Contracts and Conflict in Organizations

Canice Prendergast*

March 7, 2011

Abstract

In many organizations, the way that incentive problems are alleviated is not via contracts, but rather who is hired. This paper proposes a channel through which hiring affects incentives via a notion of career concerns called professionalism, and how its role changes as firms contract better on output. A central concern of the paper is how the hiring of biased agents (those who do not share the relative preferences of their employer) changes with contracting opportunities, giving rise to ideas of conflict, capture, and fiefdoms in firms.

*University of Chicago Booth School of Business. Email: canice.prendergast@chicagobooth.edu. Thanks to Wouter Dessein, Bob Gibbons, and seminar participants at the University of Chicago, Columbia University, Northwestern University, Harvard/MIT, UCLA, UNC, ESRI, and the NBER for helpful suggestions.

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. Similarly, finding the right tradeoff between research and teaching interests characterizes some faculty hiring in universities.[1]

An innovation of this work is the premise that firms affect effort decisions through who they hire.[2] They can choose for example, Deans who will emphasize fund raising or who will emphasize faculty research or social workers who will weight the welfare of their clients over cost saving. The central issue addressed is whether agents hired should share the preferences of their employers, and also, how those objectives will change as the firm can contract better on output.

[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]There 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.

Rather than simply posit that agents have preferences for outputs, the paper offers a fully developed interpretation of such objectives through a career concerns setting which I dub "professionalism".[3] 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.

The central building block that drives the results of the paper is very simple, and is outlined in a baseline model with two features: (i) specialization of tasks, and (ii) imperfect contracting on *aggregate* output. 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 endogenously emphasize A over 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 in the baseline model - 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.[4]

[3]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 uncontentious to posit that individual have professions - social workers, lawyers, academic economists, etc. What is less clear is how this affects behavior.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. Perhaps because it is most familiar to economists, I focus on a career concerns source of motivation, where market wages depend on prior performance.

[4]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 as important..is service to clientele" (p.66), where 83 percent of survey respondents view such service as important, compared to only 9 percent

The logic for this tradeoff, outlined in Section 1, 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.

Two results arise from the baseline model - (i) as contracting becomes more perfect, there an alignment of professional interests between the firm and the agent, and (ii) as contracting gets worse, more biased agents are hired. The remainder of the paper is concerned with extending the modeling to cases where neither of these results may be true. I do this in two parts - (i) by considering other actions that the agent that can take to affect her non-primary activity, and (ii) by allowing other contracting possibilities.

This insight about conflicts of objectives in the baseline 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 2.1, 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 activ-

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.

ity.[5] Call this activity "cooperation". In this case, I show that firms with poor contracting opportunities face an additional tradeoff - between *capture* and *fiefdoms*.

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.[6] 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.[7]

These outcomes arise in the context of a discrete cooperative activity. Section 2.1.2

[5]So, 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.

[6]So 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.

[7]As 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.

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 own - but one where the form of indifference is endogenously chosen by the principal.

The second extension of the baseline model concerns allowing other contracting possibilities. Specifically, I consider the possibility of contracting on each output separately. One striking result is obtained - in the baseline model where each effort can be equally contracted on, better contracting results in less biased agents. In the model where each output can be separately contracted on equally well, hiring is *independent* of the ability to contract. Hence, this insight is specific to the case where firms contract on aggregate output. Conditions are provided in this case for when the agent responds to uncertainty in monitoring by hiring more biased agents, as in the baseline model, but also cases are considered where the results in the baseline model do not generalize.

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 3 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.

Section 1 begins by building the benchmark model that shows the tradeoffs between hiring and monetary incentives. Following this, Section 2.1 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 3 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 4, which show that the insights are generally robust to other assumptions.

1 The Model

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 an agent who provides efforts on two inputs (tasks), 1 and 2. The outputs produced by the agent depends on both how able she is at each activity, and how much effort she exerts.

The activities differ on two important dimensions with respect to effort - (i) how specialized they are, and (ii) how well they can be monitored. I deal with each in turn.

Specialization There is an asymmetry between the two tasks, where the agent's tasks primarily affect one of the two outputs. Specifically, the agent is specialized primarily in output A.[8] 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 $\frac{e_i^2}{2}$. All effort by the agent 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.[9] Hence, tasks are partially specialized, as reflected by the parameter x. For notational convenience, let $\tilde{e}_1 = e_1 + xe_2$, and $\tilde{e}_2 = (1-x)e_2$.

Output in each of the two activities (A and B) depends not just on these 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:

$$y_i = \frac{m_i}{2} + \tilde{e}_i + \epsilon_i, i = A, B, \tag{1}$$

where $\tilde{e}_i$ is the total effort exerted on that activity.[10] The distribution of ϵ_i is assumed to be Normal with mean 0 and variance $\sigma_i^2, i = A, B$, and the noise terms are uncorrelated with each other. The objective of the firm is to maximize the sum of the two expected outputs, net of wage costs.

Monitoring The efforts exerted are inputs to the objectives of the firm. The firm is assumed to observe a contractible but noisy (unbiased) estimate of aggregate output given

[8] It is simple to add a mirror image agent who is specialized in B - so it is not the case that this reflects a firm that necessarily does primarily A.

[9] 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.

[10] The $\frac{1}{2}$ on the abilities is simply a normalization: its role will be clear below.

by

$$I_i = e_1 + e_2 + \epsilon, \tag{2}$$

where ϵ_i is Normally distributed with mean 0 and variance σ^2. For simplicity, it is assumed that the error term is uncorrelated with the error terms in (1). The first choice variable of the firm is how sensitive wages will be to this measure, i.e., *pay for performance.*

Professionalism and Career Concerns: To allow for another source of incentives here, I assume the agent exerts effort in the prospect of earning a higher future wage. 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.

Unlike the standard model of career concerns, agents here have two abilities - in A and B. I call such career concerns "professionalism". Consistent with the usual model of 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 μ_i, and variance σ_0^2. In the usual career concerns model, the prior determines the supply of the available agents. Here the supply side is given by a frontier of available agents - the firm can select agents with different perceived ability subject to a constraint that

$$\mu_A + \mu_B = M \tag{3}$$

Agents vary in their perceived ability according to (3), and the firm can costlessly choose an agent anywhere along this frontier. Conditional on an initial choice of expected abilities, true abilities are uncorrelated with each other. This is the second choice variable of the firm, namely, *who should be hired?* - should the firm hire someone perceived to be better at A or at B? This choice, and how it varies with the ability to contract, is the central issue of this paper.

The Labor Market That effort is exerted to affect external perceptions is well known from the career concerns literature, such as as Holmstrom, 1999, and Gibbons and Murphy, 1992. The innovation here is that the external constituency may not share the objectives of the principal.[11]

[11] 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.

Firms in the labor market bid for the services of the agent in period two. Firms vary in how they use skills: specifically, let each firm be indexed by τ, where a firm of type τ values ability A at τm_A and ability B at $(1-\tau)m_B$. Technically, each firm receives a τ draw at the end of period 1, where there is a continuum of firms at each τ. τ has a natural support of 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 so that the agent's wage is her expected productivity in the most efficient match. Finally, I assume for simplicity that there are no contracts in period 2 - this means that no effort is exerted in period two, which does no more than cut down on notation.

Worker Preferences The worker has exponential preferences over wages and effort, with constant absolute rate of risk aversion, r. The agent's expected utility from earning a total wage of W over the two periods and incurring effort costs of E is given by

$$-Eexp(-r(W-E)). \tag{4}$$

The timing of the game is as follows. First, the firm chooses an agent of type (μ_A, μ_B) and makes a contract offer 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 are the I_i. 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, works there and is paid. 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 making an offer to another worker.

As is common in these settings following Holmstrom and Milgrom, 1991, the firm will offer a contract to the agents that is linear in the signal:

$$w = \beta_0 + \beta I. \tag{5}$$

There are two central questions addressed in this paper - (i) what determines the choice of worker hired?, and (ii) how does that choice vary with the ability to contract upon performance? To understand the logic behind the results that follow, it is useful to begin with describing the intuition for why hiring affects efforts.

Intuition The central idea of the paper is that firms can affect effort exerted through who they hire. This occurs here through the channel of affecting the agent's next potential employer. Specifically, the assumption of efficient matching (and linearity of the production function) means that in period 2, the agent works either for a firm that (i) only values their ability at A or (ii) only values their ability at B. Then - from a purely parametric perspective for the time being - let p be the probability that the next employer values A, and $1-p$ be the probability of B being valued.

If the next employer is more likely to value, say, A, then the agent will primarily work on those things that make her look better at A. More precisely, let $s_i = \frac{\sigma_0^2}{\sigma_0^2+\sigma_{\epsilon_i}}, i = A, B$ be the signal to noise ratio for updating ability on ability i from observing y_i. Then as s_i reflects the marginal effect of output on the (out-of-equilibrium) perception of ability, equilibrium effort choices with a linear contract with sensitivity β will be given by

$$e_1 = ps_A + \beta, \tag{6}$$

and

$$e_2 = pxs_A + (1-p)(1-x)s_B + \beta. \tag{7}$$

To see the relevant tradeoffs that arise below, note that

$$\frac{de_1}{dp} = s_A, \tag{8}$$

and

$$\frac{de_2}{dp} = xs_A - (1-x)s_B. \tag{9}$$

Equations (8) and (9) provide the foundation for the results that follow. Consider the case where $s_A = s_B$. Then raising p increases effort on task 1 but decreases it on task 2. The key for the result below, however, is that increasing p increases the sum of the two efforts: $\frac{d[e_1+e_2]}{dp} = 2x > 0$, but only because $x > 0$, i.e., because tasks are specialized. However, while total effort rises with p, it does so in an increasingly unbalanced way if $p > \frac{1}{2}$, as the agent is spending too much effort on task 1 to the detriment of task 2. It is the tradeoff between increasing total effort, but in an unbalanced way, that generates the contractual interactions below.

Thus far, p has been treated as exogenous. The final building block of the model is that the firm can affect p through choosing μ - specifically, those who ex ante are perceived to be better at A are more likely to end up with A as the most valued skill next period, so by hiring someone better at A and worse at B, the agent focuses her efforts more at A. With this in mind, let $\mu = \mu_A - \mu_B$ reflect the agent's relative perceived ability at A over her ability at B. Lemma 1 show the endogenous relationship between p and μ.

Lemma 1 *Let $\Phi(.)$ be a normal distribution with mean 0 and variance $\frac{\sigma_0^2\sigma_A^2}{\sigma_0^2+\sigma_A^2} + \frac{\sigma_0^2\sigma_B^2}{\sigma_0^2+\sigma_B^2}$. Then $p = 1 - \Phi(-\mu)$, and equilibrium efforts are given by (6) and (7).*

Given this, it is useful to think of the firm choosing p rather than μ, as the actions of the firm affect efforts only through the probability of a future employer. I call the case of $p^* = \frac{1}{2}$ an unbiased agent. Now consider the incentives of the agent in period one, including marginal incentive payments of β.

Proposition 1 *The optimal choice of agent and monetary incentives are given by*

$$p^* = min\{1, \frac{[s_A(1+x) - s_B(1-x)](1-\beta^*) - [s_A x - s_B(1-x)](1-x)s_B}{s_A^2 + [s_A x - s_B(1-x)]^2}\}, \quad (10)$$

and

$$\beta^* = \frac{2 - p^*(1+x)s_A - (1-p^*)s_B(1-x)}{2(1+r\sigma_1^2)}, \quad (11)$$

In general, the outcome depends on (i) the ability to contract on each input (σ_1 and σ_2), (ii) the career concerns incentives (s_A and s_B) and (iii) the degree of specialization (x). However, in order to understand the result below, it is instructive to consider the first order condition for the choice of agent.

1.1 Benchmark Case

It is useful to begin with the benchmark case where career incentives are equal for the two activities: $s_A = s_B = s$. Then the optimal choice of hiring and contracts is given by $p^* = \frac{2x(1-\beta^*)-(2x-1)(1-x)s}{s[1+(2x-1)^2]}$ and $\beta^* = \frac{1-px-\frac{1-x}{2}}{1+r\sigma^2}$. Proposition 2 immediately follows:

Proposition 2 *If $s_A = s_B$, then the optimal choice of agent is given by*

- $p^* > \frac{1}{2}$ *for all* $x > 0$ *and* $\sigma^2 > 0$
- p^* *is increasing in* x *and in* σ^2, *and*
- $p^* = \frac{1}{2}$ *only if* $\sigma^2 = 0$ *or* $x = 0$.

In words, this proposition ties the hiring of biased agents to two issues - (i) an inability to contract on performance, and (ii) the specialization of tasks. In the absence of these issues, the firm would hire an agent with relative preferences similar to its own. Call this the *unbiasedness* result.[12] However, for any positive σ^2, monetary contracts fall, and in response,

[12]It 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. Note also that when $s_A \neq s_B$, the firm hires workers not whose objective align the limit but rather whose relative efforts align with the principal's.

$\mu^* > 0$, so that the agent is biased. Furthermore, increases in σ^2 (weakly) increase μ^*: hence, bias arises as a response to the inability to contract. Furthermore, bias is increasing both in the ability to monitor, and the extent of task specialization. Call this the *substitutes* result, as the firm responds to an inability to contract on output by substituting towards another way of inducing effort on the primary activity, A.

The logic for this simple result rests on the incentives in (6) and (7). When more biased agents are hired, total effort rises (which is good) but does so in an unbalanced way (which is bad). When contracting is good (σ^2 low) the firm does not need to incur this imbalance to induce effort exertion and so hires an agent closer to his own relative preferences and appends to this stronger monetary incentives. By contrast, when the firm cannot contract well on output, the only way it can induce effort exertion is by such imbalance, and so hires more biased agents.

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

Three Observations

1. It is an obvious outcome of the standard agency model that when contracting distortions are great, agents have little incentive to exert effort. One implication of this model, however, is that under such circumstances, agents' *relative* incentives are also distorted such that they disproportionately are more interested in one activity over the other, but are distorted as the firm chooses to choose hire biased agents. Specifically,

$$\frac{d(\frac{e_1}{e_2})}{d\sigma^2} = \frac{d\beta}{d\sigma^2}\{-\frac{e_1}{e_2}[1 - 2s^2x(2x-1) + [1 - 2s^2x]\} > 0, \tag{12}$$

so that those institutions that contract less well on output have workers increasingly focused on one activity.

2. I have focused on the incentives of an agent whose job is specialized towards activity A. But it is trivial to imagine another agent whose 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 is

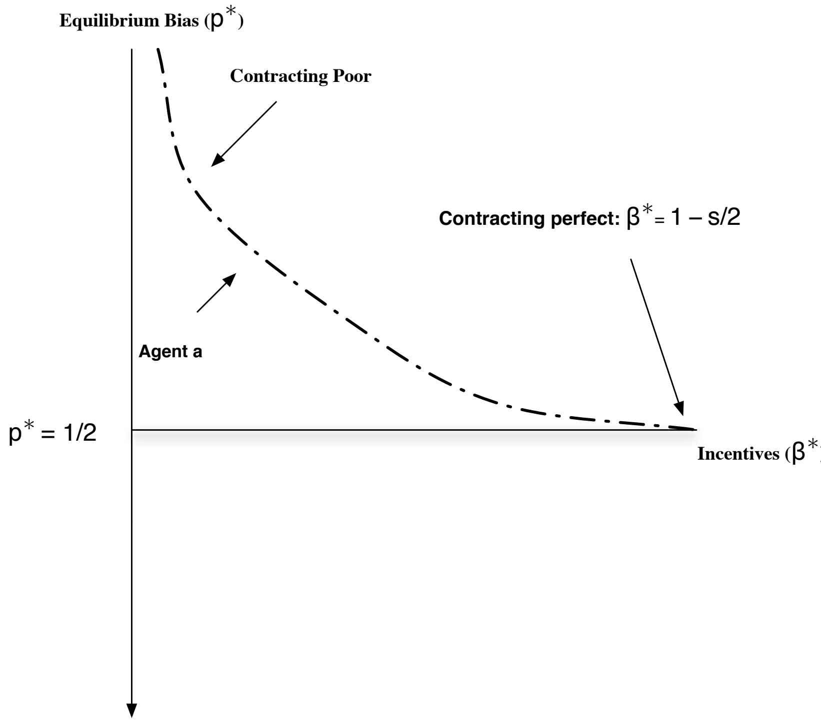

Figure 1: **Equilibrium Bias as σ changes.**

$-\mu^*$. Hence, as contracting becomes poorer, the firm hires workers with very divergent preferences.[13]

3. For what follows below, it is useful to consider the first order conditions for this problem. Let $e_{0A} = p^* s_A$ and $e_{0b} = p^* x s_A + (1 - p^*)(1 - x) s_B$ be the efforts exerted *in the absence of any monetary contracts.* Then the first order conditions for hiring is given by

$$\beta^* = \frac{1 - \frac{e_{0A} + e_{0B}}{2}}{1 + r\sigma_1^2}, \tag{13}$$

and

$$s_A(1 - e_{0A} - \beta^*) = s_B(1 - 2x)(1 - e_{0B} - \beta^*), \tag{14}$$

These have a intuitive and informative interpretation. First consider the contract β. It is common in models without any other reasons for effort for pay for performance to be

[13]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”.

given by $\beta = \frac{1}{1+r\sigma_1^2}$. The only addition here is that agents exert effort on the activities even without contracts and $1 - \frac{e_{0A}+e_{0B}}{2}$ is the residual demand for effort and so the same rule applies to this residual demand as in the standard model. Second, the s_i terms reflect how much efforts change by changing p: $\frac{de_1}{dp} = s_A$ and $\frac{de_2}{dp} = xs_A - (1-x)s_B$. But the firm also cares about the value of increasing either output. In the absence of any monetary contracts, this would be given by $1 - e_{0i}$. However, the fact that output can be contracted on reduces these marginal benefits, as the worker is already exerting effort for contractual reasons; these are the β terms. It is through this first order condition that the relevant parameters effect contracts and hiring.

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

2 Is Hiring Really a Substitiute for Monetary Incentives?

There are two central messages so far in the paper. First, firms can compensate for an inability to contract on output by strategic hiring, where hiring biased agents at least mitigates some of the problems associated with poor contracting. A natural implication of this is that firms that can contract well on output need not hire biased agents and instead hire agents whose relative preferences reflect theirs. Second, as contracting gets worse, agents become more biased, where bias is a substitute for an inability to contract. The purpose of this section is to analyze the extent to which both of these implications are, in fact, true in reasonable settings outside the basline model. To do so, I extend the model in two directions - (i) by addressing other forms of actions by the agent that affect her non-primary activity, B, which I will call interaction, and (ii) other contracting assumptions. I deal with each in turn.

2.1 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 agent and the firm grows at a rate $\frac{dp*}{d\sigma^2} \geq 0$. The purpose of this section is to show that with other forms of interaction between ouputs, 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, and second, by addressing the possibility of continuous interaction. Throughout this section, I consider the case where $s_A = s_B = s$.

2.1.1 Discrete Interation

Assume now that there is an unobserved discrete activity - "cooperation" - that the agent can engage in that benefits the non-primary output, B, at some cost to A. This activity increases the output of B by π_B at a cost of reducing the output of A by κ_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 I, 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

$$\beta(\pi_B - \kappa_A) + s(1-p)\pi_B - sp\kappa_A \geq 0, \tag{15}$$

which defines the critical value of p, called $\underline{p}$, above which the agent refuses to cooperate:

$$\underline{p} = \frac{\frac{\beta(\pi_B - \kappa_A)}{s} + \pi_B}{\pi_B + \kappa_A}. \tag{16}$$

There are two relevant issues that arise from (16). First, there is a limit to how biased the agent can be if the principal wishes to induce cooperation. Second, this limit depends on β: the better is contracting, the less need is there to distort hiring to induce efficiency. This second insight yields Proposition 3 below.

Proposition 3 *The relationship between the ability to contract and agent bias is non-monotonic in σ^2 the following way:*

- *If $\sigma^2 < \sigma_1^2$, then the principal chooses $\tilde{p}$ and $\tilde{\beta}$ as in (10) and (11), and the agent cooperates. In this region, $\frac{d\tilde{p}}{d\sigma} > 0$ and $\frac{d\tilde{\beta}}{d\sigma} < 0$*

- *If $\sigma_1^2 < \sigma^2 < \sigma_2^2$, then (16) binds, $\tilde{\beta} > \beta^*$, $\tilde{p} < p^*$, and the agent cooperates. In this region, $\frac{d\tilde{\beta}}{d\sigma} < 0$ but $\frac{d\tilde{p}}{d\sigma} < 0$*

- *If $\sigma^2 > \sigma_2^2$, then the principal chooses $\tilde{p}$ and $\tilde{\beta}$ as in (10) and (11), and the agent does not cooperate. In this region, $\frac{d\tilde{p}}{d\sigma} > 0$ and $\frac{d\tilde{\beta}}{d\sigma} < 0$.*

- *Let $S(\beta(\sigma^2), p(\sigma^2))$ define equilibrium surplus. Then σ_2^2 is finite if and only if $S(0, p_0) - S(0, p_1) \geq \pi_B - \kappa_A$, where $p_0 = min\{1, \frac{2x-(2x-1)(1-x)s}{s[1+(2x-1)^2]}\}$ and $p_1 = \frac{\pi_B}{\pi_B + \kappa_A}$.*

This proposition is easily explained and has an economically plausible interpretation. When contracting is good (σ^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 $\underline{p} = p^*(\beta^*)$ or

$$\underline{p}^*(\beta^*) = \frac{\frac{\beta^*(\pi_B - \kappa_A)}{s} + \pi_B}{\pi_B + \kappa_A}. \tag{17}$$

This is uniquely defined so let σ_1^2 be the level of difficulty of contracting at which (16) binds. At this point, further movements up the (p^*, β^*) 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 σ_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 p^* and β^* are second order. Hence there is some range over which the firm will induce the agent to cooperate by satisfying (16). It follows that the firm will choose to have (16) bind. When the cooperation constraint binds, note that

$$\frac{dp}{d\beta} = \frac{\pi_B - \kappa_A}{\pi_B + \kappa_A} < 0. \tag{18}$$

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 $S(\beta, p)$, where these depend on efforts as in (10) and (11). Then, $S(\beta^*(\sigma^2), p^*(\sigma^2))$ is the surplus when the cooperation issue is ignored, and $S(\tilde{\beta}(\sigma^2), \tilde{p}(\sigma^2))$ is the surplus produced from effort if choices are distorted to ensure cooperation occurs. Then if it exists, define σ_2^2 by

$$S(\beta^*(\sigma_2^2), p^*(\sigma_2^2)) = S(\tilde{\beta}(\sigma_2^2), \tilde{p}(\sigma_2^2)) + \pi_B - \kappa_A. \tag{19}$$

At σ_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 σ_2^2, and the firm discretely shifts by (i) reducing monetary incentives, and (ii) hiring agents with very biased preferences. Proposition 3 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.

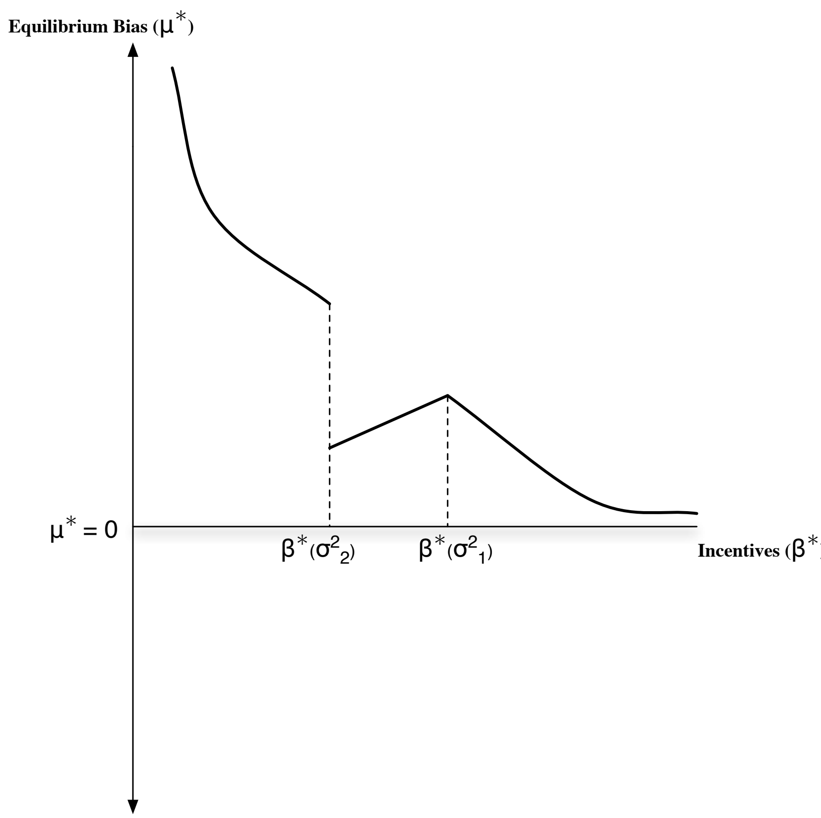

Figure 2: **Equilibrium Bias With Discrete Interation.**

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 (σ_2^2) the firm discretely switches back to the equilibrium of the last section though it involved no cooperation.

2.1.2 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, the agent now has 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, she 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 $\frac{kl^2}{2}$. These effects on output are observed in the contracted output I. Given the separability of costs, lobbying activities are chosen to maximize:

$$ps\lambda_A l - (1-p)s\lambda_B l - \frac{kl^2}{2} + \beta(\lambda_A - \lambda_B)l, \tag{20}$$

yielding the first order condition

$$kl^*(\beta, \mu_A) \geq s\lambda_A - (1-p)s(\lambda_A + \lambda_B) + \beta(\lambda_A - \lambda_B). \tag{21}$$

where (21) binds if the right hand side is positive and l^* is zero otherwise.[14]

Proposition 4 *Define σ_3^2 implicity by $p^*(\sigma_3^2) = \frac{\lambda_B + \beta^*(\sigma_3^2)(\frac{\lambda_A - \lambda_B}{s})}{\lambda_A + \lambda_B}$, where p^* and β^* are defined in (10) and (11). The relationship between the ability to contract and agent bias varies with σ^2 in the following way:*

- *If $\sigma^2 \leq \sigma_3^2$, then the principal chooses μ and β as in (10) and (11), and there is no lobbying.*

[14]This condition is intuitive - at the first best level described above, when $p = \frac{1}{2}$ and $\beta = 1 - \frac{s}{2}$, 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.

- *If $\sigma^2 > \sigma_3^2$, then agents become more (less) biased as σ^2 increases if and only if $2x > (\leq) \frac{\lambda_B^2 - \lambda_A^2}{k}$.*

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 (σ_3^2), agents begin to lobby under the old contract. When the lobbying constraint binds, it is shown in the Appendix that the optimal choice, p^{***} , is given by

$$p^{***} = min\{1, \frac{\gamma - (2x - \frac{\lambda_B^2 - \lambda_A^2}{k})\beta^{***}}{s(1 - (2x-1)^2 - 2s(\lambda_A + \lambda_B)^2)}\} \tag{22}$$

for some constant γ. 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 $2x > \frac{\lambda_B^2 - \lambda_A^2}{k}$, then as contracting becomes poorer, agents become more biased, while if $2x \leq \frac{\lambda_B^2 - \lambda_A^2}{k}$, 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) $\frac{\lambda_B^2 - \lambda_A^2}{k}$ 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 an agent whose type is no more biased than

$$p^{***} = \frac{\lambda_B}{\lambda_A + \lambda_B} > \frac{1}{2}, \tag{23}$$

while agent b has bias $1 - p^{***}$. 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 the agents is indifferent - she cares little more about their own task than the others.

The outcome here is described in Figure 3. Here once the lobbying constraint binds, at σ_3^2, 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

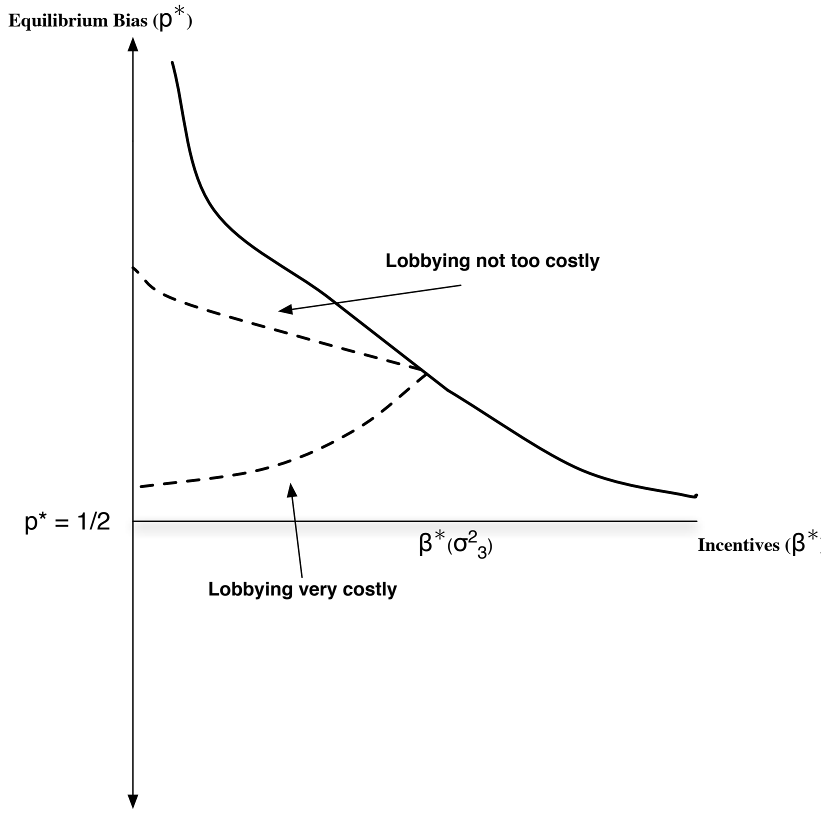

Figure 3: **Equilibrium Bias With Continuous Interaction.**

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.

2.2 Risk and Measurement Issues: Contracting on Individual Outputs

In the benchmark model above where the firm can only contract on aggregate output, two results arose. First, pay-for-performance and bias are substitutes, in the sense that as if aggregate output cannot be contracted upon well, the firm responds by hiring workers who are more biased towards their primary task. Second, as contracting gets better and better, the agent chosen converges to the relative preferences of the principal, and exert equal efforts for non-contractual reasons. Both results seem intuitive. The purpose of this section is to show that they are, though, not robust to other (reasonable) contracting technologies. To see this, two extensions are considered.

Both allow the possibility of separately contracting on each efforts. Assume now that, instead of contracting on the aggregate given by I, the firm can contract on each input separately via:

$$I_i = e_i + \epsilon_i, i = 1, 2, \tag{24}$$

where ϵ_i is Normally distributed with mean 0 and variance $\sigma_i^2, i = 1, 2$. For simplicity, it is assumed that the error terms are uncorrelated with each other and the error terms in (1). The firm will offer a contract to the agents that is linear in the signals:

$$w = \beta_0 + \beta_1 I_1 + \beta_2 I_2. \tag{25}$$

Equilibrium effort choices will now be given by $e_1 = ps_A + \beta_1$ and $e_2 = xps_A + (1-x)(1-p)s_B + \beta_2$.

Proposition 5 *Let* $\rho_i = \frac{r\sigma_i^2}{1+r\sigma_i^2}$ *and* $R = \frac{\rho_1}{\rho_2}$. *When the firm can separately contract on* I_1 *and* I_2*, the optimal choice of agent and monetary incentives are given by*

$$p^* = \frac{[s_A(R+x) - s_B(1-x)] - [s_A x - s_B(1-x)](1-x)s_B}{Rs_A^2 + [s_A x - s_B(1-x)]^2}, \tag{26}$$

and

$$\beta_1^* = \frac{1 - p^* s_A}{1 + r\sigma_1^2}, \tag{27}$$

and

$$\beta_2^* = \frac{1 - p^* s_A x - s_B(1-x)(1-p^*)}{1 + r\sigma_2^2}. \tag{28}$$

Two central results arose in the previous section. First, the no-biasedness result says that when contracting becomes more efficient, biasedness disappears in the limit, and the firm chooses $p^* = \frac{1}{2}$. Second, as contracting becomes less efficient, agents become more biased.

Non-biasedness The first of these results is no longer true here. First consider the case of common uncertainty: $\sigma_1^2 = \sigma_2^2$. Then

$$p^* = \frac{[s_A(1+x) - s_B(1-x)] - [s_A x - s_B(1-x)](1-x)s_B}{s_A^2 + [s_A x - s_B(1-x)]^2} \tag{29}$$

which is independent of monitoring noise. Hence as contracting improves here, the optimal choice of agent hired is unchanged. Furthermore, the optimal agent hired is the same agent as the benchmark model's choice of hiring *in the absence of any monetary contracts.* To say this another way, in the benchmark model, as contracting improves from no monetary incentives ($\sigma^2 = \infty$) to perfect incentives ($\sigma^2 \to 0$), the optimal choice of agent becomes less biased until in the limit of perfect contracting they become unbiased. Here, by contrast, the optimal choice of agent is unchanged from this initial level.

Why is this? The reason relates to the nature of the residual demand curve for effort, and how that differs in the two cases. Remember that $e_{0A} = p^* s_A$ and $e_{0b} = p^* x s_A + (1 - p^*)(1-x)s_B$ are the efforts exerted *in the absence of any monetary contracts.* Then in this case the first order condition for hiring is given by

$$s_A \rho_A (1 - e_{0A}) = s_B(1-2x)\rho_B(1 - e_{0B}), \tag{30}$$

This has a intuitive and informative interpretation. As before, the s_i terms reflect how much efforts change by changing p: $\frac{de_1}{dpi} = s_A$ and $\frac{de_2}{dp} = s_B(1-2x)$. But the firm also cares about the value of increasing either output. In the absence of any monetary contracts, this would be given by $1 - e_{0i}$. However, the fact that output can be contracted on reduces these marginal benefits, as the worker is already exerting effort for contractual reasons. The extent to which they are reduced from $1 - e_{0i}$ is given by the proportion $\rho_i = \frac{r\sigma_i^2}{1+r\sigma_i^2}$. Not surprisingly, this ranges from 1 when contracting is useless ($\sigma_i^2 = \infty$) to 0 when contracting is perfect ($\sigma_i^2 = 0$). Hence the first order condition equates "value of extra effort on surplus *times* marginal change in effort" for each activity. It is through this first order condition that the relevant parameters effect contracts and hiring. In the previous section, with the baseline model, these residual demand curves were given by $1 - e_{0A} - \beta^*$ and $1 - e_{0B} - \beta^*$, i.e, they were additive and not proportional.

The economic logic behind this insight is relatively simple - when only aggregate output can be contracted upon, changing the intensity of incentives has an equal effect on both efforts. But if each activity can be contracted upon separately, this is not optimal, even if each has the same ability to be monitored. Said another way, think of the case where only aggregate output is contracted on. In the limit when contracting is very efficient, the outcome will involve first best efforts of 1 for each activity. But if this is the case, it has to be the case that career incentives offer equal incentives on each output, which results in unbiasedness if $s_A = s_B$. But this is not true at the optimum when outputs can be separately contracted on and the *non-biasedness* result relies on contracting being at least partially on aggregate output.

The Effect of Monitoring Noise on Biasedness In the benchmark model, the harder it was to monitor efforts, the more biased the agent became. This result also does not generalize to other contracting technologies. Note that in the benchmark model, there was only one parameter, σ^2, to vary. Here there are more options - varying one of the two σ_i, or varying both.

First consider the case where only one variance changes, as it is most straightforward. To avoid some notation here, consider the case where $x = 0$, in which case the first order condition simplifies to

$$s_A \rho_A (1 - s_A p^*) = s_B \rho_B (1 - s_B (1 - p^*)), \tag{31}$$

and so

$$p^* = \frac{\rho_A s_A - \rho_B s_B (1 - s_B)}{\rho_A s_A^2 + \rho_B s_B^2}. \tag{32}$$

The natural intuition here is that as one input becomes harder to measure, hiring is changed in order to favor the harder to measure output. To see this is so, consider the case the effect of increasing, say σ_1^2. Then

$$\frac{dp^*}{d\sigma_1^2} = \frac{\rho_A}{\sigma_A^2} \frac{s_A(1 - s_A p^*)}{(\rho_A s_A^2 + \rho_B s_B^2)} > 0. \tag{33}$$

This seems a natural extension of the previous section.

However, note that the previous section was really about changing the overall contracting environment rather than making one action more contractible relative to another. Now consider the case where the ability to monitor *all* activities changes, which is closer in spirit to the exercise in the benchmark model. This involves changing both σ_i in some way.

First consider the simplest version of this, where $\sigma_1^2 = \sigma_2^2$. Then $\rho_1 = \rho_2$ and from (29) we know that optimal hiring is independent of monitoring, and so the substitutes result does not generalize to this case. To show how the result can be inverted, now consider a case where one task is more accurately measured than the other by a proportion γ, say where $\gamma\sigma_1^2 = \sigma_2^2$, where $\gamma < 1$, so activity B is easier to monitor. To isolate the effect of monitoring, consider the case where $x = 0$ and $s_A = s_B$. Then the first order condition is given by

$$\frac{\frac{1}{\gamma} + r\sigma_1^2}{1 + r\sigma_1^2}(1 - e_{0A}) = (1 - e_{0B}). \tag{34}$$

Let $z = \frac{\frac{1}{\gamma}+r\sigma_1^2}{1+r\sigma_1^2}$. Then increasing the ability to measure all activities - σ_1^2 only has effects through changing z. But z is increasing in σ_1^2 so $p*$ decreases in uncertainty. In words, *less* biased agents are hired as monitoring gets worse. In words, the firm favors the less well monitored activity more when monitoring is good than when it is poor. In the case where pay for performance cannot be used at all, $p^* = \frac{1}{2}$. This is the inverse of the previous section, and again points to the non-robustness of the intuitive substitutes case.

So what is the general point here? This is easy to see from the first order condition: $z = \frac{s_A(1-s_A p^*)}{(1-2x)s_B(1-s_B(1-p^*))}$. In general the ability to contract reduces the marginal importance of who is hired, as there is another way of inducing effort exertion. However, what matters here is the relative value of the residual demand for effort between the two activities. If contracting tilts the outcome in favor of B (as happens in the aggregate contracting case), less biased agents are hired, whereas in this last example, it tilts in favor of A, so more bias arises with better contracting.

2.3 Correlation between Monitoring and Career Incentives

So far, I have assumed that "noise" only affects the ability to write contracts on observed output. But remember that there is another source of noise in monitoring here - namely, the noisiness of the mapping between observed outputs and perceived ability. A reasonable extension of the model would seem to allow these to be related to one another, so that when contractible measures of inputs are noisy, perhaps also is the ability to infer ability from other measures of output.

To deal with this possibility, assume that in more uncertain settings, it is both harder to contract on output, and harder to update ability. Specifically, the following relationship operates between these: $\sigma_A^2 = \Lambda(\sigma_1^2)$, where Λ is increasing in σ_1^2. Hence more uncertain environments make it harder to provide incentives on both dimensions. Importantly, these act in opposite directions in the choice of the principal. To see this, note again the first order condition

$$s_A \rho_A (1 - e_{0A}) = s_B (1 - 2x) \rho_B (1 - e_{0B}). \tag{35}$$

Suppose we allow just the noisiness of A to increase. Previously this caused more biased agents to be hired - see (33). However, note that the weight placed on increasing surplus on A is given by $s_A \rho_A$, where s_A is the "bang for the buck" by hiring a more biased agent. Note that

$$\frac{d(s_A \rho_A)}{d\sigma_1^2} = \frac{r\sigma_0^2[\sigma_0^2 + \Lambda(.)] - \Lambda'(.)(r\sigma_A^2 + \sigma_A^4)}{1 + r\sigma_A^2)(\sigma_0^2 + \Lambda(.))}, \tag{36}$$

which cannot in general be signed. In words, making the environment more noisy now can either increase or decrease the desire to hire more biased agents. In the case where

$\sigma_A^2 = \lambda\sigma_1^2 + \eta$, this simplifies to

$$\frac{d(s_A \rho_A)}{d\sigma_1^2} = \frac{r\sigma_0^2[\sigma_0^2 + z - \lambda r \sigma_A^4]}{(1 + r\sigma_A^2)(\sigma_0^2 + \Lambda(.))}, \tag{37}$$

which is unambiguously negative for large enough λ. In other words, noisier settings reduce bias, unlike the previous section. Once again, the effect of monitoring is inverted relative to the benchmark above, for the simple reason that as contractual monitoring becomes more noisy, other incentives perhaps become even noisier, and so incentives rise.

3 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, p^* (in (10)). 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.[15] The firm then decides whether to spend these resources and attract a desired type, or else randomly select an agent. Let $S(\beta^*(\sigma^2), p^*(\sigma^2))$ denotes the surplus obtained by the principal in the benchmark model. Then by recruiting that desired type, the firm gains utility of $S(\beta^*, p^*) - k$. Alternatively, they can not incur this cost and randomly hire. This has three effects. First, on average they hire type $p = \frac{1}{2}$. Second, as they hire this type on average, they offer a contract of $\beta = \frac{1-\frac{s}{2}}{1+\sigma^2}$, as on average career incentives are $\frac{s}{2}$ - see (11). Third, there is variation in the motivation of workers hired - some have types greater than $p = \frac{1}{2}$ while others have less. The convexity of the cost function means that this variation is costly to the firm. Let σ_μ^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 $S(\frac{1-\frac{s}{2}}{1+\sigma^2}, \frac{1}{2}) - (1 + (2x-1)^2)\sigma_\mu^2$. The firm then uses targeted hiring only if

$$S(\beta^*, p^*) - K \geq S(\frac{1 - \frac{s}{2}}{1 + \sigma^2}, \frac{s}{2}) - (1 + (2x - 1)^2)\sigma_\mu^2. \tag{38}$$

$S(\beta^*, \mu^*) - S(\frac{1-\frac{s}{2}}{1+\sigma^2}, \frac{s}{2})$ is increasing in the inability to contract on output, σ^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

[15] I continue to assume that the firm maximizes surplus here.

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 σ^2 space, as seen in Proposition 6.

Proposition 6 *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 $(\beta^*(\sigma^2), p^*(\sigma^2))$ for all $\sigma^2 \geq \sigma^{2**}$, where σ^{2**} is finite if $S(0, p^*(\infty)) - k > S(0, 0) - (1 + (2x - 1)^2)\sigma_\mu^2$.*

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.

4 Robustness

The model so far has 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.

4.1 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

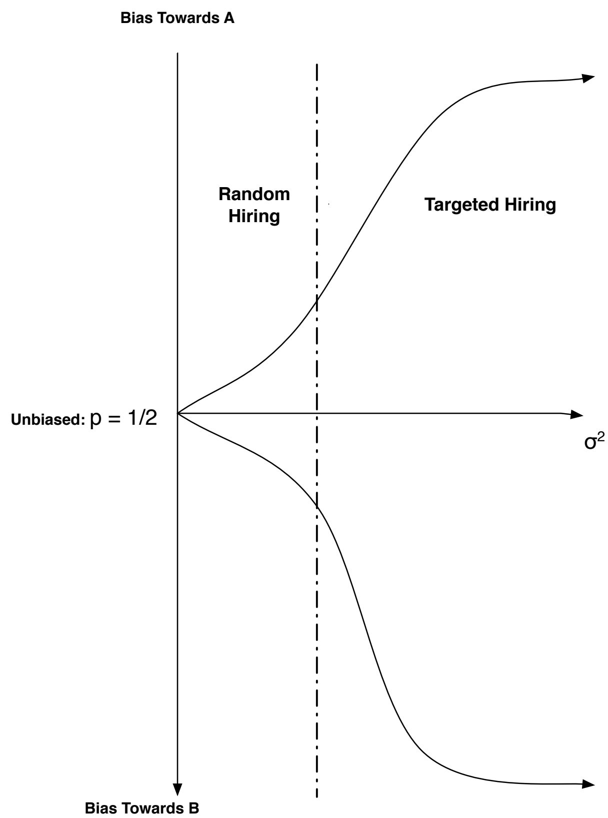

Figure 4: **Costly State Verification**

seems reasonable. Once again, the results remain unchanged when this extension is added, as shown in the Appendix.

4.2 Supply Frontier

So far, I have assumed perfect substitution between μ_A and μ_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 baseline model when $s_1 = s_2$, total effort increases in p: $\frac{d[e_1+e_2)]}{d\mu} = \phi(-\mu)2x > 0$, where ϕ is the density function for Φ. With non-perfect substitution, this is now conflated with a direct effect on expected output (holding effort constant) given by $1 - f'(\mu_A)$, because remember that output is directly affected by ability in (1).

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. The optimal choice of bias if the first order conditions continue to characterize equilibrium is now given by

$$p^* = min\{1, \frac{\frac{1-f'}{\phi(-\mu)} + [s_A(1+x) - s_B(1-x)](1-\beta^*) - [s_A x - s_B(1-x)](1-x)s_B}{s_A^2 + [s_A x - s_B(1-x)]^2}\}, \quad (39)$$

The comparative statics remain unchanged compared to the baseline model. The only outcome that changes is that in the absence of any incentives provided through career concerns, the firm will strictly prefer to hire the agent characterized by $1 = f'(\mu_A)$. The agent hired as the contract become perfectly efficient ($\sigma^2 = 0$) also does not converge to unbiased, but rather in between that point and where $1 = f'(\mu_A)$.

4.3 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 (1), 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 (1) but rather

$$y_A = \tilde{e}_A + (1+x)(\frac{m_A}{2} + \epsilon_A), \quad (40)$$

and

$$y_B = \tilde{e}_B + (1-x)(\frac{m_B}{2} + \epsilon_B). \quad (41)$$

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.[16] The import of this is extension

[16]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.

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 (11) but the optimal choice of agent is now given by

$$p^* = min\{1, \frac{2x(1-\beta^*) + \frac{2x}{\phi(-\frac{\mu^*}{s})s} - s(1-x)(2x-1)}{s(1+(2x-1)^2)}\}. \tag{42}$$

This has only one change from the basic model's outcome - the $\frac{2x}{\phi(-\frac{\mu^*}{s})s}$ 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.

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 τ. This subsection shows that the results above generalize to cases where firms value abilities differently. Specifically, this section maps into the τ interpretation as the model here has solved the case where $\tau = x$. In other words, τ 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.

4.4 Intrinsic Preferences

The model above offers one reason for exerting effort - the prospect of higher wages, either contemporaneously (though the contracted payments) or in the future (through career concerns). Yet a plausible alternative reason for exerting effort is intrinsic preference. A natural question is ask is whether the insights here continue to hold under these circumstance. To model this, assume that agents have an observed "type" (μ_A, μ_B) such that aside from the contractual payments offered by the firm, rather than have any career concerns, they simply value (y_A, y_B) at $\mu_A y_A + \mu_B y_B$, where $\mu_i \geq 0$.[17] Furthermore, assume that there is a distribution of (μ_A, μ_B) in the population of possible agents given by a frontier $\mu_A + \mu_B = M$,

[17]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

where $M < 2$. The agent has a reservation utility independent of type and normalized to 0. The timing of the game is as follows. First, the firm chooses an agent of type (μ_A, μ_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.

Under these conditions, which largely mirror those of the last section, the same two results arise in the benchmark model - namely, that (i) $p^* \geq \frac{1}{2}$, (ii) $p^* = \frac{1}{2}$ for $\sigma^2 = 0$, and (iii) p^* is increasing in x and σ^2. Details are in the Appendix. Hence the insights of the benchmark model easily extend over to the case of these intrinsic preferences.

5 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.[18]

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

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.

[18]See also MacLeod, 2003, for other work on the costs of conflict in settings with subjective performance measures.

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 other than pay as a way of aligning interests.

References

Baker, George, "Incentive Contracts and Performance Measurement", *Journal of Political Economy*, 1992, 100, 598-614.

Besley, Tim, and Maitreesh Ghatak, "Competition and Incentives with Motivated Agents," *American Economic Review*, 2005, 95(3), 616-636.

Brehm, John, and Scott Gates, *Working, Shirking, and Sabotage: Bureaucratic Responses to a Democratic Public*, 1997, University of Michigan Press, Ann Arbor.

Che, Yeon-Koo and Navin Karthik, "Opinions as Incentives," *Journal of Political Economy*, 2009, vol. 117, issue 5, pages 815-860.

Delfgaauw, Josse, and Robert Dur, "Signaling and Screening of Workers' Motivation," 2003, mimeo, Erasmus University.

Dessein, Wouter, Luis Garicano and Robert Gertner, "Organizing for Synergies: A Theory of Hybrids", mimeo, Chicago GSB, 2008.

Derthick, Martha, *Policymaking for Social Security*, Brookings Institute, Washington, 1979.

Dewatripont, Matthias, Ian Jewitt, and Jean Tirole, "Missions and the Accountability of Government Agencies," *Review of Economic Studies*, 1999, vol. 66, issue 1, pages 199-217.

Golden, Marissa, *What Motivates Bureaucrats?*, Columbia University Press, New York, 2000.

Goodsell, Charles, *The Case For Bureaucracy*, CQ Press, Washington, 2004.

Gibbons, Robert and Kevin M. Murphy, "Optimal Incentive Contracts in the Presence of Career Concerns", *Journal of Political Economy*, 1992, 100(3), 468-505 .

Holmstrom, Bengt, "Managerial Incentive Problems: A Dynamic Perspective", *Review of Economic Studies*, 1999.

Itoh, Hideshi, "Cooperation in Hierarchical Organizations: An Incentives Perspective", *Journal of Law, Economics, and Organizations*, 8(2), 321-45, 1992.

Milgrom, Paul, and John Roberts, "Employment Contracts, Influence Activities, and Efficient Organizational Design", *Journal of Political Economy*, 96(1), 42-60.

MacLeod, Bentley, "Optimal Contracting with Subjective Evaluation," *American Economic Review*, 2003, 93(1), 216-40.

Peabody, Robert, *Organizational Authority*, 1964, Atherton Press, New York.

Prendergast, Canice, "The Motivation and Bias of Bureaucrats," *American Economic Review*, 2007, 97(1), 180-96.

Pruitt, Charles, "People Doing What They Do Best: The Professional Engineers and NHTSA," *Public Administration Review*, 1979, 39(1), 363-71.

Rotemberg, Julio and Garth Saloner, "Overt Interfunctional Conflict," *RAND Journal of Economics*, 1995, 26(4), 630-53.

Van Den Steen, Eric, "Organizational Beliefs and Managerial Vision," *Journal of Law, Ecnomics and Organization*, 2005a, 21(1), 256-83.

Van Den Steen, Eric, "On the Origins of Shared Beliefs (and Corporate Culture)," mimeo, 2005b.

Van Den Steen, Eric, "Culture Clash: On the Costs and Benefits of Homogeneity", mimeo, 2004.

Wilson, James, *Bureaucracy*, Basic Books, New York, 1989.

Proof of Lemma 1 The market observes y_A and y_B at the end of the first period and updates its perception of the agent's abilities to $\hat{\mu}_A$ and $\hat{\mu}_B$ respectively.[19] As the agent matches efficiently, she will be employed in the firm where most surplus is created.The assumption of efficient matching makes this relationship between wages and perceptions very simple - the agent's utility in period 2 is $max\{\hat{\mu}_A, \hat{\mu}_B\}$. To see this, note that 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

$$(\tau^*(\hat{\mu}_A, \hat{\mu}_B))\hat{\mu}_A + (1 - \tau^*(\hat{\mu}_A, \hat{\mu}_B))\hat{\mu}_B. \tag{43}$$

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 $\tau^* = 0$.

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. This characterizes wages conditional on the realizations of first period output. When choosing effort, the agent must therefore take expectations of these future wages.

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. For the career incentives, 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 μ and variance $\frac{\sigma_0^2\sigma_A^2}{\sigma_0^2+\sigma_A^2} + \frac{\sigma_0^2\sigma_B^2}{\sigma_0^2+\sigma_B^2}$. Let $\Phi(.)$ be a normal distribution with mean 0 and variance $\frac{\sigma_0^2\sigma_A^2}{\sigma_0^2+\sigma_A^2} + \frac{\sigma_0^2\sigma_B^2}{\sigma_0^2+\sigma_B^2}$. 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 $\Phi(-\mu)$.

The agent's choice of effort Consider the incentives of the agent in period one, including marginal incentive payments of β. The reservation wage of the agent in period 2 is given by

$$prob(employer = A)w_A(\hat{\mu}_A) + prob(employer = B)w_B(\hat{\mu}_B), \tag{44}$$

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 $w_i(\hat{\mu}_i) = \hat{\mu}_i$. 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

[19] Note that I_i 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.

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 μ_i. However, this is not true out of equilibrium and instead the agent exerts effort to maximize $W_A(\mu_A) = \mu_A + s(e_1 - Ee_1) + xs(e_2 - Ee_2)$ and $W_B(\mu_B) = \mu_B + (1 - x)s(e_2 - Ee_2)$.

As a result, the agent then chooses her efforts (ignoring a constant)[20] to maximize

$$max_{\{e_1,e_2\}}\beta I + [1 - \Gamma]W_A + \Gamma W_B - \frac{e_1^2}{2} - \frac{e_2^2}{2}. \tag{45}$$

Straightforward maximization yields (6) and (7), as all the terms involving ϕ, 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 (6) and (7).

To see this, note that differentiating (45) with respect to e_1 yields

$$\beta+[1-\Phi(-\mu-(e_1-Ee_1)-(2x-1)(e_2-Ee_2))]s+\phi(-\mu-(e_1-Ee_1)-(2x-1)(e_2-Ee_2))0 = e_1, \tag{46}$$

which yields (6) in equilibrium. Differentiating with respect to e_2 similarly yields (7).

Proof of Proposition 1: The objective of the principal is to maximize output minus wages, which in the usual fashion results in maximizing expected surplus - $E[e_1 + e_2 - \frac{e_1^2}{2} - \frac{e_2^2}{2} - r\sigma^2\beta^2]$ - where $r\sigma^2\beta^2$ is the risk premium associated with the monetary contract. (The μ_i terms disappear due to the assumption of perfect substitution.) The objective of the principal is to choose $p = 1 - \Phi$, and β to maximize $E[e_1 + e_2 - \frac{e_1^2}{2} - \frac{e_2^2}{2} - r\sigma^2\beta^2]$ subject to (3), $0 \leq p \leq 1$, (6), and (7). By substitution, the principal chooses the agent's type (p) and the contract (β) to maximize

[20]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.

$$2\beta+s_Ap+xs_Ap+(1-x)s_B(1-p)-\frac{(\beta+s_Ap)^2}{2}-\frac{(\beta+xs_Ap+(1-x)s_B(1-p))^2}{2}-r\sigma^2\beta^2. \tag{47}$$

Straightforward differentiation yields (10) and (11).

Proof of Proposition 3: With the cooperation constraint, the objective of the firm is now to maximize expected surplus, subject to (3), (6), and (7), and now (15) if the firms wishes to induce cooperation. Begin by ignoring the cooperation constraint, in which case the firm's choice is given by (10) and (11). Therefore, if either (i) the firm can induce cooperation without changing from (10) and (11) or (ii) does not wish to induce cooperation, the solution remains that given by (10) and (11).

Note that for σ^2 low enough, (15) is satisfied at the equilibrium choices given by (10) and (11). To see this, note that as $\sigma^2 \to 0$, then $\beta^* \to 1-\frac{s}{2}$ and $\mu^* \to 0$ in which case (15) holds. As β^* and μ^* vary continuously with σ^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 σ^2 increases, then if there exists a point at which $\mu_A^* = \frac{\beta^*(\pi_B-\kappa_A)+\pi_B M}{\pi_B+\kappa_A}$, then the cooperation constraint is violated for all higher values of σ^2 as β^* and $-\mu^*$ are decreasing in σ^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 (10) and (11). It it does, then the firm will choose (15) 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 β is given by

$$\beta = \frac{2-2xg-(1-g)-x(1+g(1-2x))}{(1+\sigma^2)[(1-g)^2+(1+g(1-2x))^2]} \tag{48}$$

which is decreasing in σ^2, as required. Hence, if there exists a point where (15) binds, there is a range where $\tilde{\beta}$ and $\tilde{\mu}$ decline with σ^2.

Let the surplus generated by e_1, e_2 be defined by $S(\beta,\mu) = e_1(\beta,\mu)+e_2(\beta,\mu)-\frac{e_1(\beta,\mu)^2}{2}+\frac{e_2(\beta,\mu)^2}{2}-r\sigma^2\beta^2$ where e_1 and e_2 are defined in (6), and (7). Then, if it exists, define σ_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 σ_2, it must be the case that $S(\beta^*(\sigma),\mu^*(\sigma)) > S(\tilde{\beta}(\sigma),\tilde{\mu}_A(\sigma))+\pi_B-\kappa_A$ for all larger values of σ because for all $\beta > \beta^*$, $\frac{d^2S}{d\beta d\sigma^2} < 0$, and for all $\beta \leq \beta^*$, $\frac{d^2S}{d\beta d\sigma^2} = 0$. Therefore as $\tilde{\beta} > \beta^*$ and $\tilde{\mu} > \mu^*$, $S(\beta^*(\sigma),\mu^*(\sigma)) - S(\tilde{\beta}(\sigma),\tilde{\mu}(\sigma))$ is increasing in σ. As a result, for all $\sigma > \sigma_2$, the firm does not induce cooperation but instead chooses (10) and (11).

Of course, no such value of σ_2^2 may exist. Consider the limiting case as σ^2 tends to ∞. 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 σ^2 tends to ∞. As $\beta^* \to 0$, $\tilde{\mu} \to \mu_1$, where $\mu_1 = 2[\frac{\pi_B M}{\pi_B+\kappa_A} - \frac{M}{2}]$. A necessary and sufficient condition for fiefdom to exist is then

$$S(0,\mu_0) - S(0,\mu_1) \geq \pi_B - \kappa_A. \tag{49}$$

Proof of Proposition 4: The solution characterized in Section 1 continues to hold if the agents do not lobby. However, the lobbying constraint binds at $\mu^*(\sigma_1), \beta^*(\sigma_1)$ where

$$\mu^*(\sigma_1) = 2[\frac{M\lambda_B + \beta^*(\sigma_1)(\lambda_A - \lambda_B)}{\lambda_A + \lambda_B} - \frac{M}{2}]. \tag{50}$$

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

$$2\beta + (1+x)\mu_A + (1-x)(M-\mu_A) - \frac{((1+\delta)\beta + \mu_A)^2}{4} - \frac{((1-\delta)\beta + \mu_A)^2}{4} - \frac{((1+\delta)\beta + x\mu_A + (1-x)(M-\mu_A))^2}{4} - \frac{((1-\delta)\beta + x\mu_A + (1-x)(M-\mu_A))^2}{4} - (\lambda_B - \lambda_A)l - k\frac{l^2}{2} - r\sigma^2\beta^2, \tag{51}$$

subject to $\mu_A + \mu_B = M$, $0 \leq p \leq 1$, (6), and (7), and l characterized by (21) holding with equality. Maximizing this yields optimal level of incentives and preferences are given by

$$\mu_A^* = \frac{2x(1-\beta^*) - (1-x)(2x-1)M - (\lambda_A+\lambda_B)(kl^* + \lambda_B - \lambda_A)}{1+(2x-1)^2}, \tag{52}$$

As l^* is a function of β, 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

$$\mu^{***} = min\{M, \frac{2[\gamma - (2x - \frac{\lambda_B^2 - \lambda_A^2}{k})\beta^{***}]}{1+(2x-1)^2 - \frac{(\lambda_A+\lambda_B)^2}{k}} - \frac{M}{2}\} \tag{53}$$

where $\gamma = 2x - (1-x)(2x-1)M - (\lambda_A+\lambda_B)(1 - \frac{M\lambda_B}{k})$. As p in monotonically increasing in μ, Proposition 6 then follows.

Proof of Proposition 5: The firm now offers a contract where the wage - modulo a fixed payment - is given by

$$w = \beta_1\tilde{y}_1 + \beta_2\tilde{y}_2 + \beta\tilde{y}. \tag{54}$$

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.

The objective of the principal is then to choose p, β_1, and $\beta_2$to maximize

$$E[e_1 + e_2 - \frac{e_1^2}{2} - \frac{e_2^2}{2} - r\beta_1\sigma_1^2 - r\beta_2\sigma^2] \tag{55}$$

subject to $\mu_A + \mu_B = M$, $e_1 = ps_A + \beta_1$ and $e_2 = xps_A + (1-x)(1-p)s_B + \beta_2$, and $0 \leq p \leq 1$. The optimization program of the principal is now given by maximizing

$$\beta_1 + \beta_2 + s_A p + x s_A p + (1-x)s_B(1-p) - \frac{(\beta_1 + s_A p)^2}{2} - \frac{(\beta_2 + x s_A p + (1-x)s_B(1-p))^2}{2} - r\sigma_1^2\beta_1^2 - r\sigma_2^2\beta_2^2. \tag{56}$$

Straightforward differentiation yields (29), (27), and (28).

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 β and agent type μ is given by

$$\beta^* = \frac{2(1-\zeta^2)(1-\zeta) - (1-\zeta)^2 2x\mu_A^* - (1-\zeta^2)(1-x)M}{2(1-\zeta)^2(1+\sigma^2)}, \tag{57}$$

and

$$\mu_A^* = min\{M, \frac{(1-\zeta^2)(1-\zeta)2x - (1-\zeta)^2 2x\beta^* - (1+\zeta^2)(2x-1-2\zeta)(1-x)M}{[(1-\zeta(2x-1))^2 + (2x-1-\zeta)^2]}\}. \tag{58}$$

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 ($\mu^* = 0$ so $p^* = \frac{1}{2}$) and incentives given by $\beta^* = 1 - \frac{s}{2}$. Hence, the insights extend to the case where the cost functions are not independent in this way.

Intrinsic Preferences: First consider an agent a who has type (μ_A, μ_B) with marginal pay of β. She chooses efforts of

$$e_1 = \mu_A + \beta, \tag{59}$$

and

$$e_2(D) = x\mu_A + (1-x)\mu_B + \beta. \tag{60}$$

The objective of the principal is to choose μ_A, μ_B, and β to maximize

$$E[e_1 + e_2 - \frac{e_1^2}{2} - \frac{e_2^2}{2} - r\beta^2\sigma^2] \tag{61}$$

subject to $\mu_A + \mu_B = M$, (59), (60), and $M \geq \mu_i \geq 0$. By substitution, the principal chooses the agent's type (μ_A) and the contract β to maximize

$$2\beta + (1+x)\mu_A + (1-x)(M-\mu_A) - \frac{(\beta+\mu_A)^2}{2} - \frac{(\beta + x\mu_A + (1-x)(M-\mu_A))^2}{2} - r\beta^2\sigma^2. \quad (62)$$

Straightforward calculations yield the efficient level of incentives and preferences, μ^* and β^*, where

$$\mu^* = min\{M, \frac{4x(1-\beta^*) - 2(2x-1)(1-x)M}{1+(2x-1)^2} - \frac{M}{2}\} \geq 0, \quad (63)$$

and

$$\beta^* = \frac{1 - x(\frac{\mu^*+M}{2}) - (1-x)\frac{M}{2}}{1+\sigma^2} \leq 1 - \frac{M}{2}, \quad (64)$$

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

Consider the optimal choice of agent and contract. When contracting is perfect ($\sigma^2 = 0$), (63) and (64) imply that $\mu^* = 0$. In words, when contracting is perfect, the agent hired shares the preferences of the principal.