The Limits of Authority: Motivation versus Coordination
PRELIMINARY

Eric Van den Steen*
April 21, 2005

Abstract

This paper studies the micro-mechanisms of interpersonal authority in a setting where people may openly disagree, i.e., have differing priors. It first studies the emergence of authority, including what factors favor it, and who is most likely to actually get authority. It then focuses on the limitations that enforcing authority imposes on an organization.

The key conclusion of the paper is that, through a set of mutually reinforcing mechanisms, interpersonal authority induces a fundamental trade-off between motivation (effort) and coordination, as follows. If there is a high need for effort, then pay-for-performance should be high, but that weakens authority, since it gives the agent more reason to disobey when he disagrees with the principal’s orders. Weak authority leads, on its turn, to less coordination, since authority causes actions to be ‘as if’ they were taken by one person. In the other direction, a high need for coordination facilitates authority, since the principal and the agent typically incur an opportunity cost if they go their own ways, so the agent will prefer to obey, while at the same time the principal will want to make the agent obey. But strong authority usually reduces effort levels since it forces the employee to work on projects that he does not believe in. Via the choice of authority level, this set of mechanisms thus induces a trade-off between motivation (effort) and coordination. This trade-off is known as one of the key issues in organization design.

1 Introduction

Interpersonal authority, i.e., the ability of one person to tell another what to do and be (voluntarily) obeyed, is one of the corner stones of social organization. Simon (1947), for example, stated that ‘[t]he central core of many of the most important social institutions consists of a system of authority, and a set of sanctions for enforcing it,’ while Arrow (1974) considered it ‘[a]mong the most wide-spread characteristics of organizations’.

While economics often takes such interpersonal authority for granted, when assuming, for example, that the manager of a firm has all control rights, authority is actually problematic, and often difficult to achieve. Chester Barnard (1938), for example, observed that ‘[authority] is so

*MIT-Sloan School of Management (evds@mit.edu). This paper covers part of an earlier paper under a different name, ‘Interpersonal Authority: A Differing Priors Perspective’. Extensive discussions with Bob Gibbons, Paul Oyer, and Ravi Singh had an important impact on this paper. I also thank Oliver Hart and the participants in the MIT organizational economics lunch and the Harvard-MIT organizational economics seminar for the suggestions and discussions. (As yet, not all ideas and suggestions are reflected in the current version of this paper.)
ineffective that the violation of authority is accepted as a matter of course and its implications are not considered,’ and also noted that ‘[i]t is surprising (...) how generally orders are disobeyed. For many years the writer has been interested to observe this fact, not only in organizations with which he was directly connected, but in many others.’ Most managers would concur with such observations.

The reason why interpersonal authority is problematic is the fact that people in the end are masters of their own actions. They decide whether or not to obey. This is well captured in the old drill sergeant’s saying ‘The Army can’t make you do something, but it sure as hell can make you wish you had.’ Authority will work only when it is in the subject’s interest to obey, for example, because of the threat of economic or social sanctions.

Despite the importance of authority for nearly any type of social institution, and the fact that it is not self-evident, economics has paid only limited attention to its foundations. The purpose of this paper, then, is to study the micro-mechanisms of interpersonal authority in a context with differing priors, in order to understand what limitations effective authority imposes on the organization. Apart from comparative statics on what settings are most conducive to authority, and who is most likely to get it, the key conclusion of the paper is that the mechanisms behind interpersonal authority induce a trade-off between motivation and coordination, which is a well-known and important issue in organization design.

To study these issues, I focus on a cooperative project between two players. They each have to make a decision, but may openly disagree on the optimal course of action. Such open disagreement implies, by Aumann (1976), that they have differing priors, an assumption that I will discuss in more detail below. While decisions and control are non-contractible, the players can communicate their preferences, i.e., their priors, to each other (in fact ordering the other what to do). Moreover, they can contract on success or failure of the project and they can quit at any point in time (potentially after having observed the other’s decision). I show that simple contracting allows these players to endogenously generate interpersonal authority, and consider when authority is most likely to emerge, and who is most likely to get it. I show, for example, that one player is more likely to have authority over another as he is more confident about his proposed course of action, and as he has more at stake with a success. I also show that people with strong intrinsic motivation are less likely to be subject to authority. I explore the implications of these and other comparative statics.

I then study the limitations that the need to enforce authority imposes on organizations. The aforementioned trade-off between motivation and coordination is induced by a set of four self-reinforcing mechanisms, as follows.

1. A need for effort requires stronger pay-for-performance incentives. But with a higher stake in the outcome, the agent cares more about making the decision that he thinks is right, and is thus more likely to disobey the principal when the two of them disagree (Van den Steen 2005b). A need for effort thus weakens authority.

2. Authority improves coordination through the simple fact that the decisions are ‘as if’ they are made by the same person, and thus automatically coordinated. Weakened authority thus reduces coordination.

3. A need for coordination reinforces authority, since the cost of miscoordination gives the agent an extra incentive to obey and the principal an extra incentive to enforce her authority.

4. Authority leads typically to lower effort since it forces the agent to work on a project in which he doesn’t believe. In particular, Van den Steen (2005a) shows that delegation increases effort
when effort and making the right decision are complements, but lowers effort when they are substitutes.

Via the choice of authority level, this set of mechanisms thus forces a choice between motivation (effort) and coordination.

**Literature**  This paper is very closely related to two parallel papers that were already mentioned above. Van den Steen (2005b) shows that strong pay-for-performance may increase the temptation of an agent to disobey the principal, weakening the latter’s authority. This effect will play an important role throughout the paper. Van den Steen (2005a) studies the optimal allocation of control under differing priors, focusing on the question when control over different decisions should be concentrated in the hands of one person. One of its conclusions that plays an important role in this paper, is that, if there is an agent who has to spend effort that is complementary to the decision being correct, then it is often optimal to delegate that decision to that agent, i.e., to not exercise authority over the agent. I discuss in detail how this delegation effect differs from Aghion and Tirole (1997) and from Zábojník (2002). The difference can be noted from the fact that in Aghion and Tirole (1997) delegation can only make sense if there is some correlation between the valuations of the principal and those of the agent, which is not at all necessary in this model, while in Zábojník (2002) delegation is optimal even when effort and correct decisions are substitutes, which is exactly the opposite prediction from this paper.

There is a second, more indirect, link to the literature on the efficient allocation of decision authority or control (Aghion and Tirole 1997, Dessein 2002, Hart and Holmstrom 2002, Zábojník 2002, Aghion, Dewatripont, and Rey 2003, Baker, Gibbons, and Murphy 2004, Marino and Matsusaka 2004, Van den Steen 2005a). In particular, Baker, Gibbons, and Murphy (1999) criticize the assumption, implicit in much of this literature, that the principal can commit to transfer control, and study, instead, delegation in a repeated game context. Alonso and Matouschek (2004) take a similar perspective. This paper raises more or less the opposite issue: a ‘principal’ often does not have natural control over important decisions, but instead has to leverage other control rights into interpersonal authority.

This issue of how to create interpersonal authority has received limited attention in the economic literature. The most important result is the idea of efficiency wages (Shapiro and Stiglitz 1984), although the early analyses were more focused on the implications of this idea for unemployment and wages than on the mechanism itself. In response, MacLeod and Malcomson (1989) use mechanism design to explore in depth this mechanism of authority under traditional moral hazard. MacLeod and Malcomson (1998) compare the efficiency of performance pay with that of authority induced by efficiency wages.\(^1\) These papers differ from the current one both in focus and results. In particular, while these papers start implicitly from the premise that performance pay and efficiency wages are alternative solutions to the agency problem, the premise of this paper is that they may be solutions to different problems and may actually counteract each other. This is obviously closely related to the focus on differing priors rather than moral hazard as the agency problem.

The first paper to consider the conflict between effort and correct decisions in the presence of externalities is Athey and Roberts (2001). Their argument is as follows. Imagine a situation where an agent has to spend effort that increases his own output, but also has to make a decision that affects both his own output and that of others. If the firm wants to raise his effort, it needs

---

\(^{1}\)Legros and Newman (2002) also show how the ability of players to jam the signals of their court opponents may lead to authority as the optimal solution.
to increase the pay-for-performance on the agent’s own output. However, in order to get optimal
decisions in the presence of externalities, equal weight should be put on all the output-components
that are affected by that decision, i.e., incentives have to be balanced. Ideally the firm would thus
also want to raise the agent’s pay-for-performance on the others’ output. But doing so is costly
since it exposes the agent to more risk. As a consequence, increasing effort will in equilibrium lead
to more distorted decisions.

The same effect is present in Dessein, Garicano, and Gertner (2005), with the cost of giving incentives
in that model coming from the fact that giving incentives to one person automatically reduces the incentives
of others, since they impose a budget-balancing constraint and let all profits be distributed among the managers. More importantly, they show that raising pay-for-performance on the agent’s own output will also hinder communication about potential externalities, creating another important trade-off. Finally, they also consider whether it may be optimal to shift control rights to a third person.

Both these papers are very different from the current one. This is, for example, clear from the fact that in Athey and Roberts (2001) and Dessein, Garicano, and Gertner (2005), global incentives may be costly but they are never bad, they never create agency problems as they do in this paper. As a consequence, intrinsic motivation to improve the firm’s overall performance will always be good in these papers, as opposed to here. These differences in results are of course a reflection of the fundamentally different underlying mechanisms, with the core elements of this paper (differing priors and enforcement of authority) playing no role in these two papers.

The key contribution of this paper is to determine conditions that are conducive to authority and to present a new mechanism that induces the motivation-coordination trade-off, which is one the most important challenges of organization design. (I also present informal evidence that supports the mechanisms identified here as a likely source of the trade-off.)

The next section introduces the basic model, with no effort or coordination issues. Section 3 analyzes this model to derive comparative statics on when authority will be observed and who is most likely to get it. Section 4 then introduces effort and coordination issues and derives the trade-off between motivation and coordination. Section 5 concludes. Some long proofs are summarized here and presented in more detail in the extra appendices to this paper (Van den Steen 2004).

2 The Basic Model

I will first study authority in a context with no effort or coordination concerns, i.e., where the players only care about what decisions get made. Doing so allows a more transparent analysis of some of the basis issues in the emergence of authority. To that purpose, I consider a setting in which two people jointly undertake a project, and each must make some decision that affects the outcome of the project. Neither decisions nor control rights are contractible, and the players may disagree on what the right decisions are. The basic question is then whether the players can, or will, write a compensation contract such that all control is effectively in the hands of one person, and what the implications of such contract would be.

While this basic setting is essentially symmetric in the two players, it simplifies things quite a lot to focus on the question under what conditions one specific player gets control over the other. To that purpose, I will set up the model asymmetrically in some respects. I conjecture, however, that the results can be extended to the fully symmetric case.

Formally, consider a situation with two players, $P_1$ and $P_2$, who are involved in a joint project.
Hiring
1 Players 1 and 2 negotiate a contract that specifies
- \((w, \alpha)\),
- whether \(P_1\) may inspect \(P_2\)'s action in stage 3.

Beliefs
1 The prior beliefs of \(P_1\) and \(P_2\) get drawn.
2 \(P_1\) and \(P_2\) simultaneously choose whether and, if so, what message to send from \{A, B\} (announcing their preferred action).

Actions
1 Both players simultaneously choose their respective action from \{A, B\}.
2 If so agreed, \(P_1\) can inspect \(P_2\)'s action. (He then observes \(P_2\)'s action with probability \(p\).)

Payoff
1 Each player can walk away from the project and take his outside option. For \(P_2\), this outside option is 0. For \(P_1\), the outside option is a cost \(k\).
2 Project payoffs get realized.
3 Contract terms \((w, \alpha)\) get executed.

As part of the project each will have to make one decision in the form of choosing an action from the set \{A, B\}. The focus of the analysis is on the conditions under which \(P_1\) will be able to order \(P_2\) what decision to take (and \(P_2\) obeys). If so, \(P_1\) will thus behave like the principal while \(P_2\) will behave like an agent. Since this principal-agent relationship is endogenous to the model and since it does not always obtain, I will refer to the players simply as \(P_1\) and \(P_2\) instead of principal and agent.

Let \(D_i\) denote the decision that will be made by player \(P_i\). Each decision can be either right or wrong, and the project revenue will be \(R = I_1 + I_2\) where \(I_j\) denotes the indicator function that decision \(D_j\) is right. Whether a decision is right or wrong depends on its fit with the state of the world. In particular, the world can be in two states, also denoted \(A\) and \(B\), and a decision is right if and only if the decision fits the state of the world. While the state of the world is unknown, players have their own subjective beliefs about the state. Let \(\rho_i\) denote the belief of player \(i\) that the world is in state \(A\). A key element of the analysis is that players can disagree in their beliefs even though they have no private information. In other words, the players will have differing priors.\(^2\) I come immediately to how these beliefs get determined.

The timing of the game is indicated in figure 1. In stage 1, the players negotiate a contract using axiomatic Nash bargaining with bargaining power \(\lambda\) and \(1 - \lambda\). When the feasible set is non-convex, I will use the extension of Nash bargaining by Zhou (1997) in which the bargaining solution is always the point that maximizes the Nash product. The contract consists of two elements:

1. A compensation contract for \(P_2\) that consists of a wage \(w\) and share of the project revenue \(\alpha\in[0,1]\), both of which are conditional on the project getting carried out, as I will discuss below.

\(P_1\)'s compensation is then \(-w\) and the complementary share \((1 - \alpha)\) of the project revenue.

---

\(^2\)Differing priors do not contradict the economic paradigm: while rational agents should use Bayes’ rule to update their prior with new information, nothing is said about those priors themselves, which are primitives of the model. In particular, absent any relevant information agents have no rational basis to agree on a prior. Harsanyi (1968) observed that ‘by the very nature of subjective probabilities, even if two individuals have exactly the same information and are at exactly the same high level of intelligence, they may very well assign different subjective probabilities to the very same events’. For a more extensive discussion, see Morris (1995) or Van den Steen (2001).
2. An agreement whether $P_1$ will be allowed to inspect $P_2$’s action in stage 3.

I impose the no-wager condition that $\alpha \in [0, 1]$ to prevent the players from betting on the state and generating infinite utility.\(^3\) The assumption that the players can agree on whether $P_1$ can inspect $P_2$’s action is nothing out of the ordinary. While a customer generally has no right to inspect the premises or actions of a supplier, companies sometimes have such explicit agreements. Ford, for example, requires internal inspection as one of its conditions for certifying a supplier. Also, and more importantly, a firm has very broad rights to inspect its employees’ actions, so that by entering in an employment relationship such agreement is implicit. It is important to note, though, that this is one important area where the model is asymmetric. In particular, I don’t allow the players (for now) to agree that $P_2$ will be able to inspect $P_1$’s action. As mentioned earlier, this asymmetry simplifies the analysis considerably but the results seem to extend to the fully symmetric case.

At the start of stage 2, the beliefs get drawn from a random distribution. The idea of having the beliefs be drawn after the contract negotiation, is that the contentious issues, on which the players may disagree, arise only after the project has been started, so that it is only at that time that it becomes clear which beliefs are relevant. To keep the analysis transparent and tractable, I will assume a very simple degenerate distribution for these prior beliefs. In particular, for some given parameter $\nu_i \in (.5, 1)$ (which is commonly known at the start of the game), player $i$’s prior belief that the state is $A$ will be either $\nu_i$ or $1 - \nu_i$, with equal probability. In other words, $\rho_i$ is drawn from a 2-point distribution with half its weight on $\nu_i$ and half on $1 - \nu_i$. It follows that the player always has the same strength of belief, $\nu_i$, in the state that he considers most likely. Moreover, each player will believe half the time that the state is $A$, and half the time that the state is $B$. The prior beliefs will be independent draws, so that the players will disagree half the time. The expected value of $R$ according to $i$ is

$$E_i[R] = \begin{cases} 2\nu_i & \text{if both decisions are right according to } i \\ \nu_i + (1 - \nu_i) = 1 & \text{if one decision is right and one is wrong according to } i \\ 2(1 - \nu_i) & \text{if both decisions are wrong according to } i \end{cases}$$

The beliefs are originally private information. In the second step of stage 2, however, both players have a chance to tell the other what to do. In particular, they simultaneously each choose whether and, if so, what message to send from the set $\{A, B\}$. These messages can best be interpreted as ‘you should do $A$’ and ‘you should do $B$’. I will assume that players have a lexicographic preference to be obeyed. Since, as I will discuss later, the analysis will consider only pure strategy equilibria that are not Pareto dominated, the communications will reveal the players’ beliefs truthfully, if at all.

In stage 3, both players make their respective decisions. If they agreed so in their contract, then $P_1$ gets a chance to inspect $P_2$’s decision. If he decides to do so, then $P_1$ observes $D_2$ with probability $p$. Decision $D_1$ is never observed. As mentioned earlier, neither decision is contractible.

At the start of stage 4, either player can quit the project. If either player quits the project, then player $P_2$ gets his outside wage, which is normalized to zero, while $P_1$ pays a cost $k$. The simplest interpretation of $k$ is that it is the cost of firing or breaking up the contract. If neither player quits, then the state gets realized, $P_1$ receives the project’s revenue and pays $P_2$ according to the contract $(w, \alpha)$. I will, moreover, allow that the players have private benefits from success

\(^3\)This assumption could be derived endogenously if players had the ability to sabotage the project. Alternatively, it could be defended based on legal provisions that compensation should be proportional to the actions, or that forbid gambling. To maintain generality and simplify the analysis, I impose it as an assumption.
that cannot be contracted. In particular, player $i$ also gets a private benefit $\gamma_i R$, with $\gamma_i \geq 0$. \footnote{Note that the restriction $0 \leq \alpha \leq 1$ prevents players from contractually eliminating the effect of $\gamma$. Such contractual elimination would obviously defeat the purpose of introducing $\gamma$.} I will denote player $i$'s total benefit as $\alpha_i$, so that we have $\alpha_1 = 1 - \alpha + \gamma_1$ and $\alpha_2 = \alpha + \gamma_2$. These private benefits will be interpreted at a later time.

In the analysis, I will consider only pure-strategy sequential equilibria that are not Pareto dominated. Such equilibria always exist. In the presence of multiple outcome-equivalent equilibria, I will select the one that applies to the broadest range of parameters. \footnote{As far as I have analyzed, these conditions only exclude equilibria that are extremely similar to the ones I obtain here (where, for example, a player doesn’t quit for sure, but with ‘sufficiently high probability’) and equilibria that are not very realistic or not very interesting (where, for example, both players quit simply because the other quits and therefore they are indifferent between quitting or not).}

### 3 When Authority Obtains (Endogenously)

This section considers when authority will effectively emerge from the contracting between the two agents, and identifies what conditions make it more or less likely. These results are captured in proposition 1 below. In particular, the proposition says that interpersonal authority will be \textit{more} likely when the presumed ‘principal’ has stronger beliefs, more private benefits at stake, when he has an easier time observing the presumed ‘agent’s’ actions, and when it is less costly to fire. Conversely, interpersonal authority is \textit{less} likely when the presumed ‘agent’ has stronger opinions about the right course of action, when he has high intrinsic motivation or private benefits for a success, and when his actions are more difficult to observe. I will discuss the intuition for some of these results after the proof of the proposition.

To state the proposition formally, I will use $P_1 A$ ($P_1$ Authority) to describe the following equilibrium. Players $P_1$ and $P_2$ agree on a contract according to which $P_1$ will have the right to inspect $P_2$’s decision and $\alpha = 0$. In the subsequent periods,

- player $P_1$ (and $P_1$ only) sends a message to $P_2$ telling him what to do,
- $P_2$ does as $P_1$ told him while $P_1$ simply chooses the action that he believes has the highest probability of success,
- $P_1$ inspects $P_2$’s decisions, and
- $P_1$ quits if he observes that $P_2$ did not obey.

I will use $NA$ (No Authority) to describe the following equilibrium. Players $P_1$ and $P_2$ agree on a contract according to which $P_1$ will not be able to inspect $P_2$’s decision and $\alpha = 1$. In the subsequent periods,

- neither player sends a message to tell the other what to do,
- each player chooses the action that he believes has the highest probability of success,
- neither player quits in equilibrium or as a response to a deviation by the other.

The following proposition then captures the comparative statics mentioned above.

\textbf{Proposition 1} \textit{For any set of parameters, the (pure strategy, Pareto dominant) equilibrium is either $P_1 A$ or $NA$. The likelihood that the equilibrium is $P_1 A$}
1. increases in $\nu_1$ and decreases in $\nu_2$
2. increases in $\gamma_1$ and decreases in $\gamma_2$
3. increases in $\alpha$
4. decreases in $k$

Proof: Since we are looking for the Pareto optimal equilibria, it suffices to consider for each possible agreement the Pareto optimal equilibria in the subgame it induces and then apply the Nash bargaining solution. For notational convenience, I will use $X_i$ to denote the action that $P_i$ believes is best, $Y_j$ to denote the action that ends up being chosen for decision $D_j$.

A more detailed analysis in Van den Steen (2004) shows that there are two types of equilibria.

In the first type of equilibrium, both players choose the action that they believe is most likely to succeed, i.e. $P_j$ chooses $Y_j = X_j$. No beliefs get revealed and $P_1$ also does not get to inspect $P_2$’s decisions. Neither player will ever quit. This gives payoffs

\[
\begin{align*}
U_1 &= \frac{1}{2} (2(1-\alpha + \gamma_1)\nu_1 - w) + \frac{1}{2} ((1-\alpha + \gamma_1) - w) = (1-\alpha + \gamma_1) \frac{1}{2} + \nu_1 - w \\
U_2 &= \frac{1}{2} (2(\alpha + \gamma_2)\nu_2 + w) + \frac{1}{2} ((\alpha + \gamma_2) + w) = (\alpha + \gamma_2) \frac{1}{2} + \nu_2 + w
\end{align*}
\]

I will denote this the NA-equilibrium (‘do as you believe’). NA requires $w \in (-\nu_1 + \gamma_2)(\frac{1}{2} + \nu_2), (1-\alpha + \gamma_1)(\frac{1}{2} + \nu_1) + k$.

To complete the analysis of NA, we need to find the optimal $\alpha$ and the corresponding set of (Pareto optimal) payoffs that can be attained. In other words, we need to find the efficient frontier in the $u_1$-u2 space. The best starting point is to find the $\alpha$ that maximizes the joint utility and see how that constrains the allocations. Subsequently we can then consider whether other Pareto optimal allocations can be achieved by choosing other $\alpha$’s.

The joint payoff of NA is $U = \frac{1}{2} \gamma_1 \nu_1 + (1-\alpha + \gamma_1)\nu_1 + (\alpha + \gamma_2)\nu_2$ subject to $w \in (-\nu_1 + \gamma_2)(\frac{1}{2} + \nu_2), (1-\alpha + \gamma_1)(\frac{1}{2} + \nu_1) + k$. The first order derivative is $U'(\alpha) = w_1 - \nu_1$. If $\nu_1 > \nu_2$, then the optimal $\alpha = 0$ which requires $w \in (-\nu_1 + \gamma_2)(\frac{1}{2} + \nu_2), (1-\alpha + \gamma_1)(\frac{1}{2} + \nu_1) + k$. If $\nu_1 < \nu_2$, then the optimal $\alpha = 1$ which requires $w \in (-\nu_1 + \gamma_2)(\frac{1}{2} + \nu_2), 1 + (\frac{1}{2} + \nu_1) + k$. In both cases, the total utility is $U_A = \frac{1}{4} + \max(\nu_1, \nu_2) + (\frac{1}{2} + \nu_1) + \gamma_2(\frac{1}{2} + \nu_2)$ and can be freely allocated to the players by choosing $w$ appropriately. So this establishes the full efficient frontier for this equilibrium.

This set of payoff combinations is depicted in figure 2a.

In the second type of equilibrium, $P_1$ (and only $P_1$) reveals his preference, $P_2$ obeys $P_1$ so that $Y_2 = X_1$, and $P_1$ inspects. $P_1$, on the other hand, chooses $Y_1 = X_1$, so that $Y_1 = Y_2$. $P_1$ quits if $P_2$ does not choose $X_1$ (which is off the equilibrium path). The expected utilities are

\[
\begin{align*}
U_1 &= 2(1-\alpha + \gamma_1)\nu_1 - w \\
U_2 &= (\alpha + \gamma_2)\nu_2 + (1-\nu_2) + w = (\alpha + \gamma_2) \nu_2 + \frac{1}{2} + w
\end{align*}
\]

subject to $w \geq (\alpha + \gamma_2)(\theta(2\nu_2 - 1) - 2(1-\nu_2))$ and $w \in ((1-\alpha + \gamma_1) + k, 2(1-\alpha + \gamma_1)\nu_1 + k)$, where $\theta = \frac{(1-p)}{p}$. I will denote this the $P_1A$ equilibrium (‘$P_1$ Authority’).

To complete the analysis of $P_1A$, we need to find again the optimal $\alpha$ and the corresponding payoffs. I proceed as before, but the different possibilities are more complex in this case. For notational convenience, let $u_1(\alpha) = 2(1-\alpha + \gamma_1)\nu_1$ and $u_2(\alpha) = (\alpha + \gamma_2) \frac{1}{2} + \nu_2$ be $P_1$’s and $P_2$’s utilities (under the $P_1A$-equilibrium) excluding $w$.

- If $2\nu_1 \geq \nu_2 + \frac{1}{2}$, then the optimal $\alpha = 0$ for a joint utility of $U_A(0) = 2(1-\alpha + \gamma_1)\nu_1 + (\alpha + \gamma_2)(\nu_2 + \frac{1}{2})$. For $w = \max(\gamma_2\theta(2\nu_2 - 1) - 2(1-\nu_2)), (1-\alpha + \gamma_1) + k$, the set of possible payoffs in $u_1$-u2 space is the line connecting $(u_1, u_2) = (0, U_A(0))$ with $(\hat{u}_2(0) + w, \hat{u}_1(0) - w)$, as indicated in figure 2b.
- For the case that $2\nu_1 < \nu_2 + \frac{1}{2}$, let $\bar{w} = \max((1-\gamma_1)\theta(2\nu_2 - 1) - 2(1-\nu_2), \gamma_1 + k)$. If $2\gamma_1\nu_1 + k \geq \bar{w}$, then the possible combination of payoffs consists of two line-segments. On the first line segment, $\alpha = 1$ and the joint payoff is $U_A(1) = 2\gamma_1\nu_1 + (1-\gamma_2)\nu_2 + \frac{1}{2})$. This line segment connects (in $u_1$-u2 space) $(0, U_A(1))$ and $(\hat{u}_2(1) + \bar{w}, \hat{u}_1(1) - \bar{w})$. The second line segment corresponds to decreasing $\alpha$ and connects $(\hat{u}_2(1) + \bar{w}, \hat{u}_1(1) - \bar{w})$ and $(\hat{u}_2(0) + w, \hat{u}_1(0) - w)$. Note that the latter is the same point as we had when $2\nu_1 \geq \nu_2 + \frac{1}{2}$.
Figure 2: Payoff combinations in $u_1$-$u_2$ space for different equilibria
Figure 3: Effect of a decrease in $\gamma_2$ when $U_a(0) \leq U_d$ (depicted in $u_1$-$u_2$ space).

- If $2\nu_1 < \nu_2 + \frac{1}{2}$ and $2\gamma_1 \nu_1 + k < \overline{w}$, then $\alpha = 1$ is not feasible at all. Note that the largest $\alpha$ that is possible must set $u_1$ exactly to zero. So the critical $\hat{\alpha}$ is defined by
  \[(\hat{\alpha} + \gamma_2)(\theta(2\nu_2 - 1) - 2(1 - \nu_2)) = 2(1 - \hat{\alpha} + \gamma_1)\nu_1 + k\]
  with corresponding joint utility
  \[\hat{U}_a = 2(1 + \gamma_1)\nu_1 + \gamma_2(\nu_2 + \frac{1}{2}) + \hat{\alpha}(\nu_2 + \frac{1}{2} - 2\nu_1)\]

  In this case, the set of possible payoff combinations is the line segment connecting $(0, \hat{U}_a)$ with $(\hat{u}_2(0) + \overline{w}, \hat{u}_1(0) - \overline{w})$ (and this also corresponds to decreasing $\alpha$).

The proof that $P_1\lambda$ is more likely when $\nu_1$ (resp. $\gamma_1$) increases and when $\nu_2$ (resp. $\gamma_2$) decreases, works case by case. I will here only indicate what conditions imply the results, and refer again to Van den Steen (2004) for more detailed calculations. Note that what we have to prove is that, for example, when $\nu_1$ increases, the likelihood that the maximal Nash product of $P_1\lambda$ exceeds that of $NA$ increases. By way of illustration (and for later use in the proof), figure 3 represents for one set of parameter conditions the efficiency frontiers of both equilibria and the maximal Nash product curve (for one particular value of $\lambda$).

The fact that $P_1\lambda$ is more likely when $\nu_1$ (resp. $\gamma_1$) increases follows from the following results.\(^6\)

- When $2\nu_1 \geq \nu_2 + \frac{1}{2}$, then $U_a(0)/U_d$ and $(\hat{u}_1(0) - \overline{w})/U_d$ increase in $\nu_1$ (resp. $\gamma_1$).
- When $2\nu_1 < \nu_2 + \frac{1}{2}$ and $2\gamma_1 \nu_1 + k \geq \overline{w}$, then $U_a(1)/U_d$, $U_a(0)/U_d$, $(\hat{u}_1(1) - \overline{w})/U_d$, and $(\hat{u}_1(0) - \overline{w})/U_d$ all increase in $\nu_1$ (resp. $\gamma_1$).
- When $2\nu_1 < \nu_2 + \frac{1}{2}$ and $2\gamma_1 \nu_1 + k > \overline{w}$, then $\hat{U}_a/U_d$, $U_a(0)/U_d$, and $(\hat{u}_1(0) - \overline{w})/U_d$ all increase in $\nu_1$ (resp. $\gamma_1$).

The argument for $\nu_2$ is completely analogous. So consider then the statement that $P_1\lambda$ gets less likely as $\gamma_2$ increases.

The proof when $U_a(0) > U_d$ is completely analogous to the one above. For the case that $U_a(0) \leq U_d$, things are a bit more complex (since $U_a(0)/U_d$ does not decrease in $\gamma_2$), but note that we only have to consider $2\nu_1 < \nu_2 + \frac{1}{2}$ (since for the other case, the likelihood of $P_1\lambda$ does not change with $\gamma_2$). In this case, $U_d$, $U_a(0)$, and $U_a(1)$ all have the same derivative in $\gamma_2$, so that they shift in parallel, and with fixed distances among them, in the $u_1$-$u_2$ space. As indicated in figure 3, $z$ will then shift to the left if both $\hat{u}_1(1) - \overline{w}$ and $\hat{u}_1(0) - \overline{w}$ do. It follows that, also when changes are measured relative to $U_d$, a sufficient condition for $P_1\lambda$ to get less likely as $\gamma_2$ increases, is that $[\hat{u}_1(1) - \overline{w}]/U_d$ and $[\hat{u}_1(0) - \overline{w}]/U_d$ decrease in $\gamma_2$, which is indeed the case. This concludes the proof for $\nu_1$ and $\gamma_1$.

\(^6\)To see the reason for these results graphically, imagine that we change $\gamma_1$ but at the same time rescale the graph to keep $U_d$. It is then clear that the conditions indicated imply the first part of the proposition.
For the effect of $p$ and $k$, consider again the 2 equilibria and the cases in figure 2. An increase in $p$ decreases $\theta$ and thus decreases both $\bar{w}$ and $\bar{\pi}$, but does not affect any other element. It follows that the utility combination that can be reached under $U_d$ are unaffected. The utility combinations that can be reached under the different cases of $U_a$ shift towards Pareto-dominant combinations. It follows that $P_1 A$ becomes more likely as $p$ increases. As to the effect of $k$, the only thing that is affected by $k$ is the endpoint of the line segment of possible $u_1-u_2$ combinations in the $P_1 A$ equilibrium. The line segment increases when $k$ decreases. This implies again that $P_1 A$ becomes more likely. This concludes the proposition.

I now consider the intuition and implications of these comparative statics. Consider first $p$, i.e., the probability that $P_1$ observes $P_2$’s action. The fact that interpersonal authority is less likely when $p$ is low is due to a combination of 2 effects. On the one hand, it is well known from Shapiro and Stiglitz (1984) that the efficiency wage must increase as the probability of catching shirkers decreases. On the other hand, paying an extremely high wage is not subgame perfect in this model, since there is essentially employment at will. These two effects together imply the comparative static. The result implies, for example, that people in the field and highly specialized experts will be less subject to authority, which seems to fit casual observation well.

Consider next the cost of firing $k$. The comparative static that authority is more likely when $k$ is low has a clear intuition: with lower $k$, it is more credible to threaten to fire the employee upon disobedience, which makes authority more attractive. This comparative static is a well-known and widely observed phenomenon: people who can’t be fired (or otherwise punished) are more likely to disobey.

Consider next the comparative static on $\nu$, the belief strength. There are actually 4 forces that drive this result:

1. When $\nu_1$ is higher, then $P_1$ cares more about $P_2$ obeying, and is thus willing to pay a higher wage $w$ to obtain that result, making authority easier to implement.

2. When $\nu_2$ is lower, $P_2$ is more willing to obey $P_1$, making authority easier to implement.

3. When $\nu_1$ is higher, it becomes more attractive to allocate income and control to $P_1$, i.e. to increase authority.

4. When $\nu_2$ is higher, it becomes more attractive to allocate income and control to $P_2$, i.e. to decrease authority.

The latter two effects are not specific to the authority model: Van den Steen (2005a) shows that they actually follow from the optimal allocation of control.

These comparative statics on $\nu$ fit well with casual empiricism. People with strong beliefs become entrepreneurs and hire others to implement their ideas. As long as there are no private benefits or costs involved, people with weak convictions tend to easily obey orders. Managers are more likely to actively exercise authority over an issue when (they believe that) they are knowledgeable about it. Conversely, they will often let their employees more of a free rein when they feel less knowledgeable.

The intuition for the comparative static on $\gamma$ is analogous to some of the intuition for $\nu$. In particular, with high $\gamma_1$, $P_1$ always has a reason to make $P_2$ obey. Conversely, with high $\gamma_2$, $P_2$ always has a reason to want to choose the decision according to his own insights, and thus disobey $P_1$.

The interesting implications are different for $P_1$ than for $P_2$ in this case, so I consider them separately. One useful interpretation of $\gamma_1$ is that $\gamma_1$ represents the financial implications for
P_1 of P_2’s decisions that go beyond the financial means of P_2, and thus cannot be transferred contractually. For example, P_2’s decision may be a critical design choice in a $1 billion product. It is unlikely that a product designer has the financial means to contract on all the residual income. Such financial implications increase the likelihood that P_1 will have authority over P_2 in equilibrium.

The most important interpretation for γ_2 is that it represents intrinsic motivation, e.g. personal satisfaction from achieving success. In that case, proposition 1 says that people with strong intrinsic motivation are less likely to work in a setting where they are subject to authority (about what to do). Note that this is not because motivated subjects need less supervision, but because it’s costly to make them obey. A firm whose employees have strong intrinsic motivation will often need to rely on other methods to achieve coordination, such as hiring people with similar beliefs.

4 Effort and Coordination

I now introduce effort and coordination concerns in the equation. As mentioned earlier, a key result of this paper is to show that, in such context, the micro-mechanisms of authority induce a trade-off between motivation and coordination. The importance of this result derives from the fact that the dual challenge of motivation and coordination is one of the most central concerns in organization design. Indeed, many of the important design choices, such as centralization, divisionalization, incentives, or decision processes, revolve around these two goals: motivating people to do the best they can, but at the same time coordinating their actions with those of others.

It is intuitively well known that there exists such a fundamental trade-off between motivation and coordination: nearly any organization design exercise that one encounters, either in real life or in case studies, faces this trade-off (Roberts 2004). I show that this trade-off effect may be caused by the micro-mechanisms of authority, in particular the set of reinforcing mechanisms that are summarized in figure 4. The figure shows how authority and effort have a mutually negative interaction, while authority and coordination have a mutually positive interaction.

Consider first the mutually negative interaction between effort and authority. In order to motivate someone to spend effort on a project, that person often gets a stake in the outcome, for example in the form of pay-for-performance compensation or by being held ‘responsible’ for the outcome. This, however, makes the agent care about the outcome and therefore potentially disobey orders that conflict with what he or she considers the right thing to do (?). In other words, pay-for-performance reduces the agent’s ‘zone of indifference’ (Barnard 1938) or ‘zone of acceptance’ (Simon 1947). In the other direction, authority often forces a person to spend effort on a course of action that he considers suboptimal. If effort and correct decisions are complements, as they usually are, then authority will reduce the incentives to work hard.7

Consider next the mutually positive interaction, or reinforcing cycle, between coordination and authority. For the one direction, note that if one person exerts authority over another, their actions will be coordinated since it is ‘as if’ the actions were taken by the same person. In the other

7Note that when effort and success are substitutes, the opposite will be true. This latter case, however, is much less common. In fact, while situations where effort and success are complements abound, the only example of them being substitutes seems to be the case where the project simply needs to attain a fixed quality level. In that case, effort can compensate for bad decisions. The case of substitutes, however, clearly distinguishes this theory of delegation from others, such as Aghion and Tirole (1997) or Zábojník (2002).
direction, there is a double effect. On the one hand, the need for coordination means that $P_2$ will incur a coordination cost when he goes against $P_1$’s orders. Coordination thus imposes the type of cost that makes an agent obey, so that you get some level of obedience for free. In particular, I will show that it is possible to get obedience even when there is no threat of firing, as long as the cost from miscoordination is sufficiently high. The second effect of coordination on authority is that a need for coordination makes it more worthwhile for $P_1$ to impose his authority and avoid the coordination costs. So $P_1$ will be more willing to raise $w$ to make $P_2$ obey. Both effects imply that a need for coordination makes authority more likely.

Both these interactions or cycles will be even stronger if the principal can, at some cost, choose how much non-contractible effort and coordination will be required.

To study this motivation-coordination trade-off formally, I extend the game of section 2 as follows. The payoff from the project is now the sum of three parts. The first part is identical to the payoff in the basic model of section 2. The second part is a coordination cost: if the action choices on $D_1$ and $D_2$ differ, then the project incurs a cost $c$. The third part captures the potential effect of effort: player $P_2$ can spend unobservable effort $e$ on implementing his decision $D_2$ at private cost $e^2/2$, and that effort determines how much return the project generates, as specified below. Effort is

---

**Figure 4:** Interaction effects between interpersonal authority, a need for effort, and a need for coordination.
neither observable nor contractible.

The importance of the coordination part and the effort part will be parameterized by \( \beta_c \) and \( \beta_e \) respectively. Let \( I_{ij} \) denote the indicator function that player \( P_i \) thinks that \( D_j \) is the optimal decision, \( E_j[R_j] = \nu_i I_{ij} + (1 - \nu_i)(1 - I_{ij}) \) the expected revenue of decision \( j \) according to \( i \), and \( I_{nc} \) the indicator function that \( P_1 \) and \( P_2 \) are different, then the payoffs are

\[
\begin{align*}
  u_1 &= \ (1 - \alpha) \ (E_1[R_1] + E_1[R_2]) - \beta_e (1 - \alpha) c I_{nc} + \beta_e (1 - \alpha) e E_1[R_2] - w \\
  u_2 &= \ \alpha \ (E_2[R_1] + E_2[R_2]) - \beta_e \alpha c I_{nc} + \beta_e \left( \alpha e E_2[R_2] - \frac{e^2}{2} \right) + w
\end{align*}
\]

The timing of the game is indicated in figure 5. Note that the effort is spent during the implementation of the project. Moreover, to simplify the analysis, I now allow only \( P_1 \) to tell the other what to do. Finally, I will also assume that \( k \geq \beta_c c \) and \( \nu_1 = \nu_2 = \nu \).

The trade-off between motivation and coordination can be represented in a fairly graphical way, by asking the question ‘how much coordination can we get for a given level of effort?’ While this leaves out some important aspects of the problem, as I discuss below, this way of phrasing the issue puts the tradeoff in the sharpest possible terms. In particular, it generates something akin to an ‘efficient frontier’ of production theory. Figure 6 represents this ‘frontier’ in motivation-coordination space, and shows very clearly the trade-off.

The ‘frontier’ derives from the subgame-perfection constraints on the two equilibria in which \( P_1 \) has interpersonal authority over \( P_2 \). These two equilibria have been denoted ‘Automatic Obey’ (AO), i.e., \( P_2 \) obeys even though he won’t get fired if he didn’t, and ‘Obey under Threat of firing’ (OT), i.e., \( P_2 \) obeys only because he gets fired if he doesn’t. To derive the frontier, we need to find for a given level of \( \alpha \) (and thus \( c \)) the extreme \( c \) for which it is still possible to sustain authority under either of the two equilibria. Consider first the equilibrium where \( P_2 \) obeyed without \( P_1 \) threatening that he would quit unless \( P_2 \) obeyed. This equilibrium requires (with \( \eta = \beta_c c \))

\[
\eta \geq (2\nu - 1) \left( 1 + \beta_e \frac{\alpha}{2} \right)
\]

\[8\]Without the assumption that \( k \geq \beta_c c \), there would be equilibria in which one of the players quits whenever there is disagreement. We would thus have continuous turnover in a repeated version of the game. Since this is not very interesting and to simplify the analysis, I simply exclude the possibility by making this assumption.
I will represent this with the curve labelled AO (‘automatic obey’). Consider next the equilibrium where \( P_2 \) obeyed because \( P_1 \) threatened to quit otherwise. This equilibrium has two requirements

\[
(1 - \alpha) + \beta_e \alpha (1 - \alpha)(1 - \nu) \eta \nu - (1 - \alpha) \eta + k \leq (1 - \alpha) \left( 2\nu + \beta_e \alpha \nu \frac{1}{2} \right)
\]

\[
\theta \alpha \left( (2\nu - 1) \left( 1 + \beta_e \alpha \frac{1}{2} \right) - \eta \right) - \alpha (1 - \nu) \leq (1 - \alpha) \left( 2\nu + \beta_e \alpha \nu \frac{1}{2} \right)
\]

which I will represent with the curves labelled OT\(_1\) and OT\(_2\) (‘obey under threat’). Note that since both requirements are needed, the ‘frontier’ of this equilibrium is the lower of the two. The total ‘frontier’ is then the maximum of this one and AO (since either equilibrium is fine). It is indicated in bold in figure 6.

It is useful to draw attention to the way coordination is measured here. Since it is easy to get coordination when \( c \) is large, the correct measure of coordination is the lowest \( c \) for which we still get coordination. So I use as measure of coordination \( d = 1 - c/C \), where \( C \) is a parameter that I will assume to be \( C = 1 \). This measure has an attractive alternative interpretation: \( d \) is also the probability that we will get coordination if \( c \sim U[0, C] \).

![Figure 6: Frontier representing the trade-off between effort and maximal coordination (for parameters \( k = 1, \nu = .8, \theta = 15, \beta_e = \beta_c = 1 \)). Note that \( d = 1 - c \).](image)

While this representation takes indeed a form similar to the efficient frontier of production theory, there are important differences that relate to this being a pure technological trade-off with no reference to the broader problem. First, this representation only takes into account the effort and coordination components of the payoff functions, and disregards the pure decision component (i.e. the original payoff function of section 2). This pure decision component will also affect the optimal \( \alpha \) so that the final solution is not always on the frontier depicted in figure 6 but may be a bit more towards the interior. Second, the fact that we have Nash bargaining without an up-front transfer implies that the optimal solution is not necessarily the one that maximizes the joint utility. This implies again that this ‘frontier’ does not have the same implications as the efficient frontier in the neoclassical production theory. Despite these limitations, the graph of figure 6 clearly shows the technological trade-off between effort and coordination that will constrain the final solution.
There is actually quite some informal evidence that supports this story as a cause of the motivation-coordination trade-off. (Discuss Jacobs Suchard and such.)

Another way to look at the trade-off between motivation and coordination is the following proposition, which establishes the fact that a high need for coordination favors interpersonal authority while a high need for effort conflicts with interpersonal authority.

**Proposition 2** For any \( \beta_e \) and \( c \), there exists a \( \beta_c \) such that when \( \beta_c \geq \beta_c \), we always have interpersonal authority of \( P_1 \) over \( P_2 \).

For any \( \beta_c \) and \( c \), there exist \( \beta_c,1 \leq \beta_c,2 \) such that

- when \( \beta_c \leq \beta_c,1 \) we always have interpersonal authority of \( P_1 \) over \( P_2 \)

- when \( \beta_c \geq \beta_c,2 \) we never have interpersonal authority of \( P_1 \) over \( P_2 \)

**Proof:** As before, I will use \( X_i \) to denote what action player \( i \) considers most likely to succeed, and \( Y_j \) to denote the action chosen for \( D_1 \). I will also use \( \eta = \beta_e c \).

A more detailed analysis in Van den Steen (2004) shows that there are 3 possible equilibria. In the first type of equilibrium, which I denote as \( \text{NA} \) (which stands for 'automatic obey'), both players choose the action that they think has the highest likelihood to be correct and neither player quits. The optimal \( \alpha = 1 \) with a total joint utility of

\[
U = \nu + \frac{1}{2} - \frac{\eta}{2} + \beta_e \nu^2
\]

that can be allocated freely to the players. This equilibrium requires \( \eta \leq 2\nu + 1 + \beta_e \nu^2 \).

In the second equilibrium, which I denote \( \text{AO} \) (which stands for 'automatic obey'), \( P_1 \) announces his preference, and both \( P_1 \) and \( P_2 \) choose what \( P_1 \) prefers. Neither player quits in equilibrium. \( P_1 \) also doesn’t quit if \( P_2 \) would choose \( X_2 \) instead of \( X_1 \) (so there is no ‘threat of firing’). Joint utility in this case is

\[
U = 2\nu + (\beta_e \nu^2 - (2\nu - 1))\alpha + \beta_e \nu^2 + (1 - \nu)^2 - 2\nu \alpha^2
\]

If \( 2(2\nu - 1) \geq \beta_e \nu \) then the optimal \( \alpha = 0 \) so that \( U = 2\nu_1 \), else

\[
\alpha = \frac{\beta_e \nu - 2(2\nu - 1)}{\beta_e (2\nu - \nu^2 - (1 - \nu)^2)}
\]

and

\[
U = 2\nu + \frac{1}{4} \frac{(\beta_e \nu - 2(2\nu - 1))^2}{\beta_e (2\nu - \nu^2 - (1 - \nu)^2)}
\]

This equilibrium requires that either \( \alpha = 0 \) or \( \eta \geq (2\nu - 1) \cdot 1 + \beta_e \nu^2 \).

In the third equilibrium, which I denote \( P_1A \) as before, \( P_1 \) announces his preference, both \( P_1 \) and \( P_2 \) do as \( P_1 \) prefers, and \( P_1 \) quits if ever \( P_2 \) does not choose \( X_2 \). This gives total utility

\[
U = (1 - \alpha) \cdot 2\nu + \beta_e \frac{\alpha}{2} \nu + \alpha 1 + \beta_e \frac{(\nu^2 + (1 - \nu)^2) \alpha}{2}
\]

with conditions

\[
\begin{align*}
\nu &\leq \frac{(1 - \alpha) \cdot 2\nu + \beta_e \alpha \nu (\nu + (1 - \nu))}{2} \\
\nu &\geq -\frac{\alpha 2(1 - \nu) + \beta_e (1 - \nu)^2 \alpha^2}{2} \\
\nu &\geq \frac{(1 - \alpha) + \beta_e \alpha (1 - \alpha) (1 - \nu) \nu - (1 - \alpha) \eta + k}{2} \\
\nu &\geq \theta \alpha \cdot (2\nu - 1) 1 + \beta_e \frac{\alpha}{2} - \eta - \alpha (1 - \nu) \cdot 2 + \beta_e \frac{\alpha (1 - \nu)}{2}
\end{align*}
\]

I now prove the different parts of the proposition. The first part of the proposition follows from the fact that with \( \eta > 3 + \beta_e \), \( \text{NA} \)'s total expected utility is negative while \( \text{AO} \) is feasible and gives non-negative payoff.
Note next that at $\alpha = 0$ the conditions for the AO equilibrium are fulfilled, and the total utility of $U = 2\nu$ can be freely allocated to both players. This clearly Pareto dominates $NA$ at $\beta_e$ close to 0. So there exists $\tilde{\beta}_e$ such that when $\beta_e \leq \tilde{\beta}_e$, we always have $P_1$ authority. This proves the second part of the proposition.

For the third part of the proposition, note first that the condition for $NA$ that $\eta \leq 2\nu + 1 + \beta_e \nu^2$ is always satisfied for sufficiently large $\beta_e$. Moreover, the difference between the total utilities under the $NA$ equilibrium and any $P_1A$ equilibrium equals

$$U_d - U_a = \nu + \frac{1}{2} - \frac{\eta}{2} + \beta_e \nu^2 - 2\nu - (\beta_e \nu^2 - (2\nu - 1))\alpha - \beta_e \frac{\nu^2 + (1 - \nu)^2 - 2\nu \alpha^2}{2}$$

or

$$U_d - U_a = \nu + \frac{1}{2} - 2\nu - \frac{\eta}{2} + (2\nu - 1)\alpha + \beta_e \nu^2 - \frac{\nu}{2} \alpha - \frac{\nu^2 + (1 - \nu)^2 - 2\nu \alpha^2}{2}$$

The proposition now follows from the fact that

$$\frac{\nu^2}{2} - \frac{\nu}{2} \alpha - \frac{\nu^2 + (1 - \nu)^2 - 2\nu \alpha^2}{2} > 0$$

or

$$\frac{\nu^2}{2} > \frac{\nu}{2} \alpha + \frac{\nu^2 + (1 - \nu)^2 - 2\nu \alpha^2}{2}$$

which holds (as we can see by noting that the RHS is maximized at $\alpha = \frac{\nu}{(2\nu - \nu^2 - (1 - \nu)^2)^2}$, and then equals $\frac{\nu^2}{4(2\nu - \nu^2 - (1 - \nu)^2)}$).

This completes the proposition.

### 5 Conclusion

Interpersonal authority induces a trade-off between effort and coordination. One way to interpret this result is that the desirability of interpersonal authority depends on the type of incontractibility faced. If private effort and benefits are the key incontractible factors, then it is optimal to give the agent powerful pay-for-performance incentives. The cost of doing so is a loss of interpersonal authority and a loss of coordination. If, on the other hand, the key incontractible factor is making the right decision, with disagreement on the optimal course and a potential for coordination issues, then it is optimal to give the agent weak pay-for-performance incentives and to use interpersonal authority extensively. In both cases, there will be some type of disobedience or slacking off: in the ‘authority’ case, people will resist ‘working harder’ while in the ‘strong incentives’ case, people will resist taking orders about how to go about their work.

In the process, the analysis showed how authority can emerge from simple contracting, and identified the conditions under which it is most likely to emerge, and who is most likely to get it. The results matched some informal observations.

There are important limitations to this analysis, that are as many avenues for further research. For one thing, I have not considered ‘authority over effort’, which might really affect some of the conclusions. I also have focused completely on authority based on rewards and punishments, with disregard for other sources of authority, such as legitimacy. Finally, the model considered only one setup, and I can imagine that models of ‘repetition and reputation’ might modify or extend some conclusions.
References


