HARVARD | BUSINESS | SCHOOL

Authority versus Persuasion

Eric Van den Steen

Working Paper

09-085

Authority versus Persuasion

Prepared for the 2009 AEA Annual Meeting in San Francisco
Session 'Beliefs and Disagreement in Organizations'

Eric Van den Steen*

January 7, 2009

Abstract

This paper studies a principal's trade-off between using persuasion versus using interpersonal authority to get the agent to 'do the right thing' from the principal's perspective (when the principal and agent openly disagree on the right course of action). It shows that persuasion and authority are complements at low levels of effectiveness but substitutes at high levels. Furthermore, the principal will rely more on persuasion when agent motivation is more important for the execution of the project, when the agent has strong intrinsic or extrinsic incentives, and, for a wide range of settings, when the principal is more confident about the right course of action.

Managers often face a choice between authority and persuasion. In particular, since a firm's formal and relational contracts and its culture and norms are quite rigid in the short term, a manager who needs to prevent an employee from undertaking the wrong action has the choice between either trying to persuade this employee or relying on interpersonal authority.[1] Herbert Simon (1947) noted, for example, that 'when ... disagreement is not resolved by discussion, persuasion, or other means of conviction, then it must be decided by the authority of one or the other participant' and that 'in actual practice [...] authority is liberally admixed with suggestion and persuasion.' Obviously, in choosing between persuasion and authority the manager makes a cost-benefit trade-off. This paper studies that trade-off, focusing in particular on agency conflicts that originate in open disagreement, in the sense of differing priors.

*Van den Steen: Harvard Business School, 15 Harvard Way, Boston MA 02163, evandensteen@hbs.edu. This paper benefitted from the discussion by Navin Kartik and from discussions with John Roberts and Bob Gibbons. I am also grateful to Roland Benabou for organizing the session.

[1]Interpersonal authority can be defined as 'the right or power to give orders and enforce obedience.' Kenneth J. Arrow (1974) stated that 'the giving and taking of orders ... is an essential part of the mechanism by which organizations function' while Simon (1947) observed that '(o)f all the modes of influence, authority is the one that chiefly distinguishes the behavior of individuals as participants of organizations from their behavior outside such organizations.'

To that purpose, I will study a setting in which a principal and an agent are involved in a project. The project's outcome depends both on decisions and on implementation effort by the agent, i.e., on effort to execute the decisions. A key issue is that the principal and agent may openly disagree on which decisions are most likely to lead to a success even though no player has private information, i.e. the players have differing priors. For such setting, Eric Van den Steen (2002, 2004) and, independently, Yeon-Koo Che and Navin Kartik (2007) showed that open disagreement gives rise to persuasion in a very natural way: each player believes that new information will confirm her prior and thus 'persuade' the other. It is exactly this type of persuasion that I will study here.[2] Apart from such persuasion by collecting new information, I will also allow the principal to impose interpersonal authority, i.e., to make it costly for the agent to disobey an order of the principal. The sources of such interpersonal authority in a setting with open disagreement were studied in Van den Steen (2007), which showed that a firm, with its low-powered incentives and asset ownership, may be an important vehicle to convey authority to a principal. In this paper, I will use a reduced form that simply imposes a cost on the agent if he disobeys the principal.

Probably the most important result of this paper is that the principal will rely more on persuasion for projects with a high need for motivation or effort. The reason is that – under the assumption that implementation effort is a complement to correct decisions, i.e., that executing a good project is more valuable than executing a bad project – the agent will exert more effort if he believes more in the project. From the manager's perspective, persuasion will thus motivate the agent. This makes, on its turn, persuasion more attractive on projects where effort or motivation are more important.

Since persuasion can cause compliance even in the absence of authority, it seems that an increase in persuasion should lead to a decrease in the reliance on authority. This is only partially true, however: persuasion and authority can be both substitutes and complements. In particular, I will show that authority and persuasion are substitutes when authority is highly effective but complements when authority is not very effective. To see why, note that if authority alone is not sufficient to make the agent comply but the combination of authority and persuasion is, then authority is more attractive in the presence of persuasion and vice-versa, making them complements. In the other extreme, i.e. at high effectiveness, there are actually two mechanisms that make authority and persuasion substitutes. First, if both

[2]There is another natural form of persuasion in a context with differing priors. Suppose that players with differing priors may also have private information. The combination of information and priors makes observing others' beliefs insufficient to infer their private information. Communication of private information may then serve to 'persuade' others. As shown in Van den Steen (2004), players will want to communicate information that confirms their belief to 'persuade' the other and will want to hide information that contradicts their beliefs. Obviously, weak attempts at 'persuasion' will be interpreted as a negative signal. But in the context of this paper's model, persuasion would again lead to motivation.

authority and persuasion induce compliance then some of the potential (compliance) benefits of each have already been realized by the other, so that persuasion becomes less attractive in the presence of authority and vice versa. A second mechanism comes from the fact that persuasion may actually fail – when the new information contradicts the principal's belief – and then 'wake up sleeping dogs.' In particular, if the (persuasion) signal confirms the agent's view then an agent who *would* have obeyed otherwise may now decide *not* to obey. In that case, persuasion weakens authority, making authority and persuasion substitutes.

It further follows that more important effort or motivation will make the principal rely less on authority in the case that authority is very effective. Finally, authority and persuasion being substitutes also implies, from the perspective of the principal, a trade-off between motivation and cooperation. This trade-off is recognized as one of the fundamental issues in organization design (John Roberts 2004).

Another interesting, but less central, result is that the principal will rely more on persuasion (without authority) when agents have strong pay-for-performance incentives. The reason is that incentives and confidence in the project work multiplicatively. More intuitively: if the agent does not care about the outcome, then there is little gain from persuading him. Finally, there is also a positive relationship between the confidence of the principal and the use of persuasion (unless effort is not important and authority is very effective). This is caused by the fact that a more confident manager is more convinced that she will persuade the agent, making persuasion more attractive in her eyes. The reason why this relationship does not hold everywhere is that a more confident manager also cares more about the employee choosing the (subjectively) 'correct' action, which can make authority more attractive when effort is unimportant and authority is very effective.

Apart from the work already mentioned, this paper is related to a number of strands in the literature. The first is work on persuasion, such as Paul Milgrom (1981), Vincent Crawford and Joel Sobel (1982), Milgrom and Roberts (1986), or Mathias Dewatripont and Jean Tirole (1999), and work on belief formation, such as Roland Benabou and Jean Tirole (2006) or Benabou (2008). The second is work that compares different modes of decision-making related to authority and persuasion such as Philippe Aghion and Jean Tirole (1997) or Wouter Dessein (2007). Of particular interest is also Benabou and Tirole (2002) who studied the connection between confidence and motivation, although they study confidence about one's own abilities rather than confidence about the quality of a project. These two are not unrelated, however: skill and project quality both affect how effort translates into output. That relationship is reflected in the fact that both papers's results are affected by whether effort is a complement to – versus a substitute for – skill or project quality. The two do have very different interpretations and very different implications, however. Finally, the

mechanism for motivation in this paper is also related to the result in Van den Steen (2006) that delegation motivates an agent when principal and agent disagree on the optimal course of action: the agent will believe more strongly in projects that he chose himself and thus be more motivated. The difference between that paper and the current model is obviously that in this paper motivation gets induced by persuasion rather than by delegation.

The next Section lays out the model. Section 2 presents the results while Section 3 concludes. All proofs are in appendix.

1 Model

Consider a setting in which an agent executes a project for a principal. The project will be either a success or a failure. A success gives the principal and the agent respective payoffs $\gamma_P, \gamma_A \geq 0$ while their payoffs upon failure are normalized to 0. The project's probability of success (Q) depends on two decisions (D_1 and D_2) and on implementation effort (e) by the agent. The effort $e \in [0,1]$ is chosen by the agent at private cost $c(e)$, specified below. Each decision D_k is a choice from the set $D_k \in \{X, Y\}$, one and only one of which is correct as captured by the state variable $S \in \{X, Y\}$, which is the same for both decisions.

The state S is unknown, but each player i has a subjective belief about S. A key assumption is that (it is common knowledge that) players have differing priors, i.e., they can disagree about S even though neither has private information.[3] Since the players may have differing priors but have no private information about S, they will not update their beliefs when they meet someone with a different belief: they simply accept that people sometimes disagree. To keep the analysis simple, I will immediately assume that the principal and the agent disagree on S. In particular, it is common knowledge that the principal believes that $S = X$ with probability $\nu_P > .5$ while the agent believes that $S = Y$ with probability $\nu_A > .5$.[4] Note that the ν_i are the players' confidence in their beliefs.

With $d_k = I_{D_k=S}$ the indicator function that decision D_k is correct and with $\alpha, \beta \geq 0$ (with $\alpha + \beta < 1$) parameters that capture the importance of pure decision making and of implementation effort, the probability of success equals $Q = \alpha d_1 + \beta d_2 e$. This is the simplest functional form that captures all elements necessary to bring out the intuition of the paper. The first term of Q depends completely on a decision by the agent and its importance is measured by α. The second term depends on the agent's implementation effort, with

[3] See Stephen Morris (1995), Muhamet Yildiz (2000), or Van den Steen (2007) for more discussion of differing priors.

[4] This assumption is made to simplify the model. See Van den Steen (2007) for a setting where the beliefs are private information, the principal gives an 'order', and 'disobedience' is disregarding the order. The results of this paper would extend to such a context.

implementation effort a complement to the decision D_2 so that effort to implement a good project is more valuable than effort to implement a bad project, an assumption that I will discuss below. To simplify the analysis (considerably), I will also assume that the principal gets full and free compliance on D_2, so that it is as if the principal chooses D_2 and so that compliance and authority only matter for D_1. While this functional form for Q is not the most elegant, it is very transparent and will make it very clear what is driving the results.[5]

The timing of the game is very simple. First, in period 1, the principal can try to convince the agent by drawing, at a private cost c_p, a signal about the state of the world. The drawing and the signal itself are publicly observed.[6] The signal is commonly known to be correct with probability p. At the same time with her decision to draw a signal or not, the principal also decides whether to exert interpersonal authority. Exerting interpersonal authority, which comes at a private cost c_a to the principal, makes it costly for the agent to undertake an action against the will of the principal. In particular, the agent will incur a private cost c_d from choosing Y rather than X, i.e., from 'disobedience.'

In period 2, once the principal has decided on authority and persuasion, both players (simultaneously) choose their actions: the principal (essentially) chooses D_2 while the agent chooses D_1 and e. The cost of implementation effort to the agent equals $c(e) = \beta\frac{e^2}{2}$. The (only) reason to normalize effort by β is to make very clear that the results are not driven by the fact that effort would become cheaper (on a relative basis, in the absence of this normalization) when the part that depends on effort becomes more important. The decisions are non-contractible and each player is free to choose any decision he or she wants, taking into account the private and public costs and benefits.

In period 3, the state is revealed and the project outcome is realized. The players then get the benefits γ_A and γ_P upon success. No further contracting on outcomes or payoffs is possible, so that these payoffs are completely exogenously given. All players are risk-neutral and thus simply maximize the expected value of their project payoffs minus any private costs.

In terms of parameters, I will assume that $\nu_P > \nu_A$ and $\nu_P > p$. The first ensures that the principal will always follow her own beliefs while the second ensures that the signal never

[5] Very similar results obtain in a similar setting with $Q = [\alpha d_1 + \beta e]d_2$ and independent decisions. In that case, the principal's decision is a complement to the agent's full productivity. The main results also seem to hold without the assumption for the original Q that the principal gets free and full compliance on D_2, but at the cost of a considerable increase in complexity. See Rosen (1982) for a motivation why the principal's decision and the agent's effort would be complements.

[6] As pointed out by the discussant, Navin Kartik, the results would also hold if the signal was only observed by the agent. An interesting alternative formulation that fits some settings better would be to let the principal first observe the outcome of her persuasion attempt before deciding on authority.

changes the principal's mind.[7] These assumptions exclude some cases that, while sometimes interesting in their own right for different reasons, do not contribute to the analysis of this paper. To simplify the statements and analysis, I will also assume that when indifferent each player does what the other prefers, not only on the action choice but also for persuasion and authority. That implies that the principal will use persuasion when indifferent but will not use authority when indifferent. Finally, all costs are non-negative.

2 Results

Let me start by showing that authority and persuasion are complements when interpersonal authority is not very effective and substitutes when it is very effective. To state this formally, remember that exerting interpersonal authority implies that the agent incurs a cost c_d when going against the principal's beliefs. The agent is thus more likely to obey when c_d is higher, so that c_d is a good measure for the effectiveness of authority. In fact, c_d can be interpreted directly as a measure of the agent's 'zone of acceptance' or 'zone of indifference.' The following Proposition then captures the result.

Proposition 1 *Authority and persuasion are complements when* $c_d < \alpha\gamma_A(2\nu_A - 1)$ *and substitutes when* $c_d \geq \alpha\gamma_A(2\nu_A - 1)$.

The intuition for this result was explained in the introduction. The non-symmetric nature of the result, i.e., the fact that the effectiveness of authority plays a role but not the effectiveness of persuasion, may be slightly surprising. This seems to be partially due to the way that persuasion is conceptualized here. In particular, it seems that persuasion along the lines of footnote 2 would result in a more symmetric result. This, and especially its further implications, seems an interesting direction for future research.

I now turn to the most important result of the paper: that the manager will use more persuasion when employee effort or motivation is more important.

Proposition 2 *The set of parameters for which the principal uses persuasion increases in* β. *The set of parameters for which the principal uses authority decreases in* β *if* $c_d \geq \alpha\gamma_A(2\nu_A - 1)$.

The intuition is that, with effort complementary to making the right decisions, persuading the agent will increase his effort or motivation. When effort becomes more important, persuasion

[7] If $\nu_P < \nu_A$, then the principal may prefer to choose the agent's preferred action on D_2 since the motivating effect can dominate the cost of choosing the (subjectively) 'wrong' action (Van den Steen 2006). While this is an interesting observation, it would considerably complicate the analysis without, it seems, adding anything to the central arguments of this paper.

becomes more attractive and will thus be used more. The negative effect on authority when authority is relatively strong is caused by the fact that the two are substitutes in that case.

This result relies on the assumption that implementation effort is a complement to making the right decisions, i.e., that effort to execute the project is more valuable for good projects than for bad projects. (The case with substitutes is not analyzed here, but I conjecture that the result would go the other way.) This obviously raises the question whether it is indeed the case that effort and decisions are complements. An important element here is the fact that the paper has focused on effort to implement or execute the project rather than on what one could term 'corrective effort' which compensates for shortcomings in the project. While the latter is usually a substitute, the first is typically a complement. Of course, unless these two can be distinguished empirically, that only redefines the question. A more direct indication is the work of, among others, Sherwin Rosen (1982) and Michael Kremer (1993) who argued that there will be complementarities among worker (or managerial) productivities and provide empirical evidence supporting this. In fact, Rosen (1982) explicitly assumes that the quality of a manager's decision affects the output of employees multiplicatively, as in this model.

An interesting implication of this result is that there is, in the manager's eyes, a trade-off between motivation and cooperation when authority is very effective, as is clear from a graphical representation of the equilibria. This trade-off is a well-known issue in organization design (Roberts 2004). Other explanations of this trade-off include Susan Athey and John Roberts (2001), Wouter Dessein, Luis Garicano, and Robert Gertner (2005), and Van den Steen (2006). As pointed out elsewhere, sorting on beliefs ('hiring for fit') may often resolve this conflict.

A closely related result is that the manager will rely more on persuasion by itself when the agent has higher incentives γ_A. The reason is that higher incentives imply a higher base-level of effort and thus a stronger effect of persuasion. To say this in a more intuitive way: persuading someone who is indifferent about the outcome has very little effect.

Proposition 3 *The set of parameters for which the principal uses persuasion by itself increases in* γ_A.

The reason why this result only holds for 'persuasion by itself' is that a change in γ_A may also affect under which conditions the agent obeys. This may, on its turn, affect the area where the principal uses both authority and persuasion through very different mechanisms. The results would hold for persuasion in general when conditioning on 'no change in obedience.'

One would also expect a more confident manager to rely more on persuasion. In particular, a more confident principal believes more strongly that she will be able to persuade the agent, resulting in increased effort by the agent and potentially also in increased compliance.

That should make persuasion more attractive. There is, however, a counter-acting effect: a principal who is more confident about the right course of action will care more about making sure that the agent follows that course of action. Since persuasion generates at most partial compliance, a more confident principal may therefore also want to use more authority. This can make the result go the other direction when authority and persuasion are substitutes. It turns out, however, that the latter only happens when simultaneously authority is very effective and effort is not important (in a relative sense):

Proposition 4 *When $c_d < \alpha\gamma_A(2\nu_A - 1)$ then the set of parameters for which the principal uses persuasion increases in ν_P. When $c_d \geq \alpha\gamma_A(2\nu_A - 1)$, then there exists an $\epsilon \geq 0$ (which may be function of all parameters but α and β) such that the set of parameters for which the principal uses persuasion increases in ν_P when $\beta/\alpha > \epsilon$.*

One potential issue that could be raised – for the paper as a whole – is whether the absence of explicit incentives or the absence of authority over effort may be important limitations of the analysis. This does not seem to be the case. Even when the principal could also impose interpersonal authority over effort, persuasion will still play a role either as a substitute or as a complement (depending on the effectiveness of this type of authority). This would thus add more elements and trade-offs but would not undo the results. Incentives are actually a very interesting issue: while they indeed raise effort, they simultaneously create more problems for obedience (as can be easily seen from the condition under which the agent 'obeys' for D_1) since they give the agent more reason to disobey (Van den Steen 2007). But again, while effort incentives may affect the trade-off between authority and persuasion, they do not seem to undo it. Note also that the role of differing priors is to make sense of 'persuasion' as it is typically understood: with two players disagreeing, one player trying to systematically move the opinion of the other in one's own direction.

3 Conclusion

This paper studied a setting with open disagreement where a principal can use authority or persuasion to get compliance but also cares about the agent's effort in executing the decision.

The main result is that a principal will rely more on persuasion for projects with a high need for (implementation) effort. It also showed that persuasion and authority are complements when authority is relatively ineffective but substitutes when authority is very effective. This may provide a partial explanation for the well-known motivation-cooperation trade-off. Finally, the principal will also rely more on persuasion (without authority) when agents have higher pay-for-performance incentives.

The paper focused on persuasion by means of collecting new information, but also pointed to persuasion mechanisms by means of existing information. This seems to be an interesting avenue for future research.

A Appendix

Subsection A.1 determines the equilibria and, based on that, Subsection A.2 gives the proofs of the Propositions.

A.1 Equilibria

Persuasion and Beliefs Before starting the backwards induction, let me determine the players' beliefs when the principal does collect new information. Remember that ν_P was the principal's belief in X while ν_A was the agent's belief in Y. Let the updated beliefs be denoted $\nu_i^- < \nu_i^+$ when the signal respectively contradicted or confirmed i's belief. (Note that while $.5 < \nu_i < \nu_i^+$, ν_i^- can be smaller or larger than .5.) Bayesian updating implies that

$$\nu_i^+ = \frac{p\nu_i}{p\nu_i + (1-p)(1-\nu_i)}$$

and

$$\nu_i^- = \frac{(1-p)\nu_i}{(1-p)\nu_i + p(1-\nu_i)}$$

Furthermore, the fact that $\nu_P > p$ implies that $\nu_P^- > .5$ so that the principal always believes that X is more likely to succeed.

The Decision and Effort of the Agent I now start the backwards induction. The fact that the project payoff is additively separable in the agent's effort and decision allows me to treat them independently. For the agent's effort, let $\mu_A(d_2)$ be the agent's belief in the action that the principal chose for D_2. In choosing effort, the agent then solves

$$\max_e \gamma_A \beta e \mu_A(d_2) - \beta \frac{e^2}{2}$$

So that the optimal effort $\hat{e} = \gamma_A \mu_A(d_2)$. Note that this will depend on the principal's decision (through $\mu_A(d_2)$).

The agent's choice of D_1 trades off the benefit from following his own belief against the cost of disobedience. Let $\tilde{\nu}_A$ be the agent's belief that Y is most likely to succeed. Taking into account that P always prefers X and that the agent does as the principal wants when indifferent, the agent will choose X (i.e., obey) iff

$$\alpha\gamma_A(1-\tilde{\nu}_A) \geq \alpha\gamma_A\tilde{\nu}_A - c_d$$

or

$$c_d \geq \alpha\gamma_A(2\tilde{\nu}_A - 1)$$

or

$$\tilde{\nu}_A \leq \frac{1}{2} + \frac{c_d}{2\alpha\gamma_A}$$

Note that this is completely independent from the principal's decision on D_2. One implication is that the agent always complies when $\tilde{\nu}_A \leq \frac{1}{2}$ which can happen when $\nu_A^- \leq \frac{1}{2}$ (which is the case when $p \geq \nu_A$).

The Decision of the Principal Since the principal chooses his decision D_2 simultaneously with the agent's decision and with the agent's choice of effort, she takes the level of effort as given. It is then a dominant strategy for the principal to choose X (given that $.5 < \nu_P^- < \nu_P < \nu_P^+$ and that her expected payoff from D_2 equals $\beta\gamma_P\mu_P(d_2)e$ where $\mu_P(d_2)$ is the principal's belief in the action she chooses for D_2).

It then follows that the subgame starting in period 2 is uniquely determined by the parameters. In particular, the principal always chooses $D_2 = X$, the agent always chooses $\hat{e} = \gamma_A(1 - \tilde{\nu}_A)$, and the agent chooses $D_1 = X$ iff $\tilde{\nu}_A \leq \frac{1}{2} + \frac{c_d}{2\alpha\gamma_A}$ and $D_1 = Y$ otherwise.

To determine now the equilibria, it is useful to distinguish different cases along two dimensions. The first is the strength of persuasion. The key here is whether $\nu_A^- \leq .5$ or not. In particular, if $\nu_A^- > .5$ then the agent always chooses Y absent authority ('weak persuasion'), but if $\nu_A^- \leq .5$ then the agent does choose X, even absent authority, when the signal confirms the principal's belief ('strong persuasion'). The second dimension is the strength of authority. There are 4 cases that matter here:

1. $c_d \geq \alpha\gamma_A(2\nu_A^+ - 1)$ so that the agent always obeys ('strong authority'),
2. $\alpha\gamma_A(2\nu_A^+ - 1) > c_d \geq \alpha\gamma_A(2\nu_A - 1)$ so that the agent obeys unless the principal collected information and that confirmed the agent's view ('ν_A authority'),
3. $\alpha\gamma_A(2\nu_A - 1) > c_d \geq \alpha\gamma_A(2\nu_A^- - 1)$ so that the agent obeys only when the principal collected information and that confirmed the principal's view ('ν_A^- authority').
4. $\alpha\gamma_A(2\nu_A^- - 1) > c_d$ so that the agent never obeys ('no authority').

This gives a total of 8 cases. However, strong persuasion and no obedience requires both $\nu_A^- > \frac{1}{2} + \frac{c_d}{2\alpha\gamma_A}$ and $\nu_A^- \leq .5$, which is impossible. This leaves 7 cases to be considered.

I consider now the equilibrium case by case. For the further analysis, it is useful to define:

$$\begin{aligned}
\theta &= \beta\gamma_A\gamma_P\left\{[p\nu_P + (1-p)(1-\nu_P)]\nu_P^+(1-\nu_A^-) + [p(1-\nu_P) + (1-p)\nu_P]\nu_P^-(1-\nu_A^+)\right\} \\
&= \beta\gamma_A\gamma_P\nu_P(1-\nu_A)\left\{\frac{p^2}{(1-p)\nu_A + p(1-\nu_A)} + \frac{(1-p)^2}{p\nu_A + (1-p)(1-\nu_A)}\right\}
\end{aligned}$$

$$\begin{aligned}
\Delta_p &= \theta - \beta\gamma_A\gamma_P\nu_P(1-\nu_A) \\
&= \beta\gamma_A\gamma_P\nu_P(1-\nu_A)\left\{\frac{(2p-1)^2\nu_A^2}{((1-p)\nu_A + p(1-\nu_A))(p\nu_A + (1-p)(1-\nu_A))}\right\}
\end{aligned}$$

$$\Delta_a = \alpha\gamma_P(2\nu_P - 1)$$

and

$$\delta_a = \alpha\gamma_P(\nu_P + p - 1)$$

Note that since $.5 < p < \nu_P$, $0 < \delta_a < \Delta_a$.

The choice for persuasion and/or authority now completely depends on the principal's expected utility. I will use U^P, U_a^P, U_p^P, and U_{ap}^P for the principal's utility (excluding the costs of persuasion or authority) when, respectively, she uses neither authority nor persuasion, she uses only authority, she uses only persuasion, and she uses both authority and persuasion. I now study case by case.

Weak persuasion ($\nu_A^- > .5$), strong authority Remember that in this case, the agent always obeys upon authority but never chooses X absent authority.

If the principal does not use either authority or persuasion, the agent chooses Y and the principal's expected utility becomes:

$$U^P = \alpha\gamma_P(1-\nu_P) + \beta\gamma_A\gamma_P\nu_P(1-\nu_A)$$

If she uses 'authority only', her expected utility becomes:

$$U_a^P = \alpha\gamma_P\nu_P + \beta\gamma_A\gamma_P\nu_P(1-\nu_A)$$

It follows that her gain from authority (excluding c_a) equals

$$U_a^P - U^P = \alpha\gamma_P(2\nu_P - 1) = \Delta_a$$

If she uses 'persuasion only', her expected utility becomes

$$\begin{aligned} U_p^P &= [p\nu_P + (1-p)(1-\nu_P)]\left\{\alpha\gamma_P(1-\nu_P^+) + \beta\gamma_A\gamma_P\nu_P^+(1-\nu_A^-)\right\} \\ &\quad +[p(1-\nu_P) + (1-p)\nu_P]\left\{\alpha\gamma_P(1-\nu_P^-) + \beta\gamma_A\gamma_P\nu_P^-(1-\nu_A^+)\right\} \\ &= \alpha\gamma_P(1-\nu_P) + \beta\gamma_A\gamma_P\nu_P(1-\nu_A)\left\{\frac{p^2}{(1-p)\nu_A + p(1-\nu_A)} + \frac{(1-p)^2}{p\nu_A + (1-p)(1-\nu_A)}\right\} \\ &= \alpha\gamma_P(1-\nu_P) + \theta \end{aligned}$$

Now we can express the gain from persuasion as

$$\begin{aligned} U_p^P - U^P &= \beta\gamma_A\gamma_P\nu_P(1-\nu_A)\left\{\frac{p^2}{(1-p)\nu_A + p(1-\nu_A)} + \frac{(1-p)^2}{p\nu_A + (1-p)(1-\nu_A)} - 1\right\} \\ &= \Delta_p \end{aligned}$$

Furthermore, when choosing between purely authority and purely persuasion, the gain from persuasion is

$$U_p^P - U_a^P = (U_p^P - U^P) - (U_a^P - U^P) = \Delta_p - \Delta_a$$

If the principal uses both authority and persuasion, her expected utility becomes

$$\begin{aligned}
U_{ap}^P &= \alpha\gamma_P\nu_P + [p\nu_P + (1-p)(1-\nu_P)]\left\{\beta\gamma_A\gamma_P\nu_P^+(1-\nu_A^-)\right\} + [p(1-\nu_P) + (1-p)\nu_P]\left\{\beta\gamma_A\gamma_P\nu_P^-(1-\nu_A^+)\right\} \\
&= \alpha\gamma_P\nu_P + \beta\gamma_A\gamma_P\nu_P(1-\nu_A)\left\{p\frac{p}{(1-p)\nu_A + p(1-\nu_A)} + (1-p)\frac{(1-p)}{p\nu_A + (1-p)(1-\nu_A)}\right\} \\
&= \alpha\gamma_P\nu_P + \theta
\end{aligned}$$

The gain from persuasion is then

$$\begin{aligned}
U_{ap}^P - U_a^P &= \alpha\gamma_P\nu_P + \theta - \alpha\gamma_P\nu_P - \beta\gamma_A\gamma_P\nu_P(1-\nu_A) \\
&= \Delta_p
\end{aligned}$$

so that

$$U_{ap}^P - U_a^P - U_p^P + U^P = \Delta_p - \Delta_p = 0$$

Since the criteria for persuasion and authority are completely independent, the equilibria for this case are thus:

- The principal will use persuasion whenever $c_p \leq \Delta_p$.
- The principal will use authority whenever $c_a < \Delta_a$.

Strong persuasion ($\nu_A^- \leq .5 < \nu_A$), strong authority. The outcomes without persuasion are obviously identical to the ones above:

$$\begin{aligned}
U^P &= \alpha\gamma_P(1-\nu_P) + \beta\gamma_A\gamma_P\nu_P(1-\nu_A) \\
U_a^P &= \alpha\gamma_P\nu_P + \beta\gamma_A\gamma_P\nu_P(1-\nu_A) \\
U_a^P - U^P &= \alpha\gamma_P(2\nu_P - 1) = \Delta_a
\end{aligned}$$

The outcome with both persuasion and authority is also identical to the one above:

$$\begin{aligned}
U_{ap}^P &= \alpha\gamma_P\nu_P + [p\nu_P + (1-p)(1-\nu_P)]\left\{\beta\nu_P^+\gamma_A\gamma_P(1-\nu_A^-)\right\} \\
&\quad + [p(1-\nu_P) + (1-p)\nu_P]\left\{\beta\nu_P^-\gamma_A\gamma_P(1-\nu_A^+)\right\} \\
&= \alpha\gamma_P\nu_P + \theta \\
U_{ap}^P - U_a^P &= \Delta_p
\end{aligned}$$

If she uses 'persuasion only', the principal's expected utility is

$$\begin{aligned} U_p^P &= [p\nu_P + (1-p)(1-\nu_P)]\left\{\alpha\gamma_P\nu_P^+ + \beta\gamma_A\gamma_P\nu_P^+(1-\nu_A^-)\right\} \\ &\quad +[p(1-\nu_P) + (1-p)\nu_P]\left\{\alpha\gamma_P(1-\nu_P^-) + \beta\gamma_A\gamma_P\nu_P^-(1-\nu_A^+)\right\} \\ &= [p\nu_P + (1-p)(1-\nu_P)]\left\{\alpha\gamma_P(1-\nu_P^+) + \beta\gamma_A\gamma_P\nu_P^+(1-\nu_A^-)\right\} \\ &\quad +[p\nu_P + (1-p)(1-\nu_P)]\alpha\gamma_P(2\nu_P^+ - 1) \\ &\quad +[p(1-\nu_P) + (1-p)\nu_P]\left\{\alpha\gamma_P(1-\nu_P^-) + \beta\gamma_A\gamma_P\nu_P^-(1-\nu_A^+)\right\} \\ &= \alpha\gamma_P p + \theta \end{aligned}$$

The gain from persuasion is then

$$\begin{aligned} U_p^P - U^P &= \alpha\gamma_P p + \theta - \alpha\gamma_P(1-\nu_P) - \beta\gamma_A\gamma_P\nu_P(1-\nu_A) \\ &= \delta_a + \Delta_p \end{aligned}$$

Furthermore, when choosing between purely persuasion and purely authority, the gain from persuasion is

$$\begin{aligned} U_p^P - U_a^P &= \alpha\gamma_P p + \theta - \alpha\gamma_P\nu_P - \beta\gamma_A\gamma_P\nu_P(1-\nu_A) \\ &= \delta_a - \Delta_a + \Delta_p < \Delta_p \end{aligned}$$

The gain from authority when using persuasion is

$$U_{ap}^P - U_p^P = \alpha\gamma_P\nu_P + \theta - \alpha\gamma_P p - \theta = \Delta_a - \delta_a$$

It follows that

$$U_{ap}^P - U_a^P - U_p^P + U^P = \Delta_p - (\delta_a + \Delta_p) < 0$$

so that persuasion and authority are substitutes.

To find the equilibria, note that if $c_p \leq \Delta_p$ then there will always be persuasion. If $\Delta_p < c_p \leq \delta_a + \Delta_p$ then there will be 'persuasion only' absent authority. And if $\delta_a + \Delta_p < c_p$ then there will never be persuasion.

Furthermore, if $c_a < \Delta_a - \delta_a$ then there will always be authority. If $\Delta_a - \delta_a \leq c_a < \Delta_a$ then there will be 'authority only' when there is no persuasion. Finally, when $c_a \geq \Delta_a$ then there will be no authority at all.

The one issue remaining is what happens when authority and persuasion are both possible but exclusive. In that case, the principal uses persuasion iff

$$U_p^P - U_a^P = \delta_a - \Delta_a + \Delta_p \geq c_p - c_a$$

The equilibria are thus as follows:

- If $c_p \leq \Delta_p$, then there is always persuasion. Moreover, there is also authority iff $c_a < \Delta_a - \delta_a$.
- If $\Delta_p < c_p \leq \Delta_p + \delta_a$ then we have the following:

  - If $c_a < c_p - (\Delta_p - \Delta_a + \delta_a)$ then the principal uses authority.
  - If $c_a \geq c_p - (\Delta_p - \Delta_a + \delta_a)$ then the principal uses persuasion.

- If $c_p > \Delta_p + \delta_a$ then there is never persuasion. If $c_a < \Delta_a$ then there is authority.

Weak persuasion, ν_A-authority The payoffs for U^P, U_p^P, and U_a^P are unchanged from the case with weak persuasion and strong authority:

$$\begin{aligned} U^P &= \alpha\gamma_P(1-\nu_P) + \beta\gamma_A\gamma_P\nu_P(1-\nu_A) \\ U_p^P &= \alpha\gamma_P(1-\nu_P) + \theta \\ U_a^P &= \alpha\gamma_P\nu_P + \beta\gamma_A\gamma_P\nu_P(1-\nu_A) \\ U_p^P - U^P &= \Delta_p \\ U_a^P - U^P &= \alpha\gamma_P(2\nu_P - 1) = \Delta_a \end{aligned}$$

When using both authority and persuasion, the employee only obeys when the signal confirms the principal's belief, so that the principal's expected utility equals:

$$\begin{aligned} U_{ap}^P &= [p\nu_P + (1-p)(1-\nu_P)]\left\{\alpha\gamma_P\nu_P^+ + \beta\gamma_A\gamma_P\nu_P^+(1-\nu_A^-)\right\} \\ &\quad +[p(1-\nu_P) + (1-p)\nu_P]\left\{\alpha\gamma_P(1-\nu_P^-) + \beta\gamma_A\gamma_P\nu_P^-(1-\nu_A^+)\right\} \\ &= \alpha\gamma_P p + \theta \end{aligned}$$

Finally,

$$\begin{aligned} U_{ap}^P - U_p^P &= \alpha\gamma_P(p + \nu_P - 1) = \delta_a \\ U_{ap}^P - U_a^P &= \alpha\gamma_P p + \theta - \alpha\gamma_P\nu_P - \beta\gamma_A\gamma_P\nu_P(1-\nu_A) \\ &= \Delta_p - \Delta_a + \delta_a \end{aligned}$$

so that $U_{ap}^P - U_a^P - U_p^P + U^P = \Delta_p - \Delta_a + \delta_a - \Delta_p < 0$.

The gain from using (only) persuasion rather than (only) authority is

$$\begin{aligned} U_p^P - U_a^P &= U_p^P - U^P - (U_a^P - U^P) \\ &= \Delta_p - \Delta_a \end{aligned}$$

The equilibrium is then as follows:

- If $c_p \leq \Delta_p - \Delta_a + \delta_a$ then always persuasion. Moreover, the principal will also use authority if $c_a < \delta_a$.

- If $\Delta_p - \Delta_a + \delta_a < c_p \leq \Delta_p$ then she will use persuasion if $c_p \leq c_a + \Delta_p - \Delta_a$ and authority otherwise.

- If $\Delta_p < c_p$ then she will never use persuasion but she will use 'authority only' if $c_a < \Delta_a$.

Strong persuasion, ν_A-authority Under persuasion, only ν_A^- would obey when exerting authority, but ν_A^- would comply anyways since persuasion is strong, so that $U_{ap}^P = U_p^P$. It

follows that authority and persuasion are completely exclusive. When using only authority or only persuasion, this case is identical to the case with strong authority (and strong persuasion).

The equilibrium is then as follows:

- If $c_p \leq \Delta_p + \delta_a$ and
  - if $c_a < c_p - (\Delta_p - \Delta_a + \delta_a)$ then authority.
  - If $c_a \geq c_p - (\Delta_p - \Delta_a + \delta_a)$ then persuasion.
- If $c_p > \Delta_p + \delta_a$ then there is never persuasion. If $c_a < \Delta_a$ then there is authority.

It also follows that $U^P_{ap} - U^P_a - U^P_p + U^P = -\Delta_a < 0$.

Weak persuasion, ν^-_A-authority In this case, authority when used alone is never obeyed, so that $U^P_a = U^P$ and authority alone cannot be optimal. On the other hand, the payoffs for U^P and U^P_p are the same as in the case with weak persuasion but strong authority:

$$\begin{aligned} U^P &= \alpha\gamma_P(1-\nu_P) + \beta\gamma_A\gamma_P\nu_P(1-\nu_A) \\ U^P_p &= \alpha\gamma_P(1-\nu_P) + \theta \\ U^P_p - U^P &= \Delta_p \end{aligned}$$

The principal's utility when using both persuasion and authority equals

$$\begin{aligned} U^P_{ap} &= [p\nu_P + (1-p)(1-\nu_P)]\left\{\alpha\gamma_P\nu^+_P + \beta\gamma_A\gamma_P\nu^+_P(1-\nu^-_A)\right\} \\ &\quad +[p(1-\nu_P) + (1-p)\nu_P]\left\{\alpha\gamma_P(1-\nu^-_P) + \beta\gamma_A\gamma_P\nu^-_P(1-\nu^+_A)\right\} \\ &= \alpha\gamma_P p + \theta \end{aligned}$$

so that

$$\begin{aligned} U^P_{ap} - U^P &= \Delta_p + \delta_a \\ U^P_{ap} - U^P_p &= \alpha\gamma_P p - \alpha\gamma_P(1-\nu_P) = \delta_a \end{aligned}$$

and thus

$$U^P_{ap} - U^P_p - U^P_a + U^P = \delta_a > 0$$

So the equilibrium is then as follows:

- If $c_p \leq \Delta_p$ then the principal always uses persuasion. Moreover, the principal will also use authority if $c_a < \delta_a$.
- If $c_p > \Delta_p$, then the principal will use both authority and persuasion if $c_p + c_a < \Delta_p + \delta_a$. Otherwise the principal uses neither of the two.

Strong persuasion, ν_A^--authority In this case, authority has no effect (since an agent with belief ν_A does not obey, while an agent with ν_A^- would obey but complies anyways even without authority), and is thus never optimal. It also follows that $U_{ap}^P = U_p^P$ and $U_a^P = U^P$ so that $U_{ap}^P - U_p^P - U_a^P + U^P = 0$. The utilities for U^P, U_p^P, and thus $U_p^P - U^P$ are the same as in the case of 'strong persuasion, strong authority'. It follows that the equilibrium is that the principal will use persuasion iff $c_p \leq \Delta_p + \delta_a$.

Weak persuasion, no authority With no authority, authority will never be used and $U_{ap}^P = U_p^P$ and $U_a^P = U^P$ so that $U_{ap}^P - U_p^P - U_a^P + U^P = 0$. The utilities for U^P, U_p^P, and thus $U_p^P - U^P$ are the same as in the case of 'weak persuasion, strong authority'. It follows that the equilibrium is that the principal will use persuasion iff $c_p \leq \Delta_p$.

A.2 Proofs

Proof of Proposition 1: It suffices to show for each of the cases with $c_d < \alpha\gamma_A(2\nu_A - 1)$ that $U_{ap}^P - U_p^P - U_a^P + U^P \geq 0$ and for each of the cases with $c_d \geq \alpha\gamma_A(2\nu_A - 1)$ that $U_{ap}^P - U_p^P - U_a^P + U^P \leq 0$. These follow from the derivations of the equilibria. That concludes the proof. ■

Proof of Proposition 2: Note first that the boundaries between the different cases are independent of β. It thus suffices to show this case by case. Note also that Δ_p increases in β, while Δ_a and δ_a are independent of β. It follows that we can simply consider the effect of an increase in Δ_p. In what follows, I will assume that c_p changes along the horizontal axis and c_a changes along the vertical axis when describing regions in the parameter space.

The result is straightforward for the cases with weak persuasion and either strong authority or no authority and for the case with strong persuasion and ν_A^--authority.

Consider next the case with strong persuasion and strong authority. Points that are originally in the $c_p \leq \Delta_p$ region are not affected by an increase in Δ_p. Points that are originally in the $c_p > \Delta_p + \delta_a$ either remain in that region and are not affected or become part of the middle region. But any such point with $c_a \geq \Delta_a$ then falls in the persuasion part. Any such point with $c_a < \Delta_a$ either remains 'authority only' or becomes 'persuasion only'. It follows that the Proposition holds for these two regions. Consider finally points that are in the middle region. Any such point that satisfied $c_a \geq c_p - (\Delta_p - \Delta_a + \delta_a)$ and that remains in the middle region will still satisfy that condition. Any such point that satisfied $c_a < c_p - (\Delta_p - \Delta_a + \delta_a)$ and that remains in the middle region either remains 'authority only' or becomes 'persuasion only'. The result thus also holds for such points. Consider finally any point from the middle region that goes to the left region. Any such point that satisfies $c_a \geq c_p - (\Delta_p - \Delta_a + \delta_a)$ will then satisfy $c_a \geq \Delta_a - \delta_a$ (since at the boundary between the left and middle region, $c_p = \Delta_p$) and thus remain in the 'persuasion only' part. Any such point that satisfies $c_a < c_p - (\Delta_p - \Delta_a + \delta_a)$ either becomes 'persuasion only' or persuasion and authority. It follows that the Proposition holds for such points. The proof

for the cases with ν_A-authority are completely analogous. That proves the Proposition. ■

Proof of Proposition 3: Since Δ_p increases in γ_A while Δ_a and δ_a are independent of γ_A, the proof that this holds true on a case by case basis is analogous to the proof of Proposition 2.

However, γ_A does affect the boundaries between the different authority cases, so that I now have to show that the Proposition also holds when going from one case to another. An increase in γ_A reduces authority. Consider first the case with weak persuasion. With strong authority, the principal uses 'persuasion only' when $c_p \leq \Delta_p$ and $c_a \geq \Delta_a$. With ν_A-authority, she will never use persuasion when $c_p > \Delta_p$. Moreover, when $c_p \leq \Delta_p$, she will use 'persuasion only' at least when $c_a \geq \Delta_a$ and she will never use 'persuasion only' when $c_a < \delta_a$. With ν_A^--authority, she will never use 'persuasion only' when $c_p > \Delta_p$. Moreover, when $c_p \leq \Delta_p$, she will use 'persuasion only' iff $c_a \geq \delta_a$. Finally, with no authority she will use only persuasion iff $c_p \leq \Delta_p$. This proves the result for all transitions between cases with weak persuasion. The argument for strong persuasion is analogous. That proves the Proposition. ■

Proof of Proposition 4: Note first that the boundaries between the different cases are independent of ν_P. It thus suffices to show this case by case.

The statement for $c_d < \alpha\gamma_A(2\nu_A - 1)$ follows directly from the fact that Δ_p, Δ_a, and δ_a all increase in ν_P.

The statement for $c_d \geq \gamma_A(2\nu_A - 1)$ in the case with weak persuasion and strong authority also follows directly from the fact that Δ_p increases in ν_P. Consider next the case with strong persuasion and strong authority. Points that are originally in the $c_p \leq \Delta_p$ region remain there and thus keep persuasion. Points that are originally in the $c_p > \Delta_p + \delta_a$ can only increase in persuasion. Consider finally points that are in the middle region. Any such point that goes to the left region increases in persuasion. Any such point that remains in the middle region and that satisfied $c_a \geq c_p - (\Delta_p - \Delta_a + \delta_a)$ will still satisfy that condition if

$$\frac{d(\Delta_p - \Delta_a + \delta_a)}{d\nu_P} = \beta\gamma_A\gamma_P(1-\nu_A)\left\{\frac{(2p-1)^2\nu_A^2}{((1-p)\nu_A + p(1-\nu_A))(p\nu_A + (1-p)(1-\nu_A))}\right\} - \alpha\gamma_P \geq 0$$

or if

$$\frac{\beta}{\alpha} \geq \frac{((1-p)\nu_A + p(1-\nu_A))(p\nu_A + (1-p)(1-\nu_A))}{\gamma_A(1-\nu_A)(2p-1)^2\nu_A^2}$$

so that the result holds for that case if

$$\epsilon \geq \frac{((1-p)\nu_A + p(1-\nu_A))(p\nu_A + (1-p)(1-\nu_A))}{\gamma_A(1-\nu_A)(2p-1)^2\nu_A^2}$$

Given the similarity of their equilibria, the same condition also works for strong persuasion and ν_A-authority. So consider finally weak persuasion and ν_A-authority. The above condition guaranteed that $\Delta_p - \Delta_a + \delta_a$ increases, so that points originally in the $c_p \leq \Delta_p - \Delta_a + \delta_a$

remain there and thus keep persuasion. Points that are originally in the $c_p > \Delta_p$ can only increase in persuasion. Points that are originally in the $\Delta_p - \Delta_a + \delta_a < c_p \leq \Delta_p$ and move to the left region also can only increase in persuasion. All that remains is to consider points in that middle region that stay in that middle region. Any such point that satisfied $c_a \geq c_p - (\Delta_p - \Delta_a)$ will still satisfy that condition if

$$\frac{d(\Delta_p - \Delta_a)}{d\nu_P} = \beta\gamma_A\gamma_P(1-\nu_A)\left\{\frac{(2p-1)^2\nu_A^2}{((1-p)\nu_A + p(1-\nu_A))(p\nu_A + (1-p)(1-\nu_A))}\right\} - 2\alpha\gamma_P \geq 0$$

or if

$$\frac{\beta}{\alpha} \geq \frac{2((1-p)\nu_A + p(1-\nu_A))(p\nu_A + (1-p)(1-\nu_A))}{\gamma_A(1-\nu_A)(2p-1)^2\nu_A^2}$$

so that the result holds if

$$\epsilon \geq \frac{2((1-p)\nu_A + p(1-\nu_A))(p\nu_A + (1-p)(1-\nu_A))}{\gamma_A(1-\nu_A)(2p-1)^2\nu_A^2}$$

which also implies the earlier condition. This concludes the Proposition. ■

References

Aghion, P., and J. Tirole (1997): "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105(1), 1– 29.

Arrow, K. J. (1974): *The Limits of Organization*. W. W. Norton & Co, New York.

Athey, S., and J. Roberts (2001): "Organizational Design: Decision Rights and Incentive Contracts," *American Economic Review, Papers and Proceedings*, 91, 200– 205.

Benabou, R. (2008): "Ideology," *Journal of the European Economic Association*, 6(2), 321– 352.

Benabou, R., and J. Tirole (2002): "Self-Confidence and Personal Motivation," *Quarterly Journal of Economics*, 117(3), 871– 915.

——— (2006): "Belief in a Just World and Redistributive Politics," *Quarterly Journal of Economics*, 121(2), 699– 746.

Che, Y.-K., and N. Kartik (2007): "Opinions as Incentives," Working Paper.

Crawford, V., and J. Sobel (1982): "Strategic Information Transmission," *Econometrica*, 50, 1431– 1452.

Dessein, W. (2007): "Why a Group Needs a Leader: Decision-making and Debate in Committees," Working Paper.

Dessein, W., L. Garicano, and R. Gertner (2005): "Organizing for Synergies: Communication, Incentives, and Centralized Conflict Resolution," Mimeo, University of Chicago.

Dewatripont, M., and J. Tirole (1999): "Advocates," *Journal of Political Economy*, 107(1), 1– 39.

Kremer, M. (1993): "The O-Ring Theory of Economic Development," *Quarterly Journal of Economics*, 108(3), 551– 575.

Milgrom, P. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12, 380– 391.

Milgrom, P., and J. Roberts (1986): "Relying on the Information of Interested Parties," *Rand Journal of Economics*, 17(1), 18– 32.

Morris, S. (1995): "The Common Prior Assumption in Economic Theory," *Economics and Philosophy*, 11, 227– 253.

Roberts, J. (2004): *The Modern Firm: Organizational Design for Performance and Growth*. Oxford University Press, New York.

Rosen, S. (1982): "Authority, Control, and the Distribution of Earnings," *Bell Journal of Economics*, 13(2), 311– 323.

Simon, H. (1947): *Administrative Behavior*. Free Press, New York.

Van den Steen, E. (2002): "Disagreement and Information Collection," Working paper.

——— (2004): "The Costs and Benefits of Homogeneity, with an Application to Culture Clash," Working paper.

——— (2006): "The Limits of Authority: Motivation versus Coordination," Sloan Working Paper 4626-06.

——— (2007): "Interpersonal Authority in a Theory of the Firm," Sloan Working Paper 4667-07.

Yildiz, M. (2000): "Essays on Sequential Bargaining," PhD Dissertation, Stanford University.