Contracts and Conflict in Organizations

Canice Prendergast*

March 9, 2010

Abstract

In many organizations, the way that incentive problems are alleviated is not via contracts, but rather who is hired. This paper offers a theory of targeted hiring, and how its role changes as contracting becomes poorer.

*University of Chicago, GSB. Email: canice.prendergast@gsb.uchicago.edu. Thanks to Wouter Dessein, Bob Gibbons, and seminar participants at the University of Chicago, Columbia University, Northwestern University, Harvard/MIT, UCLA, UNC, and the NBER for helpful suggestions.
1 Introduction

Agency theory has largely been about using compensation to align interests. Yet pay is often a very poor way of providing incentives, and a more relevant tool in many settings is instead who to hire. This paper shows how hiring can reduce agency issues, and how its role changes as it becomes more difficult to contract on performance. Specifically, the focus is on the endogenous creation of conflict as a response to poor contracting, where this conflict takes the form of an organization hiring people whose intrinsic objectives differ from its own. The paper’s primary conclusion is that whenever jobs are specialized, workers should generally not share the objectives of the organization. Furthermore, this divergence depends on the ability to contract on performance.

A motivating example might be useful - a university President hiring a new Dean. I have served on a number of search committees for such Deans. These committees spent (literally) no time on how the Dean should be paid given the obvious problems in finding and aggregating an appropriate set of measures. Instead, much of its concern was on scrutinizing the background and previous activities of the candidates, as these were felt to be indicative of issues that they would emphasize on the job. Some might be better at fostering research, while others seemed more interested in fundraising or keeping students and alumni happy. Finding the “right” person on this spectrum was how these committees alleviated some agency concerns.1

The central building block that drives the results of the paper is very simple. Imagine that a worker in a firm carries out two tasks, A and B, but A is her primary responsibility (tasks are specialized). Her employer values both outputs equally. Normally agency problems are resolved by paying the worker based on how much she produces. The innovation here is that in addition to this, potential hires vary in how much they intrinsically care about A and B - to use the example above, some Dean candidates care more about faculty research and others care about student well-being.

A natural starting point would be to imagine that since the firm cares equally about outputs A and B, it would hire a worker who shares those preferences. However, this turns out not to be true - only in the limiting case where the firm can contract perfectly on output will this be the optimal outcome. Instead, whenever contracting is imperfect, the firm will choose an agent who cares more for A than B - i.e, an agent who is biased. Furthermore, this bias increases as the ability to contract on output gets worse, or as tasks become more specialized.2

---

1 Similar tradeoffs arise when choosing someone to run a government department, charitable institution, police department, company board, or - indeed - country. As a good example of these issues, see Golden, 2000, for a discussion of how changing senior government officials during the Reagan administration changed policy.

2 So for example, the difficulty in measuring the output of a social services provider results in agencies hiring social workers whose objective is excessively to serve their clients even though they are insufficiently motivated to control the costs of doing so. In a survey on the preferences of social workers, Robert Peabody, 1964, notes that “by far the most dominant organizational goal perceived
The logic for this tradeoff, outlined in Section 2, is straightforward. Hiring an agent whose preferences are biased in favor of $A$ over $B$ results in more effort devoted to $A$ at the expense of $B$. When would a firm favor such a tradeoff? All else equal, when the agent’s task doesn’t involve much $B$ but instead is largely devoted to producing $A$. In the model below, total output produced increases in the agent’s bias towards $A$ only because of such task specialization (if tasks were not specialized, the decline in $B$ would exactly offset the increase in $A$). Hence the demand for bias, and how it depends on the degree of specialization. So what role does contracting play? Hiring biased agents also has a downside - biased agents devote too much effort to one activity over the other. Hence there is an imbalance in efforts. If contracting is poor, the only way to increase output is to put up with this imbalance, and so hiring very biased agents is efficient. On the other hand, if contracting is good, the firm can use pay for performance to induce high levels of effort without significant imbalance. How do they avoid imbalance? By hiring less biased agents - in words, those who more closely share the objectives of the firm. Hence, specialization leads to a demand for more biased agents, while the efficiency of contracting pushes in the opposite direction. Simple though these observations are, they illustrate a cost to using hiring as a tool to alleviate agency concerns - namely, the endogenous creation of conflicts in objectives, where equilibrium conflict depends on the quality of performance measures.

So far, I have described one agent in this firm, the agent who is specialized towards carrying out activity $A$. Yet another agent may be specialized in $B$. So, for example, in a university setting, faculty primarily carry out research and teaching, and the Dean primarily deals with alumni and fund raising. As enunciated so far, the prediction of the model would be that as the ability to contract on the outputs of a university disimproves, the inherent objectives of the faculty and the dean would diverge more and more. In this sense, poor contracting leads to worse conflicts of interest.

This insight about conflicts of objectives derives from a very simple interaction of two important features of organizations - an inability to contract on performance, and the specialization of tasks. However, this posited relationship between worse contracting and more divergence in intrinsic preferences may not be true when other issues are considered, and much of the remainder of the paper is concerned with understanding these other effects. In Section 3, I extend the basic model to consider a case where one party has a discrete idea that could benefit the “other” activity, but at some cost to her primary activity.² Call this activity “cooperation”. In this case, I show that firms with poor contracting opportunities face an additional tradeoff - between capture and fiefdoms.

²as important “is service to clientele” (p.66), where 83 percent of survey respondents view such service as important, compared to only 9 percent who see “obligation to taxpayers” or “assistance to the public in general” as important concerns affecting their decisions. Derthick, 1979, also provides some evidence on such conflicts for social workers when they were asked by the SSA to be instrumental in denying coverage to applicants.

³So, for example, a lawyer at the FTC may have information useful to the economists about bringing a case against a firm, or a social worker has information on clients making false claims.
To understand how this extension affects outcomes, begin by considering what would happen if the firm continued to hire in the manner alluded to above. Then when it is easy to contract on performance, monetary incentives are strong and the agent has intrinsic incentives close to those of the principal. In that case, it is clearly straightforward to induce cooperation. However, the impact of worse contracting from above is to both reduce monetary incentives and to hire agents biased towards increasing the output of their primary activity. Both of these make cooperation less likely - formally, there is a point at which cooperation does not occur without changing either compensation or who is hired. In this case, the optimal response is to deviate from the original behavior by hiring an agent with more desire to increase their non-primary output. I call this phenomenon capture, because the firm hires agents who on average are biased towards one of the two activities, even though the firm values both of their outputs equally. This tendency towards capture becomes stronger as contracting opportunities initially gets worse.

However, the optimal response to the possibility of interaction need not be capture. As contracting on performance continues to get worse, the cost of distorting hiring to induce cooperation can become too great, and the firm discretely reverts to the outcome of the original model, where the agents’ preferences are very divergent. This has the advantage that the agent works hard on her primary task, but at the cost of giving up on cooperation. This outcome I term fiefdoms, as it results in highly motivated agents in each position, yet where their motivations are so divergent from each other than they will not carry out activities that increase the common good.

These outcomes arise in the context of a discrete cooperative activity. Section 3.2 addresses the continuous case. In this case, the outcomes change more continuously than above, where monetary and intrinsic incentives are either substitutes over the whole relevant parameter space, or are complements over the whole parameter space. In the case where the marginal benefit of the cooperative activity is high, I show that when contracting is poor, both agents hired have relatively similar objectives, even though their jobs are very different (one is specialized in A and the other in B, which previously led to very divergent agents hired). In this sense, the paper offers a theory of indifferent agents - where they care almost as much about the other’s task as their

---

4So for instance, if interactions between a faculty and a dean are sufficiently important, it may necessary to hire a faculty-friendly dean, even if it involves the dean ignoring important aspects of the job. Or consider staffing the National Highway Traffic Safety Administration, whose charge is to reduce accidents. For a long period of time, the predominant practice of the NHTSA was to hire engineers, using their professional interest in finding scientific solutions to reduce accidents. This gave rise to significant criticism of the agency where it always sought engineering solutions to safety problems (such as better seat belts, air bags, etc.) to the detriment of changing (for example) attitudes towards dangerous driving. See Pruitt, 1979, for details.

5As an example of this, note (i) the discussion of the Federal Trade Commission in Wilson (p.61), where the preferences of the economists hired were often at variance of those of the lawyers, with resulting tension, or (ii) Goldner’s, 2000, description of the standoff between Reagan political appointees and the staff of the Equal Employment and Opportunities Commission.
own - but one where the form of indifference is endogenously chosen by the principal.

All the observations above simply identify the kinds of workers that institutions would like to hire. Yet another intuition in this setting might be that when contracting is good, who cares who is hired? To address this, in Section 4 I consider a case where firms must incur a cost to identify intrinsic objectives. Not surprisingly, those firms that cannot contract well on output have the greatest reason to incur these costs, as they rely a great deal on intrinsic incentives. Those firms that can contract well on output have less to gain from finding the right employee, and do no incur these costs. Hence, with such costly state verification, those who contract poorly ultimately match best to their needs, whereas those who contract well randomly hire from the population.

The ideas proposed here are based on two premises. First, workers exert effort for reasons beyond contracted payments. Second, who firms hire affects what they do. Such variation in preferences can occur for many reasons. In Section 6 I endogenize these preferences in one plausible way - the avenue proposed here for why workers vary in their willingness to exert effort is “professionalism” - a commonly stated reason in the literature on motivation. Professionalism is modeled here as a career concerns issue, where future wages depend on current performance measures. The innovation here is that the external constituency may not share the objectives of the principal. So, for example, a lawyer in the federal government may exert effort based on the prospect of getting a job in the private legal sector. Using this lens, it is shown that who is hired affects what they do, with the implications outlined above.

Section 2 begins by building the benchmark model that shows the tradeoffs between hiring and monetary incentives. Following this, Section 3 illustrates how interaction across activities leads to the notions of capture, fiefdom, and indifference that are the choices facing firms that cannot contract well on output. Section 4 highlights problems that arise when identification of talents is costly. In each of these sections, various simplifying assumptions are made regarding contracts and technology. These are relaxed in Section 5, which show that the insights are generally robust to other assumptions. Section 6 endogenizes the preferences of the agent through the lens of professionalism.

---

6There is a large empirical literature on why workers often do far more than would be predicted by the standard economic model of agency. Much of this literature is in public administration or political science (such as Goodsell, 1998, and Brehm and Gates, 1997). As an extreme case, note that the US Post Office has ontime delivery rates of mail in the region of 98%, and in less than 3% of cases do government officials fail to give enough benefits to welfare recipients (Goodsell, 1998). This is surely not because these organizations tie pay to the performance of their employees - in many of these settings, pay is pretty much independent of performance.
2 The Model With Exogenous Preferences

An institution produces two outputs, A and B. For concreteness, let A be the provision of service to clients, and B be cost control. As an input to these objectives, it hires two agents who each provide efforts on two inputs (tasks), 1 and 2. There is an asymmetry between the two tasks, where each of the two workers is primarily charged with one of the two outputs. I focus initially on the agent who is specialized primarily in output A - agent a - considering another agent largely specialized in B - agent b - below. (To avoid complication, I do not make notation for efforts and outcomes agent specific: everything that follows unless otherwise stated refers to agent a.)

Let effort on task i be given by $e_i$, $i = 1, 2$, where to keep matters simple, the costs of effort on task i is $e_i^2$. All effort by agent a on task 1 increases the returns solely of output A. By contrast, effort on task 2 has a shared benefit. A fraction $x$ of effort on task 2 benefits output A, while the remaining $(1 - x)$ benefits output B.\(^7\)

Hence, tasks are partially specialized, as reflected by the parameter $x$. Output i, $y_i$, is therefore given by $y_A = e_1 + xe_2$, and $y_B = (1 - x)e_2$. These uncorrupted measures of performance cannot be contracted on. To avoid the possibility of corner solutions, I assume $x < \frac{1}{2}$. Assume for the moment that the principal maximizes the surplus created by both agents,\(^8\) so for agent a the firm maximizes $E[y_A + y_B - e_1^2 - \frac{e_2^2}{2}] = E[e_1 + e_2 - \frac{e_1^2}{2} - \frac{e_2^2}{2}]$.

There are two reasons to exert effort in the model: (i) pay for performance contracts, and (ii) intrinsic incentives. I describe each in turn.

Contracts Available performance measures are imperfect. Following Baker (1992), I assume that the principal can contract upon an unbiased but imperfect signal of total effort (and output):

$$\tilde{y} = (1 + D)e_1 + (1 - D)e_2,$$

where $D$ takes on value $\delta$ and $-\delta$ with equal probability. (I ignore agent subscripts for simplicity.)

The parameter $\delta$ thus measures the extent to which effort can be effectively contracted upon, and its realization is privately observed by (only) the agent after contracts are signed. At one extreme, $\delta = 0$ and performance measures are perfect while at the other extreme $\delta = \infty$, they are useless. This abstract contracting technology is used simply to illustrate distortions while retaining the inherent symmetry of the problem. The firm can condition the agent’s pay on observed output $\tilde{y}$. For ease of exposition, I consider linear contracts where the agent is given a fraction of output, $\beta \tilde{y}$ and a fixed payment.\(^9\)

\(^7\)The most natural interpretation of this is that output 2 involves cost containment which has benefits across the entire organization as cost savings are shared.

\(^8\)This assumption is analyzed in Section 4.

\(^9\)Optimal contracts are considered in Section 5.
**Intrinsic Preferences**  For the moment, assume that agents have an intrinsic preference over the two outputs. To model this, assume that agents have an observed “type” \((\mu_A,\mu_B)\) such that aside from the contractual payments offered by the firm, they value \((y_A,y_B)\) at \(\mu_A y_A + \mu_B y_B\), where \(\mu_i \geq 0\).  

**Supply of Intrinsic Preferences**  There is a distribution of \((\mu_A,\mu_B)\) in the population of possible agents. If that distribution includes the point \((1,1)\), then the first best is attainable by simply choosing that individual and offering no payment based on output - in that case they would choose \(e_1^* = 1\) and the problem is solved. I ignore this trivial solution by assuming that there exists no such individual and preferences are characterized by the line \(\mu_A + \mu_B = M\), where \(M < 2\). Note that supply is characterized by perfect substitution, an assumption relaxed below. The agent has a reservation utility independent of type and normalized to 0.

**Timing**  The timing of the game is as follows. First, the firm chooses an agent of type \((\mu_A,\mu_B)\) and makes a wage offer \(\beta_0 + \beta \tilde{y}\). If the agent accepts the offer, the agent exerts efforts, output is realized, and the agent is paid. If the agent rejects the wage offer, the game begins again with the firm making an offer to another worker.

**Effort Choices**  Consider an agent \(a\) who has type \((\mu_A,\mu_B)\) with marginal pay of \(\beta\). She chooses efforts of  

\[
e_1(D) = \mu_A + (1 + D)\beta, \tag{2}
\]

and  

\[
e_2(D) = x\mu_A + (1 - x)\mu_B + (1 + D)\beta. \tag{3}
\]

For notational convenience, let \(\mu = \mu_A - \mu_B\) be the bias of the agent. To see the relevant tradeoff, note that  

\[
\frac{de_1(D)}{d\mu} = 1, \tag{4}
\]

and  

\[
\frac{de_2(D)}{d\mu} = 2x - 1 < 0. \tag{5}
\]

\[^{10}\text{Sometimes such information comes from direct observation of agent’s actions. In other cases, relevant information can come from other activities that candidates engage in. For example, during the Reagan administration, potential political appointees were asked about their membership in societies that they felt were relevant for determining allegiance to that administration’s preferences, both positive (Federalist Society) and negative (the Sierra Club). While this is new to the agency literature, the notion that matching preferences to the needs of employers is already well established in studies on efficiency in the public sector. Specifically, there is a field of research in public administration called “representative democracy” which deals with the idea that - since compensation cannot be used to align incentives effectively - the bureaucracy of the U.S. should resemble the population of the country in terms of education, voting behavior, and attitudes to social issues. See Goodsell, 2004, for details.}\]
Equations (4) and (5) provide the foundation for the results that follow: raising $\mu$ increases effort on task 1 but decreases it on task 2. The key for the result below, however, is that increasing $\mu$ increases the sum of the two efforts: $\frac{d[e_1(D)+e_2(D)]}{d\mu} = 2x > 0$, but only because $x > 0$, i.e., because tasks are specialized. However, while total effort rises with bias, it does so in an unbalanced way, as the agent is spending too much effort on task 1 to the detriment of task 2. When the firm can contract well on output, it need not incur this unbalancing cost to get efficient effort, and so does not want biased agents. By contrast, when performance measures are poor, the firm has no other way of increasing effort, and so hires more biased agents.

Before describing the outcomes of the model, it is worthwhile outlining a benchmark where the agent shares the relative preferences of principal. This arises when the agent values each output equally, namely when $\mu = 0$ (or $\mu_i = M/2$). I refer to this agent as unbiased. Intrinsic incentives then give the worker a reason to exert effort of $M/2$ on each task. In the absence of any contracting distortions, the firm could then “top up” incentives by offering $\beta = 1 - M/2$, and so the agent exerts efficient effort on each task.

Straightforward calculations (see Appendix) yield the efficient level of incentives and preferences, $\mu^*$ and $\beta^*$, where

$$\mu^* = \min\{M, \frac{4x(1-\beta^*) - 2(2x-1)(1-x)M}{1+(2x-1)^2} - \frac{M}{2}\} \geq 0,$$

and

$$\beta^* = \frac{1-x(\mu^*+M)-(1-x)\frac{M}{2}}{1+\sigma^2} \leq 1 - \frac{M}{2},$$

where $\sigma^2 = \text{var}(1+D) = \frac{(1+\delta)^2 + (1-\delta)^2}{2} - 1$.

Consider the optimal choice of agent and contract. When contracting is perfect ($\sigma^2 = 0$), (6) and (7) imply that $\mu^* = 0$. In words, when contracting is perfect, the agent hired shares the preferences of the principal. However, for any positive $\sigma^2$, monetary contracts fall, and in response, $\mu^* > 0$, so that the agent is biased. Furthermore, increases in $\sigma^2$ (weakly) increase $\mu^*$: hence, bias arises as a response to the inability to contract.

So far, I have focused on solely the incentives of the agent whose job is specialized towards activity $A$. The other agent’s tasks are specialized towards $B$. To see how specialization of tasks in this way affects organizations, consider an agent, $b$, who is the mirror image of the agent considered above, where the parameter $x$ represents how her activities are specialized towards $B$. Then the optimal choice of that agent

---

11It is worth noting that this is the unique outcome, so it is not the case when incentive contracting is efficient, who to hire is irrelevant. The reason is that the principal only observes aggregate output, not the individual components, so there remains an issue of ensuring the output is produced in the optimal fashion.
is \(-\mu^*\). Hence, as contracting becomes poorer, the firm hires workers with very divergent preferences.\(^{12}\)

The outcome with two specialized agents is described in Figure 1, where equilibrium pairs of \(\beta^*\) and \(\mu^*\) are plotted by the hashed line, and each point represents a different value of \(\sigma\). The point \(\mu^* = 0, \beta^* = 1 - \frac{M}{2}\) is the outcome when there are no contracting distortions. As \(\sigma\) increases, the equilibrium pairs of incentives and preferences are plotted, with the negative slope reflecting the substitutability of monetary and other incentives.

This completes the description of the basic model. It offers an intuitively plausible outcome - where intrinsic and monetary incentives are substitutes - yet there is a cost to relying on intrinsic motivation as efforts become unbalanced. Its novelty is in offering a tradeoff for firms that find contracting on output to be difficult, namely, it

\(^{12}\)So, to give an example, it would suggest a department of social work where the social workers spend time helping clients, but have little time for the objectives of their supervisors to save on costs. See Brehm and Gates, 1997, for evidence on the resulting distrust between social workers and their superiors, who they feel are only interested in the “bottom line”.

Figure 1: Equilibrium Bias as \(\sigma\) changes.
hires agents whose preferences closely align with the task they primarily carry out, but at the cost of having them ignore other aspects of their jobs.

3 Interaction

Thus far, the model has the seeds of a theory of conflict within firms, where as outcomes become harder to contract on, the divergence in objectives between the two agents grows at a rate \( \frac{d\mu}{d\sigma} \geq 0 \). The purpose of this section is to show that with other forms of interaction between outputs, the choice of the principal becomes more complicated, and can result in the creation of fiefdoms, capture, or indifference. I do this in two ways - first, by considering the possibility of discrete interaction by one party, and second, by addressing the possibility of continuous interaction by both agents.

3.1 Discrete Interation

Assume now that there is an unobserved discrete activity - “cooperation” - that one party, agent \( a \), can engage in that benefits the other party’s primary output, \( B \), at some cost to \( A \). This activity increases the output of \( B \) by \( \pi_B \) at a cost of reducing the output of \( A \) by \( \kappa_A \). It is assumed that \( \pi_B > \kappa_A \) so the principal would like this cooperative activity carried out. Agent \( a \) is critical for the implementation of that idea. These profits are included in the contracted measure of output \( \bar{y} \), so that the agent has reason to cooperate for monetary reasons.

There is now an additional incentive constraint, namely, that if the firm wishes this activity carried out, the agent must want to do so. The agent only cooperates if

\[
\beta(\pi_B - \kappa_A) + \mu_B \pi_B - \mu_A \kappa_A \geq 0,
\]

which defines the critical value of bias \( \mu \), called \( \mu(\beta) \), above which the agent refuses to cooperate:

\[
\mu(\beta) = 2\left[\frac{\beta(\pi_B - \kappa_A) + \pi_B M}{\pi_B + \kappa_A} - \frac{M}{2}\right].
\]

There are two relevant issues that arise from (9). First, there is a limit to how biased the agent can be if the principal wishes to induce cooperation. Second, this limit depends on \( \beta \): the better is contracting, the less need is there to distort hiring to induce efficiency. This second insight yields Proposition 1 below.

**Proposition 1** The relationship between the ability to contract and agent bias is non-monotonic in \( \sigma^2 \) the following way:

- If \( \sigma^2 < \sigma^2_1 \), then the principal chooses \( \bar{\mu} \) and \( \bar{\beta} \) as in (6) and (7), and the agent cooperates. In this region, \( \frac{d\bar{\mu}}{d\sigma} > 0 \) and \( \frac{d\bar{\beta}}{d\sigma} < 0 \)
• If $\sigma_1^2 < \sigma^2 < \sigma_2^2$, then (8) binds, $\tilde{\beta} > \beta^*$, $\tilde{\mu} < \mu^*$, and the agent cooperates. In this region, $\frac{d\tilde{\mu}}{d\sigma} < 0$ but $\frac{d\tilde{\beta}}{d\sigma} > 0$.

• If $\sigma^2 > \sigma_2^2$, then the principal chooses $\tilde{\mu}$ and $\tilde{\beta}$ as in (6) and (7), and the agent does not cooperate. In this region, $\frac{d\tilde{\mu}}{d\sigma} > 0$ and $\frac{d\tilde{\beta}}{d\sigma} < 0$.

• Let $S(\beta(\sigma^2), \mu(\sigma^2))$ define equilibrium surplus. Then $\sigma_2^2$ is finite if and only if $S(0, \mu_0) - S(0, \mu_1) \geq \pi_B - \kappa_A$, where $\mu_0 = \min\{M, \frac{3(2\pi - 2x - 1)(1-x)M}{1-(2x-1)^2} - M\}$ and $\mu_1 = \frac{2\pi B M}{\pi B + \kappa_A} - M$.

This proposition is easily explained and has an economically plausible interpretation. When contracting is good ($\sigma^2$ low), the agent has good incentives to cooperate as she has enough monetary incentive to do so, and, in any case, has close enough preferences to the principal. In this range, as contracts become less efficient, the principal responds by choosing more biased agents just as before. However, as contracting gets worse, the previously optimal contract no longer induces the agent to cooperate: this arises when $\underline{\mu}(\beta^*) = \mu^*(\beta^*)$ or

$$\underline{\mu}^*(\beta) = 2\left[\frac{\beta^* (\pi_B - \kappa_A) + \pi_B M}{\pi_B + \kappa_A} - \frac{M}{2}\right].$$

(10)

This is uniquely defined so let $\sigma_1^2$ be the level of difficulty of contracting at which (10) binds. At this point, further movements up the $(\mu^*, \beta^*)$ frontier in Figure 1 cause the cooperation constraint to be violated because of lower monetary incentives and more biased agents.

Two issues then arise - (i) does the principal want to induce cooperation at this level of incentives? and (ii) how can she do so? The answer to the first question is yes at $\sigma_1^2$, for the reason that benefits to inducing cooperation at that point are first order but the costs from marginally distorting effort away from $\mu^*$ and $\beta^*$ are second order. Hence there is some range over which the firm will induce the agent to cooperate by satisfying (8). It follows that the firm will choose to have (8) bind. When the cooperation constraint binds, note that

$$\frac{d\mu}{d\beta} = \frac{2(\pi_B - \kappa_A)}{\pi_B + \kappa_A} < 0.$$  

(11)

In words, as contracting becomes more imprecise, the principal responds by choosing less biased agents, the opposite of the previous section. It is in this sense that the model exhibits capture by one group, where an agent specialized in $A$ has preferences that get closer to those of agent $b$.

Yet there is a third possible outcome. This arises when the cost of inducing cooperation becomes too large to make it worthwhile. Specifically, let surplus produced be defined by $S(\beta, \mu)$, where these depend on efforts as in (2) and (3). Then, $S(\beta^*(\sigma^2), \mu^*(\sigma^2))$ is the surplus when the cooperation issue is ignored, and $S(\tilde{\beta}(\sigma^2), \tilde{\mu}(\sigma^2))$
is the surplus produced from effort if choices are distorted to ensure cooperation occurs. Then if it exists, define $\sigma_2^2$ by

$$S(\beta^*(\sigma_2^2), \mu^*(\sigma_2^2)) = S(\tilde{\beta}(\sigma_2^2), \tilde{\mu}(\sigma_2^2)) + \pi_B - \kappa_A.$$  \hspace{1cm} (12)

At $\sigma_2^2$, the value of cooperation just matches the cost of distorting both incentives and hiring to induce cooperation. Up to that point, the firm strictly prefers to induce the agent to cooperate. This is no longer true beyond $\sigma_2^2$, and the firm discretely shifts by (i) reducing monetary incentives, and (ii) hiring agents with very biased preferences. Proposition 1 provides a necessary and sufficient condition for this region to exist. In this region, which I term *fiefdom*, each division holds diametrically opposed preferences to each others and does not cooperate. Note that such endogenous creation of fiefdoms arises for those institutions that are least able to contract on output.

![Equilibrium Bias With Discrete Interaction](image)

**Figure 2: Equilibrium Bias With Discrete Interaction.**

The outcome of this section is described in Figure 2 for the case where fiefdoms arise. At both extremes the outcome is exactly as in Figure 1 because (i) when contracting is very good, there is no reason not to cooperate, and (ii) when contracting
is very poor, the cost of inducing cooperation is too high, and so the firm does not do so. It is in the intermediate range where the outcome differs. In this region, less efficient contracting causes agents to become less biased, as it is the most efficient way to induce cooperation. Yet this is costly to effort exerted on the primary task: hence at some point \((\sigma^2)\) the firm discretely switches back to the equilibrium of the last section though it involved no cooperation.

### 3.2 Continuous Interaction

In this section, I consider the impact of allowing more continuous interaction between the outputs on the equilibrium choice of contracts and preferences. Here it is shown that the outcomes vary more continuously than above. Specifically, both agents now have access to a technology that can inefficiently transfer resources to her primary output at the expense of the other activity, similar to Milgrom and Roberts, 1988. For simplicity, call this a “lobbying activity”. Specifically, agent \(a\) chooses an intensity of lobbying \(l\), which increases \(y_A\) by \(\lambda_A l\) but reduces \(y_B\) by \(\lambda_B l\), where \(\lambda_A < \lambda_B\). Lobbying involves a personal cost \(kl^2\). (Agent \(b\)’s incentives are the mirror image.) These effects on output are observed in the contracted output \(\tilde{y}\). Given the separability of costs, lobbying activities are chosen to maximize:

\[
\mu_A \lambda_A l - \mu_B \lambda_B l - \frac{kl^2}{2} + \beta(\lambda_A - \lambda_B)l, \tag{13}
\]

yielding the first order condition

\[
kl^*(\beta, \mu_A) \geq \mu_A(\lambda_A + \lambda_B) - M\lambda_B - \beta(\lambda_A - \lambda_B). \tag{14}
\]

where (14) binds if the right hand side is positive and \(l^*\) is zero otherwise. This condition is intuitive - at the first best level described above, when \(\mu = 0\) and \(\beta = 1 - \frac{M}{2}\), then \(l^* = 0\). In words, when incentives are high, and agents are not biased, they value maximizing aggregate output, and so do not lobby. However, when the constraint above binds, then lobbying is increasing in agent bias and decreasing in monetary incentives.

**Proposition 2** Define \(\sigma^2_3\) implicity by \(\mu^*(\sigma^2_3) = 2[\frac{M\lambda_B + \beta^*(\sigma^2_3)(\lambda_A - \lambda_B)}{\lambda_A + \lambda_B} - \frac{M}{2}]\), where \(\mu^*\) and \(\beta^*\) are defined in (6) and (7). The relationship between the ability to contract and agent bias varies with \(\sigma^2\) in the following way:

- If \(\sigma^2 \leq \sigma^2_3\), then the principal chooses \(\mu\) and \(\beta\) as in (6) and (7), and there is no lobbying.
- If \(\sigma^2 > \sigma^2_3\), then agents become more (less) biased as \(\sigma^2\) increases if and only if \(2x > (\leq) \frac{\lambda^2_A - \lambda^2_B}{k}\).
The first part of this is hardly surprising - if measurement is sufficiently good, lobbying does not arise and so nothing changes from the basic model. However, at some point \( \sigma_3 \), agents begin to lobby under the old contract. When the lobbying constraint binds, it is shown in the Appendix that the optimal choice, \( \mu^{***} \), is given by

\[
\mu^{***} = \min\left\{ M, \frac{2[\gamma - (2x - \frac{\lambda_B - \lambda_A}{k})\beta^{***}]}{1 + (2x - 1)^2 - \frac{(\lambda_A + \lambda_B)^2}{k}} - \frac{M}{2} \right\}
\]

(15)

where \( \gamma = 2x - (1 - x)(2x - 1)M - (\lambda_A + \lambda_B)(1 - \frac{M}{k}) \). Therefore the effect of reduced reduced monetary incentives on optimal bias is now linear - caused by the quadratic cost assumption - but can increase or decrease depending on parameter values. In particular, if \( 2x > \frac{\lambda_B - \lambda_A}{k} \), then as contracting becomes poorer, agents become more biased, while if \( 2x \leq \frac{\lambda_B - \lambda_A}{k} \), institutions with less ability to contract on output will result in less biased agents. The intuition here is straightforward - when the marginal cost of lobbying (normalized by its responsiveness to incentives) \( \frac{\lambda_B - \lambda_A}{k} \) is large, the efficient outcome is to deter that activity by choosing less biased agents when contracting is poorer, but if the cost of lower effort (the \( x \) term) is high, then this is reversed.

Many institutions use close to no formal pay for performance as the incentives for dysfunctional responses are so large. How then can incentives for lobbying be deterred? The only way to deter lobbying \( (l^* = 0) \) when \( \beta^* = 0 \) is to choose agent \( a \) whose type is no more biased than

\[
\mu^{***} = 2M\left[ \frac{\lambda_B}{\lambda_A + \lambda_B} - \frac{1}{2} \right] > 0,
\]

(16)

while agent \( b \) has bias \(-\mu^{***}\). Note that this will be the solution chosen by the principal as \( k \to 0 \), as otherwise the costs of lobbying become very large.

This last observation offers a view of agent selection rather different from that in previous sections, in that it is not capture by one group but rather both agents have preferences that move towards \( \mu = 0 \). Instead, it offers a notion of indifferent agents, where despite poor contracting, the principal hires agents to carry out specialized jobs whose preferences look close to his. It is in this sense that the agents are indifferent - as they care little more about their own task than the others.

The outcome here is described in Figure 3. Here once the lobbying constraint binds, \( \sigma_3 \), the outcomes change in a more continuous way, and can either decline or increase (albeit more slowly than without lobbying) than in the benchmark model.

At a more general level, in this and the previous section, another cost to specializing agents was added to the basic model. In the last example, that cost was discrete - the agent would discretely choose not to cooperate at some point. It is that discreteness that caused the unambiguous move towards capture. More generally, the effect of such activities on hiring depends on the marginal benefit of the efficient activity (more effort on the primary task) relative to that of the marginal cost on the “other”
task. This is shown here by considering another activity where marginal costs are more continuous.

4 Costly Identification of Talent

I have to this point considered only which kinds of workers firms would like to hire, by simply assuming that the firm can identify type without cost. Yet this is often not true. Suppose instead that firms incur a fixed cost $K > 0$ to identify their desired type of agent, $\mu^*$ (in (6)). If they do not incur this cost, the firm randomly hires from the population of all agents. I also assume here that $E\mu = 0$ in the population of applicants.\(^\text{13}\) The firm then decides whether to spend these resources and attract a desired type, or else randomly select an agent. Remember that $S(\beta^*(\sigma^2), \mu^*(\sigma^2))$ denotes the surplus obtained by the principal above. Then by recruiting that desired type, the firm gains utility of $S(\beta^*, \mu^*) - k$. Alternatively, they can not incur this cost

\(^{13}\text{I continue to assume that the firm maximizes surplus here.}\)
and randomly hire. This has three effects. First, on average they hire type $\mu = 0$.
Second, as they hire this type on average, they offer a contract of $\beta = \frac{1 - M}{1 + \sigma^2}$, as on average intrinsic incentives are $\frac{M}{2}$ - see (7). Third, there is variation in the motivation of workers hired - some have types greater than $\mu = 0$ while others have less. The convexity of the cost function means that this variation is costly to the firm. Let $\sigma_{\mu}^2$ be the variance of the distribution of supply of worker types. Then the firm’s utility from randomly selecting agents is easily shown to be

$$S(\frac{1 - M}{1 + \sigma^2}, \frac{M}{2}) - (1 + (2x - 1)^2)\sigma_{\mu}^2.$$  

The firm then uses targeted hiring only if

$$S(\beta^*, \mu^*) - K \geq S(\frac{1 - M}{1 + \sigma^2}, \frac{M}{2}) - (1 + (2x - 1)^2)\sigma_{\mu}^2. \quad (17)$$

$S(\beta^*, \mu^*) - S(\frac{1 - M}{1 + \sigma^2}, \frac{M}{2})$ is increasing in the inability to contract on output, $\sigma^2$. The reason is intuitive. When contracting is good, the desired agent is close to $\mu^* = 0$, the same type as is hired on average by randomly hiring. As contracting on output becomes worse, the firm optimally hires more specialized agents, and so random hiring results in a hire far from the desired agent. Additionally, the ability to compensate for random hiring - by offering large pay for performance - becomes attenuated as contracts become more costly. As a result, the relative merits of targeted versus random hiring cross once in $\sigma^2$ space, as seen in Proposition 3.

**Proposition 3** Assume that it costs firms $k$ to identify the type of its candidate employees and that all workers earn rents from the job. Then:

- If $K < (1 + (2x - 1)^2)\sigma_{\mu}^2$, the firm always targets hiring.
- If $K \geq (1 + (2x - 1)^2)\sigma_{\mu}^2$, then for all $\sigma^2 < \sigma^{2*}$, the firm hires randomly, but targets hiring on $(\beta^*(\sigma^2), \mu^*(\sigma^2))$ for all $\sigma^2 \geq \sigma^{2*}$, where $\sigma^{2*}$ is finite if $S(0, \mu^*(\infty)) - k > S(0, 0) - (1 + (2x - 1)^2)\sigma_{\mu}^2$.

The reason for this section is simple - to capture another intuition about the role of hiring as contracting varies. Specifically, when contracting is good, it is not difficult to orient the actions of agents as pay for performance is not so costly. Hence, who cares who is hired? This section formalizes this, simply showing that firms in good contracting environments are content to devote little resources to recruiting, whereas those who find contracting difficult will be willing to incur costs to find the right person.

The outcome is described in Figure 4, where those who can contract well randomly hire workers, whereas those who cannot target hiring.

## 5 Robustness

The model so far has made a series of potentially restrictive assumptions to generate its results. In this section, I relax many of these assumptions. In some of
these cases - where the results remain largely unchanged - I relegate all substantive analysis of the problem to the Appendix. In those cases where results change in any substantive way, I address these changes here.

### 5.1 Contracting on Individual Output

The model thus far only allows contracting on aggregate output. As a result, the only way to orient the agent towards one activity over the other is to change who you hire. But couldn’t this be done instead by contracting on the individual components of output, thus eliminating the need for the hiring practices outlined above? In order to identify the robustness of the insights, a natural extension would be to allow contracting on the individual efforts. To address this, assume that in addition to contracting on $\bar{y}$, the principal can observe an imperfect measure of $y_i$, $i = 1, 2$, given by

$$\bar{y}_i = (1 + \delta_i)e_i,$$

(18)
where $\delta_i$ takes on a values $\delta^*$ and $-\delta^*$ with equal probability. (So, for example, noisy information could be obtained both on the quality of service provision to customers, in addition to the costs of doing so.) The $\delta_i$ variables are uncorrelated with each other and with the distortion on the aggregate signal, $\delta$.\textsuperscript{14} In the Appendix, it is show that allowing this extension does not change the results in any qualitative way - bias and the efficiency of contracting still remain substitutes, and the preferences of the agent tend towards unbiasedness as contracting becomes perfectly efficient. See the Appendix for details.

### 5.2 Cost of Effort

In the sections above, the costs of effort on the two tasks were assumed to be independent. This was done to simplify the analysis but does not change the essential logic of the paper. To see this, assume that the cost function for effort is now given by $C(e_1, e_2) = \frac{e_1^2}{2} + \zeta e_1 e_2 + \frac{e_2^2}{2}$, where $\zeta < 1$. Hence effort on one task increases the marginal cost of the other effort, as seems reasonable. Once again, the results remain unchanged when this extension is added, as shown in the Appendix.

### 5.3 State-Contingent Contracts

By assumption, contracts have not been made state contingent. In order to identify the robustness of the earlier results, consider a setting where the principal can offer a different monetary contract when $D = \delta$ compared to when $D = -\delta$.\textsuperscript{15} It is shown in the Appendix that the optimal monetary contract will indeed be state dependent, but average incentives are unchanged from the basic model. Furthermore, the bias of the agent is negatively correlated with the efficiency of average incentives in the same way as above. Hence this extension does not change the results of the paper.

### 5.4 The preferences of the firm

Thus far, I have assumed that that the firm’s objective is to maximize social surplus, which ties down the optimal choices of $\mu$ and $\beta$ in (6) and (7). Although agency problems usually consider firms as profit maximizers, such surplus maximization problems are ultimately commonplace in this literature, and reflect the fact that all effort costs must be ultimately paid in the form of higher expected wages to workers.

\textsuperscript{14}Allowing covariance between the error terms does not yield any interesting insights. As the costs of effort are independent, correlation between $\delta_1$ and $\delta_2$ changes no results. What does change are the outcomes when there is correlation between the $\delta_i$ and $D$, the aggregate signal, as correlation between the signals exacerbates expected distortions in effort. However, this extension changes results in the obvious way - namely, increased correlation reduces incentives.

\textsuperscript{15}Another conceivable extension would be where both the type of agent hired and the contract can be conditioned on the state. I do not deal with this because at a practical level, the idea that potential applicants to a firm privately observe contracting distortions seems far-fetched.
However, there is one additional issue here. Specifically, agents in this paper care about what is produced and so the wage accepted may depend on output \textit{per se}, much like a compensating differential. This has been ignored thus far. The justification for ignoring this effect, and thereby maximizing surplus, derives from the assumption that agents cares about outputs, not \textit{their contributions} to those outputs. Therefore, all that matters for computing reservation wages is output produced, from whatever its source. To say this another way, an agent considering an offer from a firm also considers the counterfactual of what output would be produced if she were to turn down the job. Specifically, if the agent perceives that in that state, someone of the same type will ultimately accept the job, there is no difference in outputs produced if she takes the job or not. Hence if she turns down the job, her only loss is the expected wage minus effort costs. If this is the case, then so long as the firm offers an expected wage at least equal to these effort costs, the agent will accept and the principal is effectively maximizing surplus, as has been assumed so far.

Remember that in the game used above, if an agent says no to the job, the principal will offer it to someone else. What matters for reservation wages is what happens if such an additional offer is made. The implicit equilibrium restriction that justifies the principal maximizing surplus is that if any agent rejects an offer \textit{off the equilibrium path}, the principal will then offer the same contract to the same type of person, and that contract will be accepted. If this is the case, the agent sees no reason to take a discount for output produced - even though she cares about it - as she believes that the next person offered the job will produce just as much output. This assumption allows me to focus on the case where the outcome is generated solely by the demands of the firm, as the reservation wage is independent of type.

In order to determine the robustness of these outcomes, consider an extreme alternative assumption where, if the agent rejects the offer, the firm can make no more offers. This is the opposite extreme, as it implies that the agent now realizes that if she says no, something she cares about (outputs) cannot now be produced. The difference between this and the case studied above is that the worker will no longer accept an expected wage that offers at least $\frac{c_1^2}{2} + \frac{c_2^2}{2}$ (remember that the wage elsewhere has been normalized to 0) but will accept any offer at least as great as $\frac{c_1^2}{2} + \frac{c_2^2}{2} - (\mu_A y_A + \mu_B y_B)$. As a result, the firm’s preferences now become to maximize $E[(1 + \mu_A)y_A + (1 + \mu_B)y_B - \frac{c_1^2}{2} - \frac{c_2^2}{2}]$, whereas before it was to maximize $E[y_A + y_B - \frac{c_1^2}{2} - \frac{c_2^2}{2}]$: in effect the weight placed on each output increases as there are benefits both in terms of output and reducing wages. Maximizing this objective function has exactly the same features as in the standard model, and the comparative statics are the same. As a result, the results above generalize to cases where efforts affect the supply price of agents.
5.5 Preference Frontier

So far, I have assumed perfect substitution between $\mu_A$ and $\mu_B$ on the supply side. Assume now that instead of $\mu_A + \mu_B = M$, the frontier of expected ability is given by $\mu_B = -f(\mu_A)$ where $f' > 0$. Hence the abilities need not be perfect substitutes.

In the basic model, total effort exerted is affected by bias because jobs are specialized: remember that in the model above, $\frac{d[e_1(D)+e_2(D)]}{d\mu} = 2x > 0$ so hiring more biased agents increased total effort. With non-perfect substitution, this is now conflated with an effect on the supply side and is given by $\frac{d[e_1(D)+e_2(D)]}{d\mu} = 1 + x - f'(\mu_A)(1-x)$, which can no longer be signed.

The central comparative static of the model - that worse contracting results in more biased agents - continue to hold so long as $1 + x - f'(\mu_A)(1-x) > 0$. Let $\eta = 1 + x - f'(\mu_A)(1-x)$. Then the optimal choice of bias if the first order conditions continue to characterize equilibrium is now given by

$$\mu_A^* = \min\{M, \frac{\eta(1-\beta^*) - (\eta - 1)(1-x)(f(\mu_A) - \mu_A)}{1 + (2x - 1)(\eta - 1)}\},$$

(19)

The analog to $\eta$ in the previous section was $2x$, and the analog to $f(\mu_A) - \mu_A = -(\mu_A + \mu_B)$ was $-M$. If $\eta > 0$, bias is negatively related to the ability to contract, so the result is not affected by this change, subject to the first order approach continuing to hold.\(^{16}\) What does change, of course, is that in the limit as contracting becomes perfect, it is no longer the case that the desired preferences of the agent converge to those of the principal. In the case where $\eta < 0$, the results obviously flip as hiring more biased agents then causes more effort on task 2 over task 1.

6 Where Do “Intrinsic” Incentives Come From?

So far, I have simply posited preferences for the agent. I now provide one justification for these via a notion of career concerns that I call professionalism. According to Wilson, 1989, “professionals are those employees who receive some significant portion of their incentives from organized groups of fellow practitioners located outside the agency” (p.60). It seems uncontroversial to posit that individuals have professions - social workers, lawyers, academic economists, etc. What is less clear is how this affects behavior.\(^{17}\) Perhaps because it is most familiar to economists, I focus on a career

---

\(^{16}\)Equilibria can jump discretely here when parameters are perturbed. This is not the central focus of the paper, so I ignore it here by assuming that marginal changes can be imputed with the usual first order approach.

\(^{17}\)A number of mechanisms have been proposed in the literature. First, professional training inculcates norms of behavior into individuals, where enforcement is largely internal - so, for instance, social workers are taught to care for the welfare of their clients and not doing so leads to a sense of guilt or failure. Second, peer pressure from fellow professionals may enforce certain behavior, such as where a soldier puts herself in the face of danger so as not to look bad in front of other soldiers. Finally, there is the career concerns route taken here.
concerns source of motivation, where market wages depend on prior performance. That effort is exerted to affect external perceptions is well known from the career concerns literature, such as Holmstrom, 1999, and Gibbons and Murphy, 1992. The innovation here is that the external constituency may not share the objectives of the principal. So, for example, a lawyer in the federal government may exert effort based on the prospect of getting a job in the private legal sector.

Assume now that output in each of the two activities (A and B) depends not just on efforts, but also the abilities of the agent. Reflecting the fact that the firm produces two outputs, the agent has two abilities - those in area A and area B - given by $m_A$ and $m_B$ respectively, and output produced is the sum of this ability, total effort on the activity, and noise:

$$y_i = \frac{m_i}{2} + \tilde{e}_i + \epsilon_i, \quad i = A, B,$$

where $\tilde{e}_i$ is the total effort exerted on that activity.\(^{18}\) Remember from the description above, all the agent’s $e_1$ increases output on A, as does a fraction $x$ of $e_2$. Hence $\tilde{e}_A = e_1 + xe_2$, and $\tilde{e}_B = (1 - x)e_2$. The distribution of $\epsilon_i$ is assumed to be Normal with mean 0 and variance $\sigma^2$ and the noise terms are uncorrelated with each other. The objective of the firm is to maximize the sum of the two outputs, net of wage costs.

**Career Concerns:** Firms cannot perfectly observe the abilities of workers. Instead there is symmetric uncertainty, where at the point where the agents are hired, the distribution of $m_i$ is assumed by all to be Normal with mean $\mu_i$, and variance $\sigma_0^2$. The firm can select agents with different perceived ability subject to a constraint that

$$\mu_A + \mu_B = M$$

Agents vary in their perceived ability according to (21), and the firm can costlessly choose an agent anywhere along this frontier.\(^{19}\) Conditional on an initial choice of expected abilities, true abilities are uncorrelated with each other.

Career concerns models are premised on the assumption that observed performance reflects ability, which is reflected in future wages. To model this, assume that there is another (undiscounted) period - period 2 - in which the agent will be employed.

The agent may change firms for period 2 based on observed outputs. The reason for this is that firms vary in how they use skills: specifically, let each firm be indexed by $\tau$, where a firm of type $\tau$ values ability A at $\tau m_A$ and ability B at $(1 - \tau)m_B$. All other assumptions are unchanged re productivity.\(^{20}\) $\tau$ has a natural support of

\(^{18}\)The $\frac{1}{2}$ on the abilities is simply a normalization: its role will be clear below.
\(^{19}\)Note that $\hat{y}$ is an unbiased measure of total output given this assumption.
\(^{20}\)Technically, each firm receives a $\tau$ draw at the end of period 1, where there is a continuum of firms at each $\tau$ ex post.
0 to 1 so at the two extremes, the firms use only one of the two skills, whereas all others use at least some of each. Further assume that in the second period, there is efficient matching of workers to jobs and the agent’s reservation wage is her expected productivity in the most efficient match.

The market observes $y_A$ and $y_B$ for each agent and updates its perception of the agent’s abilities to $\hat{\mu}_A$ and $\hat{\mu}_B$ respectively.\(^{21}\) As the agent matches efficiently, she will be employed in the firm where most surplus is created, including whatever surplus is created through incentive payments in that period.

**Modified Timing** The timing of the modified game is as follows. First, the firm chooses an agent of type $(\mu_A, \mu_B)$ and makes a wage offer $\beta_0 + \beta \tilde{y}$ to that agent. The market observes all of these. If the agent accepts the offer, the agent exerts efforts, and outputs $y_A$ and $y_B$ are realized and observed by all parties, as is $\tilde{y}$. The agent is then paid. After this, firms can make an offer to the agent for period 2. The agent can accept at most one offer in that period. Contracting and payment in period 2 are exactly as in period 1, and the game ends after the agent is paid in period 2. If the agent rejects the wage offer in either period, the game begins again with the firm makes an offer to another worker.

The usual one dimensional career concerns model of ability implies that the worker’s wage depends on perceived ability. The natural extension of this here is that this wage depends on some function of perceptions of both abilities. The assumption of efficient matching makes this relationship between wages and perceptions very simple, as seen in Lemma 1.

**Lemma 1** The agent’s utility in period 2 is $\max\{\hat{\mu}_A, \hat{\mu}_B\}$ and a constant independent of ability.

Hence, all that matters for future pay is which ability is higher. Effort is exerted in period 1 to maximize both contemporaneous monetary payments and these future wages. Note that Lemma 1 characterizes wages conditional on the realizations of first period output. When choosing effort, the agent must therefore take expectations of the future wages characterized by Lemma 1.

For notational convenience, let $\mu = \mu_A - \mu_B$, and $\hat{\mu} = \hat{\mu}_A - \hat{\mu}_B$ be the difference in expected abilities before and after observing first period output. These will be the shorthand for bias in what follows. For the career incentives in Lemma 1, all that matters is which ability is higher and so the agent estimates the likelihood that $\hat{\mu} > 0$. Routine calculation shows that the distribution of $\hat{\mu}$ is Normal with mean $\mu$ and variance $\frac{\sigma^4}{\sigma^2_0 + \sigma^2_\tilde{y}}$. Let $\Phi(.)$ be a normal distribution with mean 0 and variance

\(^{21}\)Note that $\tilde{y}$ carries no additional information on ability, so updating is solely based on $y_A$ and $y_B$. For other recent work on how career concerns affect incentives in similar settings, see Dewatripont et al, 2003.
Then the equilibrium probability that $\hat{\mu} > 0$ - i.e., that skill $A$ determines pay in period 2 - is given by $1 - \Phi(-\mu)$, and the probability that his reservation wage is determined by $B$ ($\hat{\mu} \leq 0$) is given by $\Phi(-\mu)$.

It is worthwhile pausing here, as $[1 - \Phi(-\mu)]$ and $\Phi(-\mu)$ provide the key to the insights below. When the firm chooses a worker whose $\mu$ is high, that worker is perceived to have greater talent in activity $A$ than in $B$. This is relevant here only because this makes activity $A$ the most likely next employer, and so the agent orientates her efforts more in that direction, to the detriment of activity $B$.

**The agent’s choice of effort** Consider the incentives of the agent in period one, including marginal incentive payments of $\beta$. Let $\sigma^2 = \frac{\sigma^2}{\sigma_0^2 + \sigma_c^2}$ be the signal to noise ratio for updating ability.

**Lemma 2** Equilibrium efforts are given by

$$e_1(D) = s[1 - \Phi(-\mu)] + (1 + D)\beta,$$

and

$$e_2(D) = xs[1 - \Phi(-\mu)] + (1 - x)s\Phi(-\mu) + (1 - D)\beta.$$

The incentives of the agent are intuitive (although the proof is a little more complicated as out-of-equilibrium actions affect both perceived ability conditional on a job, and which job is attained). First, the monetary part of this is familiar - agents exert effort to increase pay, where the parameter $D$ induces a distortion as in Baker, 1992. Less familiar are the terms involving $\Phi$. $1 - \Phi(-\mu)$ is the probability of $A$ being the next employer - in which case increasing output $A$ is beneficial - and $\Phi(-\mu)$ is the likelihood that the next employer uses $B$ instead, in which case the incentive is to increase $y_B$.

**Proposition 4** The optimal choice of agent and monetary incentives are given by

$$1 - \Phi(-\mu^*) = \min\{1, \frac{2x(1 - \beta^*) - s(1 - x)(2x - 1)}{s(1 + (2x - 1)^2)}\} \geq \frac{1}{2},$$

and

$$\beta^* = \frac{2 - (1 + x)s + 2xs\Phi(-\mu^*)}{2(1 + \sigma^2)} \leq 1 - \frac{s}{2},$$

where $\sigma^2 = \text{var}(1 + D) = \frac{(1 + \delta)^2(1 - \delta)^2}{2} - 1$.

These outcomes closely mirror those of the model with exogenous preferences. First, when contracting is perfect ($\sigma^2 = 0$), (24) implies that $\Phi(-\mu^*) = \frac{1}{2}$ or $\mu^* = 0$. Second, for any positive $\sigma^2$, monetary contracts fall, and in response, $1 - \Phi(-\mu^*) > \frac{1}{2}$, so that the agent is biased. Finally, increases in $\sigma^2$ (weakly) increase $\mu^*$. Hence the outcomes in the basic model above can easily be interpreted through the lens of this form of career concerns.

---

22Note that some monetary contracts are always necessary to induce efficiency as the career incentives can never sum to more than $s < 1$. 

6.1 Ability and Specialization

One of the primary motivations for the paper is that tasks tend to be specialized in firms. So, for example, I primarily do research and my dean primarily deals with alumni, students, and large donors. The reflection of this is the parameter $x$, such that efforts disproportionately affect activity $A$ in the basic model. However, the model of this section incorporates both effort and ability, yet only the effects of effort decisions are specialized. Said another way, if my efforts primarily affect research, shouldn’t my abilities do likewise? So far, I have ignored this though the technology in (20), which implies that holding effort fixed, the firm values both abilities equally.

To capture this possibility of a direct effect of ability on expected output, I now add a symmetry between specialization in efforts and abilities, where the parameter $x$ affects not only the marginal effect of increasing efforts, but also the marginal effect of ability. Specifically, assume now that output is no longer given by (20) but rather

$$y_A = \tilde{e}_A + (1 + x)\left(\frac{m_A}{2} + \epsilon_A\right),$$  \hspace{1cm} (26)

and

$$y_B = \tilde{e}_B + (1 - x)\left(\frac{m_B}{2} + \epsilon_B\right).$$  \hspace{1cm} (27)

This technology is now symmetric in its treatment of ability and effort - in words, if efforts have a biased effect on output, now ability does likewise. The import of this is extension that there is now an additional reason to bias hiring - namely, even if efforts are zero, the agent should be biased towards that task at which she is specialized as that ability matters more. Straightforward calculations reveal that the optimal choice of contract is still given by (25) but the optimal choice of agent is now given by

$$1 - \Phi\left(-\frac{\mu^s}{s}\right) = \min\{1, \frac{2x(1 - \beta^s) + \frac{x}{\frac{2\phi(-\frac{\mu^s}{s})}{s}} - s(1 - x)(2x - 1)}{s(1 + (2x - 1)^2)}\}. \hspace{1cm} (28)$$

This has only one change from the basic model’s outcome - the term was previously not relevant. This term offers an additional reason for bias because ability has a greater direct effect on $A$ than on $B$ (though $x$). Hence if ability has no effect on effort (because $s = 0$), then the firm would hire the most biased agent available. It remains the case that this bias arises even in the limit where $\sigma^2 = 0$. Subjects to these caveats, though, the results of the paper are robust to this extension as all comparative statics remain unchanged.

---

23 There is one issue that is ignored here concerning learning. Specifically, by multiplying the error terms by $x$, it ignores the possibility that speed of learning may be faster for one ability than the other. This is ignored here because my interest in this paper is not in instances where it is easier to provide incentives on one task than the other. Instead, I retain symmetry across tasks throughout the paper in order to avoid the obvious insight that when one task is easier than the other to provide incentives on, hiring should be weighted towards those who will exert effort on the other task without monetary incentives.

24 It is also the case here that the first best is not attained in the limit here, and there are
An Observation: Note that without this extension, the firm being studied valued each ability equally yet other firms in period two had a distribution of types given by \( \tau \). This sub-section shows that the results above generalize to cases where firms value abilities differently. Specifically, this section maps into the \( \tau \) interpretation as the model here has solved the case where \( \tau = x \). In other words, \( \tau \) comes from the specialization of tasks, a natural interpretation. To say this another way, this section generalized the results of the earlier section to show how it does not matter for the comparative statics the mechanism by which agents are matched to firms in the first period as all firms have similar demands on the margin.

7 Conclusion

The central premise of this work is that institutions may function best when their employees do not share a common objective, but where this divergence in objectives depends on the ability to contract and the specialization of tasks. There are other papers that share some of its insights, though on other dimensions. First, Itoh, 1992, and Dessein, Garicano, and Gertner, 2008, Rotemberg and Saloner, 1995, show how monetary contracts can be designed to tradeoff efforts on primary and secondary tasks, where the bias is generated though the contracts. The contributions of Che and Karthik, 2009, and Van Den Steen, 2004, 2005a, 2005b, are closest to this work, in that they show how hiring agents with biased beliefs about the marginal effects of their efforts can improve efficiency. Finally, Prendergast, 2007, offers another reason for the benefits of bias, based on agent’s altruism towards clients.25

Perhaps the central problem for the economics literature on agency theory is that in a wide range of situations, tying pay to performance simply does not help. This paper argues that in such settings, a useful line of research may be to consider recruitment based on the preferences or skills of potential employees. Here the tradeoffs become somewhat different to normal - in the basic model, the price of not being able to contract on output is that there will be a divergence of preferences across different parts of the organization because the firm may end up hiring workers with (often) radically different interests. When direct interaction between agents was considered, the cost of this is either fiefdoms, where agents refuse to help each other yet are zealous about their own tasks, or capture, where the institution ends up recruiting agents who are (more) similar, even though they carry out very different jobs. As an example, a likely cost of operating say a non-profit institution is the possibility of difficulties in integrating different aspects of what the firms does. If nothing else, it at least raises both these issues in these firms, and considers the use of an instrument unbalanced efforts between the two tasks. This arises because the firm cannot separately identify \( y_A \) and \( y_B \). When they can, the limiting case involved first best efforts, but agents remain biased towards their primary task.

25See also MacLeod, 2003, for other work on the costs of conflict in settings with subjective performance measures.
other than pay as a way of aligning interests.
REFERENCES


Optimal $\mu$ and $\beta$  The objective of the principal is to choose $\mu_A, \mu_B,$ and $\beta$ to maximize

$$E[e_1 + e_2 - \frac{e_1^2}{2} - \frac{e_2^2}{2}], \quad (29)$$

subject to $\mu_A + \mu_B = M$, (2), (3), $\mu_i \geq 0$, $D = \delta(-\delta)$ with probability $\frac{1}{2}$. By substitution, the principal chooses the agent’s type ($\mu_A$) and the contract $\beta$ to maximize

$$2\beta + (1 + x)\mu_A + (1 - x)(M - \mu_A) - \frac{((1 + \delta)\beta + \mu_A)^2}{4} - \frac{((1 - \delta)\beta + \mu_A)^2}{4}$$

$$- \frac{(1 + \delta)\beta + x\mu_A + (1 - x)(M - \mu_A))^2}{4}$$

$$- \frac{(1 - \delta)\beta + x\mu_A + (1 - x)(M - \mu_A))^2}{4}. \quad (30)$$

Straightforward differentiation yields (6) and (7).

Proof of Proposition 1: With the cooperation constraint, the objective of the firm is now to maximize expected surplus, subject to (2), (3), and now (8) if the firms wishes to induce cooperation. Begin by ignoring the cooperation constraint, in which case the firm’s choice is given by (6) and (7). Therefore, if either (i) the firm can induce cooperation without changing from (6) and (7) or (ii) does not wish to induce cooperation, the solution remains that given by (6) and (7).

Note that for $\sigma^2$ low enough, (8) is satisfied at the equilibrium choices given by (6) and (7). To see this, note that as $\sigma^2 \to 0$, then $\beta^* \to 1 - \frac{M}{2}$ and $\mu^* \to 0$ in which case (8) holds. As $\beta^*$ and $\mu^*$ vary continuously with $\sigma^2$, this implies that there is a range over which the firm does not change its choice of agent with the cooperation constraint. However, as $\sigma^2$ increases, then if there exists a point at which $\mu_A^* = \beta^*(\pi_B - \kappa_A + \pi_B M), \pi_B + \kappa_A$, then the cooperation constraint is violated for all higher values of $\sigma^2$ as $\beta^*$ and $-\mu^*$ are decreasing in $\sigma^2$.

The optimal solution ($\tilde{\beta}, \tilde{\mu}$) then depends on whether the firm wishes to induce cooperation. If it does not, then the solution continues to be characterized by (6) and (7). It does, then the firm will choose (8) to bind, in which case the firm chooses combinations of $\tilde{\beta}$ and $\tilde{\mu}$ such that $\frac{d\mu}{d\beta} = \frac{2(\pi_B - \kappa_A)}{\pi_B + \kappa_A} \equiv g > 0$. Then straightforward calculations show that the optimal choice of $\tilde{\beta}$ is given by

$$\beta = \frac{2 - 2xg - (1 - g) - x(1 + g(1 - 2x))}{(1 + \sigma^2)((1 - g)^2 + (1 + g(1 - 2x))^2]} \quad (31)$$

which is decreasing in $\sigma^2$, as required. Hence, if there exists a point where (8) binds, there is a range where $\tilde{\beta}$ and $\tilde{\mu}$ decline with $\sigma^2$.

Let the surplus generated by $e_1, e_2$ be defined by $S(\beta, \mu) = e_1(\beta, \mu) + e_2(\beta, \mu) - \frac{e_1(\beta, \mu)^2}{2} + \frac{e_2(\beta, \mu)^2}{2}$ where $e_1$ and $e_2$ are defined in (2) and (3). Then, if it exists, define $\sigma_2$ by $S(\beta^*(\sigma_2), \mu^*(\sigma_2)) = S(\tilde{\beta}(\sigma_2), \tilde{\mu}(\sigma_2)) + \pi_B - \kappa_A$. At this point, the benefits
of cooperation are just matched by the costs in terms of distorted contracting and hiring. If this condition holds for any $\sigma_2$, it must be the case that $S(\beta^*(\sigma), \mu^*(\sigma)) > S(\bar{\beta}(\sigma), \bar{\mu}(\sigma)) + \pi_B - \kappa_A$ for all larger values of $\sigma$ because for all $\beta > \beta^*$, $\frac{d^2 S}{d\beta^2} < 0$, and for all $\beta \leq \beta^*$, $\frac{d^2 S}{d\beta^2} = 0$. Therefore as $\bar{\beta} > \beta^*$ and $\bar{\mu} > \mu^*$, $S(\beta^*(\sigma), \mu^*(\sigma)) - S(\bar{\beta}(\sigma), \bar{\mu}(\sigma))$ is increasing in $\sigma$. As a result, for all $\sigma > \sigma_2$, the firm does not induce cooperation but instead chooses (6) and (7).

Of course, no such value of $\sigma_2^2$ may exist. Consider the limiting case as $\sigma_2$ tends to $\infty$. Then $\beta^* \to 0$ and $\mu^* \to \mu_0$, where $\mu_0 = 2[\frac{2x-(2x-1)(1-x)M}{1+(2x-1)^2} - \frac{M}{2}]$. This is the optimal level of bias implemented if no cooperation arises. Similarly consider the return to inducing cooperation as $\sigma_2$ tends to $\infty$. As $\beta^* \to 0$, $\bar{\mu} \to \mu_1$, where $\mu_1 = 2[\frac{\pi_BM}{\pi_B+\kappa_A} - \frac{M}{2}]$. A necessary and sufficient condition for fiefdom to exist is then

$$S(0, \mu_0) - S(0, \mu_1) \geq \pi_B - \kappa_A.$$  

(32)

**Proof of Proposition 2:** The solution characterized in Section 2 continues to hold if the agents do not lobby. However, the lobbying constraint binds at $\mu^*(\sigma_1), \beta^*(\sigma_1)$ where

$$\mu^*(\sigma_1) = 2[\frac{M\lambda_B + \beta^*(\sigma_1)(\lambda_A - \lambda_B)}{\lambda_A + \lambda_B} - \frac{M}{2}].$$  

(33)

When the lobbying constraint binds, the firm maximizes expected surplus, which us is now given by

$$2\beta + (1 + x)\mu_A + (1 - x)(M - \mu_A) - \frac{(1 + \delta)\beta + \mu_A}{4} - \frac{(1 - \delta)\beta + \mu_A}{4}$$

$$- \frac{(1 + \delta)\beta + x\mu_A + (1 - x)(M - \mu_A)}{4} - \frac{(1 - \delta)\beta + x\mu_A + (1 - x)(M - \mu_A)}{4}$$

$$- (\lambda_B - \lambda_A)l - k\frac{l^2}{2},$$  

(34)

subject to $\mu_A + \mu_B = M$, (2), (3), and $D = \delta(-\delta)$ with probability $\frac{1}{2}$, and $l$ characterized by (14) holding with equality. Maximizing this yields optimal level of incentives and preferences are given by

$$\mu^*_A = \frac{2x(1 - \beta^*) - (1 - x)(2x - 1)M - (\lambda_A + \lambda_B)(kl^* + \lambda_B - \lambda_A)}{1 + (2x - 1)^2}.$$  

(35)

As $l^*$ is a function of $\beta$, substitution is necessary to determine the total effect of changing monetary incentives on intrinsic motivation. Substituting for this and noting that $\mu = 2[\mu_A - \frac{M}{2}]$, equilibrium bias is given by

$$\mu^{***} = \min\{M, \frac{2[\gamma - (2x - \frac{\lambda_A^2 - \lambda_B^2}{k})\beta^{***}]}{1 + (2x - 1)^2} - \frac{M}{2}\}$$  

(36)

where $\gamma = 2x - (1 - x)(2x - 1)M - (\lambda_A + \lambda_B)(1 - \frac{M\lambda_B}{k})$. Proposition 3 then follows.
Contracting on Individual Outputs The firm now offers a contract where the wage - modulo a fixed payment - is given by

\[ w = \beta_1 \bar{y}_1 + \beta_2 \bar{y}_2 + \beta \bar{y}. \]  

(37)

Hence the firm can now influence the relative choice of \( e_1 \) and \( e_2 \) directly through contracts rather than only through the preferences of the agents that they hire. Let

\[ \eta^2 = \text{var}(1 + \delta^*) = \frac{(1+\delta^*)^2+(1-\delta^*)^2}{2} - 1. \]

The objective of the principal is then to choose \( \mu_A, \beta_1, \beta_2 \) and \( \beta \) to maximize

\[ E[e_1 + e_2 - \frac{e_1^2}{2} - \frac{e_2^2}{2}] \]

subject to \( \mu_A + \mu_B = M, e_1 = \beta_1(1+\delta) + \beta(1+D) + \mu_A, e_2 = \beta_2(1+\delta_2) + \beta(1-D) + x\mu_A + (1-x)(M-\mu_A), \mu_i \geq 0, D = \delta(-\delta) \) with probability \( \frac{1}{2} \), and \( \delta_2 = \delta^*(-\delta^*) \) with probability \( \frac{1}{2} \). The optimization program of the principal is now given by maximizing

\[
2x\mu_A + (1-x)M + \beta_1 + \beta_2 + 2\beta - \frac{((1+\delta)\beta + (1+\delta^*)\beta_1 + \mu_A)^2}{8} \\
- \frac{((1+\delta)\beta + (1-\delta^*)\beta_1 + \mu_A)^2}{8} - \frac{((1-\delta)\beta + (1+\delta^*)\beta_1 + \mu_A)^2}{8} \\
- \frac{((1-\delta)\beta + (1-\delta^*)\beta_1 + \mu_A)^2}{8} - \frac{((1-\delta)\beta + (1+\delta^*)\beta_2 + x\mu_A + (1-x)(M-\mu_A))^2}{8} \\
- \frac{((1+\delta)\beta + (1-\delta^*)\beta_2 + x\mu_A + (1-x)(M-\mu_A))^2}{8} \\
- \frac{((1+\delta)\beta + (1-\delta^*)\beta_2 + x\mu_A + (1-x)(M-\mu_A))^2}{8}. \]

(39)

The optimal choice of four relevant variables - \( \{\beta_1, \beta_2, \beta, \mu_A\} \) - are simply characterized by

\[ \beta_1^* = \frac{1 - \mu_A^* - \beta^*}{1 + \eta^2}, \]

(40)

\[ \beta_2^* = \frac{1 + (2x-1)\mu_A^* - (1-x)M - \beta^*}{1 + \eta^2}, \]

(41)

\[ \beta^* = \frac{2 - 2x\mu_A^* - (1-x)M - \beta_1^* - \beta_2^*}{1 + \sigma^2}, \]

(42)

and

\[ \mu_A^* = \frac{2x(1-\beta^*) - (1-x)(2x-1)M - \beta_1^* - (2x-1)\beta_2^*}{(1 + (2x-1)^2)}. \]

(43)

However, note that from (40) and (41) that \( \beta_1^* = (1-2x)\beta_2^* \) so (42) and (43) become

\[ \mu_A^* = \frac{2x(1-\beta^*) - M(1-x)(2x-1)}{s(1 + (2x-1)^2)}. \]

(44)
and
\[ \beta^* = \frac{\eta^2}{\eta^2 + \sigma^2 + \eta^2 \sigma^2} (1 - \frac{M(1 + x)}{2} - \mu^*_A). \] (45)

The equilibrium choice of agent \( a \) hired can then be characterized in terms of the exogenous parameters of the model as
\[ \mu^*_A = \frac{2x - (2x - 1)(1 - x)M - 2xz(1 - \frac{M(1-x)}{2})}{1 + (2x - 1)^2 - 2xz}, \] (46)
where \( z = \frac{\eta^2}{\eta^2 + \sigma^2 + \eta^2 \sigma^2} \), as required. The two instruments \( \beta \) and \( \mu \) vary exactly as above. The previous section is analogous to a case where \( \eta^2 = \infty \). But for any \( \eta^2 < \infty \), \( \beta^* \) is lower and so equilibrium bias increases. Hence the case considered in the previous section offers least bias, so that contracting on individual outputs strengthens the results.

**Interactions in the Cost of Effort**

Straightforward calculations then show that in the basic model of Section 2 - but where there is an additional cost of \( \zeta e_1 e_2 \) in the agent’s cost function - the optimal choice of incentives \( \beta \) and agent type \( \mu \) is given by
\[ \beta^* = \frac{2(1 - \zeta^2)(1 - \zeta) - (1 - \zeta)^2 2x \mu^*_A - (1 - \zeta^2)(1 - x)M}{2(1 - \zeta)^2 (1 + \sigma^2)}, \] (47)
and
\[ \mu^*_A = \min \{ M, \frac{(1 - \zeta^2)(1 - \zeta)2x - (1 - \zeta)^2 2x \beta^* - (1 + \zeta^2)(2x - 1 - 2\zeta)(1 - x)M}{[(1 - \zeta(2x - 1))^2 + (2x - 1 - \zeta)^2]} \}. \] (48)

This equilibrium has exactly the same features as in the basic model. Financial incentives and bias are substitutes, with bias increasing as the ability to contract on output becomes worse. Similarly, the limiting case of perfect contracting still results in the unique outcome of unbiased agents (\( \mu^* = 0 \)) and incentives given by \( \beta^* = 1 - \frac{x}{2} \). Hence, the insights extend to the case where the cost functions are not independent in this way.

**Optimal Contracts**

As there are only two states of nature, a maximum of two different contracts are necessary. Call the marginal monetary payment in these states \( \beta \) and \( \beta \) respectively.\(^{26}\)

Initially assume that the principal knows the state of nature. Then simple computation shows that she will choose
\[ \beta^* = \frac{1 - \frac{(1-x)M}{2} - 2x \mu^*_A - \delta(1-x) \mu^*}{1 + \sigma^2}, \] (49)

\(^{26}\)As there is no residual uncertainty, the assumption of linearity is without loss of generality, and merely specifies a unique mapping from the relevant contract into a \( y \) for that state.
and
\[ \beta^* = \frac{1 - \frac{(1-x)M}{2} - 2x\mu_A^* + \delta(1-x)\mu^*}{1 + \sigma^2}, \]  
(50)

and
\[ \mu_A^* = \min\{1, \frac{2x(1 - E\beta^*) - M(1-x)(2x-1) - 2x\delta^2(1-x)}{1 + (2x-1)^2 - 4x\delta^2(1-x)}\}. \]  
(51)

First note that the optimal monetary contract depends on the state, but that average monetary incentives are unchanged \((\bar{\beta}^* + \bar{\beta}^*) = \beta^*)\). Comparative statics here are as in the basic model, where average incentives and bias become substitutes as before. Furthermore note that there is no truth-telling problem here, as it is trivial to show that the agent prefers \(\beta^*(\bar{\beta}^*)\) when the state is \(\delta(-\delta)\) to the other contract, and hence this will be the outcome when contracting is state contingent.

**Proof of Lemma 1:** By assumption, there is efficient matching of workers to posts. As no effort is exerted for career concerns reasons, the agent matches to the firm that offers the highest value of
\[ (\tau^*(\hat{\mu}_A, \hat{\mu}_B))\hat{\mu}_A + (1 - \tau^*(\hat{\mu}_A, \hat{\mu}_B))\hat{\mu}_B. \]  
(52)

This has a very simple allocation for the second period - if \(\hat{\mu}_A \geq \hat{\mu}_B\), then \(\tau^* = 1\), while if \(\hat{\mu}_A < \hat{\mu}_B\), then \(\tau^* = 0\). But the firm can, of course induce effort exertion by contracting on output. Simple computations show the optimal pay for performance coefficient in period 2 is given by \(\tilde{B} = \frac{2}{1+\sigma^2}\) which is independent of perceived ability. Hence, the agent earns \(\max\{\hat{\mu}_A, \hat{\mu}_B\}\) and a constant independent of ability.

**Proof of Lemma 2:** The reservation wage of the agent in period 2 is given by
\[ \text{prob(employer} = A)w_A(\hat{\mu}_A) + \text{prob(employer} = B)w_B(\hat{\mu}_B), \]  
(53)

where “employer = \(i\)” means that the employer only uses skill \(i\), and \(w_i(\hat{\mu}_i)\) is the expected wage offered in period 2 in that job, where \(w_i(\hat{\mu}_i) = \hat{\mu}_i + [f_1 + f_2]\), and \(f_i\) is second period effort on task \(i\), which is independent of ability (from Lemma 1). In equilibrium, the probability that the employer uses only skill \(A\) is given by \([1 - \Phi(-\mu)]\). However, out of equilibrium deviations could change this, and it is only true that in equilibrium that this is treated as parametric. Specifically, the probability that the agent is employed by \(A\) is given by \(\Gamma(e_1, e_2) = [1 - \Phi(-\mu + (e_1 - Ee_1) + (2x-1)(e_2 - Ee_2))]\) and the probability of employer \(B\) being relevant is \(1 - \Gamma\) (the expectations are those held by the market). Conditional on a second period match, the marginal value of a unit increase in the relevant first period output on second period wages is \(s\), as is familiar in models of career concerns. The martingale property of these learning problems implies that when exerting effort in period one, the expected value of \(w_i(\hat{\mu}_i)\) is given by \(\mu_i + [f_1 + f_2]\). However, this is not true out of equilibrium and instead the
agent exerts effort to maximize \( W_A(\mu_A) = \mu_A + s(e_1 - Ee_1) + xs(e_2 - Ee_2) + [f_1 + f_2] \) and \( W_B(\mu_B) = \mu_B + (1 - x)s(e_2 - Ee_2) + [f_1 + f_2] \).

As a result, the agent then chooses her efforts (ignoring a constant)\(^{27}\) to maximize
\[
\max_{\{e_1(D), e_2(D)\}} \tilde{\gamma} + [1 - \Gamma] W_A + \Gamma W_B - \frac{e_1^2}{2} - \frac{e_2^2}{2}.
\]

Straightforward maximization yields (22) and (23), as all the terms involving \( \phi \), the derivative of the probability of being employed, are evaluated at \( \phi(-\mu) \) at which point \( \hat{\mu} = 0 \) so that \( w_A = w_B \). Hence, these terms can be ignored - in words, by exerting more effort on task 1, the agent is more likely to be employed in a firm that uses task 1. However, since it is a marginal change, the agent is indifferent between which of the two firms to work for, and so efforts are given by (22) and (23).

To see this, note that differentiating (54) with respect to \( e_1 \) yields
\[
(1 + D) \beta + [1 - \phi(-\mu - e_1 - Ee_1) - (2x - 1)(e_2 - Ee_2)] s + \phi(-\mu - (e_1 - Ee_1) - (2x - 1)(e_2 - Ee_2)) 0 = e_1,
\]
which yields (22) in equilibrium. Differentiating with respect to \( e_2 \) similarly yields (23).

**Proof of Proposition 4:** The objective of the principal is to maximize output minus wages, which in the usual fashion results in maximizing expected surplus - \( E[e_1 + e_2 - \frac{e_1^2}{2} - \frac{e_2^2}{2}] \) - as the agent earns her reservation utility in expectation. The objective of the principal is to choose \( \mu \), and \( \beta \) to maximize \( E[e_1 + e_2 - \frac{e_1^2}{2} - \frac{e_2^2}{2}] \) subject to (21), (22), (23), and \( D = \delta(-\delta) \) with probability \( \frac{1}{2} \). By substitution, the principal chooses the agent’s type \( (\mu) \) and the contract \( (\beta) \) to maximize
\[
2\beta + (1 + x)s[1 - \phi(-\mu)] + (1 - x)s\phi(-\mu) - \frac{((1 + \delta)\beta + s[1 - \phi(-\mu)])^2}{4} - \frac{((1 - \delta)\beta + s[1 - \phi(-\mu)])^2}{4} - \frac{((1 + \delta)\beta + xs[1 - \phi(-\mu)] + (1 - x)s\phi(-\mu))^2}{4} - \frac{((1 - \delta)\beta + xs[1 - \phi(-\mu)] + (1 - x)s\phi(-\mu))^2}{4}.
\]

Straightforward differentiation yields (24) and (25).

\(^{27}\)The constant includes the returns from second period efforts and first period returns from perceived ability, neither of which is affected by first period efforts, and so can be treated as a constant.