Online appendix to *Moral Hazard, Discipline, and the Management of Terrorist Organizations*

This appendix provides proofs of all results in *Moral Hazard, Discipline, and the Management of Terrorist Organizations* and describes several insights the existing agency theory literature provides on why terrorist groups so often engage in seemingly risky paperwork and managerial oversight.

1 Proofs of all Results

Proof of Lemma 1: Define $\mathbf{m} = \{q^{i,1}_{t+1}, V^{i,1}_{t+1}, -q^{i,0}_{t+1}, -V^{i,0}_{t+1}\}$. As far as M_i is concerned, expectations of future hiring rules are unaffected by his choice of a_t today; indeed the causal arrow goes in the other direction, and B will try to use present and past hiring rules to produce expectations of future hiring rules that properly condition agents' actions now. Further, since his action now only determines the expectation of which state of the world (which element of V^{i,s_t}_{t+1}) will obtain at time $t+1$, a_t only influences future utility in so far as it determines the likelihood of reaching each state of the world. Thus, we can treat both continuation values and future rehiring rules as independent of a_t for the purposes of M_i's decision. Consequently, as M_i's expected utility after being hired (the only time he will get to take action) is single crossing in a_t and each element of $\mathbf{m}$, the Milgrom-Shannon Monotonicity Theorem provides that $a^*_t = \arg\max_{a_t} E[V^{i,1}_t]$ is nondecreasing in these elements. ■

Proof of Proposition 1: As noted in the text, M's strategy is optimal at every point so no unilateral deviation is beneficial in equilibrium. B's strategy is clearly optimal whenever either M succeeds—firing the better M would weaken his conditioning, as per Lemma 1, and result in the use of a worse agent for a period for no reason. Nor can it be beneficial for B to retain the worse agent given that agent's strategy. Thus, the only question is whether the payoff from deviating after a failure (i.e., rehiring the better agent rather than firing him) is positive.

We begin with the expected utility on the path. Here we work from the text's equation (3), noting the simplifications deriving from our hiring rule and stationarity. We are interested in the expected utility $E[U^{1,0}]$, indicating that the better agent was used but failed in his attack. In equilibrium the worse agent is used in any period after the better agent fails, followed by the better agent regardless of M_2's result. This implies that $E[U^{1,0}] = \tilde{p_2} + \delta_B E[U^{1,1}]$. Whenever the better agent is used, with probability $\hat{p_1}$ he succeeds and is used again, and with probability $1-\hat{p_1}$ he fails and is not used. Thus $E[U^{1,1}] = \hat{p_1} + \delta_B(\hat{p_1}E[U^{1,1}] + (1-\hat{p_1})E[U^{1,0}])$. Some algebra yields the necessary expected utility:

$$E[U^{1,0}] = \frac{\tilde{p_2} + \delta_B \hat{p_1}(1-\tilde{p_2})}{1-\delta_B\hat{p_1} - \delta_B^2(1-\hat{p_1})}. \tag{1}$$

Now consider the benefit to B from deviating. First, the better M plays his myopic strategy in this period. Second, which state of the world obtains next period depends on the outcome of this period, given the better agent's myopic action. After this, due to the strictures of one-period memory (and the one-shot deviation principle), we assume that B plays his equilibrium hiring strategy from then on, and M responds according to equilibrium play. Thus $E[U_{\text{rehire}}] = \tilde{p_1} + \delta_B(\tilde{p_1}E[U^{1,1}] + (1-\tilde{p_1})E[U^{1,0}])$. Using the continuation values in the previous step, and performing some algebra, yields the expected utility from a one-period deviation:

$$E[U_{\text{rehire}}] = \frac{\tilde{p_1}(1-\delta_B(1-\delta_B)\tilde{p_2} - \delta_B^2) + \delta_B(1-\delta_B\hat{p_1})\tilde{p_2} + \delta_B^2\hat{p_1}}{1-\delta_B\hat{p_1} - \delta_B^2(1-\hat{p_1})}. \tag{2}$$

Whenever equation (1) exceeds (2), there exists no beneficial deviation, and we have an equilibrium. More algebra produces the condition:

$$\tilde{p_1} \leq \frac{\tilde{p_2}(1-\delta_B)(1-\delta_B\hat{p_1}) + \delta_B(1-\delta_B)\hat{p_1}}{1-\delta_B(1-\delta_B)\tilde{p_2} - \delta_B^2}. \tag{3}$$

Consider the worst possible bad agent, who puts forth no effort. Further assume that $p(0) = 0$. Clearly if it were ever beneficial to use this agent for a period, it would be beneficial to use any agent for a period. Plugging $\tilde{p_2} = 0$ into equation (3) and performing a little algebra produces the simple equation $\tilde{p_1} \leq \frac{\delta_B \hat{p_1}}{1+\delta_B}$. Since the right side is always positive, it is always possible to sustain a one-period trigger equilibrium regardless of the other agent, for sufficiently bad myopic outcomes from the good agent. The rest of the proposition follows directly from taking derivatives of (3), and noting that in equilibrium the probability of success is weakly greater following a failure than following a success by Lemma 1. ■

Proof of Lemma 2: Note first that, as in Lemma 1, on the equilibrium path M_1's decision in period t only affects which continuation value holds; the values themselves are effectively exogenous in this respect. Now consider the myopic equilibrium. With no conditioning at all, M_1 has the smallest incentive possible to put weight on $\tilde{p_1}$. Since $v(w-a)$ is decreasing in a, this implies that $\tilde{a_1}$ can be no greater than any action taken in the presence of a rehiring rule that allows for conditioning of the agents through differentially hiring after a success and after a failure.

Now turn to the comparison between a_F and a_S. Similar logic to that in Lemma 1 implies that whenever $V_F \geq V_2$, we will have $a_F \geq a_S$ in equilibrium, given the incentives on M_1. To show that $V_F \geq V_2$, note first that by the assumptions on its components, $\frac{\partial^2 \nu}{\partial a^2} < 0$. Since by definition $\tilde{a_1}$ maximizes ν and is no greater than a_F and a_S, ν is decreasing away from its maximum and toward a_F and a_S. If $a_F \geq a_S$, then, $\nu_F \leq \nu_S$, and vice-versa. So, to create a contradiction assume that $V_F < V_2$, which by the above logic implies $a_S \geq a_F$ and $\nu_F \geq \nu_S$. By equations (10) and (11) of the text, $V_F < V_2$ implies that (after some algebra) $\frac{\nu_F}{\delta_M(1-p_F)} < \frac{\nu_S}{1+\delta_M(1-p_S)}$. However, as the denominator of the left hand side is weakly less than 1, while that on the right is weakly greater than one, $\nu_F \geq \nu_S$ contradicts this inequality. Thus, $V_F < V_2$ cannot be true, and so in any equilibrium $V_F \geq V_2$. ■

Proof of Proposition 2: Both M's strategies are optimal at every point so no unilateral deviation is beneficial in equilibrium. B's strategy is clearly optimal whenever either M succeeds—firing the better M would weaken his conditioning, as per Lemma 2, and result in the use of a worse agent for a period for no reason. Nor can it be beneficial for B to retain the worse agent given M_2's strategy. Lemma 2 implies that it cannot be beneficial to fire after one failure in this equilibrium. Thus, the only question is whether the payoff from deviating after two successive failures is positive. In this case, if B rehires M_1 then M_1 reverts to myopic play for as long as he remembers the deviation. Again, this is akin to learning that B will not follow through on threats and so they lose their conditioning value. Suppose this deviation occurs after a second failure in period $t-1$. In period t M_1 is rehired and plays myopically. Having made this one-shot deviation, B returns to equilibrium play in period $t+1$, firing M_1 if he failed and rehiring him otherwise.

Suppose M_1 fails in t. When he returns in period $t+2$ he notes only that he was rehired after a failure followed by a period off, and has no reason (given his memory) to play myopically. Suppose then that M_1 succeeds in t and is rehired in $t+1$. In $t+1$ M_1 remembers that he was rehired after the failure in $t-1$ and then rehired again after a success in period t. This is exactly what M_1 would see had he succeeded in $t-2$ and failed in $t-1$ and so has no reason to play myopically. Thus, even with two period memory M_1 only reverts to myopic play for one period.[1]

This argument implies that the utility for B from deviating will be: $U_{rehire} = \tilde{p_1} + \delta_B[\tilde{p_1}U_S + (1-\tilde{p_1})U_2]$. B's analogues for the same three states of the world as in the text's equations (8)-(10) are: $U_S = p_S + \delta_B(p_S U_S + (1-p_S)U_F)$, $U_F = p_F + \delta_B(p_F U_S + (1-p_F)U_2)$, and $U_2 = \tilde{p_2} + \delta_B U_S$.

Whenever U_2 exceeds U_{rehire}, there exists no beneficial deviation, and we have an equi-

[1] As in Proposition 1, we could consider the case where M_1 knows his previous period's actions, and so is able to play myopically forever after a deviation. As discussed earlier, the conditions for an equilibrium are *easier* to obtain when M remembers his own actions as well as outcomes.

librium. More algebra produces the condition:

$$\tilde{p_1} \leq \frac{\tilde{p_2}(1-\delta_B p_S - \delta_B^2 p_F(1-p_S)) + \delta_B(p_s + \delta_B p_F(1-p_S))}{1+\delta_B(1-\tilde{p_2}) + \delta_B^2(1-\tilde{p_2})(1-p_S)}. \tag{4}$$

As in the proof of Proposition 1, consider the worst possible bad agent. Plugging $\tilde{p_2}=0$ into the previous equation yields the simple equation $\tilde{p_1} \leq \frac{\delta_B(p_s+\delta_B p_F(1-p_S))}{1+\delta_B+\delta_B^2(1-p_S)}$. Since the right side is always positive, it is always possible to sustain the gentle-trigger equilibrium regardless of the other agent, for sufficiently bad myopic outcomes from the good agent. The rest of the the proposition follows directly from taking derivatives of (4) and can be seen readily in the simplified form of the inequality when $\tilde{p_2}=0$. ■

2 Bureaucracy, Moral Hazard, and Adverse Selection

The paper focuses on the particular need for bureaucracy when there are only a finite number of usable agents for the terrorist boss to employ as this scenario is both relevant for terrorist organizations and, to our knowledge, has not been explored well in the agency theory literature. Even when this limitation is not present and there are many potential agents to use, bureaucracy and record keeping can be useful for the terrorist group trying to solve its organizational challenges. In this section of the appendix, we briefly discuss two uses for bureaucracy that, unlike those described in the paper, follow from existing agency theory literature. The first subsection considers how bureaucratic monitoring can help organizations reduce their moral hazard problem by providing additional signals of agents' performance. The second subsection considers how record-keeping can ameliorate the adverse selection problem that arises when the quality of the agents is unknown.

2.1 Moral Hazard

Consider the utility of bureaucratic monitoring in obtaining an additional signal of agent success. [2] Existing work shows that when enforceable contracts are available, reducing the variance of the signal a principal receives of her agent's effort–which a second independent signal does–(weakly) improves outcomes for the principal (Holmström, 1979). To see how improved bureaucratic monitoring works in our specific context we modify the Shapiro and Siegel (2007) model by adding a second signal, $\hat{\Phi}$, that can take two values: 0, indicating weak effort; or 1, indicating high effort. The probability of a high effort signal is $\Phi_t = Prob(\hat{\Phi}_t = 1|a_t; \phi)$ which we assume is increasing in a_t. The parameter ϕ dictates how sensitive the signal is to action taken by the agent.

In plain language, establishing a bit of bureaucracy with rules to check up on agents' performance can provide the data necessary to better distinguish "lucky" successes from "unlucky" failures. Lucky agents should be fired more often than their success rate would indicate, while unlucky ones should be retained more often. Agents, knowing the data to make such determinations are available, will be incentivized to higher levels of effort. This comes at a cost in security, however. To assess when accepting the fixed cost of an additional signal is worth it to B we need to compare equilibrium outcomes with and without the signal.

Figure 1 displays a conditional plot of B's utility in equilibrium as a function of ϕ, γ, and the budget constraint, given stationary strategies, under the assumptions that $\Phi = \frac{e^{\phi(a-\alpha)}}{1+e^{\phi(a-\alpha)}}$, with $v = b(w-a)$, $H = cw$, $p = \frac{e^{\beta(a-\alpha)}}{1+e^{\beta(a-\alpha)}}$, $\delta_M = .83$, $\alpha = 600$, $\beta = .005$, $b = .00088675$, and $c = .0004167$.[3] The vertical axis displays B's utility, each horizontal axis displays variation

[2]Shapiro (2008) discusses the fact that reporting requirements also serve as a signal to subordinates about leaders' preferences.

[3]These parameter values were used and justified in previous work by the authors, with the changes to γ and a discussed in the text. Note that, unlike in the text or in the following subsection, we allow B to choose a level of w here, subject to a budget constraint we vary. As inShapiro and Siegel (2007), closed-form equilibrium best response functions are not available in the model with two signals and, since it is not a

in ϕ from 0 to 0.01, and each plot corresponds to a particular combination of γ and the budget constraint. The budget constraint increases vertically, taking values 600, 900, 1200, and 1500. M's type, γ, increases horizontally, taking values 0, 0.1, 0.2, 0.3, and 0.4. For higher values of γ, given these parameter values, the second signal does not benefit B.

A few trends are apparent from the figure. First, the more sensitive the second signal, the more useful it is to B. Second, the additional signal of agent action appears to be more needed the more constrained B is either by budget or by the poor quality of the agent. When the budget is very low, the second signal allows B to use agents that otherwise would not be used, $\gamma = 0.4$ for example, leading to a broader range of parameters in which attacks happen. As the budget increases, the second signal becomes useful for worse agents, until a leader unconstrained by budget (the top row) benefits from the second signal only for the lowest levels of γ. Third, the biggest benefit for B occurs when the second signal allows for the use of poor-quality agents who would not otherwise be used in equilibrium. Therefore, the additional signal not only improves B's utility; it also leads to additional terrorist attacks.

2.2 Adverse Selection

An additional signal via monitoring is a straightforward way in which bureaucracy can aid the boss, but it is not the only one. Another important role of bureaucracy is in the record keeping required to store information about agents' past actions. Answering the question of how this can be useful to the terrorist boss is more complex than in the case of two signals,

---

focus of this paper, we settle for illustration of the point via Figure 1 of this appendix. However, one can determine implicitly conditions under which B would make use of the signal, either (i) retaining the agent after a failed attack but a high signal, or (ii) not retaining after a successful attack but a low signal. Letting $\zeta = \frac{1-\gamma}{\gamma}$, (i) happens when $\frac{\Phi'}{\Phi} > \frac{\frac{-v'}{v} + \left(\frac{\alpha v+1}{\zeta v}\right)\left(\frac{p'}{1-p}\right)}{1+\frac{p}{\zeta v}}$, while (ii) happens when $\frac{\Phi'}{1-\Phi} > \frac{\frac{v'}{v} + \frac{p'}{p}}{1+\frac{p}{\zeta v}}$. Though the right hand side of each inequality is not particularly intuitive, the left is clear. The more sensitive the second signal is to the action M takes at the equilibrium level of effort, the more useful the signal is in improving B's utility, and so the more often B will choose to employ the signal and the bureaucracy it entails.

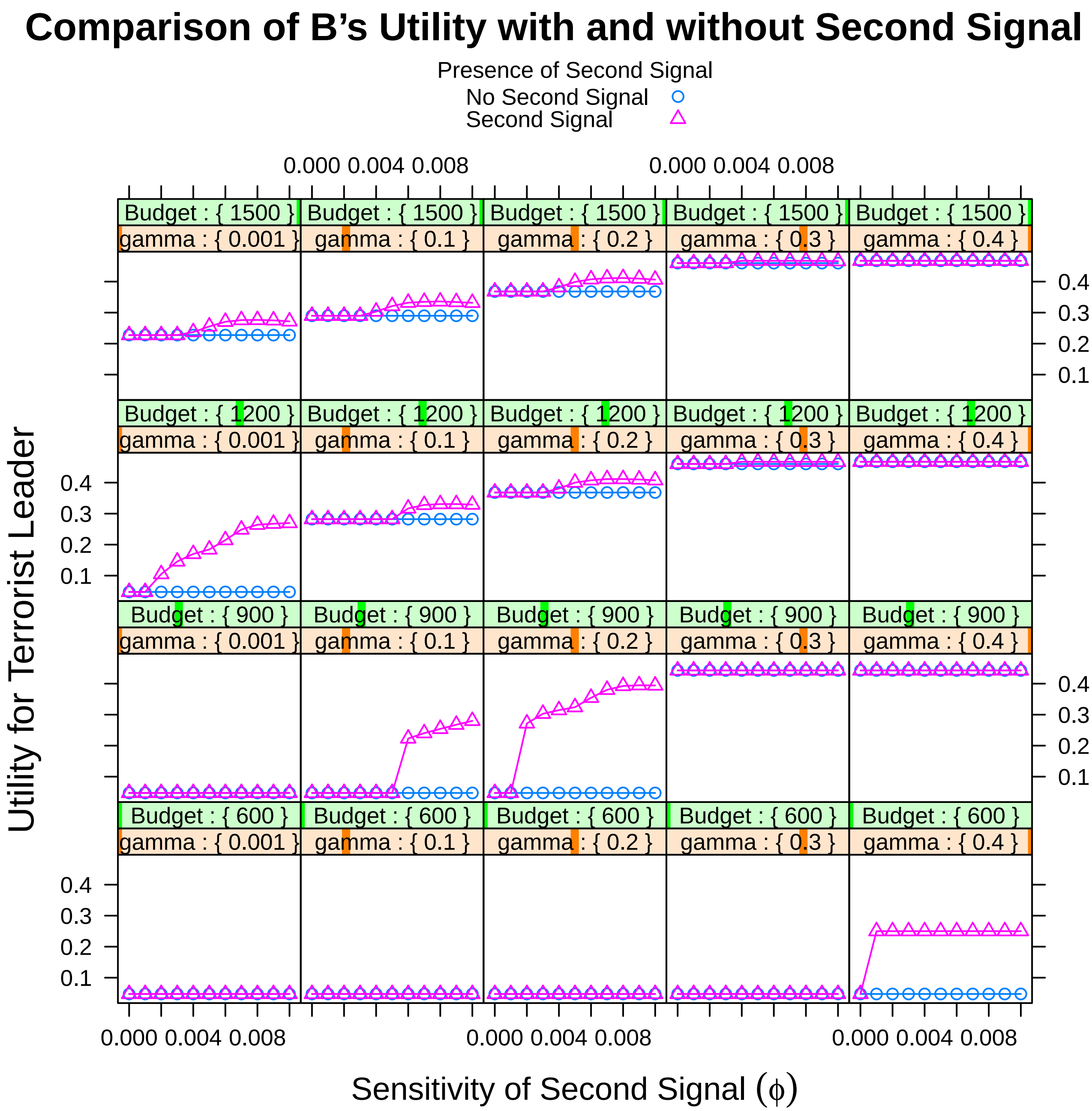

Figure 1: Comparison of B's Utility with and without Second Signal

however. Consider a variant of the basic model from Shapiro and Siegel (2007) with an infinite pool of heterogeneous agents where agent types are unknown to the terrorist boss. The principal in this scenario faces both the adverse selection problem associated with agents' incentives to suggest they will perform better than their types would imply and the moral hazard problem associated with not being able to perfectly monitor or control the agents' actions. The combination of both issues turns the scenario into a particularly intractable version of what is known as a "multi-armed bandit problem," in which not only must B discover the best "arm to pull" (i.e., agent to use), but the chance that each "arm" will pay off depends on B's strategy for testing arms. There are, to our knowledge, no extant optimal solutions for this problem; indeed, considerably simpler versions of this problem (e.g., the "restless bandit" from Whittle (1988)) have been shown not to have straightforward optimal solutions.

For certain specializations of interest, however, the extant agency theory literature does provide some insight into the problem, and it is worth addressing them briefly. Specifically, consider two simplifications: (i) only the greediest agents are available (i.e., all agents have small γ_i), and (ii) only the least greedy agents are available (large γ_i). If we can identify uses for bureaucracy at both extremes, we should expect to see such uses for more varied agent types as well.

Consider the case of greedy agents first. When material gain trumps devotion to the cause, M's utility function is decreasing in the action taken, and the only incentive for agents is the material benefit they can gain by being rehired. In this case, a simplified version of the model discussed above reduces to that presented in Banks and Sundaram (1993).[4] Their

[4]Their assumptions A1-A3 (p. 297) on the reward distribution follow directly from defining the probability of success of an attack as $F(r_t|a_t;\theta)$, the Bernoulli density function induced by $p(a;\theta)$. In particular, though there does not exist a reward in the interior of $[0,1]$, the rest of A3 holds, and so we could trivially define an interior reward with a vanishingly small probability of occurring in order to satisfy the assumption fully without altering our results. (Also see the bottom of page 308.) With fixed $w > 0$, M's utility is strictly

Proposition 1 implies there must exist sequential equilibria that are anonymous and time-consistent in which all agents with $\gamma > 0$ take actions strictly better than the lowest available, leading to strictly better outcomes for the terrorist boss.

One key to these equilibria is that the boss updates his beliefs about the "incumbent" agent in use. As a practical matter, in any reasonably complex organization such updating requires some form of record keeping and the attendant bureaucracy. Thus bureaucracy allows otherwise worthless agents—the most greedy of the bunch—to be conditioned to take higher actions. It is also worth noting the role of moral hazard here. Because B needs to condition all agents, he does better by playing an equilibrium strategy that sometimes fires agents whom he views as a better type.[5]

Now turn to the least greedy agents, those who more closely share the boss's interests. Does bureaucracy help at this other extreme? Here Banks and Sundaram (1998) provide relevant results subject to the simplification that agents can be rehired at most one time.[6]

---

positive and decreasing in action. As a technical assumption we also require for this example that the equilibrium value of a_t be in the interior of the action space. This is empirically reasonable given a fixed w, which guarantees some level of utility to the agent, and a rehiring strategy increasing in a_t.

[5]However, note that while adverse selection in this setting implies the use of bad agents sometimes, it does not imply the repeated use of an agent known to be bad, as in the case of a finite agent pool.

[6]Though equilibrium conditions are difficult to derive for multi-period agent use, assuming they exist, any such equilibrium would require additional bureaucracy, and would be weakly better for the boss, since he could always revert to the strategy of firing after every period. Thus, in a sense we offer the weakest case for bureaucracy here. After restricting an agent's use to two periods, identifying the connection between our model with multiple types and theirs requires a bit of manipulation. Let $V(a, w, q, \gamma) = \sum \delta_M^t q_t \left(-\frac{(1-\gamma)}{\gamma} v(a_t) + r_t\right)$. For sufficiently good agents (those with high γ) the utility is still positive and supermodular in a_t and γ. With these changes the model is now a special case of that offered in Banks and Sundaram (1998). Assumptions A1-A4 of that paper are satisfied directly by our $F(r_t|a_t;\theta)$. If we let $z(a_t, \gamma) = \frac{(1-\gamma)}{\gamma} v(a_t)$, then an M who is hired receives per-period expected utility $p(a_t;\theta) - z(a_t;\gamma)$, with $z(a_t, \gamma)$ submodular and decreasing in γ. Assumptions A5-A7 are therefore satisfied as well, for sufficiently high γ. These changes allows us to apply their Propositions 1-7 directly to our problem without

Because no M may be rehired more than once in this game, all M myopically maximize their single-period utility in the second period they are hired. Our question about the utility of bureaucracy thus boils down to a particularly simple form. If keeping track of the abilities of agents in the first period allows a terrorist boss to achieve better outcomes, it may be beneficial for that boss to use bureaucratic tools to keep track of these actions. Banks and Sundaram (1998) show that for a pool of sufficiently good agents who may only be rehired once each, there exists a cutoff strategy such that all agents who succeed are rehired, and take weakly better actions in the first period of employ than in their second. This conditioning induces a better situation for the boss, who prefers to rehire a successful agent even though that agent will surely act myopically in the next period, because the boss expects that rehired agent to be of a better type. Once again, by helping B keep track of agents' actions, bureaucracy enables (weakly) better outcomes for the boss.[7]

Interestingly, the boss can even benefit in some circumstances by the uncertainty in agent type. A pure moral hazard version of this problem, in which all agent types are identical and known, leaves no room for conditioning—all agents always take their myopically optimal actions. A small amount of uncertainty, however, allows the boss to employ credible disciplinary strategies that motivate better agents to higher effort.

Tracking of agents types also proves useful in Schwabe (2009), which considers a more general infinitely repeated model in which agents choose a level of effort and principals condition them by using reputation-dependent cutoff rules, i.e. firing agents whose performance drops below a set level, where that level depends on the principals' belief about an agent's type. Schwabe proves the existence of a class of equilibrium that are time-consistent and anonymous in the sense we define and discusses some of their characteristics. What is relevant for this paper is that as a practical matter, all of these equilibria require tracking

substantively altering the model.

[7]This result can be made strict by disallowing the absolutely committed agents (i.e., those with $\gamma = 1$) and assuming certain technical regularity conditions.

agents' reputations, and thus imply bureaucracy in any reasonably large organization.

References

Jeffrey S. Banks and Rangarajan K. Sundaram. Adverse selection and moral hazard in a repeated elections model. In William A. Barnett, Melvin J. Hinich, and Norman J. Schofield, editors, *Political economy: Institutions, competition, and representation*. New York: Cambridge University Press, 1993.

Jeffrey S. Banks and Rangarajan K. Sundaram. Optimal retention in agency problems. *Journal of Economic Theory*, 82:293–323, 1998.

Bengt Holmström. Moral Hazard and Observability. *Bell Journal of Economics*, 10(1):74–91, 1979.

Rainer Schwabe. Reputation and accountability in repeated elections. Unpublished manuscript, Princeton University, 2009.

Jacob N. Shapiro. Bureaucracy and control in terrorist organizations. Manuscript, Princeton University, 2008.

Jacob N. Shapiro and David A. Siegel. Underfunding in terrorist organizations. *Interational Studies Quarterly*, 51(2):405–429, 2007.

Peter Whittle. Restless bandits: Activity allocation in a changing world. *Journal of Applied Probability*, 25:287–298, 1988.