 - \beta} (V_0 - V_1), \quad V_0, V_1 \in [V_P, V_F].
\end{aligned}$$

Defining R as $-R_+$ and using the fact that $F(R_+) = 1 - F(-R_+)$, we can rewrite

(A.3) as

$$\begin{aligned}
& \hat{J}_+(\hat{V} + \eta) \\
= & \max_{V_0, V_1} (1 - \beta) \int_{R_+} x dF(x) + \beta [(1 - F(R))J_+(V_1 + \eta) + F(R)J_+(V_0 + \eta)], \\
\text{s.t. } & \hat{V} + \eta = (1 - \beta) \int_{R_+} (x + \eta) dF(x) + \beta [(1 - F(R))V_1 + F(R)V_0 + \eta] \\
& R = \eta - \frac{\beta}{1 - \beta} (V_1 - V_0), \quad V_0, V_1 \in [V_P, V_F].
\end{aligned} \tag{A.4}$$

Using the symmetry of the quality distribution and the definition of R , we can now rewrite (A.4) as

$$\begin{aligned}
& \hat{J}_+(\hat{V} + \eta) \\
= & \max_{V_0, V_1} (1 - \beta) \int_{R_-} x dF(x) + \beta [(1 - F(R))J_+(V_1 + \eta) + F(R)J_+(V_0 + \eta)], \\
\text{s.t. } & \hat{V} = (1 - \beta) \int_{R_-} (x - \eta) dF(x) + \beta [(1 - F(R))V_1 + F(R)V_0] \\
& R = \eta - \frac{\beta}{1 - \beta} (V_1 - V_0), \quad V_0, V_1 \in [V_P, V_F].
\end{aligned} \tag{A.5}$$

In the objective function, we used the fact that $\int_{R_+} x dF(x)$ equals $\int_{-R_+} x dF(x)$ and the definition $R = -R_+$. In the promise-keeping constraint, we collected η on the right-hand side and used the fact that $\int_{R_+} x dF(x)$ equals $\int_{R_-} x dF(x)$.

For $\hat{V} + \eta$, the functional equation (4.22) can be written as

$$\begin{aligned}
J_+(V + \eta) &= \max_{p_+, \hat{V}_+} p_+ J_P^+ + (1 - p_+) \hat{J}_+(\hat{V}_+), \\
\text{s.t. } & V + \eta = p V_P^+ + (1 - p) \hat{V}_+, \\
& p_+ \in [0, 1], \quad \hat{V}_+ \in [V_L^+, V_F^+].
\end{aligned} \tag{A.6}$$

Let $\hat{V}$ be defined as $\hat{V}_+ - \eta$. Since $\hat{V}_+$ belongs to the interval $[V_L^+, V_F^+]$, $\hat{V}$ belongs to the interval

$$\begin{aligned}
[V_L^+ - \eta, V_F^+ - \eta] &= [(1 - \beta)V_P^+ + \beta V_F^+ - \eta, V_F^+ - \eta] \\
&= [(1 - \beta)(V_P^+ - \eta) + \beta(V_F^+ - \eta), V_F^+ - \eta] \\
&= [(1 - \beta)V_P + \beta V_F, V_F] = [V_L, V_F].
\end{aligned} \tag{A.7}$$

The third line makes use of the fact that

$$V_P^+ - \eta = -\eta = V_P. \tag{A.8}$$

The third line also makes use of the fact that

$$\begin{aligned}
V_F^+ - \eta &= \int_{-\phi} (x + \phi) dF(x) - \eta \\
&= \int_{-\phi} x dF(x) + \phi(1 - F(-\phi)) - \eta \\
&= \int_{\eta} (x - \eta) dF(x) = V_F.
\end{aligned} \tag{A.9}$$

Using the definition of $\hat{V}$ and the above observations, we can rewrite (A.6) as

$$\begin{aligned}
J_+(V + \eta) &= \max_{p, \hat{V}} pJ_P^+ + (1 - p)\hat{J}_+(\hat{V} + \eta) \\
\text{s.t. } V + \eta &= pV_P^+ + (1 - p)(\hat{V} + \eta), \\
p &\in [0, 1], \quad \hat{V} \in [V_L, V_F], .
\end{aligned} \tag{A.10}$$

Rearranging terms in the objective function and using the fact that $J_P^+ = J_P$, we can rewrite (A.10) as

$$\begin{aligned}
J_\phi(V + \eta) &= \max_{p, \hat{V}} pJ_P + (1 - p)\hat{J}_\phi(\hat{V} + \eta) \\
\text{s.t. } V &= pV_P + (1 - p)\hat{V}, \\
p &\in [0, 1], \quad \hat{V} \in [V_\ell, V_F], .
\end{aligned} \tag{A.11}$$

The value functions $\hat{J}_+(\hat{V} + \eta) = \hat{J}(\hat{V})$ and $J_+(V + \eta) = J(V)$ solve the functional equations (A.5) and (A.11), since $\hat{J}(\hat{V})$ and $J(V)$ are the solution to (2.4) and (2.1), (A.5) is identical to (2.4), and (A.11) is identical to (2.1). Parts 2, 3 and 4 of the proposition directly follow from the change in choice variables implemented in the derivation of (A.5) and (A.11). ■