

WORK INCENTIVES, MOTIVATION, AND IDENTITY[†]

Intrinsic Motivation and Incentives

By CANICE PRENDERGAST*

Workers act in the interests of their employers for a host of reasons. Sometimes money is the motivation, but often it is because they care about what they do. Despite the importance of such intrinsic motivation, the economic literature has offered little in terms of understanding relevant trade-offs when alignment of inherent preferences (rather than monetary interests) is what motivates people.¹ Instead, the literature has focused largely on the efficiency gains that arise from agents sharing the preferences of their employers.² This paper offers such a theory of intrinsic motivation, where firms partially solve agency problems by hiring agents with particular preferences. Unlike the previous literature, I show that firms, in general, hire agents who do not share their interests—instead, agents are disproportionately motivated to carry out a subset of what the firm cares about. Specifically, I show that the institution hires the agent with similar preferences to himself only in the limiting case where the agent's preferences are irrelevant—namely, when output can be perfectly contracted on. Instead, the optimal response of the institution is to hire biased agents, with the degree of bias increasing as contracting measures get worse.

[†] *Discussant:* Robert Gibbons, Massachusetts Institute of Technology.

* Graduate School of Business, University of Chicago, 1101 East 58th Street, Chicago, IL 60637 (e-mail: canice.prendergast@gsb.uchicago.edu). Any errors are my own. Thanks to Bob Gibbons for very helpful comments.

¹ The idea that something other than money must motivate people seems clear, a point often made in studies of the public sector. For example, even the US Post Office—an institution with little link between employee performance and pay—has on-time delivery rates of mail in the region of 98 percent. See Charles Goodsell (1998) for a review of this literature.

² For the relevant literature, see Mathias Dewatripont, Ian Jewitt, and Jean Tirole (1999), Bruno Frey and Reto Jegen (2001), Roland Benabou and Jean Tirole (2003), George Akerlof and Rachel Kranton (2005), Tim Besley and Maitreesh Ghatak (2005) and Prendergast (2007).

The reason for this is rather simple—that jobs within firms are specialized. So, for example, some people make things, others design them, yet other sell them, and so on. Similarly, academics do research, deans raise money, administrators do the books, social workers provide services to clients while their superiors are charged with cost control, etc. The principal novelty of this paper is to show how hiring practices—and specifically the intrinsic motivation of those hired—changes as the ability to contract worsens in a world where tasks are specialized. In doing so, it paints a picture of firms responding to noncontractibility by hiring agents with very extreme interests, at the cost of their ignoring other aspects of their jobs. So, for example, if the output of a social worker cannot be measured, a natural response will be to hire agents who are highly motivated to get benefits to clients, at the cost of their ignoring things like cost control. Simple though this observation is, it illustrates a cost to using intrinsic motivation—namely, the endogenous hiring of those who increasingly care only about one aspect of their job. In this way, the simple model outlined here lays the foundation for a view of “poor contracting” firms as fiefdoms, where those with extreme competing interests reside, a point I elaborate on elsewhere (Prendergast 2008).

For a theory of intrinsic motivation to be useful, firms must have some control over it. The central assumption here is that firms have some ability to hire employees based on these intrinsic preferences. I build a model of an institution that carries out two activities—say, service provision and cost control. It employs two agents to do so. The tasks are somewhat specialized, in that one agent primarily (but not exclusively) does service provision and the other primarily does cost control. The institution has an objective that trades off these two components and has a performance measure that can reward agents. The firm also chooses whom to hire. Potential

hires have preferences over outcomes, and the firm can recruit based on these preferences. It is in the context of this setup that bias arises.

I. Model

Consider an institution that carries out two tasks, *A* and *B*. The institution employs two agents to carry out these tasks. The agents are partially specialized, in that one agent (agent *a*) primarily does activity *A*, while the other (agent *b*) primarily does activity *B*. Initially, consider the actions of agent *a* who is primarily carrying out activity *A*. Agent *a* provides efforts on two tasks—1 and 2—that generate benefits for the firm. The agent exerts effort e_1 and e_2 on these activities, each of which yields marginal benefits of 1. For the sake of concreteness, consider effort 1 as yielding services to clients, and effort 2 as doing so at lower cost. To keep matter simple, the costs of effort on task *i* is $e_i^2/2$. (There is no linkage between efforts in the cost function to avoid usual multitasking issues.) Assume that the preferences of the principal are to maximize the standard notion of surplus: $E[e_1 + e_2 - e_1^2/2 - e_2^2/2]$. (See Prendergast (2008) for a discussion of this.) As a result of these simplifying assumptions, the first-best level of effort is given by $e_i^* = 1$.

There are two difficulties in inducing efficient effort. The first is familiar—that effort cannot be directly contracted upon where available performance measures are imperfect in a way that the agent can take advantage of. Following George Baker (1992), I assume that the principal can observe an unbiased but imperfect signal of surplus:

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

where D takes on values δ and $-\delta$ with equal probability. The parameter δ thus measures the extent to which effort can be effectively contracted upon, and is privately observed by the agent after contracts are signed. At one extreme, $\delta = 0$ and performance measures are perfect, while at the other extreme $\delta = \infty$, they are useless. This abstract contracting technology is used simply to illustrate contracting distortions, while retaining the inherent symmetry of the problem.

The second issue that becomes relevant to incentive provision is that there is an asymmetry

between the two efforts. Specifically, while all effort by agent *a* on task 1 increases the returns solely of activity *A*, effort on task 2 has shared benefit. Specifically, a fraction x of e_2 benefits the activity in which the agent is engaged, while the remaining $(1 - x)$ benefits the other activity. The most natural interpretation of this is that activity 2 involves cost containment, which has benefits across the entire organization as cost savings are shared, such as where a benefit officer disqualifies a candidate for those benefits. For agent *b*, his output is the mirror image with all his e_1 increasing activity *B*, and $(1 - x)$ of his e_2 increasing output *A*.

So far, there is little conceptually novel about constraints on efficiency—it offers a specific parameterization of multitasking problems which would offer the obvious implication of lower-powered incentives when δ gets larger. The novelty concerns the role of intrinsic motivation, where agents care about outcomes. Specifically, I assume that agents have *observable* preferences over activities *A* and *B*, given by (μ_A, μ_B) , so that the welfare that they receive from outcomes (y_A, y_B) is $\mu_A y_A + \mu_B y_B$.³ As a result of this, the principal chooses not only what contracts to offer to agents, but also what type of agent to hire. Ignore the participation decision of the agent, as it is subsumed in the preferences of the principal.

To avoid the facile observation of “hire those who care more,” I assume that (a) no agent exists who internalizes all benefits without at least some monetary incentives, and (b) on the relevant frontier there is a trade-off between how much they care about the two activities. If that distribution includes the point (1,1), then the first best is attainable by simply choosing that individual and offering no payment based on output. I ignore this trivial solution by assuming that $\mu_A + \mu_B = M$, where $M < 2$ and $\mu_i \geq 0$.⁴ It is worth discussing this frontier briefly. First, it is negatively sloped with the point $M/2 < 1$. As a result, it introduces a role for monetary contracts, as the unbiased point $(M/2, M/2)$ does not yield efficient outcomes. (This point is unbiased

³ y_A is the sum of $e_1 + xe_2$ of agent *a* and $(1 - x)$ of e_2 by agent *b*, with y_B defined accordingly.

⁴ The assumption of μ_A and μ_B being perfect substitutes implies that there are no supply-side reasons for bias, as the trade-off between the supply of agents is identical to how the principal trades off the two activities.

as it trades off the two activities in the same way as the principal.) Second, the frontier is negatively sloped, so there is a potential trade-off that firms must make—should they hire those who care more about *A* or about *B*?

The firm can condition a wage paid to an agent on observed output \tilde{y} . (Ignore agent subscripts for simplicity.) For ease of exposition, consider these as linear contracts where the agent is given a fraction of output, $\beta\tilde{y}$, and a fixed payment.⁵

First, consider the behavior of agent *a* who has type (μ_A, μ_B) with marginal linear incentives of β . She chooses effort of $e_1 = \mu_A + (1 + D)\beta$, and $e_2 = x\mu_A + (1 - x)\mu_B + (1 + D)\beta$. Thus, effort is increasing in intrinsic motivation and in monetary incentives. The objective of the principal is then to choose μ_A, μ_B and β to maximize $E[e_1 + e_2 - e_1^2/2 - e_2^2/2]$ subject to e_i defined above, $\mu_A + \mu_B = M, \mu_i \geq 0$, and $D = \delta(-\delta)$ with probability 1/2. Straightforward calculations yield the efficient level of incentives and bureaucratic preferences as

$$(2) \quad \mu_A^* = \min \left\{ M, \frac{2x - (2x - 1)(1 - x)M - \beta^*}{1 - (2x - 1)^2} \right\},$$

and

$$(3) \quad \beta^* = \frac{1 - x\mu_A^* - (1 - x)\frac{M}{2}}{1 + \sigma^2},$$

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

First, consider the optimal contract offered to the agent, β^* . The efficient level of effort is $e_i^* = 1$. In the case where there are no distortions in the output measure, $\delta = 0$, the contract ensures that this level of effort arises. When $\delta = 0$, the firm chooses $\mu_1^* = M/2$ and $\beta^* = 1 - M/2$. In words, in firms where output measures are good, there is no need to bias the selection of the agent, and effort is at the first-best level, where incentives are used solely to “top

up” those incentives that derive from intrinsic motivation.

A second natural result is that, as σ^2 rises, β falls for the familiar reason that it mitigates distortionary effort decisions. Third, as intrinsic incentives become more aligned with those of the principal, monetary incentives fall. This can be seen in two ways. First, as M rises, the agents have intrinsic preferences closer to the first-best level, and so β^* falls with M . Second, the parameter x measures the extent to which activity *A* captures all the benefits of the agent’s efforts. It is the fact that x deviates from 1 that leads the firms not to choose the most biased agent. Hence, it follows that, as x rises, these external effects decrease, and so β^* falls (since $\mu_A^* \geq M/2$ —see below) as intrinsic incentives are more aligned with the wishes of the principal.

Now consider who is hired—the optimal bias of agents. Consider agent *a*. First, $\mu_A \geq M/2$, with the equality strict if $\delta > 0$ and $x > 0$, so that bias is optimal—this follows from the specialization of tasks. Second, as x tends to 0, the agent’s tasks are not specialized and hence it is efficient to choose an agent with representative preferences, $\mu_A = M/2$, while as x tends to 1, the agent’s efforts affect only activity *A* and so μ_A tends toward M . Of most importance, note that bias $(\mu^* - M/2)$ decreases in β as monetary and intrinsic incentives are substitutes. Hence, firms respond to poor contractibility of output (σ^2) by hiring agents with more extreme preferences.

Remember, however, that there is another agent, *b*, whose tasks are the mirror image of those of agent *a*. The optimal choice of agent *b* is simply the mirror image of μ_A^* : $\mu_B^* = M - \mu_A^*$. As a result, the view that this paper offers of agencies where output is hard to contract on is not the usual one where agents simply stop exerting effort, but instead the firm responds partially by hiring biased agents whose equilibrium biases move in opposite directions to each other. So, for instance, social work departments would be populated with social workers who are highly motivated to provide benefits to clients, yet who show little interest in controlling costs (see Martha Derthick (1979) and John Brehm and Scott Gates (1977), for evidence on this) and superiors who show excess interest in cost control. For instance, in a survey on the preferences of social workers, Robert Peabody (1964, 66) notes that “by far the most dominant

⁵ As there is no aggregate uncertainty about observed output in this model, this is without loss of generality. The interpretation of a linear contract is merely to illustrate the extent to which marginal pay depends on observed performance.

organizational goal perceived as important... is service to clientele,” where 83 percent of survey respondents view such service as important, compared to only 9 percent who see “obligation to taxpayers” or “assistance to the public in general” as important concerns affecting their decisions. By contrast, those institutions that can contract on output have interests aligned, not just because of monetary concerns, but also because their preferences are more similar.

II. Conclusion

There are many instances where the use of significant monetary incentives is likely to backfire. Yet firms still seek some way to motivate their employees, and spend considerable resources thinking about how to do so. This paper argues that a useful line of research may be to consider sorting based on the preferences of potential employees. It shows that a likely price to be paid from not being able to contract on output is that there will be a divergence of preferences across different parts of the organization, not only because it is hard to use money to align interests, but also because the firm may end up hiring workers with radically different interests. As a result, a likely cost of operating a, say, nonprofit organization is the possibility of strife across different parts of the firm, as shown in Prendergast (2008).

The central premise of the paper is that firms can at least partially observe the preferences of their employees. The most obvious example of this is in the political process, where voters make a determination of the candidate’s preferences based on campaigns. Firms also devote considerable resources to identifying such fit issues. For example, social welfare departments attempt to identify the extent to which social workers identify with their clients, and there is evidence that such altruism affects performance (see Prendergast 2007 for details). In another realm, the Reagan Administration routinely used questionnaires with potential appointees, asking them about their political views before hiring them, and used associations with groups such as the Federalist Society (positive) or Sierra Club (negative) as indicators of such preferences (Marissa Goldin 2000). In other instances, such information may be gleaned from interviews—a considerable part of the purpose of interviews is to identify how well candidates match firms. An

extreme case of the use of “interviews” to determine preferences arises from congressional hearings on political appointees, such as to the Supreme Court.

While this endogenous alignment of preferences 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 diversity,” which deals with the idea that—since compensation cannot be used to align incentives effectively—the US bureaucracy should resemble the population of the country in terms of education, voting behavior, and attitudes toward social issues.⁶ The basis of this literature is little more than the following: if you cannot use money to align incentives, then aligning preferences might be a useful alternative.

REFERENCES

- Akerlof, George, and Rachel Kranton.** 2005. “Identity and the Economics of Organizations.” *Journal of Economic Perspectives*, 19(1): 9–32.
- Baker, George.** 1992. “Incentive Contracts and Performance Measurement.” *Journal of Political Economy*, 100(3): 598–614.
- Bénabou, Roland, and Jean Tirole.** 2003. “Intrinsic and Extrinsic Motivation.” *Review of Economic Studies*, 70(3): 489–520.
- Besley, Timothy, and Maitresh Ghatak.** 2005. “Competition and Incentives with Motivated Agents.” *American Economic Review*, 95(3): 616–36.
- Brehm, John, and Scott Gates.** 1997. *Working, Shirking, and Sabotage: Bureaucratic Response to a Democratic Public*. Ann Arbor: University of Michigan Press.
- Berthick, Martha.** 1979. *Policymaking for Social Security*. Washington, DC: Brookings Institution Press.

⁶ Two results typically arise from this literature (see Goodsell, 2004, for details). First, overall bureaucrats look quite like the population. Second, there is variation across departments in preferences, where those in “traditional” agencies such as Commerce, Agriculture, Treasury etc., tend to be more conservative than those in more “activist” agencies such as Health and Human Services and Housing and Urban Development (Stanley Rothman and Robert Lichter, 1983).

- Dewatripont, Mathias, Ian Jewitt, and Jean Tirole.** 1999. "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies." *Review of Economic Studies*, 66(1): 199–217.
- Frey, Bruno, and Reto Jegen.** 2001. "Motivation Crowding Theory." *Journal of Economic Surveys*, 15(5): 589–611.
- Golden, Marissa.** 2000. *What Motivates Bureaucrats?* New York: Columbia University Press.
- Goodsell, Charles.** 2004. *The Case For Bureaucracy.* Washington, DC: CQ Press.
- Peabody, Robert.** 1964. *Organizational Authority.* New York: Atherton Press.
- Prendergast, Canice.** 2007. "The Motivation and Bias of Bureaucrats." *American Economic Review*, 97(1): 180–96.
- Prendergast, Canice.** 2008. "The Price Of Intrinsic Motivation." Unpublished.
- Rothman, Stanley, and Robert Lichter.** 1983. "How Liberal Are Bureaucrats?" *Regulation*, 7(1), 16–22.

Copyright of *American Economic Review* is the property of American Economic Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.