Designing Promotion and Hiring Procedures with Biased Evaluators

Christopher Avery and Margaret Meyer *

April 1, 2011

Abstract

We study settings where hiring or promotion decisions are based on recommendations by informed, but potentially biased, evaluators. Evaluators may be biased in favor of those they are evaluating for either psychological or self-interested reasons. We examine to what extent partisan evaluators can be effectively disciplined by the knowledge that their recommendations today will affect how much their advice is relied on in the future. Formally, we analyze a repeated cheap-talk game with one principal, one evaluator, and in each of two periods, one (passive) candidate, whom the principal must choose whether or not to hire, on the basis of the evaluators report. The evaluator privately observes the current candidates ability at the start of each period. The principal begins the game uncertain about the evaluators degree of bias, which may be either low or high. We find that a concern with preserving a reputation for objectivity will induce both types of evaluator to be 'tougher' in their early evaluations; however, while this raises the value to the organization of the advice from the more biased evaluator, it may lower the value of the advice from the less biased one. Furthermore, and paradoxically, we show that these reputational incentives can actually reduce the principals ability to make inferences from early recommendations about the evaluators degree of bias. We demonstrate that the overall effect of making evaluators care about their reputations can be either beneficial or detrimental to the organization, and we identify which features of the environment make each of these possibilities more likely.

*Avery: Harvard Kennedy School, Christopher_Avery@hks.harvard.edu; Meyer: Nuffield College and Dept. of Economics, Oxford, margaret.meyer@nuffield.ox.ac.uk.

1 Introduction

Recommendations are the lifeblood of hiring and promotion decisions in a variety of industries, from the military to professional service firms. Many careers require recommendations at a variety of stages. Gaining tenure at an American university requires recommendation letters at no fewer than four critical points: applying to university, to graduate school, for a position as assistant professor, and finally for tenure. Furthermore, those recommendations are often paramount to a succesful decision. In a 2001 survey, 60 United States federal court of appeals judges ranked recommendation letters from familiar professors as one of the three most important factors in their choices of clerks - far more than the 34 who ranked membership on the law review as one of the three most important factors in that decision. (Avery, Jolls, Posner, Roth, 2001)

Despite the importance of recommendation letters, there is no universal standard for assessments. Further, evaluators tend to be biased, particularly when they have trained the very candidates that they are also evaluating. And some evaluators are more biased than others. Yet, there is also no central repository of information about a recommenders past history of evaluations in most job markets. Long-time members of a faculty recruiting committee may notice if a well-known professors evaluates one student each year as "my best student in the past ten years," but the process of accounting for bias in recommendation letters is haphazard at best.

In this paper, we study the effect of bias and reputation in a generalized recommendation process. We analyze a repeated cheap-talk game with one principal and one evaluator. The evaluator assesses the ability of a different candidate in each of two periods, and makes a recommendation to a principal. The principal then decides, on the basis of the evaluator's recommendation, whether or not to hire the current candidate. This is the simplest setting that allows for reputational considerations to affect recommendations.

Keeping track of past evaluations provides an obvious advantage to the principal because it enables her to identify the most partisan evaluators and then discount their recommendations. Yet, once the principal is known to keep track of past evaluations, recommenders will change their actions to preserve their reputations for objectivity. In a world where recommenders tend to be biased in favor of the candidates that they evaluate, there is a reputational cost for making a positive evaluation. This is so, even if the principal updates her beliefs based only on the evaluation and not on any additional information about the candidate's actual ability. As a consequence, early in their careers (i.e., in the first period of the two-period model), recommenders will be systematically tougher than they would be if reputational considerations were absent. Intrinsically unbiased evaluators will tend to apply more rigorous standards than the principal would desire, and even very biased evaluators may become too tough from the principal's standpoint.

While updating beliefs about evaluators' objectivity based on their early behavior enables the principal to better interpret later recommendations from an evaluator, evaluators' anticipation of such updating can reduce the informativeness of recommendations in the first period in two ways. First, the use of tougher first-period standards by evaluators may reduce the principal's ability to distinguish good and bad candidates in the first period. Second, the evaluators' reputational concerns may, paradoxically, reduce the informativeness of their first-period recommendations about their intrinsic bias. That is, the standards applied by more- and less-biased evaluators may become more similar, as well as tougher, when reputational considerations are introduced. Thus, the reputational concerns of evaluators may induce conflicting effects on the quality of hiring or promotion decisions.

Given these conflicting effects of reputational concerns, the key question we address with our dynamic cheap-talk model is: In what environments is the overall quality of the principal's decision-making improved, and in what environments is it worsened, by the principal's tracking of evaluators' past recommendations? In our baseline dynamic model, the only respect in which the two periods might differ ex ante is in the relative weight assigned to them by the principal and the evaluator. In this baseline model, we find that, regardless of the relative weights assigned to the two periods by the evaluator, the principal's payoff in each period in the unique equilibrium is at least as large as in the equilibrium of a one-period version of the model, in which reputational concerns are absent.

We then extend our analysis to allow either the evaluator's bias or the ability distribution of the candidate to vary across periods. In these more general environments, we find that the result from the baseline dynamic model may be overturned: If the intertemporal correlation of the evaluator's bias is sufficiently low, or the distributions of the candidates' abilities in the two periods are sufficiently different, then evaluators' reputational concerns can result in a weakly lower payoff for the principal in each period than if he did not track past recommendations. The lessons from these more general environments are that whether or not evaluators' reputational concerns improve the overall quality of the principal's decision-making depends both on the *absolute* strength of reputational concerns and on the strength of the less biased evaluator's reputational concerns *relative to* those of the more biased evaluator. It is only when both of these measures are sufficiently large that the costs of reputational incentives can outweigh their benefits.

Our model is broadly related to the previous literature on reputation and career concerns, where the agents' reputations, based on actions early in their careers, determine their opportunities and rewards later in their careers. A central theme in this literature is that when "good agents" are characterized by patterns of both outcomes and actions, reputational considerations provide incentives for agents not only to "do well" but also to "look good". When the probablistic rewards are larger for "failing gracefully" than for "succeeding awkwardly", reputational equilibrium outcomes are often inefficient. (See, for example, Scharfstein and Stein (1990), Brandenburger and Polak (1996), Prendergast and Stole (1996), Levy (2004), Prat (2005)). Our model produces career concerns of a different sort, as we assume that agents differ in their underlying motives rather than in their expertise. Reputational considerations induce agents to choose actions that signal preferences that are in accord with the preferences of the principal. Other models that study signaling of preferences in a reputational setting include Sobel (1985), Benabou and Laroque (1992), Ely and Valimaki (2003), and Gentzkow and Shapiro (2006). Our model is most closely related to that of Morris (2001), who considers incentives for agents to take politically correct positions to signal that their underlying preferences are socially acceptable. In contrast to many existing studies of reputational concerns, which model the reputational component of payoffs in a reduced-form fashion, our model explicitly derives the reputational component from the fundamentals of the decision-making environment. This allows us to perform our explicit comparison of the overall quality of decision-making with and without the presence of reputational incentives and to identify how the comparison is affected by the fundamentals of the environment.

Given that the information contained in evaluator recommendations is "soft", and given the issues of confidentiality surrounding many types of job recommendations, it is difficult to find appropriate datasets for testing the implications of our model. Nevertheless, a recent study by Fafchamps and Moradi (2009), using recruitment and performance data from the British colonial army in Ghana, provides evidence of referee opportunism and specifically of the impact of referees' career concerns on the degree of such opportunism. We discuss their work in more detail in Section 6.

The paper is organized as follows. Section 2 describes the model and the equilibrium for a single-period interaction. Section 3 analyzes the equilibrium of the baseline dynamic game with two periods and reputational incentives operative in the first period. Section 4 interprets that equilibrium and compares the principal's welfare in the dynamic game to her welfare if reputational incentives were absent. Section 5 analyzes more general dynamic environments and shows how and when conclusions about the overall welfare effects of reputational incentives can be altered. Section 6 discusses Fafchamps and Moradi's (2009) evidence on referee opportunism, and Section 7 concludes.

2 The Model

We study a repeated cheap-talk model where, in each of two periods, an agent (A) evaluates a candidate for a job that involves working for a principal (P). In each period, the principal decides whether or not to hire the current candidate and is unconstrained in the second period by the decision made about the first-period candidate. The principal and agent both know that the period-t candidate's ability, s_t, is drawn from a uniform distribution on [0,1] and that s_1 and s_2 are independent. At the start of each period, the agent observes the ability of the current candidate precisely and sends a cheap-talk message (i.e. a costless but unverifiable report) about the candidate's ability to the principal. We do not restrict the form of the message. We will use the terms "message" and "recommendation" synonymously, and for clarity, we assume that the principal is female and the agent male.

In each period, the principal has a reservation value of r, the ability of a default worker. We assume that $r > 1/2$, so that the principal is inclined not to hire a candidate without additional positive information about that candidate. The principal receives a payoff of $\pi_t^P = s_t$ if she hires the candidate and $\pi_t^P = r$ if she does not. The agent is biased in favor of the candidate, with bias c_i. The agent receives a payoff of $\pi_t^A = s_t + c_i$ if the candidate is hired and $\pi_t^A = r$ if the candidate is not hired. The agent privately observes his bias, c_i, at the start of the first period. For now, we assume that both c_i and r remain constant across periods: this is our baseline dynamic model. In Section 5, we relax this assumption.

The agent may have either a high bias c_H or a low bias c_L, where $0 \leq c_L < c_H \leq r$. There is prior probability p_0 that the agent has a low bias and $1-p_0$ that the agent has a high bias. Given this payoff structure, the principal prefers the worker to be hired if and only if $s_t \geq r$, while the agent prefers the worker to be hired if and only if $s_t \geq r - c_i$. Thus, the incentives of the principal and agent are aligned for workers of extreme abilities: they both prefer the worker to be hired if $s_t > r$, and both prefer the worker not to be hired if $s_t < r - c_i$. Their interests conflict, however, for workers with abilities in the range $(r - c_i, r)$, where the bias of the agent overcomes the shortfall in the worker's ability relative to the reservation value r.[1]

We assume for the rest of the paper that $c_L < 1 - r < c_H$, which provides incentives for the principal to distinguish between type-H and type-L agents. In a one-period game with no reputational considerations, each type of agent would recommend candidates with abilities of $r - c_i$ or higher, corresponding to conditional expected abilities of $(1 + r - c_i)/2$ for a candidate recommended by an agent of type-i. The assumption $c_L < 1 - r < c_H$ implies that $(1-r+c_L)/2 > r > (1-r+c_H)/2$ so that the principal prefers to hire a worker recommended by a type-L agent, but prefers not to hire a worker recommended by a type-H agent.

Both principal and agent are risk neutral. The agent's aggregate payoff for the two periods

[1]Contrast this to Benabou and Laroque, etc., where the agent does not have this overlap of values with the principal.

is a weighted sum of his per-period payoffs: $\pi_1^A + \delta\pi_2^A$. The weighting factor δ represents the importance of the second-period decision for the agent relative to the first-period decision and may lie anywhere in the interval $[0, \infty)$. The principal's aggregate payoff for the two periods is also a weighted sum of per-period payoffs $\pi_1^P + \delta_P \pi_2^P$, with $\delta_P \in [0, \infty)$.

3 The Static Equilibrium

The static equilibrium (which will be played in the second period of a two-period game) is simpler than the equilibrium in the well-known Crawford and Sobel (1982) model. The principal in our model faces a binary decision (hire or not hire) while the principal in the Crawford and Sobel model faces a continuous decision.

Proposition 1 *There is a cutoff value $\overline{p}(r, c_H, c_L)$ for the existence of an informative equilibrium of the static game.*

(1) If $p_0 \geq \overline{p}$, then there is an equilibrium of the static game with binary recommendations where an agent of type i makes a positive recommendation "Y" if $s \geq r - c_i$ and a negative recommendation "N" if $s < r - c_i$, and the principal hires a candidate if and only if she receives a positive recommendation. All other informative equilibria of the static game are equivalent in outcome to this binary recommendation equilibrium.[2]

(2) If $p_0 < \overline{p}$, there is no informative equilibrium of the static game.

Proof: See Appendix.

The agent's role in the static game is quite simple since he has a strict preference for hiring the candidate if $s > r - c_i$ and a strict preference against hiring if $s < r - c_i$. Given any set of positive recommendations, the agent will make the most positive recommendation ("Yes" or Y) if $s > r - c_i$ and the most negative recommendation ("No" or N) if $s < r - c_i$. We denote the type-specific thresholds for positive recommendations as $z_H = r - c_H$ and $z_L = r - c_L$.

[2]If $s = r - c_i$, then an agent of type i is indifferent between recommending and not recommending the candidate. We assume that the agent recommends the candidate in this case.

The principal's decision rule depends on the underlying parameters $r, c_H, c_L,$ and p_0. But since we assume that $r > 1/2$, the principal never hires the candidate after a negative evaluation. The default decision for the principal without further information is not to hire the candidate. Thus,:if the agent is not enthusiastic about hiring the candidate, then the principal certainly does not want to do so. With uncertainty about the agent's type, the principal would hire after a positive recommendation in the static game if the probability of a type-L agent is sufficiently high ($p_0 \geq \overline{p}$) to ensure $E(s|Y) \geq r$. We solve directly for $\overline{p}$ in the proof of Proposition 1 in the appendix: $\overline{p} = \frac{c_H^2-(1-r)^2}{c_H^2-c_L^2}$. If $p_0 < \overline{p}$, then a positive recommendation by the agent is not sufficiently credible to influence the principal's decision. In this case, all equilibria of the static game involve babbling (i.e. uninformative recommendations) by the agent and probability zero that the principal hires the candidate.

4 Dynamic Equilibrium

In the two-period model, the agent evaluates one candidate in each of two successive periods and sends successive messages m_1 and m_2 to the principal. We assume that the agent does not know the ability of the period-2 candidate at the time that he sends message m_1 to the principal. The principal observes each message and chooses in turn whether to hire each of the two candidates. This dynamic model produces reputational considerations for both principal and agent because the principal can update her assessment of the agent's type after observing the agent's recommendation for the first candidate.[3] We use the terms "message" and "recommendation" interchangeably to describe the information provided by the agent to the principal in each period. Both principal and agent anticpate that in period 2, the agent will use the static equilibrium thresholds.

Reputation in this model is summarized by the principal's updated probability, p_1, that the agent has low bias after observing the first period recommendation. In particular, reputation is important if p_1 is above $\overline{p}$ after (at least) one first-period recommendation by the

[3] We assume that the principal does not observe the true ability of candidate 1 until after making a hiring decision for candidate 2. Section 5.2 discusses an extension of the model where the principal does observe the ability of candidate 1 prior to period 2.

agent, and below $\overline{p}$ after (at least) one other first-period recommendation. Then the principal follows the agent's advice in the second period after some first-period recommendations but ignores that second-period advice after other first-period recommendations.

The agent's concern for his reputation derives from the influence of that reputation on the principal's decision rule in the second period. As a result, reputation can be precisely valued in terms of an option value Δ_i equal to the δ-weighted gain in the agent's expected payoff when the principal changes decision rule from not following the agent's advice to following the agent's advice with probability 1 in the second period. The ex ante probability a candidate's being hired when the agent uses the static reporting thresholds is $1-z_i = 1-r+c_i$, producing an average increase in payoff of $(1-r+c_i)/2$ to the agent when the candidate is hired.[4] The (unweighted) option value to the agent is then $\Delta_i = \frac{(1-r+c_i)^2}{2}$. Not surprisingly, a type-H agent places more value on reputation than does a type-L agent. Proposition 2 uses this option value to characterize the nature of dynamic equilibrium in the game.

Proposition 2 *In any informative equilibrium of the dynamic game:*

(1) Each type of agent chooses one of two distinct first-period messages: "Positive" (Y_1) and "Negative" (N_1).

(2) Each type of agent follows a threshold strategy for first-period recommendations. A type-i agent makes a negative first-period recommendation if $s_1 < z_i$ and a positive first-period recommendation if $s_1 > z_i$, where s_1 represents the ability of the first-period candidate and $z_i \geq r - c_i$ represents the threshold for positive first period message for type i.

(3) The principal hires the first-period candidate with positive probability after Y_1 and does not hire the first-period candidate after N_1.

(4) $z_L \geq z_H$, so a positive first-period recommendation causes the principal to lower his assessment that the agent's bias is low, while a negative first-period recommendation causes her to raise it.

Proof: See Appendix.

[4] The conditional distribution of the candidate's ability given a positive report in static equilibrium is $U(z_i, 1)$, producing a conditional expected ability of $(1+z_i)/2 = (1+r-c_i)/2$. The agent gains an additional c_i whenever the candidate is hired. Relative to the reservation value, then, the agent gains an average of $(1+r-c_i)/2 + c_i - r = (1-r+c_i)/2$ per positive report.

Proposition 2 characterizes equilibrium behavior in the two-period model, indicating that there are at most two distinct messages in each period. The principal will not hire after N_1, since the expected ability of the period-1 candidate, conditional on N_1, is less than 1/2 which type of agent she is facing. In any informative equilibrium, then, the principal must hire with positive probability after Y_1. In fact, when $p_0 \geq \overline{p}$, the principal hires with probability 1 after Y_1, since each type of agent adopts a weakly higher threshold for a positive first-period recommendation than in the equilibrium of the static game. A positive recommendation in the first period serves as a noisy signal of a type-H (more biased) agent and a negative recommendation in the first period serves as a noisy signal of a type-L (less biased) agent, so in an informative equilibrium, $p_1 < p_0$ after recommendation Y_1 and $p_1 > p_0$ after recommendation N_1.

Proposition 3 *There is at most one informative equilibrium in the dynamic game for any combination of parameters (δ, r, p, c_L, c_H). Reputational incentives arise in the first period if and only if $p_l \leq p_0 < p_u$, where these values do not depend on δ. For $p_0 \geq p_u$, equilibrium behavior in each period matches the static equilibrium, while for $p_0 < p_l$, there is no informative equilibrium at all.*

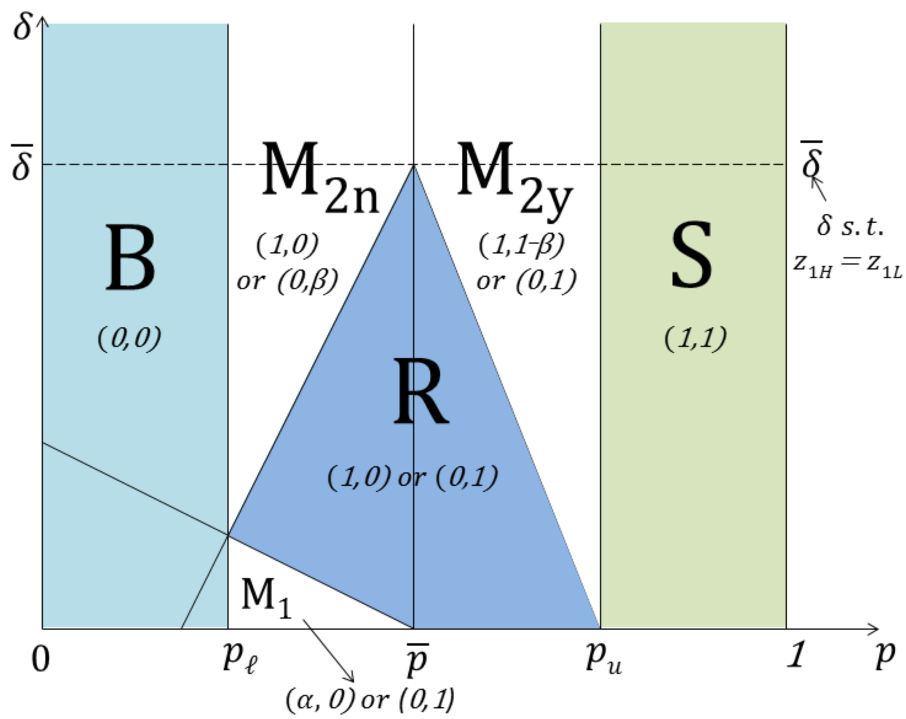


Figure 1 demonstrates the nature of equilibrium as a function of δ and p. From the agent's perspective, the principal's strategy can be summarized by two parameters:

$\alpha =$ P(Hire in Period 1 | Y_1)

$\beta =$ P(Hire in Period 2 | N_1, Y_2)- P(Hire in Period 2 | Y_1, Y_2).

Given these parameters, the agent's reputational gain from N_1 is equal to $\beta\delta\Delta_i$, while the agent's immediate gain in payoff from Y_1 is $\alpha(s_1 + c_i - r)$. Comparing these values, a type i agent is indifferent between Y_1 and N_1 if $\alpha(s_1 + c_i - r) = \beta\delta\Delta_i$, producing first-period ability threshold $z_i = (\beta/a)\,\delta\Delta_i + r - c_i$, where the term $(\beta/a)\,\delta\Delta_i$ represents the agent's adjustment in threshold due to reputational considerations. Of course, the principal's decision rule varies endogenously with the agent's strategies, so the parameters α and β must be compatible with the principal's incentives induced by thresholds z_L and z_H. Since the effect of reputation on thresholds is determined by the multiplier $\beta\delta/a$, we define a separate variable for the reputational multiplier $K = \beta\delta/\alpha$. Formally, we write the principal's updated probability after observing the first-period recommendation as $p_1 = p^*_K(p_0, m_1)$ where m_1 is the message sent by the agent in period 1.

We now discuss the incentives for the principal and agent and the nature of equilibrium play in each of the separate regions of Figure 1.

Region S

In an extreme case where prior probability p_0 is very close to 1 ($p_0 > p_u$), the principal's strong inclination to follow advice outweighs the information revealed by first period recommendation based on static equilibrium cutoffs. Then there is a pure strategy equilibrium with no reputational considerations ($\alpha = 1$,$\beta = 0$) and agent thresholds equal to static equilibrium thresholds in each period. This case corresponds to Region S in Figure 1, where $p^*_{K=0}(p_0, Y_1), p^*_{K=0}(p_0, N_1) \geq \bar{p}$ so that the principal follows advice in the second period regardless of the recommendation in the first period. The left-hand boundary of Region S is implicitly defined by the equation $p^*_{K=0}(p_0, Y_1) = \bar{p}$, corresponding to $p_0 = p_h$ since the agent's thresholds don't vary with δ. (Note that since p_u is defined by $p^*_{K=0}(p_u, Y_1) = \bar{p}$, $p_u > \bar{p}$ so it is optimal for the principal to follow the agent's advice in the first period in this region.)

Region R, $p_0 \geq \overline{p}$

By contrast, if $p_0 > \overline{p}$, but p is quite close to $\overline{p}$, then the principal has incentive to vary the second period decision rule with the first-period recommendation, folllowing the agent's advice in the second period after observing N_1, but ignoring the agent's advice in the second period after observing Y_1. Then there is a pure strategy equilibrium with reputational considerations ($\alpha = 1, \beta = 1$). This case corresponds to the right-hand part of Region R in Figure 1 (with $p_0 \geq \bar{p}$). In Region R, $p^*_{K=\delta}(p_0, Y_1) \leq \bar{p}, p^*_{K=\delta}(p_0, N_1) \geq \bar{p}$ so that the principal will only follow the agent's advice in the second period after a negative first-period recommendation. The right-hand boundary of Region R is implicitly defined by the equation $p^*_{K=\delta}(p_0, Y_1) = \bar{p}$. As shown in Figure 1, this boundary varies with δ since the agent's first-period thresholds vary with δ.

Existence of a unique informative equilibrium for $p_0 \geq \overline{p}$ results from an important technical property of the game: an increase in multiplier K causes recommendations in a reputational equilibrium to be less informative about an agent's type. In words, an increase in K (so that types H and L adjust their first period thresholds by greater degrees for reputational purposes) reduces the degree of updating by the principal about the agent's type based on the first-period recommendation so that $p_0 - p^*_K(p_0,Y_1)$ and $p^*_K(p_0,N_1) - p_0$ are both declining in K.

Region M$_{2Y}$

At the boundary of Region S, $p^*_{K=0}(p_u, Y_1) = \overline{p}$, so $p^*_{K=\delta}(p_u, Y_1) > \bar{p}$ for any $\delta > 0$. Further, for values of p just below p_u (i.e. $p_0 = p_u - \varepsilon$), $p^*_{K=0}(p_0, Y_1) < \overline{p}$ and $p^*_{K=\delta}(p_0, Y_1) > \bar{p}$. That is, for $p_0 = p_u - \varepsilon$, first-period thresholds based on multiplier $K = 0$ induce "too much" updating by the principal to allow a pure strategy equilibrium in region S, while first-period thresholds based on $K = \delta$ induce "too little" updating by the principal to allow a pure strategy equilibrium in region R. Thus, as shown in Figure 1, there is a gap between regions R and S for any $\delta > 0$. But since $p^*_{K=0}(p_0, Y_1) < \overline{p}$ and $p^*_{K=\delta}(p_0, Y_1) > \bar{p}$, there is a unique value K_M between 0 and δ such that $p^*_{K=K_M}(p, Y) = \overline{p}$. This value K_M corresponds to parameters ($\alpha = 1$, $\beta = K_M/\delta$), which in turn results from a mixed strategy where the principal follows the agent's advice in the second period with probability K_M/δ after a positive recommendation in the first period. Thus, there is a unique mixed strategy equilibrium in region M_{2Y} for each $(p_{0,}\delta)$ in the area between regions R and S in Figure 1.

Region B

Equilibrium analysis follows a similar pattern for $p_0 < \overline{p}$. In an extreme case where prior probability p is very close to 0, the principal's strong inclination not to follow advice outweighs the information revealed by first period recommendation based on static equilibrium cutoffs. Then there is a pure strategy equilibrium with no reputational considerations ($\alpha = 0$,$\beta = 0$) and agent thresholds equal to static equilibrium thresholds in each period. This case corresponds to Region B in Figure 1, where $p^*_{K=0}(p_0, Y_1), p^*_{K=0}(p_0, N_1) \leq \overline{p}$ so that the principal does not follows advice in the second period after observing either of the two possible recommendations in the first period. But since the principal never hires a candidate in either period under these conditions, the equilibrium is uninformative and is equivalent to babbling.

Region R, $p_0 < \overline{p}$

If $p_0 < \overline{p}$, but p_0 is quite close to $\overline{p}$, then once again, there is a pure strategy equilibrium with reputational considerations ($\alpha = 1, \beta = 1$). This case corresponds to the left-hand part of Region R in Figure 1 where $p^*_{K=\delta}(p_0, Y_1) \leq \overline{p}, p^*_{K=\delta}(p_0, N_1) \geq \overline{p}$. One part of the left-hand boundary of Region R is implicitly defined by the equation $p^*_{K=\delta}(p_0, N_1) = \overline{p}$.

Regions $\mathbf{M}_1$,$\mathbf{M}_{2N}$

One additional complexity with $p_0 < \overline{p}$ is that the principal will not follow advice based on the static equilibrium cutoffs. Thus, an informative equilibrium also has to satisfy an additional condition that the first-period recommendations are sufficiently informative so that the principal is willing to hire the candidate after a positive first-period recommendation: $E(s_1|Y_1, p_0, K) \geq r$ (where δ and reputational multiplier K are sufficient to identify the first-period thresholds for each type of agent). This constraint creates an additional lower left-hand boundary for Region R: $E(s_1|Y_1, p_0, K) = r$. As shown in Figure 1, when $p_0 < \overline{p}$, the reputational region R consists of those combinations of parameters that lie to the right of the boundary $E(s_1|Y_1, p_0, K) = r$ and to the left of the boundary $p^*_{K=\delta}(p_0, N_1) = \overline{p}$ For parameters that lie to the right of one of these two boundary lines, but not the other, there is again a unique mixed strategy equilibrium where the principal adopts a mixed strategy in period 1 (in Region M_{2N}) or in period 2 after a negative recommendation in period 1 (in Region M_{2Y}).

4.1 Welfare Analysis

Our interst in studying the welfare properties of the equilibrium of the dynamic game is to compare the principal's payoffs for in the two-period dynamic equilibrium and two successive periods of static equilibrium play. That is, if the principal is going to interact with the same agent for two successive periods, does the principal gain in payoff by updating the probability of the agent's type based on first period play (and making clear to the agent that she is doing so)? If $p_0 < p$, then the static equilibrium produces babbling in each period, so any informative dynamic equilibrium must be an improvement for the principal. For this reason, we concentrate on the case $p_0 \geq \bar{p}$ in the rest of our welfare analysis.

The principal gains information about the agent's type By observing the first-period recommendation in a dynamic equilibrium; this information can only improve the principal's expected value for second-period hiring decisions. In fact, we also find that the princpal gets higher expected first period payoff in a dynamic equilibrium than in the corresponding static equilibrium. Thus, the principal unambigiously prefers the dynamic case to the twice-repeated static outcome, indicating that the principal can only gain from tracking the agent's reputation in this model.

Proposition 4 *For any combination of (δ, p_0), the principal's equilibrium expected payoff in each period of the baseline dynamic model is at least as large as in the equilibrium of the static model.*

Intuitively, if $c_L > 0$, small increases in reputational multiplier K from $K = 0$ (e.g. by an increase in δ from $\delta = 0$) rein in each type of agent from positive recommendations for candidates with abilities in some range of values below r, the principal's reservation value, which can only increase the principal's first period payoff. If K is sufficiently large, however, both types of agents will set first-period thresholds for positive recommendations that are above the principal's reservation value r, indicating that reputational considerations can cause the agent to set standards that are too tough rather than too lax. As shown in Figure 1, the combination of possible values (δ, p_0) that produce a pure strategy dynamic equilibrium is limited by the constraint $p_K^*(p_0, Y_1) < \overline{p}$. As δ increases, first-period recommendations based on the pure-strategy thresholds $r - c_i + \delta\Delta_i$ become less informative about the agent's type, and so there is no pure strategy equilibrium in region M_{2y}. Our proof of Proposition 4, presented in the Appendix, demonstrates that the combination of values (δ, p_0) that produce a pure-strategy equilibrium is sufficiently limited that the principal actually prefers the first-period outcome in a dynamic equilibrium in region R to the static equilibrium for the same p_0. (We demonstrate further that any mixed-strategy dynamic equilibrium in region M_{2y} yields an first-period expected payoff for the principal that is equivalent to the payoff in a specific pure strategy equiibrium in region R, so that the principal cannot do better with the static equilibrium than with a particular mixed-strategy dynamic equilibrium in region M_{2y}.)

4.2 Agent Discretion and Rules-of-Thumb Management

The dynamic equiliibrium can be summarized by the number of candidates that the principal will hire in expectation if the agent provides positive recommendations in each period. In Region S, the principal follows advice in each period, so we say that the agent has authority discretion to hire (up to) two candidates or "discretion equal to 2". Similarly, in region R, the agent has discretion equal to 1 and in region B, the agent has discretion equal to 0. In mixed strategy region M_1, the principal hires in the first period with probabilty α after a positive recommendation (but will not hire in the second period after a positive recommendation in the first period) and so we say that the agent has discretion equal to α.(even though the principal will follow advice in the second period with probabliity 1 after a negative first period recommendation). Similarly in mixed strategy region M_{2Y}, the principal hires in the second period with probability 1-β after positive recommendations in both periods, so we say that the agent has discretion equal to 2-β. Finally, in mixed strategy region M_{2N}, the principal hires in the first period with probability 1 after a positive recommendation and hires in the second period with probability β after a positive recommendation in the first period and a negative recommendation in the second period. For the purposes of comparative static analysis within region M_{2N}, we say that the agent has discretion β in this region, even though this is not consistent with our definition of discretion in the other regions.

Proposition 5 *For $\delta > 0$, the agent's discretion is weakly increasing in p_0.*

Returning to Figure 1, as p_0 increases from 0 to 1, equilibrium play typically moves from Region B (where the agent has discretion 0) to either mixed strategy region M_1 or M_{2N} (where the agent has discretion between 0 and 1) to Region R (where the agent has discretion 1), to region M_{2Y} (where the agent has discretion between 1 and 2) to Region S (where the agent has discretion 2). Each change of regions as p_0 increases produces a natural increase in agent discretion, so it remains only to verify that discretion is weakly increasing in p_0 within each region. Our comparative static analysis verifies this property and also demonstrates that discretion is continuous in p_0 for $p_0 > p_l$.

Our definition of discretion and the finding that discretion is weakly increasing in p_0 produces an intutive property of dynamic equilibrium in this model. The principal is willing to place more weight on advice from an agent with stronger prior reputation than on advice from an agent with weaker prior reputation as summarized by p_0. The connection between informative dynamic equilibrium and agent discretion also suggests a possible connection to behavior in practice. While we certainly do not expect recommenders and employers to understand the dynamics of equilibrium play as summarized in Figure 1, it is natural to imagine that employer-recommender dynamics could follow rule-of-thumb interactions, corresponding to the numerical discretion level given to the recommender in any dynamic equilibrium.

The choice of dynamic equilibrium in Figure 1 need not, however, match the principal's optimal choice of discretion in a contracting game where the principal chooses the level of discretion to grant the agent prior to the start of period 1. Given the choice of contracting discretion, the principal would necessarily do at least as well or better than in repeated static interactions, just as in Proposition 3. But the choice of contracting discretion also gives the principal an additional advantage, essentially allowing the principal to commit in advance to suboptimal second-period rules.

For example, with $p_0 = p_u$, the agent adopts the same threshold for a positive recommendation in the first period in the dynamic equilibrium as in the static equilibrium. Then the principal is indifferent between following advice and not following advice in the second period after a positive recommendation in the first period, corresponding to discretion 2 for the agent. But if the principal could choose the level of discretion to give to the agent, he would prefer to set discretion less than 2 with $p_0 = p_u$, as this would cause each type of agent to increase the threshold for a positive recommendation in the first period. An incremental reduction in discretion with $p_0 = p_u$ instantaneously improves the principal's first-period payoff, but has no marginal effect on second-period payoff since the principal is indifferent in the second period after a first period positive recommendation.

The dynamic equilibrium corresponds to the optimal choice of discretion by the principal when the principal cannot commit in advance to a second-period decision rule. In this sense, tracking the reputation of the agent adds only a single option for the principal when $p_0 \in (p_l, p_u)$. The principal can follow advice in both periods (twice repeated static equilibrium), follow advice in neither period (twice repeated babbling), or adopt the decision rule corresponding to the dynamic equilibrium indicated in Figure 1. Proposition 4 confirms that the dynamic equilibrium can only improve the principal's expected payoff from the baseline of twice-repeated static equilibrium – i.e. that the dynamic equilibrium can be viewed as the solution to the principal's constrained maximization in the choice of agent discretion.

5 Extensions of the Model

5.1 Changes in the Game from Period 1 to Period 2

We first consider two changes in the game so that period 2 need not reproduce the structure of the stage game in period 1. Here, we find that seemingly small changes in the model can overturn the result of Proposition 4, so that inducing reputational incentives by tracking past evaluations can hurt the principal. In these instances, the cost of inducing overly-stringent standards in the first period may outweigh the principal's gain from knowing more about the agent's bias in the second period. This holds in (at least) two natural extensions of the model: The first involves different reservation values in the two periods: $r_1 \neq r_2$. The second involves imperfect correlation in the agent's bias across periods 1 and 2, so that a type-H agent is not certain to be a type-H agent in the second period.

Proposition 6 *Assume $r_2 > 1/2$ and $0 \leq c_L < 1 - r_2 < c_H \leq r_2$. Then there exists $\hat{r}_1(r_2, c_H, c_L)$ such that*

1. For $r_1 \in [c_H, \hat{r}_1]$ (i.e. for r_1 not too different from r_2) and for all (p_0, δ), the principal's expected payoff in each period of the dynamic equilibrium is weakly larger than in the equilibrium of the corresponding one-period game.

2. For $r_1 >, \hat{r}_1$, we can find, for all δ_P, (p_0, δ) such that the principal's overall expected payoff in the dynamic equilibrium is smaller than in the corresponding sequence of two one-period games.

When $r_1 \neq r_2$ but the two values are not too disparate, (i.e. $r_1 < \hat{r}_1$), then the dynamic equilibrium takes the same qualitative form as in Figure 1, and the welfare comparison from Proposition 4 continues to hold. In this case, the principal clearly benefits from tracking reputational incentives in each of the two periods of the dynamic game.

By contrast, when $r_1 > \hat{r}_1$, reputational incentives cause the first-period recommendation to become more informative about the agent's type rather than less so. Although $\Delta_H > \Delta_L$ continues to hold – i.e. there are stronger reputational incentives for type H than for type L, $\Delta_H/\Delta_L < (1 - r_1 + c_H)/(1 - r_1 + c_L)$ so that the proportional difference in reputational incentives is not as large as the proportional difference of probabilities of positive recommendations in the static equilibrium. As a result, $r_1 > \hat{r}_1$ induces an important technical change in the nature of the game. For δ sufficiently large and $p_0 > \bar{p}_2$, there is a pure strategy reputational equilibrium where $z_L = r_1 - c_L + \delta\Delta_L$ approaches 1 while $z_H = r_1 - c_H + \delta\Delta_H$. That is, if reputational incentives are strong enough, type L will not make a positive recommendation in period 1, which in turn supports a pure strategy reputational equilibrium with $p_0 > \bar{p}_2$ since a positive recomendation is a near-perfect signal of type H.

Extending this reasoning, with a larger value of δ, type L will choose first period threshold of 1 (will absolutely never make a positive recommendation) and type H will choose first period threshold just slightly below 1. This reputational equilibrium provides near certain rejection of the first period candidate, and also provides the principal with infinitesimal information about the agent's type. The principal clearly loses in payoff in the first period of this dynamic equilibrium by comparison to the static equilibrium and gains only marginally in the second period of hte dynamic equilibrium by comparison to the static equilibrium. Combining these two effects, the principal prefers the twice-repeated static game to the dynamic game (i.e. prefers not to track agent reputation) when $r_1 > \hat{r}_1$, $p_0 > \bar{p}_2$ and δ sufficiently large.

With imperfect correlation of types (but the same two possible types in each period), we define transition probabilities $P(c_2 = c_j | c_1 = c_i) = p_{ij}$. Given prior probability $P(c_1 = c_L) = p_0$, we assume that $P\left(c_2 = c_L\right) = p_0$, so the unconditional expected level of agent bias is the same in each period. Given this assumption, the four transition probabilities p_{ij} are all determined by the two parameters p_0 and κ, where κ is the correlation between c_1 and c_2.

Proposition 7 *Suppose that $r_1 = r_2 = r$ and assume that c_1 and c_2 are imperfectly positively correlated with $\kappa = corr(c_1, c_2) > 0$. Then there exists $\hat{\kappa}(r, c_H, c_L) \in (0, 1)$ such that*

1. For $\kappa \geq \hat{\kappa}$ and for all (p_0, δ), the principal's expected payoff in each period of the dynamic equilibrium is weakly larger than in the equilibrium of the corresponding static model.

2. For $\kappa < \hat{\kappa}$, we can find, for all δ_P, (p_0, δ) such that the principal's overall expected payoff in the dynamic equilibrium is smaller than in the twice-repeated equilibrium of the corresponding static model.

Imperfect correlation of bias across periods changes the nature of first-period play through a change in the reputational values Δ_H and Δ_L. With imperfect correlation, Δ_H and Δ_L

are a probabilistic mixture of the original values $(1-r+c_H)^2/2$ and $(1-r+c_L)^2/2$ since the agent is not certain in period 1 whether he will remain the same type in period 2:

$$\Delta_L = p_{LL}(1-r+c_L)^2/2 + (1-p_{LL})(1-r+c_H)^2/2$$

$$\Delta_H = p_{HH}(1-r+c_H)^2/2 + (1-p_{HH})(1-r+c_L)^2/2$$

When κ is relatively large, p_{LL} and p_{HH} are close to 1 and the incentives in the game are much like those that gave rise to Figure 1. Under these conditions ($\kappa \geq \hat{\kappa}$), the welfare result of Proposition 4 continues to hold. However, when κ is relatively small, $\Delta_H - \Delta_L$ becomes small as well, and once again, reputational incentives cause the first-period recommendation to become more informative about the agent's type. Then as in the case where $r_1 \neq r_2$, for δ sufficiently large, there are pure-strategy reputational equilibria where type L never makes a positive recommendation in period 1 and type H only makes a positive recommendation with infinitesimal probability in period 1. Since the principal learns almost nothing about the agent's type or about the first period candidate's ability, the principal prefers the twice-repeated static game to the dynamic game under these conditions.

5.2 Non-Uniform Distribution of Candidate's Ability

Our primary reason for assuming a uniform distribution for the candidate's ability is analytical tractability. In particular, the uniform distribution supports concise formulae for reputational value and straightforward computation of probability thresholds for the principal (e.g. such as $\bar{p}$). We suspect that many other distributions of ability would produce similar qualitative results, but with much more complicated calculations required. At the same time, assuming a uniform distribution of abilities does, no doubt, rule out some properties of equilibrium that might occur with other distributions.

One important technical property of the uniform distribution, as discussed above, is that it produces likelihood ratios that are monotonic in the importance of reputation (as captured by the reputational multiplier K), with the implication that the first-period recommendation becomes less informative about the agent's type as the reputational multiplier increases. Beyond the case of the uniform distribution, the likelihood ratio need not be monotonic in the reputational multiplier, which gives rise to the possibility of multiple equilbria. However, we also find that all distributions share with the uniform distribution the feature that $\bar{z} < 1$, i.e. that the threshold values $z_L(K) = r - c_L + K\Delta_L$ and $z_H(K) = r - c_H + K\Delta_H$ cross at or below 1 [5] Thus, in general, first-period recommendations become less informative as the importance of reputation increases, ruling out the types of examples that we used with imperfect correlation of types and $r_1 \neq r_2$ to show that reputational incentives for the agent can be harmful for the principal's payoff. We can produce examples with a discrete set of possible abilities with multiple equilibria, where the principal does better in one dynamic equilibrium, but worse in another dynamic equilibrium, than in the twice-repeated static equilibrium, but these examples tend to be knife-edge cases with $c_L = 0$ and δ close to 0. We conjecture that the principal gains from tracking the agent's type for almost all continuous distributions of agent ability.

5.3 Revelation of the Abliity of the First-Period Candidate

We have assumed that the principal only learns information about the ability of the first-period candidate from the recommendation of the agent prior to hiring. This excludes the possibility that the principal can learn more information about the ability of the first period candidate during that candidate's subsquent career, and thus can infer additional information about the agent's type. We now consider a natural extension to the game where first-period candidate's ability is revealed after the hiring decision in the first period, but before the start of the second period. In this extension of the game, the principal's updated probability that the agent is type-L is a function of both the agent's first period recommendation and the realized ability of the first-period candidate.

We focus on threshold equilibria where type H uses threshold z_{1H} and type L uses threshold z_{1L} for a positive recommendation in period 1. When both types use threshold strategies, updating by the principal is particularly simple. For candidates of abilty $z < \min(z_{1H}, z_{1L})$ both types of agents make negative recommendations, while for candidiates of ability $z \geq \max(z_{1H}, z_{1L})$, both types of agents make positive recommendations in period

[5] Specifically, for any distribution of abilities and associated reputational option values Δ_H and Δ_L, there exists K^* such that $z_H(K) = z_L(K) \leq 1$ at $K = K^*$.

1. In either case, the principal learns nothing about the agent's type from observing the combination of recommendation and candidate's ability, so the principal assigns updated probability $p_1 = p_0$ on the equilibrium path. However, for candidates with ability z satisfying $\min(z_{1H}, z_{1L}) \leq z < \max(z_{1H}, z_{1L})$, the two types of agents make different recommendations . Then, the principal can identify the type of agent perfectly from observing the combination of recommendation and candidate's ability, so the principal assigns probability 1 (correctly) to one type of agent in period 2.

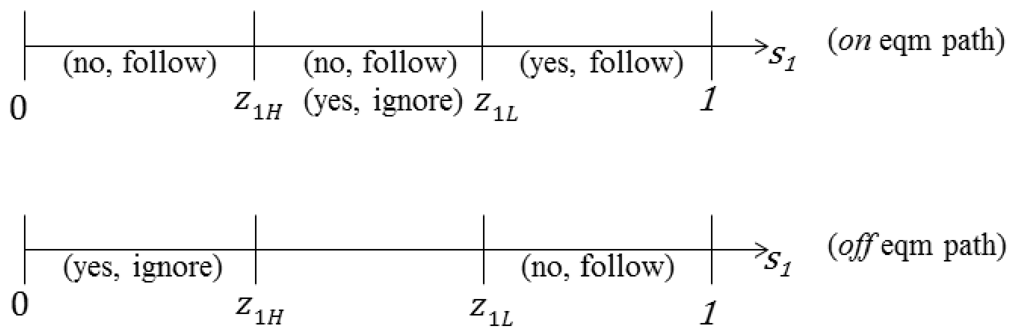
Under these conditions, the game remains a noisy signaling game, but it takes on more of the character of a Spence-signaling game. Given $p_0 \geq \bar{p}$, the principal will follow advice in the second period unless first-period play reveals unfavorable information about the agent's type. Then a type-H agent has incentive to incur costs to prevent revelation of his type and equilibrium play takes on a pooling form. In a two-recommendation pure strategy equilibrium, type H will sacrifice up to $\delta\Delta_H$ in first period payoff in order to pool with type L. Intuitively, then there is an equilibrium where type L adopts the threshold $z_{1L} = z_{2L} = r - c_L$, with no adjustment in threshold for reputational purposes, and where type H adopts the threshold $z_{1H} = \min(r - c_L, r - c_H + \delta\Delta_H)$. That is, type H adjusts the threshold upwards by as much as $\delta\Delta_H$ to to pool with type L, but it is not necessary for type H to adjust threshold to any value higher than z_{1L} in order to pool with type L.

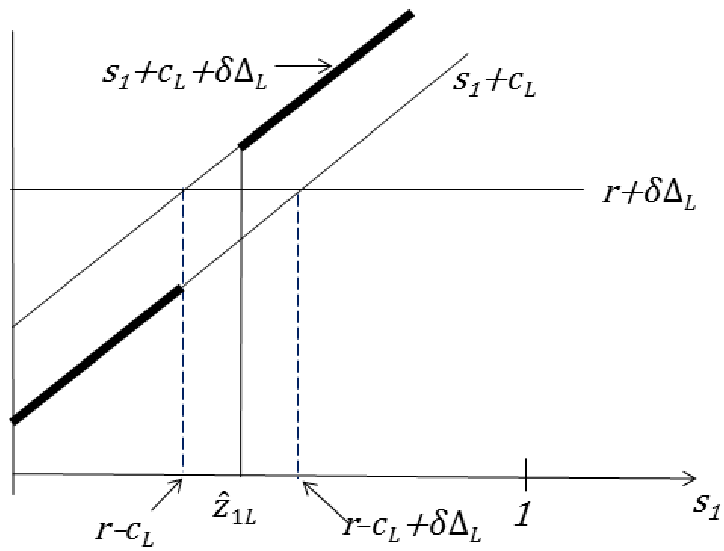
Proposition 8 *If $p_0 > \bar{p}$ and $r - c_H + \delta\Delta_H < r - c_L$, there is a range of threshold equilibria with $z_{1H} = r - c_H + \delta\Delta_H$ and $z_{1L} \in [r - c_L, r - c_L + \delta\Delta_L]$. If $p_0 > \bar{p}$ and $r - c_H - \delta\Delta_H \geq r - c_L$, there is a range of threshold equilibria with $z_{1H} = r - c_H + \delta\Delta_H$ and $z_{1L} \in (r - c_H + \delta\Delta_H, r - c_L + \delta\Delta_L]$ and there is also a range of pooling equilibria with $z_{1H} = z_{1L} \in [r - c_L, r - c_L + \delta\Delta_L]$.*

Proposition 8 demonstrates that there is a range of threshold equilibria whenever $p_0 > \bar{p}$. As discussed above, type H will adjust the first-period threshold up from $r - c_H$ in order to (partially or fully) pool with type L. By contrast, type L can only gain from revelation of type and seemingly has no incentive with $p_0 > \bar{p}$ to change the first period threshold from the static equilibrium value of $r - c_L$ This logic points to the equilibrium with $z_{1L} = r - c_L$,

$z_{1H} = \min(r - c_H + \delta\Delta_H, r - c_L)$ where type L does not adjust the threshold for reputational purposes and type H increases threshold as necessary (up to $r - c_H + \delta\Delta_H$) to pool with type L. In the other equilibria described in Proposition 8, type L increases the threshold from $r - c_L$ as the result of self-fulfilling beliefs. If the principal believes that only a type H agent would make a positive report for a candidate of ability $z \in [r - c_L, z^*)$where $z^* < r - c_H + \delta\Delta_H$, then type L agent will choose $z_{1L} \geq z^*$, thereby justfiying those beliefs. (See Figure 2.)

Our particular interest is in the welfare implications of these equilibria with $p_0 > \bar{p}$. Given the existence of multiple equilibria, we focus on those with $z_{1L} = r - c_L$, and $z_{1H} = \min(r - c_L, r - c_H + \delta\Delta_H)$ to match the intution that there is no need for the preferred type (in this case type L) to deviate from most-preferred single-period action when there are incentives for the less-preferred type (in this case type H) to pool. Revelation of first-period ability in combination with these thresholds is a straightforward improvement for the principal over the twice-repeated static equilibrium. Type L chooses the same first-period threshold while type H sets a higher first-period threshold than in the static equilibrium. Here, type H́'s desire to pool with type L induces negative first-period recommendations for $r - c_H \leq a_1 < z_{1H} \leq r$ which is an improvement for the principal in terms of first-period payoff. Further, the principal can only gain in second period payoff given additional information about the agent's type. So, period by period, the principal prefers the equilibrium of Propostion 7 to twice repeated static play. Welfare comparisons between the scenarios where the first-period candidate's ability is and is not revealed after period 1 are more ambiguous, as described in Proposition 8.





Proposition 9 *As $p_0 \to 1$, the principal achieves higher equilibrium payoff in the equilibrium of the reputational game where the candidate's first-period ability is revealed and $z_{1L} = r - c_L$, and $z_{1H} = \min(r - c_L, r - c_H + \delta\Delta_H)$ than in the reputational game where the candidate's first-period ability is not revealed. But as $p_0 \to \bar{p}$, then if $2c_L/\Delta_L > (c_H - c_L)/\Delta_H$, there is a range of values of δ producing equilibria where the principal achieves lower equilibrium payoff in the reputational game where the candidate's first-period ability is revealed than in the reputational game where the candidate's first-period ability is not revealed.*

When the candidate's first-period ability will be revealed prior to the second period, the character of equilibrium play does not depend on the value of p_0 given that $p_0 > \bar{p}$. But as $p_0 \to 1$, the equilibrium in the game without revelation of first period ability falls into Region S in Figure 1, where both types simply adopt their static thresholds in each period.

Anticipating the revelation of first period ability, type H adjusts its threshold upwards but continues to choose a threshold that is too low (below r) from the perspective of the principal. This adjustment for type H improves the principal's first period payoff. In addition, since the first period recommendation can reveale information about the agent's type, the principal does at least as well in the second period by comparison to equilibrium in region S.

For values of p_0 sufficiently close to $\bar{p}$, equilibrium in the game without revelation of first-period ability falls into region R, where both types of agents adjust their thresholds upwards for reputational purposes in the first period.[6] By contrast, in the game where first period ability is revealed prior to the second period, then only type H will adjust first period threshold for reputational purposes. Further, when δ approaches $(c_H - c_L)/\Delta_H$, type H's threshold approaches type L's threshold and both types choose identical thresholds if $\delta \geq (c_H - c_L)/\Delta_H$. Paradoxically, the revelation of first period ability eliminates the transmission of information about the agent's type for δ sufficiently close to $(c_H - c_L)/\Delta_H$ and for δ greater than $(c_H - c_L)/\Delta_H$. The principal necessarily prefers the reputational equilibrium without revelation of first period ability whenever the principal's first period payoff from type L is less with threshold $r - c_L$ than with threshold $r - c_L + \delta\Delta_L$ – i.e. when $\delta < 2c_L/\Delta_L$. So long as $2c_L/\Delta_L > (c_H - c_L)/\Delta_H$ (which holds for c_L close to c_H), then there is a range of values of δ so that the revelation of first period ability reduces the principal's expected equilibrium payoff.

Intuitively, when $p_0 > \bar{p}$, the knowledge that the candidate's first period ability will be revealed before the second period reduces incentives for type L to change threshold for reputational purposes. Depending on the values of c_L and δ, this change in incentives fortype L can either increase or reduce the principal's expected first-period payoff. However, for δ sufficiently large, the change in incentives for type L is unambiguously bad for the principal in terms of second period payoff, as then the two types of agents choose similar or equal first-period thresholds, suppressing any revelation of information about the agent's type..

[6]The one caveat is that δ cannot be so large that $r - c_H + \delta\Delta_H \geq r - c_L + \delta\Delta_L$, which leads to a mixed strategy equilibrium in region M_{2N} as $p \to \bar{p}$. At $\delta = (c_H - c_L)/\Delta_H$, $r - c_H + \delta\Delta_H = r - c_L < r - c_L + \delta\Delta_L$, so this condition for δ holds so long as δ is sufficiently close to $(c_H - c_L)/\Delta_H$.

6 References

Avery, Christopher, Christine Jolls, Richard Posner and Alvin Roth, “The Market for Federal Judicial Law Clerks,” *University of Chicago Law Review*, (68), 2001,p. 793-902.

Benabou, Roland and Guy Laroque, “Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility,” *Quarterly Journal of Economics*, (107), 1992, p. 921-958.

Brandenburger, Adam and Ben Polak, “When Managers Cover Their Posteriors: Making the Decisions the Market Wants to See, “ *Rand Journal of Economics*, 27(3), 1996, p. 523-541.

Crawford, Vincent and Joel Sobel, ”Strategic Information Transmission,” *Econometrica*, (50), 1982, p. 1431-1451.

Ely, Jeffrey and Juuso Valimaki, “Bad Reputation,” *Quarterly Journal of Economics*, 118 (3), 2003, p. 785-814.

Fafchamps, Marcel and Alexander Moradi, “Referral and Job Performance: Evidence from the Ghana Colonial Army,” C.E.P.R. D.P. No. 7408, Aug. 2009.

Gentzkow, Matthew and Jesse Shapiro, “Media Bias and Reputation,” *Journal of Political Economy*, 2006, (114), p. 280-316.

Levy, Gilat, “Anti-Herding and Strategic Consultation,” *European Economic Review*, (48), 2004, p. 503-525.

Morris, Stephen, “Political Correctness,” *Journal of Political Economy*, (109), 2001, p. 231-265.

Prat, Andrea, ”The Wrong Kind of Transparency,” *American Economic Review,* (95), 2005, p. 862-877.

Prendergast, Canice and Lars Stole, “Impetuous Youngsters and Jaded Oldtimers,” *Journal of Political Economy*, (104), 1996, p. 1105-1134.

Scharfstein David and Jeremy Stein, “Herd Behavior and Investment,” *American Economic Review*, (80), 1990, p. 465-479.

Sobel, Joel, “A Theory of Credibility,” *Review of Economic Studies*, (52), 1985, p. 557-573.

7 Appendix

Proof of Proposition 1 Suppose that the set of possible messages for the agent is $(m_1, m_2, ..., m_L)$, and that the principal hires the candidate with probability $P(m_j)$ in response to recommendation m_j. Suppose in addition that we index these messages so that $P(m_1) \leq P(m_2) ... \leq P(m_L)$. Since an agent of type-i has strict preference for the candidate to be hired if $s > r - c_i$ and strict preference for the candidate not to be hired if $s < r - c_i$, it is optimal for the agent to choose message m_1 if $s < r - c_i$ and optimal for the agent to choose message m_L if $s \geq r - c_i$. Therefore, at most two distinct messages are chosen with positive probability in any static equilibrium. One of these messages is "Negative", corresponding to the range of abilities $[0, r - c_i)$ for an agent of type-i, while the other message is "Positive", corresponding to the range of abilities $(r - c_i, 1]$ for an agent of type-i.

After a negative recommendation from a type-i agent, the conditional ability of the candidate is uniform on the range $(0, r - c_i)$, meaning that the candidate is known to have ability below the principal's reservation value. Since this is true for each possible type of agent, the principal's uncertainty about the agent's type does not affect the principal's optimal decision – it is optimal not to hire the candidate.

After a positive recommendation from a type-i agent, the conditional ability of the candidate is uniform on the range $(r - c_i, 1]$, corresponding to conditional expected ability $E(s|Y, c_i) = (1 + r - c_i)/2$. Absent specific knowledge of an agent's type, the principal's conditional expected value given a positive recommendation is a weighted average for the two possible types of agent:

$$E(s|Y) = \frac{P(L)P(Y|L)E(s|Y, L) + P(H)P(Y|H)E(s|Y, H)}{P(L)P(Y|L) + P(H)P(Y|H)},$$

Defining $z_i = r - c_i$, this simplifies to

$$E(s|Y) = \frac{p(1 - z_L)(1 + z_L)/2 + (1 - p)(1 - z_H)(1 + z_H)/2}{p(1 - z_L) + (1 - p)(1 - z_H)}.$$

Differentiating with respect to p, we find that $E(s|Y)$ is increasing in p. Thus, there is a critical value, $\overline{p}$ for the prior probability such that the principal will hire a candidate given a positive evaluation if $p_0 > \overline{p}$, but will not hire the candidate even given a positive evaluation if $p_0 < \overline{p}$. (Note that if $p_0 < \overline{p}$, there can be no informative equilibrium.) This value $\overline{p}$ is defined implicitly by the equation $ES(Y) = r$. That is,

$$\frac{\overline{p}(1 - (r - c_L)^2)/2 + (1 - \overline{p})(1 - (r - c_H)^2)/2]}{\overline{p}(1 - r + c_L) + (1 - \overline{p})(1 - r + c_H)} = r,$$

which simplifies to

$$\overline{p} = \frac{c_H^2 - (1 - r)^2}{c_H^2 - c_L^2}.$$

Since we have assumed that $c_L < 1 - r < c_H$, this formula provides a value of $\overline{p}$ that is between 0 and 1.□

Proof of Proposition 2

In any subgame perfect equilibrium of the dynamic game, the principal and agent follow the static equilibrium in the second period of play, with the caveat that the prior probability for the agent's type in the second period, p_1, is updated based on first period play. Regardless of the value of p_1, the principal will not hire a candidate after a negative recommendation in the second period since $r > 1/2$.

Now suppose that the set of possible first-period messages for the agent is $(m_{1,0}, m_{1,1}, m_{1,2}, ..., m_{1,N})$, that each of these messages is selected with positive probability by at least one type of agent, and that the principal hires the first-period candidate with probability α_j in response to $m_{1,j}$. Once again, we index these messages so that $a_0 \leq \alpha_1 ... \leq \alpha_N$. The choice of first period message cannot affect the principal's action after a negative second period recommendation (the principal will not hire in this case), but may influence the principal's action after a positive second period recommendation. Thus, we define β_j as the probability that the principal hires the candidate in the second period after message m_j in the first period and a subsequent positive message in the second period. If $\alpha_j = \alpha_{j+1}$ for any j and $\beta_j < \beta_{j+1}$, then $m_{1j\ j}$ is dominated by message $m_{1j\ j+1}$ and would never be chosen by either agent.. Thus, we can strictly order the α's: $\alpha_0 < \alpha_1 < ... < \alpha_N$ (If any two messages have identical hiring probablities in both periods, we view them as the same message.). Further, if $\beta_{j+k} \geq \beta_j$ for any $k > 0$ then $m_{1j\ j}$ is dominated by $m_{1j\ j+K}$. Since we assume that all messages are selected with positive probability by at least one type of agent, and $\beta_0 > \beta_1 > ... > \beta_N$.

Step One: Threshold Strategies

Given a first-period candidate of ability s_1, the expected payoff for an agent of type i for message $m_{1,j}$ is $\pi_i\left(m_{1,j}, s_1\right) = \alpha_j(s_1 + c_i) + (1 - \alpha_j)r + \delta[r + \beta_j\Delta_i]$. This payoff is a linear function of s_1 with coefficient α_j, and the single-crossing property applies to the payoffs for any pair of messages. Since $\alpha_N = \max(\alpha_j)$, the slope of the payoff function in s_1 for message $m_{1,N}$ is greater than the slope of the payoff function in s_1 for any other message. Thus, if message N is chosen with positive probability by type i, it must be chosen for a range of highest abilities: $(s_{i,N}, 1]$. By inductive application of this logic, each type of agent must follow a threshold strategy for recommendations, with recommendation 0 corresponding to abilities $(0,\, s_{i,1})$, recommendation j corresponding to abilities $(s_{i,j}, s_{i,j+1})$, and recommendation N corresponding to abilities $(s_{i,N}, 1]$ for a type-i agent.

Comparing the payoffs for adjacent messages, an agent of type-i prefers $m_{1,j}$ to $m_{1,j-1}$ if $s_1 \geq r - c_i + \delta\Delta_i(\beta_{j-1} - \beta_j)/(\alpha_j - \alpha_{j-1})$. Defining $K_j = \delta(\beta_{j-1} - \beta_j)/(\alpha_j - \alpha_{j-1})$, the threshold values s_{ij} are given by $s_{i,j} = r - c_i + \Delta_i K_j$. For each recommendation to be chosen with positive probability by at least one type of agent, the K_j coefficients must also be strictly increasing in j.

We now demonstrate that the $0 < s_{i,0} < s_{i,N} < 1$ for each i, indicating that each type chooses each message with positive probability.[7] First consider the possibility that message

[7]We demonstrate above that K_j must be strictly increasing in j for each message to be chosen with positive probability by at least one type of agent. The conclusion that both types of agents choose each

0 is not chosen with positive probability by both types of agents. Since $r > c_i$, $s_{i,0} < 0$ is only possible if $K_1 < 0$, which implies in turn that $s_{H,0} = r - c_i + \Delta_H K_0 < r - c_i + \Delta_L K_0$ $= s_{L,0}$ since $\Delta_H > \Delta_L$ and $K_0 < 0$. That is, if $m_{1,0}$ is chosen with probability zero by type L it is also chosen with probability zero by type H. Since we assume that each message is chosen with positive probability by at least one type, it must be that message 0 is chosen with positive probablity by type L. If type H chooses message $m_{1,0}$ with probability 0, then observing message 0 would cause the principal to conclude that the agent is type L and to set $\beta_0 = 1$. But then it would be optimal for type H to choose message $m_{1,0}$ for $s_1 < r - c_H$ – a contradiction.

Next consider the possibility that $m_{1,N}$ is not chosen with positive probability by both types of agents. Define $\bar{K} = (c_H - c_L)/(\Delta_H - \Delta_L)$ so that $r - c_H + \bar{K}\Delta_H = r - c_L + \bar{K}\Delta_L = \bar{z}$. By definition, if $K_j < \bar{K} = \frac{2}{2(1-r)+c_H+c_L}$, then $\bar{z} > s_{L,j} > s_{H,j}$,while if $K_j > \bar{K}$, then $\bar{z} < s_{L,j} < s_{H,j}$. Substituting for Δ_H and $\bar{K}$ gives the formula.

$$\bar{z} = r + \frac{(1-r)^2 - c_L c_H}{2(1-r) + c_H + c_L}.$$

Thus, $\bar{z}$ is decreasing in both c_H and c_L. Further if $c_H = c_L = 0$, $\bar{z} = (1+r)/2 < 1$, so $\bar{z} < 1$. So $s_{i,j} > 1$ requires $s_{i,j} > 1$ and is only possible if $s_{H,j} > s_{L,j}$. That is, if message N is chosen with probability zero by type L, it is also chosen with probability zero by type H. Once again, since we assume that each message is chosen with positive probability by at least one type, it must be that message N is chosen with positive probablity by type L. If type H chooses message N with probability 0, then observing message N would cause the principal to conclude that the agent is type L and to set $\beta_N = 1$. But then it would be optimal for type H to choose message N for $s_1 > r - c_H$ – a contradiction.

Step Two: Two Messages Chosen with Positive Probability by Both Types

Suppose that there are three or more distinct messages ($N \geq 3$) Then, for $0 < j < N$, P(recommendation $j\,|$ type i) $= \Delta_i(K_{j+1} - K_j)$. This implies that the likelihood ratio $L(m_{1,j}) = P(m_{1,j}\,|$ type $L)$ / $P(m_{1,j}$—type $H) = \Delta_L/\Delta_H$. Further, $P(m_{1,N}\,|$type $i) =$ 1-$s_{i,N} = 1 - \bar{z} + (\bar{z} - s_{i,N})$. This implies that the likelihood ratio $L(m_{1,N}) = [(1-\bar{z}) + (\bar{z} - s_{L,N})]/[(1-\bar{z}) + (\bar{z} - s_{H,N})]$. The first terms are identical for numerator and denominator, while the second terms in numerator and denominator follow the ratio Δ_L/Δ_H. Thus, the likelihood ratio for $m_{1,N}$ falls between Δ_L/Δ_H and 1 and is greater than the likelihood ratio for recommendations 1, ... $N-1$.[8] As a result, the updated probability p_1 will be higher after $m_{1,N}$ than after $m_{1,N-1}$, implying $\beta_N \geq \beta_{N-1}$. Then $m_{1,N}$ yields a higher probability of hiring in period 1 and a higher probability of hiring after a positive recommendation in period 2 than does $m_{1,N-1}$. This implies that each type of agent will choose first-period message $m_{1,N}$ if $s_1 \geq r - c_i$. But then any other first period message would indicate that the first period candidate's ability is less than r, with the result that the principal would set $\alpha_j = 0$ for any $m_{1,j} \neq m_{1,N}$ so that there would be at most two distinct messages.

message with positive probability follows from the fact that K_j strictly increasing in combination with $s_{1,0} > 0$ and $s_{1,N} < 1$.

[8]An exception is if $s_{H,N} > s_{L,N} > \bar{z}$, which also implies $\beta_N > \beta_{N-1}$ and leads to a similar contradiction as in the text.

From the analysis above, the first-period cutoff between positive and negative recommendation for an agent of type-i is given by

$$z_i = r - c_i + \delta(\beta_0 - \beta_1)\Delta_i/(\alpha_1 - \alpha_0)$$

We know that $c_H > c_L$ and $\Delta_H > \Delta_L$, so if $\beta_0 \leq \beta_1$, then $s_{H,1} < s_{L,1}$. If $\beta_1 > \beta_0$, however, it must be that $P(\text{type } L \mid \text{recommendation } 1) > P(\text{type } L \mid \text{recommendation } 0)$, which also requires $z_H < z_L$.□

Proof of Proposition 3

Step One: Each recommendation becomes less informative about the type of agent as K increases.

Case 1: Positive first-period recommendation: The likelihood ratio for type-L to type-H given a positive recommendation is

$$L(Y_1) = P(Y_1|L)/P(Y_1|H) = \frac{1 - K\Delta_L - r + c_L}{1 - K\Delta_H - r + c_H}.$$

The sign of the derivative of the likelihood ratio $L(Y_1)$ with respect to K is the same as the sign of

$$-\Delta_L(1 - K\Delta_H - r + c_H) + \Delta_H(1 - K\Delta_L - r + c_L).$$

Simplifying this expression, $dL(Y_1)/dK$ takes the same sign as $\Delta_H(1 - r + c_L) - \Delta_L(1 - r + c_H)] = (1 - r + c_L)(1 - r + c_H)(c_H - c_L)/2 > 0$. Then, since $L(Y_1)$ is increasing in K, the updated value $p_K^*(p_0, Y_1)$ is increasing in K as well. The informativeness of a positive recommendation can be represented as $p_0 - p_K^*(p_0, Y_1)$, the magnitude with which the principal updates prior probability p_0 after a first period recommendation of Y. Since $p_K^*(p_0, Y_1)$ is increasing in K, $p_0 - p_K^*(p_0, Y_1)$ is decreasing in K indicating that a recommendation of Y_1 becomes less informative about the type of agent as K increases.

Case 2: Negative first-period recommendation: The likelihood ratio for type-L to type-H given a negative recommendation is

$$P_1(N|L)/P_1(N|H) = \frac{K\Delta_L - r + c_L}{K\Delta_H - r + c_H}.$$

Differentiating with respect to K, $dL(N_1)/dK$ takes the same sign as

$$\Delta_L(K\Delta_H + r - c_H) - \Delta_H(K\Delta_L + r - c_L) = \Delta_L(r - c_H) - \Delta_H(r - c_L) < 0.$$

This means that the updated value for p after a negative recommendation is falling in K, i.e. $p_K^*(p_0, N_1) - p_0$ is decreasing in K, so that a negative recommendation is less informative as δ increases.

Step Two: The Conditional Expected Value for First Period Ability given Positive recommendation is Increasing in K

The conditional expected value given a positive recommendation is
$E(s_1|Y_1) = p^*_K(p_0, Y_1)\ E(s_1|\text{ type L}, Y, K) + (1 - p^*_K(p_0, Y_1))\ E(s_1|\text{ type H}, Y_1, K)$.

As K increases, there are two separate effects: (1) $E(s_1|$ type i, $Y_1, K)$ increases; (2) $p^*_K(p_0, Y_1)$ increases (since as shown in Step 1, the informativeness of the recommendation declines as K increases). Each of these has unmbiguous positive effect on $E(s_1|Y_1)$. (To sign the second effect, note that Proposition 2 demonstrates that $z_L > z_H$, implying $E(s_1|$ type L, $Y_1) > E(s_1|$ type H, $Y_1)$).

Step Three: If $p_0 \geq \bar{p}$, then in any informative dynamic equilibrium, the principal will hire after a positive recommendation, hire with probability 0 after a negative recommendation, and will give discretion to the agent after a negative recommendation ($\alpha_0 = 0, \alpha_1 = 1, \beta_0 = 1$) in the first period.

A negative recommendation N_1 corresponds to cutoffs (0, z_L) and (0, z_H), with conditional expected value for each agent $s_{1i}/2 < 1/2 < r$. Thus, the principal will not hire in the first period after a negative recommendation: $\alpha_0 = 0$. Further, since $z_L > z_H$, P(N_1|type L, K) > P(N_1 |type H, K) and therefore $p^*_K(p_0, N_1) > p_0 > \bar{p}$ (assuming $K > 0$), indicating that the principal will give discretion to the agent in the second period: $\beta_0 = 1$.

Since $\beta_0 = 1$, agent i will use cutoff $s_{i,1} = r - c_i + [\delta(1 - \beta_1)\Delta_i]/\alpha_1 \geq r - c_i$ for a positive recommendation in the first period, corresponding to multiplier $K_1 > 0$. By the result of Step 2, since $K_1 > 0$ the principal will be even more inclined to hire the candidate after a positive recommendation based on reputational multiplier K_1 than after a positive recommendation based on static cutoffs. For $p_0 \geq \bar{p}$**,** the principal would be willing to hire after a positive recommendation based on the static cutoffs, and so would also be willing to hire based on a positive recommendation and tougher first-period cutoffs: $\alpha_1 = 1$.

Step Four: There exists a value p_u such that the principal is indifferent between giving the agent discretion or not in the second period after a positive recommendation based on the static cutoffs in the first period: $p^*_{K=0}(p_u, Y_1) = \bar{p}$. The value p_u satisfies $\bar{p} \leq p_u \leq 1$.

There exists a value p_{ll} such that the principal is indifferent between giving the agent discretion or not in the second period after a negative recommendation based on the static cutoffs in the first period: $p^*_{K=0}(p_{ll}, N_1) = \bar{p}$. The value p_u satisfies $0 \leq p_u \leq \bar{p}$

If the agents use their static cutoffs ($K = 0$), $P(Y_1|$ type $H) = 1 - r - c_H$, whereas $P(Y_1|$ type $L) = 1 - r - c_L$. By Bayes' Rule,

$$p^*_{k=0}(p_0, Y_1) = \frac{p_0(1 - r + c_L)}{1 - r + p_0 c_L + (1 - p_0)c_H}$$

Define p_u implicitly to be the probability such that $p^*_{K=0}(p_0, Y_1) = \bar{p}$. Since p^* is continuous and increasing in p_0, $p^* = 0$ at $p_0 = 0$, and $p^* = 1$ at $p_0 = 1$, there exists a unique value p_u satisfying this equation. We can also solve directly for p_u.

$$p_u = \frac{\bar{p}(1 - r + c_H)}{1 - r + c_L + \bar{p}(c_H - c_L)}.$$

A similar argument applies to demonstrate the existence of a unique p_{ll} such that $p^*_{K=0}(p_0, N_1) =$

$\bar{p}$.

Step Five: There is a unique dynamic equilibrium for $\mathbf{p}_0 > \bar{p}$.

We know from step three that if $p_0 \geq \bar{p}$, then $\alpha_0 = 0, \alpha_1 = 1, \beta_0 = 1$, implying that $K = \delta(1-\beta_1)$ i.e. agent i uses cutoff $s_{i,1} = r - c_i + [\delta(1-\beta_1)\Delta_i]$. From step 1, $p^*_K(p_0, Y_1)$ is increasing in K. There are three possible forms of dynamic equilibrium:

1.$\beta_1 = 1$, corresponding to $p^*_{K=0}(p_0, Y_1) \geq \bar{p}$.

2.$\beta_1 = 0$, corresponding to $p^*_{K=\delta}(p_0,Y_1) < \bar{p}$.

3.$0 < \beta_1 < 1$, corresponding to $p^*_{K=\delta\beta_1}(p_0,Y_1) = \bar{p}$.

To see that these three possibilities are mutually exclusive and collectively exhaustive, first observe that condition 1, $p^*_{K=0}(p_0, Y_1) \geq \bar{p}$, implies $p^*_{K>0}(p_0, Y_1) > \bar{p}$ since $p^*_K(p_0, Y_1)$ is increasing in K. That is, condition 1 is incompatible with either of conditions 2 or 3. Next observe that condition 2, $p^*_{K=\delta}(p_0,Y_1) < \bar{p}$ implies $p^*_K(p_0,Y_1) < \bar{p}$ for all $K < \delta$, which means that condition 2 is incompatible with condition 3. Finally, note that either condition 1 holds or condition 2 holds unless $p^*_{K-0}(p_0, Y_1) < \bar{p}$ and $p^*_{K=\delta}(p_0, Y_1) > \bar{p}$ In this case, since $p^*_K(p_0, Y_1)$ is continuous and strictly increasing in K, there is a unique value $K^* \in (0, \delta)$ such that $p^*_{K^*}(p_0, Y_1) = \bar{p}$, and thus a unique dynamic equilibrium with $\beta_1 = (\delta - K^*)/\delta$.

Step Six: If $p_0 < \bar{p}$, then in any informative dynamic equilibrium, the principal will hire with probability 0 after a negative recommendation, and will not give discretion to the agent after a positive recommendation ($\alpha_0 = 0, \beta_1 = 0$).

As in Step Three, a negative recommendation corresponds to cutoffs (0, z_{1L}) and (0, z_{1H}), with conditional expected value for each agent $z_{1i}/2 < 1/2 < r$. Thus, the principal will not hire in the first period after a negative recommendation: $\alpha_0 = 0$. In addition, $p^*_K(p_0, Y_1) \leq p_0 < \bar{p}$, so the principal will not give discretion to the agent if $p_0 < \bar{p}$.

Step Seven: There is a unique dynamic equilibrium for $p_0 < \bar{p}$..

The equilibrium analysis for $p_0 < \bar{p}$ follows the same form as for $p_0 \geq \bar{p}$, with the additional complication that we know the values for only two parameters ($\alpha_0 = 0, \beta_1 = 0$) instead of three parameters when $p_0 > \bar{p}$. If there is an informative equilibrium, the threshold between positive and negative recommendation is given by $r - c_i + \delta\Delta_i(\beta_0 - \beta_1)/(\alpha_1 - \alpha_0)$. Further, since $\alpha_0 = 0, \beta_1 = 0$, the threshold must be greater than $r - c_i$ for type i (i.e. multiplier $K \geq 0$). From step 1, $p^*_K(p_0, N_1) > p$ is decreasing in K since the signals become less informative about the agent's type as K increases. There are five possible forms of dynamic equilibrium.

1. $\beta_0 = 0$, corresponding to $p^*_{K-0}(p_0, N_1) \leq \bar{p}$

2.$\beta_0 = 1$, $\alpha_1 = 1$, corresponding to $p^*_{K=\delta}(p_0, N_1) \geq \bar{p}$ and $E(s_1|p_0, Y_1, K) \geq r$.

3. $\beta_0 = 1$, $0 < \alpha_1 < 1$, corresponding to $p^*_{K=\delta/\alpha_1}(p_0, N_1) \geq \bar{p}$ and $E(s_1|p_0, Y_1, K) = r$.

4.$0 < \beta_0 < 1$, $\alpha_1 = 1$, corresponding to $p^*_{K=\delta\beta_0}(p_0, N_1) = \bar{p}$ and $E(s_1|p_0, Y_1, K) \geq r$.

5.$0 < \beta_0 < 1$, $0 < \alpha_1 < 1$, corresponding to $p^*_{K=\delta\beta_0/\alpha_1}(p_0, N_1) = \bar{p}$ and $E(s_1|p_0, Y_1, K) = r$.

When condition 1 holds, even the static equilibrium strategies are not sufficiently informative for the principal to give discretion after a negative recommendation. Since $p^*_K(p_0, N_1)$ is decreasing in K, $p^*_K(p_0, N_1) < \bar{p}$ for all possible combinations of α_1 and β_0, implhying $\beta_0 = 0$. That is, the principal does not give discretion to the agent in the second period in any case, indicating that no equilibrium with $\beta_0 > 0$ is possible.

Condition 5 only holds in the non-generic case where there exists a reputational multiplier K_5 such that $p^*_{K=K_5}(p_0, N_1) = \bar{p}$ and $E(s_1|p_{0,}Y_1, K_5) = r$. This possibility occurs at just a single point in Figure 1 at (p_l, δ_5). In this case, all combinations of (α_1, β_0) such that $\alpha_1/\beta_0 = K_5$ will produce a dynamic equilibrium. (If in fact, the parameters are such that $K_5 = \delta$, it is even possible to produce a reputational equilibrum satisifying any of conditions 2, 3, 4 at $K_5 = \delta$) Having noted this non-generic possibility, we assume that it does not hold in the remaining analysis below.

When condition 2 holds, $p^*_{K=\delta}(p_0, N_1) \geq \bar{p}$ and $E(s_1|p_0, Y_1, K = \delta) \geq r$ (since condition 2 corresponds to $\beta_0 = \alpha_1 = 1$). Then since $E(s_1|p_0, Y_1, K)$ is increasing in K while $p^*_K(p_0, N_1)$ is decreasing in K, $E(s_1|p_{0,}Y_1, K) = r$ requires $K \leq K_2$, while $p^*_K(p_0, N_1) = \bar{p}$ requires $K \geq K_2$. That is, condition 2 is incompatible with each of conditions 3, 4, and 5. By similar reasoning, conditions 3 and 4 are incompatible.

Proof of Proposition 4

Proposition 4 is a special case of Proposition 7, part 1 (with $r_1 < \hat{r}_1$) as proven below.

Proof of Proposition 5

The text explains why discretion increases as a result of each change of regions associated with an increase in p_0. Discretion is constant within the pure strategy equilibrium regions B,R , and S. It remains only to show that discretion is weakly increasing in p_0 within the mixed strategy regions M_1, M_{2Y}, M_{2N}.

On the interior of region M_1, the reputational multplier K is implicitly defined by $E(s_1|Y_1, p_0, K) = r$. An increase in p_0 requires a corresponding reduction in K (reducing the informational content of a positive recommendation) to maintain $E(s_1|Y_1, p_0, K) = r$. In region M_1,$\beta = 1$, so a reduction in $K = \beta\delta/\alpha$ requires an increase in α. That is, on the interior of M_1, an increase in p_0 requires a corresponding increase in α, indicating an increase in agent discretion.

On the interior of region M_{2N}, the reputational multplier K is implicitly defined by $p^*_K(p_0, N_1) = \bar{p}$. In this region, $p_0 < \bar{p}$, so an increase in p_0 means that the first period recommendation Y_1should be less informative about agent type in order to maintain the condition $p^*_K(p_0, N_1) = \bar{p}$ as p_0 increases. The first-period recommendation becomes less informative about agent type as K increases. In region M_{2N}, $\alpha = 1$, so an increase in K requires an increase in β which in turn indicates an increase in agent discretion, since discretion is defined to be equal to β in region M_{2N}.

On the interior of region M_{2Y}, the reputational multplier K is implicitly defined by $p^*_K(p_0, Y_1) = \bar{p}$. In this region, $p_0 > \bar{p}$, so an increase in p_0 means that the first period recommendation N_1should be more informative about agent type in order to maintain the condition $p^*_K(p_0, Y_1) = \bar{p}$ as p_0 increases. The first-period recommendation becomes more informative about agent type as K decreases. In region M_{2Y}, $\alpha = 1$, so an reduction in K requires an reduction in β which in turn indicates an increase in agent discretion, since discretion is defined to be equal to 2-β in region M_{2N}.

Proof of Proposition 7

In Step 1 of the proof of Proposition 2, we found that the first-period thresholds z_L and z_H become equal at the point $\overline{z} = r + \frac{(1-r)^2 - c_L c_H}{2(1-r)+c_H+c_L}$. When $r_1 \neq r_2$, this value $\bar{z}$ depends on both of those values, $\bar{z}(r_1, r_2) = r_1 + \frac{(1-r_2)^2 - c_L c_H}{2(1-r_2)+c_H+c_L}$, Define $\hat{r}_1 = 1 - \frac{(1-r_2)^2 - c_L c_H}{2(1-r_2)+c_H+c_L}$.

Then, for $r_1 < \hat{r}_1$, $\bar{z}(r_1, r_2) < 1$, and the logic of the proof of Proposition 3 continues to hold. In particular, when $r_1 < \hat{r}_1$, $dL(Y_1)/dK$ and $dL(N_1)/dK$ take the same signs as when $r_1 = r_2$, implying that the first period recommendation becomes less informative about the agent's type as the reputational multiplier $K = \beta\delta/\alpha$ increases. Thus, dynamic equilibria continue to take the same qualitative form as in Figure 1 when $r_1 < \hat{r}_1$, and we use the same terminology for the regions of play as in Figure 1.

Consider an equilibrium In region M_1 with discount factor δ_0. Then the first-period hiring probability $\alpha_1(\delta_0)$and the associated reputational multiplier are determined by the condition $E(s_1|p_0, Y_1, K = \delta/\alpha_1(\delta_0)) = r_1$. The resulting thresholds $z_i = r - c_i + K\Delta_i$ are identical to the pure strategy thresholds for an equilibrium on the boundary of region R when the discount factor is given instead by $\delta = \delta_0/\alpha_1(\delta_0)$. Similarly, the mixed strategies thresholds in regions M_{2Y} and M_{2N} are identical to the thresholds for pure strategy equilibria on the boundary of region R. Thus, for purposes of comparison of the principal's first period expected payoff with $r_1 \neq r_2$ to the principal's expected payoff in the static equilibrium, it suffices to consider pure strategy dynamic equilibria with $r_1 \neq r_2$and $r_1 < \hat{r}_1$.

Define $\pi_{1D}(p_0, K)$ to be the principal's first-period payoff in dynamic equilibrium based on thresholds $z_i = r_1 - c_i + K\Delta_i$ and define $\pi_S(p_0)$ to be the principal's expected payoff in the equilibrum of the one period static game. The principal's payoff for not hiring a marginal worker z_i is $r_1 - z_i$, with $d(z_i - r_1)/dK < 0$. Since the distribution of worker abilities is uniform, $d(z_i - r_1)/dK < 0$ implies $d^2\pi_{1D}/dK < 0$, so for any p_0, $\pi_{1D}(p_0, K)$ takes minimum value on the range $K \in [0, \bar{\delta}]$ at either $K = 0$ or $K = \bar{\delta}$. Further, since $\pi_{1D}(p_0, K = 0) = \pi_S(p_0)$,we know that $\pi_{1D}(p_0, K) < \pi_S(p_0)$ is only possible if $\pi_{1D}(p_0, K = \bar{\delta})$ $< \pi_S(p_0)$. Both types of agents use the same thresholds for a positive recommendation at $\delta = \bar{\delta}$, so $\pi_{1D}(\bar{\delta}, p)$ is constant in p. in contrast, $\pi_{1S}(p)$ is increasing in p. Since there are no reputational equilibria for $p_0 > p_u$, $\pi_{1S}(p)$ is bounded above by $\pi_{1S}(p_u)$ It suffices to show that $\pi_{1D}(p_u, K = \bar{\delta}) \geq \pi_{1S}(p_u)$ to show that dynamic first-period profits are greater than static first-period profits for the principal for any (δ, p_0).

With $r_2 \neq r_{1,}$ p_u is defined implicitly by the equation $(1 - r_1 + c_L)p_u/[(1 - r_1 + c_L)p_u + (1 - r_1 + c_H)(1 - p_1) = \bar{p}_2$ (where $\bar{p}_2 = (c_H^2 - (1 - r_2)^2)/(c_H^2 - c_L^2))$ so that the principal will update beliefs from p_u to $\bar{p}_2$ following a positive recommendation based on static thresholds in the first period. It is algebraically easiest to compare the principal's expected loss due to biased recommending in static equilibrium to the principal's expected loss due to use of reputational cutoffs z_H and z_L - in each case using the principal's ideal cutoff r for a positive recommendation as a benchmark for comparison. The static equilibrium loss at $p_0 = p_u$ is given by $\theta_S(p_u) = [p_u c_L^2 + (1 - p_u)c_L^2)\ ]/2$. Substituting for p_u and then differentiating with respect to r_1 verifies $d\theta_S(p_u)/dr_1 \leq 0$. Further, since $\bar{z} = *$, the expected dynamic loss of $(* - r_1)^2/2$ is constant in r_1. Thus, $\theta_S(p_u) - \theta_D(p, \delta)$ takes maximum value at the maximium possible value $r_1 = \hat{r}_1$.

Proof of Proposition 8: There cannot be an equilibrium with $z_{1H} > z_{1L}$. If so, then type H makes a negative recommendation and type L makes a positive recommendation for a candidate of ability $z \in (z_{1L}, z_{1H})$, and therefore, the principal will follow advice in the second period after a positive recommendation. Thus a type H would gain the reputational value $\delta\Delta_H$ by deviating from the proposed equilibrium strategy to make a positive recommendation, indicating that $z_{1H} \leq r - c_H - \delta\Delta_H$. Since $\delta\Delta_H > \delta\Delta_L$ and $c_H > c_L$, type L would lose more than the reputational value $\delta\Delta_L$ in the first period by making a positive recommendation for a candidate with ability z_{1H} (and causing that candidate to be hired), so type L would gain by a deviation to $z_{1L} \geq z_{1H}$.

Next look for equilibria with $z_{1H} < z_{1L}$. In such an equilibrium, the principal hires the first period candidate after a positive report. Further, if $z \in (z_{1H}, z_{1L})$, $p_1(Y, z) = 0$; $p_1(N, z) = 1$, since the principal can identify the type of agent from the combination of recommendation and candidate's ability. Then for $z \in (z_{1H}, z_{1L})$, a type H agent gains $z - r + c_H$ in payoff by making a positive recommendation (so that the candidate is hired) and $\delta\Delta_H$ in (expected) payoff by making a negative recommendation. Thus the threshold must satisfy $z_{1H} \geq r - c_H + \delta\Delta_H$ for the agent to make positive recommendations for $z \in (z_{1H}, z_{1L})$. Similarly, for a type H agent to make negative recommendations for $z < z_{1H}$, it must be that $\delta\Delta_H > z + c_H - r$ (assuming out of equilbrium beliefs that a positive report for $z < z_{1H}$ indicates a type H agent) so that $z_{1H} \leq r - c_H + \delta\Delta_H$. Combining these observations, $z_{1H} = r - c_H + \delta\Delta_H$ in any threshold equiliibrium with $z_{1H} < z_{1L}$. Type L will adjust the threshold for a positive first period recommendation up from the static equilibrium threshold of $r - c_L$ by up to $\delta\Delta_L$ in order to induce the principal to follow advice in the second period. Therefore, each combination of thresholds $(z_{1H} = r - c_H + \delta\Delta_H, z_{1L})$ produces an equilibrium so long as $z_{1L} > z_{1H}$ and $z_{1L} \leq r - c_L + \delta\Delta_L$.

Finally, look for equilibria with $z_{1H} = z_{1L}$. Since type H will adjust the threshold up from $r - c_H$ by as much as $\delta\Delta_H$ and type L will adjust the threshold up from $r - c_L$ by as much as $\delta\Delta_L$, there can be an equilibrium with $z_{1H} = z_{1L}$ so long as $r - c_H \leq z_{1H} \leq r - c_H + \delta\Delta_H$ and $r - c_L \leq z_{1L} \leq r - c_L + \delta\Delta_L$. So if $r - c_H + \delta\Delta_H < r - c_L$, there are no equilibria with $z_{1L} = z_{1H}$, but if $r - c_H + \delta\Delta_H \geq r - c_L$, then there are a range of equilibria with $z_{1L} = z_{1H} \in [r - c_L, r - c_H + \delta\Delta_H]$.