Incomplete Contracts versus Communication*

Andreas Blume, Inga Deimen, and Sean Inoue

August, 2022

Abstract

We consider a principal's choice between either controlling an agent's action through an incomplete contract or guiding him through non-binding communication. The principal anticipates receiving private information and must hire an agent to take an action on her behalf. Contracts can only specify a limited number of actions as a function of the state. The principal is at liberty not to specify actions for some of the states. States not covered by the contract induce a communication game. Contract clauses create gaps in the state space of the communication game, which can be used to generate distance between communication events. This relaxes incentive constraints for communication, helping enable and structure influential communication. We find that close alignment of interests favors communication and, thus, ceding authority to the agent, while strong misalignment favors reliance on contracts. In the uniform-quadratic environment, optimal contracts that induce influential communication split the communication region: there are at least two communication actions separated by contract actions. For sufficiently closely aligned interests, it is also the case that communication splits the contract region: there are at least two contract actions separated by a communication action.

JEL: D83, D82
Keywords: *strategic communication, cheap talk, incomplete contracts*

*Blume: University of Arizona and CEPR, ablume@arizona.edu, Deimen: University of Arizona and CEPR, ideimen@arizona.edu, Inoue: Colgate University, sinoue@colgate.edu. University of Arizona, Eller College of Management, Department of Economics, 1130 E. Helen St, Tucson, AZ 85721. We are grateful for comments from Pierpaolo Battigalli, Francesc Dilmé, Joel Sobel, Joel Watson, Bill Zame, and seminar audiences at the University of Arizona, ASSA (2021), Bonn Berlin Micro Theory Seminar Series, CRETA Economic Theory Conference (Warwick), University of Düsseldorf, GAMES 2020, Organisationsökonomischer Ausschuss, International Conference on Game Theory 2020 (Stony Brook), Pennsylvania State University, QuantLaw Conference, SWET (UCSD), Tinbergen Institute, UCSD Theory Foundations Conference.

1 Introduction

Principals (firms, organizations, researchers, innovators) frequently have to rely on a skilled agent to perform a task. Structuring that relationship is challenging when the task is complex and the principal does not have all relevant information at her disposal at the time the agent is hired. Contracts tend to be incomplete and the principal has to decide how much discretion to leave to the agent. Conflicts of interest favor limiting discretion. The flexibility afforded by being able to respond to newly arriving information with non-binding communication favors discretion. This motivates our investigation of when and how cheap talk is used as a substitute for contractual control of the flow of information.

Typical contracts we have in mind are between a service provider (agent) and a client (principal) in situations with specification uncertainty: a production company hires an IT firm to develop customized software; a property owner relies on a contractor to oversee construction; or, a researcher employs a laboratory technician to run experiments. These services are performed in highly complex environments that require a specialized language that makes it hard to write complete contracts understandable to third parties.[1] They are prone to modifications, with new information becoming available over time: software development needs to be adopted to changes in demand; property owners have changing needs and preferences that may require alterations in building plans and materials; new research results may require the redesign of experiments. While it is the agent who executes the task, the new information is typically first observed by the principal who has advance knowledge of the overall situation.[2]

We present a parsimonious model of the interplay of incomplete contracting and non-binding communication. A principal contracts with an agent for an action she cannot take herself. For each state of the world, the principal's preferred action differs from that of the agent. Prior to the realization of the state, the principal offers a contract that consists of a bounded number of clauses. Each clause specifies an action to be taken for a subset of the state space. These *contract actions*, i.e., actions that are mandated by a contract clause, are enforceable. Once the state realizes, if it is covered by a contract clause, the principal issues an instruction to perform the action the contract prescribes for that state. Ex post it can be verified that the instruction was appropriate for the realized state and that the agent followed the instruction.

The principal has the option not to have the contract cover the entire state space. At

[1]Crocker and Reynolds (1993) observe in their study of airforce engine procurement that "In practice, . . . , the costs of identifying contingencies and devising responses increase rapidly in complex or uncertain environments, placing economic limits on the ability of agents to draft and implement elaborate contractual agreements." As a result, one sees "agreements that are left intentionally incomplete with regard to future duties or contingencies." Regarding the difficulty of writing contracts that third parties can comprehend, Banerjee and Duflo (2000) observe that in the customized software industry "The extent to which contracts can protect . . . is limited by the fact that the desired end-product tends to be complex and difficult to describe ahead of time in a way that a court . . . would understand . . . ."

[2]In the context of the construction industry, for example, Chakravarty and MacLeod (2006) refer to the owner as the principal and the contractor as the agent.

states not covered by the contract a communication game is played once the state realizes. In this communication game, the state is the principal's private information, her type. After learning her type, she sends a message to the agent, who then takes an action. *Communication actions*, i.e., actions that are induced by non-binding communication, are not enforceable. Communication is cheap talk as in Crawford and Sobel (1982) (henceforth CS).[3]

The model captures four key features of the interaction between a client (principal) and a service provider (agent): (i) The principal anticipates receiving private information relevant to the action that will be taken by the agent. (ii) Hiring and explaining the task to the agent takes time and the agent may need to invest in task-specific capabilities. Hence, the principal cannot defer hiring the agent until all information is available and then offer a spot contract. Instead, she offers a long-term contract that creates a framework for delivering information to the agent. (iii) There are exogenous constraints on the principal's ability to contractually tie the agent's action to the state. As a result, contracts are never *fully detailed complete.* (iv) The principal may choose contracts that are *obligationally incomplete*, i.e., do not cover all states.[4]

We have in mind the following type of situation (motivated by a case study conducted by Mayer and Argyres (2004), see also Eckfeldt, Madden and Horowitz (2005)): A firm, the principal, contracts with a software developer, the agent, for a project. At the time the contract is agreed upon there is uncertainty about the needs of the firm. The firm faces uncertainty due to changing technology, consumer tastes, and competitor behavior in its downstream market. The needs of the firm determine the state of the world. The firm and the software developer disagree about the scope of the project. For any given need, the firm prefers more functionality to address that need, while it is costly for the software developer to add functionality. This biases the firm relative to the software developer: for any given state the firm prefers a higher action than the software developer.

The firm anticipates that during the course of the project it will be learning about the changes in technology, consumer tastes, and competitor behavior in its downstream market, which then become its private information. The firm's contractual relationship with the software developer regulates the way in which this information is brought to bear on the software developer's work: some of it formally, as articulated in the contract, and some informally, left to non-binding communication: In the case examined by Mayer and Argyres the contract contains a "statement of work" (SOW) with instructions for the work to be performed by the software developer. Recognizing the uncertainties facing the project, the SOW includes limited forms of contingency planning in the form of an "engineering change procedure" for changes initiated by the firm. As Mayer and Argyres point out, this procedure is not always followed, but creates a framework that structures formal and informal communication.

In the model, we make a number of simplifying assumptions: (1) Contract negotiation is

[3]Since it can be verified ex post whether an instruction that was issued was appropriate for the realized state, the principal cannot disguise a cheap-talk message as an instruction and hope thereby to induce an action specified in the contract.

[4]Ayres and Gertner (1992) introduce this terminology.

reduced to having the principal make a take-it-or-leave-it offer; (2) limited forms of contingency planning are captured by bounding the number of contract clauses; and, (3) informal communication takes the form of a single message sent by the principal whenever a state is not covered by a contract clause. This makes it possible for us to maintain a tight focus on the tradeoff between (formal) contracting and (informal) communication, and the way contracts are used to structure communication.

We are interested in how the interplay of contracting and (non-binding) communication impacts the scope and structure of contracting. Obligational incompleteness in our model is endogenous. When determining the extent of this incompleteness, the principal faces a tradeoff: a more complete contract gives her more direct control over the agent but sacrifices flexibility that could be exploited through communication.

In our analysis, we focus on *principal-optimal subgame-perfect equilibria*, i.e., equilibria that maximize the principal's payoff among all subgame-perfect equilibria. We refer to the contracts offered in those equilibria as *(principal) optimal contracts*. For these optimal contracts, we are interested in the size and structure of the *contract region*, the set of states covered by the contract, and the *communication region*, the set of states that induce non-binding communication.

Under general distributional and payoff assumptions, we find that optimal contracts have three intuitive properties: (i) They use the maximal number of available clauses and thus are as detailed as possible. (ii) As the number of available clauses grows without bound, the contract region converges toward occupying the entire the state space, i.e., contracting crowds out communication. (iii) If we let the interests of principal and agent converge toward perfect alignment, the communication region converges toward occupying the entire the state space, i.e., communication drives out contracting.

For the uniform-quadratic specification of the model we can be more specific.[5] If the conflict of interest between principal and agent is large relative to the available number of clauses, it is optimal to write an obligationally complete contract – there is no communication. In contrast, with closely aligned incentives it is optimal to leave a gap in the contract – there is communication. Furthermore, whenever there is communication, it induces at least two actions, i.e., communication is influential.

The principal's payoff is strictly decreasing in the bias for low values of the bias and constant in the bias for high values; it is strictly increasing in the number of clauses. To get a sense of the comparative statics of the size of the contract region, we derive an upper and a lower bound for that size in optimal contracts. We examine how these bounds vary with the level of conflict and the number of available clauses. We find that the lower bound is nondecreasing in the bias. The upper bound is nondecreasing in the number of available clauses and (in our numerical illustrations) in the bias. This is in line with the intuition that both an increasing bias and a larger number of available clauses favor contracting over communication.

Our primary interest is in the structure of optimal equilibria and optimal contracts. We

[5]The version of the model with quadratic payoff functions, a uniform distribution over state, and a constant bias matches CS's leading example, which has been a workhorse in the applied cheap-talk literature.

show that optimal equilibria are interval partitional and monotonic: Every set of types inducing a common action, be it a contract action or a communication action, forms an interval (for contract actions this holds by fiat). Higher types induce higher actions. We obtain two "splitting results": (i) Whenever an optimal contract induces influential communication, it splits the communication region: there are at least two communication actions separated by contract actions. (ii) For a modest lower bound on the number of available contract clauses and sufficiently closely aligned interests, communication splits the contract region: there are at least two contract actions separated by a communication action.

The rationale for our splitting results can be illustrated with only three states. Suppose the principal anticipates that she will learn the realization of one of three ordered states. The preferences of the principal and any agent she may hire to take the action are misaligned: the agent prefers the action to match the state, whereas the principal prefers the action to match the nearest higher state. This makes truthful communication of all states impossible. Suppose the principal can impose an incomplete contract that (only) mandates that the principal's favorite action is taken for the intermediate state. With such a contract in place, the remaining two states can now be truthfully communicated; this explains the splitting of the communication region. In essence, splitting the communication region this way generates common interest for communicating the extreme states. With more states, it becomes tempting and sometimes feasible to replicate this effect for other groupings of states; this explains the splitting of the contract region.

In general, the principal can use contract clauses to separate events that induce distinct communication actions. This is beneficial because it relaxes incentive constraints in the communication subgame. In the uniform-quadratic specification of the model, this relaxation of incentive constraints makes it possible to make the sizes of communication intervals more equal. Since quadratic loss functions disproportionately penalize actions that are more distant from players' ideal points, and equalizing the sizes of communication intervals amounts to substituting smaller for larger distances from ideal points, this increases expected payoffs. This highlights the dual role of contracting as both substituting for and facilitating communication – the principal uses contracts not only to impose her favorite actions, but also to structure communication.

The paper is organized as follows. Section 2 presents a simple example. In Section 3, we introduce the model and discuss our modeling choices. In Section 4, we study properties of optimal equilibria in a general framework. We derive further properties of these equilibria in Section 5 under the assumption of having a uniform distribution over states and quadratic payoff functions with a constant bias. Section 6 offers additional examples of optimal contracts and extensions. We discuss related literature in Section 7. In the final section, we summarize our findings and suggest possible directions for future work. All proofs are in the appendix.

2 A simple example

This section presents the simplest version of our contracting game. It illustrates some of the key features of the interplay between incomplete contracting and communication: there is an equilibrium in which the principal offers a contract that leaves gaps for, structures, and improves communication.

A principal needs to hire an agent to take an action on her behalf. The ex post payoffs of the principal, $U^P(y, \theta, b) = -(\theta + b - y)^2$, and the agent, $U^A(y, \theta) = -(\theta - y)^2$, depend on the agent's action, $y \in \mathbb{R}$, the state of the world, θ, and a parameter $b > 0$ that measures the principal's bias relative to the agent. The state of the world is commonly known to be uniformly distributed on $[0, 1]$.

Prior to learning the state of the world, the principal specifies a contract consisting of a single condition and an instruction. The condition C has to be an interval in $[0, 1]$ and the instruction x an action in $\mathbb{R}$. If it turns out that the state belongs to the condition C, then the contract requires the agent to execute the instruction by taking the action $y = x$. Otherwise, the principal sends a recommendation, which is a cheap-talk message, to the agent who responds with an action y of his choice.

Suppose that $b = 1/3$. With a bias of this size, we know from Crawford and Sobel (1982) (p. 1441) that if the principal had to rely on cheap talk alone, she would be unable to influence the agent's action: in every equilibrium of the cheap-talk game the agent would take a single action, irrespective of the state of the world. This action, $y = 0.5$, would minimize the agent's expected quadratic loss given the prior. In that case, the principal's expected payoff would be -0.194.

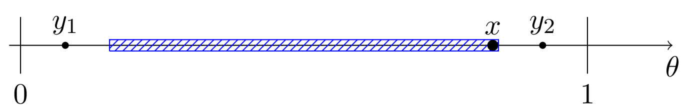

Figure 1: The optimal contract with one contract clause for a bias $b = \frac{1}{3}$.

In contrast, with the ability to write a contract, as described, it becomes possible to achieve influential communication and raise the principal's payoff. With $b = 1/3$, the contracting game has an equilibrium in which the principal offers a contract with the condition $C = [0.157, 0.843]$ and the instruction $x = 0.833$. In this equilibrium, if $\theta < 0.157$ the principal sends a cheap talk message that induces the agent to take action $y_1 = 0.079$; if $\theta = [0.157, 0.843]$ the contract compels the agent to take action $x = 0.833$, and if $\theta > 0.843$ the principal sends a cheap talk message that induces the agent to take action $y_2 = 0.921$. See Figure 8, for an illustration; the derivation and additional details are in Section 6. The principal's expected payoff given that contract and equilibrium equals -0.062, which exceeds -0.194, the maximal payoff achievable with cheap talk alone.

This equilibrium is optimal for the principal. The contract that achieves optimality is unique. It enables influential communication by creating a wedge between sets of states that are sufficiently far apart. With the contract in place, it is in the principal's and agent's common interest to communicate whether the state is above or below the contract's condition.

The example captures the following stylized features of a firm, the principal, contracting with a software developer, the agent, for a project. The project could be a software tool, where the adequate degree of functionality is uncertain ex ante. For a given adequate functionality (the state θ), the software developer has an incentive to provide that degree of functionality through the action y. He may, for example, suffer from a reputational loss if there is a considerable mismatch between the two. The principal prefers more functionality, up to a point – functionality that will be rarely used or requires excessive training does not add value. Together, these two features of preferences imply that there is a bias, which is bounded. In our example this bias has a constant value, b. The firm's incentive to exaggerate her need for functionality creates a barrier for communication. The bias is large enough (b is greater than $1/4$) that it is impossible to link the chosen functionality to the adequate functionality via communication alone. The contract fixes functionality for a medium range of states. As a consequence, only the extremely low and high states remain possible for communication. For extremely low states, it is unattractive for the firm to pretend that the state is extremely high, because the bias is bounded. Extremely high states can be communicated, because the firm is biased towards even higher states. As a result it becomes possible to link functionality to the state: chosen functionality varies, depending on whether the state is low, medium, or high. This improves payoffs relative to trying to rely solely on communication or having a contract that fixes a single functionality that applies to all states.

3 Model

We consider a game between a principal, P, and an agent, A. They interact in two phases. In the first phase, prior to receiving private information about the state of the world, the principal writes a contract. The contract determines how information is dealt with in the second phase. It consists of a list of contract clauses. Each clause specifies a subset of the state space and an action: the agent must take that action whenever the state realization belongs to the subset of states specified by the clause.[6] We assume that there is a finite bound on the number of clauses in any contract. The principal may elect to write a contract that does not cover all states of the world. In the second phase, for any state covered by the contract, the agent takes the action that is prescribed by the clause covering that state. For states not covered by the contract, a communication game is played.

The payoff and information structure closely follows CS. Players' payoffs, $U^P(y, \theta, b)$ for the principal and $U^A(y, \theta)$ for the agent, depend on the agent's action $y \in \mathbb{R}$, the state of the world $\theta \in [0, 1]$, and a parameter $b > 0$ that measures the divergence of preferences between the principal and the agent. For notational convenience, we sometimes suppress the dependence of the principal's payoff on the bias parameter b and write $U^P(y, \theta)$ instead of

[6]Note that we assume that there are no transfers, and therefore there is no incentive provision through contingent transfers. We briefly discuss non-contingent transfers in an example in Section 6.3.

$U^P(y, \theta, b)$. The state is drawn from a common prior distribution F with continuous density f that is positive everywhere; $f(\theta) > 0$ for all $\theta \in [0, 1]$. The payoff functions U^i, $i = P, A$, are assumed to be twice continuously differentiable. Denoting derivatives by subscripts, we assume that the payoff functions are strictly concave in the agent's action, $U^i_{11} < 0$; the sorting conditions $U^i_{12} > 0$ hold; and, for all θ, there are actions $y^P(\theta, b)$ and $y^A(\theta)$ such that $U^P_1(y^P(\theta, b), \theta) = 0$ and $U^A_1(y^A(\theta), \theta) = 0$, respectively. We assume that the principal's ideal point exceeds the agent's ideal point, i.e., $y^P(\theta, b) > y^A(\theta)$ for all $\theta \in [0, 1]$; the sorting conditions ensure that these ideal points are strictly increasing in the state.

At the beginning of the *contract-writing game* G, the principal writes a *contract* $\mathcal{C} = \{(C_k, x_k)\}_{k=1}^K$. The contract specifies K *clauses* (C_k, x_k), $k = 1, \ldots, K$. There is an exogenous maximal number of clauses $\widehat{K}$.[7] Each clause (C_k, x_k) consists of a *condition* $C_k \subseteq [0, 1]$ and an *instruction* $x_k \in \mathbb{R}$. The interpretation is that if condition C_k holds – i.e., $\theta \in C_k$ is realized – then the agent is instructed to take the action $y = x_k$. Contracts must satisfy: $C_{k'} \cap C_{k''} = \emptyset$ for all $k' \neq k''$ (to avoid contradictions); C_k is an interval for each $k = 1, \ldots, K$ (motivated by keeping contracts simple);[8] and the *contract region* $\bigcup_{k=1}^K C_k$ is a closed set. Denote the lower (upper) endpoint of the interval C_k by $\underline{C}_k$ ($\overline{C}_k$). We allow for an empty contract without clauses, in which case we adopt the convention that $K = 0$. An *obligationally complete contract* covers the entire state space, in which case $\bigcup_{k=1}^K C_k = [0, 1]$. Denote the set of all contracts by $\mathfrak{C}$. Sometimes, it will be convenient to highlight the maximal number of clauses and the principal's bias, in which case we make the dependence on these parameters explicit and write $G(\widehat{K}, b)$ for the contract-writing game.

After the contract $\mathcal{C}$ is written and observed by the agent, the state θ is realized and privately observed by the principal. For any state covered by the contract – e.g., $\theta \in C_{k'}$ – the instruction stipulated for that state, $x_{k'}$, is implemented. For any state not covered by the contract $\mathcal{C}$, the principal sends a message $m \in M$ to the agent, where M is an infinite measurable space. After observing the principal's message, the agent takes an action $y \in \mathbb{R}$.

Every contract $\mathcal{C}$ induces a *communication subgame*, $\Gamma^{\mathcal{C}}$, in the event that the state θ belongs to the *gap* $\mathcal{L}(\mathcal{C}) := [0, 1] \setminus \bigcup_{k=1}^K C_k$ in the contract, i.e., $\theta \in \mathcal{L}(\mathcal{C})$. In this communication subgame, the commonly known type distribution $F^{\mathcal{C}}$ is the prior F concentrated on the set $\mathcal{L}(\mathcal{C})$. If the contract $\mathcal{C}$ is empty, we denote the resulting communication subgame by Γ^0. The game Γ^0 is simply a CS game. A (behavior) strategy $\sigma : \mathcal{L}(\mathcal{C}) \to \Delta(M)$ of the principal in the communication subgame $\Gamma^{\mathcal{C}}$ maps states to distributions over messages. A strategy $\rho : M \to \mathbb{R}$ for the agent in $\Gamma^{\mathcal{C}}$ maps messages to actions. Given the strict concavity of the agent's utility in his action, the restriction to pure agent strategies is without loss of generality. A strategy of the principal $\left(\mathcal{C}; (\sigma^{\mathcal{C}'})_{\mathcal{C}' \in \mathfrak{C}}\right)$ in the contract-writing game G specifies a contract $\mathcal{C}$ and for every possible communication subgame $\Gamma^{\mathcal{C}'}$ a strategy $\sigma^{\mathcal{C}'}$. A strategy of the agent $\left((\rho^{\mathcal{C}'})_{\mathcal{C}' \in \mathfrak{C}}\right)$ in the game G specifies a strategy $\rho^{\mathcal{C}'}$ for every possible communication subgame $\Gamma^{\mathcal{C}'}$. We are interested in subgame-perfect equilibria of the contract-writing game

[7] This corresponds to a limiting case of having a writing cost function that is increasing in the number of clauses (see, e.g., Dye (1985)). Writing costs are zero for the first $\widehat{K}$ clauses and prohibitive thereafter.

[8] Sets other than intervals require more detailed and, therefore, more costly descriptions.

$G(\widehat{K}, b)$ that maximize the principal's payoff among all subgame-perfect equilibria. We refer to any such equilibrium $e(\widehat{K}, b)$ as an *optimal equilibrium* and the contract chosen in that equilibrium as an *optimal contract*.

3.1 Discussion of modeling choices

Contractual incompleteness in our model originates in the difficulty of clarifying language to the point where it becomes comprehensible to third parties. Thus, our focus differs from the frequently invoked observable-but-not-verifiable rationale for contracts being incomplete. Indeed, we assume that adherence to contract clauses that are written in a language that the court understands can be verified and will be enforced by the court. Contracts in our model are not constrained by lack of verifiability per se, but by the cost of creating the appropriate language and putting in place measurement and record keeping systems that make the language meaningful. In this subsection, we discuss how we implement this perspective.

Timing, observability, and verifiability. At the ex ante stage, information is symmetric. The state is neither observable nor verifiable. Instead, the principal and the agent have a common prior over its future realizations. The principal writes a contract prior to observing the state. There are two reasons for doing so: (a) The agent needs to be present when the action has to be taken, and the hiring process takes time. (b) Principal and agent may need to make relationship-specific investments prior to the realization of the state that enable the agent to execute the principal's instructions. These could include the principal providing the agent with job-specific training or with firm-specific information, and the agent investing in project-specific capabilities. The contract helps protect such investments. Both of these factors create incentives for having a long-term contract, rather than waiting for the state to be realized and then writing a spot contract.[9]

Having to face ex ante uncertainty is characteristic of many procurement problems, including jet engine procurement (Crocker and Reynolds (1993)), the customized software industry (Banerjee and Duflo (2000)), and the construction industry (Bajari and Tadelis (2001)). According to Bajari and Tadelis "...uncertainty about many important design changes that occur after the contract is signed and production begins ..." is a common feature of procurement contracting. They note that because of this design uncertainty "the procurement problem is primarily one of ex post adaptations rather than ex ante screening."

Our contracts govern how the principal's anticipated information is brought to bear on the agent's future action. The agent has no private information and the principal anticipates no problem observing the agent's action. Hence, there are no adverse selection or moral hazard problems.

[9]Spot contracting is also not free from language concerns. Therefore, it may not be possible to write a satisfactory spot contract at a moment's notice. If instead the language for the spot contract is developed in advance, it will be necessary to endow it with the ability to prescribe different actions in advance. This raises the complexity of spot contracts, making them more akin to the types of long-term contracts we consider here.

At the interim stage, information becomes asymmetric: The principal privately learns the state, e.g., changes in technology or conditions in its downstream market. The state remains unverifiable and to the agent unobservable. The principal, after observing the state, either issues one of the instructions specified in the contract or sends a cheap-talk message to the agent. Knowing the contract, the players can differentiate between these. After observing the principal's instruction or message, the agent takes an action. At this stage, the principal would breach the contract if she did not issue an instruction required by the contract given the state or tried to mimic an instruction even though the state is not covered by the contract. The agent would breach the contract if he failed to taken an action matching an instruction given by the principal that is prescribed by the contract.

We assume that ex post, the principal and the agent again have symmetric information. The agent learns the realized state and the principal learns the action taken by the agent. Thus both state and action become observable. The agent learns the state through his payoff and through having completed the task (taken the action), which helps him understand the problem the task was meant to address.[10] A software developer (the agent), for example, in addressing his client's (the principal's) needs, learns about the environment in which the client is operating. Likewise the client learns the software developer's action through the product she receives.

While action and state are observable to principal and agent, they are not necessarily verifiable by a court. What is verifiable is a function of the contract. If a court gets involved, it can verify whether the state is covered by a clause in the contract, and in that case whether that clause was adhered to or a breach occurred. The writing of the contract establishes a language that makes conditions and instructions meaningful for the court; record keeping systems document contract relevant details of the work relationship. The contract may, for example, have provisions for how to deal with changes in the client's downstream market, e.g., competitors offering new products, features, or technologies that need to be matched. If these provisions are sufficiently carefully formulated a court will be able to verify whether the agent accommodated those needs. Shortfalls in meeting these contract requirements could be documented through error messages in the client's IT system, issues revealed through prescribed audits, a paper (email) trail linked to required routine progress reports, or failure to deliver interim versions of the final product. Verifiability is coarse in that the court can only verify whether the state is covered by a clause and by which clause. The court generally cannot verify individual states, and verifiability increases with how detailed the contract is written.

Both the principal and the agent may be tempted to breach the contract. In the simple example of Section 2, the agent prefers an action that is lower than the instruction, conditional on the inference about the state that he draws from receiving the instruction. Similarly, in that example, if the state for which the contract mandates issuing the instruction is sufficiently high, the principal strictly prefers sending a cheap-talk message instead.

[10] This is more than what we need given the simplicity of our contracts, in which conditions are defined by their boundaries. All the agent needs to observe is whether the state is above or below these thresholds; he does not have to observe the exact state. In this sense, simplicity facilitates observability.

Principal and agent in our model do not yield to these temptations for three reasons: (a) Since the state and the action are ex post observable, it becomes commonly known between the principal and the agent if a breach occurs. (b) In the event that either party goes to court, the court can verify whether a breach occurred. Principal and agent know that the clarification of language in the contract and the establishment of record keeping systems will enable the court to determine whether a breach occurred in the event of a dispute. This allows them to anticipate the court's decision should there be a dispute. (c) We assume that the anticipated court's action in the event of a breach is incentive enough for the principal and agent to go to court when they observe a breach by the other party.[11] In summary, the threat of ending up in court serves as a deterrent. In our model, we take this as being sufficient to ensure that the stipulations of the contract are followed. Since the deterrent is effective, disputes and dispute resolution occur only off the equilibrium path.[12]

In contrast to the prescriptions of the contract, cheap-talk messages sent in communication subgames are not verifiable, either because they are insufficiently documented or the language is not shared with the court. Thus, it is the contract that renders states governed by it verifiable. States not governed by the contract remain unverifiable. Since the contract determines what is verifiable and what is not, we have endogenous verifiability.[13]

Contractual incompleteness. The language barrier that gives rise to contractual incompleteness in our model is an issue vis-à-vis third parties but not between the principal and the agent. Imagine the principal hiring the agent for a research project. Principal and agent share a common terminology that is not familiar to the population at large. Then it takes resources to clarify details of the research in a language that is understandable for third parties. It is difficult to describe precisely which actions are to be taken, and for which states. In addition, it is costly to put in place measurement and record keeping systems that make the language meaningful and verification possible, which is needed to make the contract enforceable. As a result the principal economizes on contract writing.

A key feature of our model is that the principal can vary the degree of obligational completeness by choosing which contingencies to cover by the contract. In the terminology of Crocker and Reynolds (1993), our principal may choose "agreements that are left intentionally incomplete." Bajari and Tadelis (2001) and Banerjee and Duflo (2000) suggest that in procurement contracting it is frequently the case that design uncertainty gives rise to endogenous incompleteness.

The contracts we allow are moderately complex, allowing for coarse conditioning on states and specifying state-contingent actions in advance. Alternative incomplete contracts that

[11] Alternative dispute resolution mechanisms (e.g., mediation or arbitration), involving a third independent party, may serve as substitutes for courts. Such mechanisms can be designed to minimize the costs of resolving disputes as well as to deter them. The contract may also include further deterrents, such as specifying that the loser in a dispute covers the prevailing party's attorney's fees.

[12] Disputes and dispute resolution are also off path in Hart and Moore (1988)'s study of incomplete contracts and renegotiation.

[13] Kvaløy and Olsen (2009) note that that "careful contracting" can improve verifiability. Like in our model, endogenous verifiability is a source of endogenous contractual incompleteness. In their model private information is absent and hence, only verifiability of actions is an issue.

do not specify actions in advance, in which instead the principal first hires the agent and at a later stage (once she learns the state) dictates the action to the agent, would leave the agent at the whim of the principal. Such contracts might be appropriate in settings with well defined limits on the principal's possible needs and requirements, or where it is easy for the agent to walk away from the contract if the principal makes excessive demands. When, however, the agent has to make substantial relationship-specific investment, there is an incentive to use the contract in part to limit the scope of what the principal can ask the agent to do. Relationship-specific investments, including the provision of training or information, also help explain why bargaining power shifts to the agent in those states not covered by the contract: it makes the agent more valuable after being hired.

Contracts in our model are lists of clauses, prescribing instructions for intervals of states. Instructions are precise; we do not consider the difficulty of writing detailed instructions. The principal does, however, face a trade-off when choosing the sizes of conditions to which instructions apply. Given the finite bound on the number of instructions, having an instruction apply to a narrowly defined condition implies sacrificing accuracy for other states.

Having conditions be intervals captures the intuition that it is easier to write contracts that treat similar circumstances similarly. Intervals are fully described by their boundaries. This motivates our focus on the number of conditions, rather than their size. Our main results, in Section 5, on splitting the communication region and splitting the contract region would not change if we assumed a writing cost function that is increasing in the number of clauses, starts from a low level, and is unbounded; this could also accommodate some dependence of costs on the size of conditions. Our results on splitting the communication region would remain equally unaffected if we permitted limited precision of instructions in the form of noisy instructions or instructions that are sets of actions.

Role assignments. In our model, the party who expects private information is the principal who makes a contract offer at the ex ante state that the agent can either accept or reject. This role assignment is guided by the main applications we have in mind and it helps us organize results. It is not crucial for some of the key results that characterize the form of optimal contracts. Our results on splitting the communication region go through if we switch who makes the initial contract offer or have the contract designer maximize social surplus. The same is true for a version of our result on splitting the contract region (we indicate in the paper where this is the case).

4 General features of optimal equilibria

4.1 Communication subgames

Since the type distributions for communication subgames $\Gamma^{\mathcal{C}}$ are the prior F concentrated on the sets $\mathcal{L}(\mathcal{C})$, they do not have full support on the state space $[0, 1]$. With that exception, communication subgames satisfy all the assumptions of CS. As a result, key characteristics of equilibria that do not depend on the full support assumption carry over. Since we make use of these characteristics throughout, we list them here for convenience.

For a strategy profile $(\sigma^{\mathcal{C}}, \rho^{\mathcal{C}})$ in communication subgame $\Gamma^{\mathcal{C}}$, we say that a *communication action* y is *induced* by that profile if there is a type θ and a message m in the support of $\sigma^{\mathcal{C}}(\theta)$ such that $\rho^{\mathcal{C}}(m) = y$. If, in addition, $(\sigma^{\mathcal{C}}, \rho^{\mathcal{C}})$ is an equilibrium profile, we say that action y is *induced in equilibrium.* Given any measurable set $\Phi \subseteq [0, 1]$ with $\text{Prob}(\Phi) > 0$, we define $y^{*i}(\Phi) := \arg\max_y \int_\Phi U^i(y, \theta)dF(\theta)$, the optimal action for $i = A, P$ given prior beliefs concentrated on Φ.

As in CS, because of the conflict of interest between principal and agent, full revelation of information is ruled out in equilibrium. Moreover, since the principal's bias is everywhere bounded away from zero, there has to be a minimal distance between equilibrium actions: types of the principal meant to induce the lower of two sufficiently close equilibrium actions otherwise would have an incentive to deviate and induce the higher action. Hence, there is a finite upper bound on the number of equilibrium actions that does not depend on the contract $\mathcal{C}$ and the type distribution $F^{\mathcal{C}}$ that the contract induces.[14]

If the actions that are induced in equilibrium are $0 < y_1 < y_2 < \ldots < y_{n-1} < y_n < 1$, there are $n+1$ *critical types* $0 = \theta_0 < \theta_1 < \theta_2 < \ldots < \theta_{n-1} < \theta_n = 1$ such that type θ_j is indifferent between actions y_j and y_{j+1} for $j = 1, \ldots, n-1$. We follow the convention of referring to the indifference requirement for critical type θ_j, $j = 1, \ldots, n-1$, as that type's *arbitrage condition*

$$U^P\left(y^{*A}\left((\theta_{j-1}, \theta_j) \cap \mathcal{L}(\mathcal{C})\right), \theta_j, b\right) - U^P\left(y^{*A}\left((\theta_j, \theta_{j+1}) \cap \mathcal{L}(\mathcal{C})\right), \theta_j, b\right) = 0.$$

In an equilibrium that induces n actions, we refer to the interval (θ_{j-1}, θ_j) as *step* j, for $j = 1, \ldots, n$. We call an equilibrium that induces n actions an *n-step equilibrium* and say that it is *influential* if $n > 1$.

Two facts about the set of types who induce a communication action y are worth noting. First, this set, unlike in CS, is not necessarily bounded by critical types, since critical types may belong to the interior of a condition. Second, also unlike in CS, this set of types need not be an interval: there may be types both below and above a condition of the contract who induce the same communication action y. For these reasons, we find it convenient to focus on intervals of a different kind. For any communication action y, the associated *communication interval* is the smallest interval that *contains* the set of types who induce action y. Note that for an arbitrary contract, a communication interval may contain one or more conditions of the contract in its interior. We will show later (Proposition 2) that this cannot be the case for optimal equilibria of the contract-writing game. We refer to communication intervals by their endpoints. These endpoints may or may not be included: When a communication interval is bounded by conditions, the interval is open. Adjoining communication intervals are separated by a critical type, which may belong to either interval.

Figure 2 illustrates different kinds of communication intervals, using a contract with five conditions C_i and instructions x_i, $i = 1, \ldots, 5$. The communication intervals are $[0, \theta_1)$, $(\theta_1, \underline{C}_1)$, $(\overline{C}_1, \theta_3]$, and $(\theta_3, \underline{C}_4)$. The first of these is bounded solely by critical types; the others have bounds some of which belong to conditions. Note that critical type θ_2 belongs

[14]This is formalized as Lemma A.1 in the appendix.

to the interior of condition C_1 and the set of types inducing action y_3 is a strict subset of the step (θ_2, θ_3). The set of types inducing action y_4 is a strict subset of the communication interval $(\theta_3, \underline{C}_4)$.

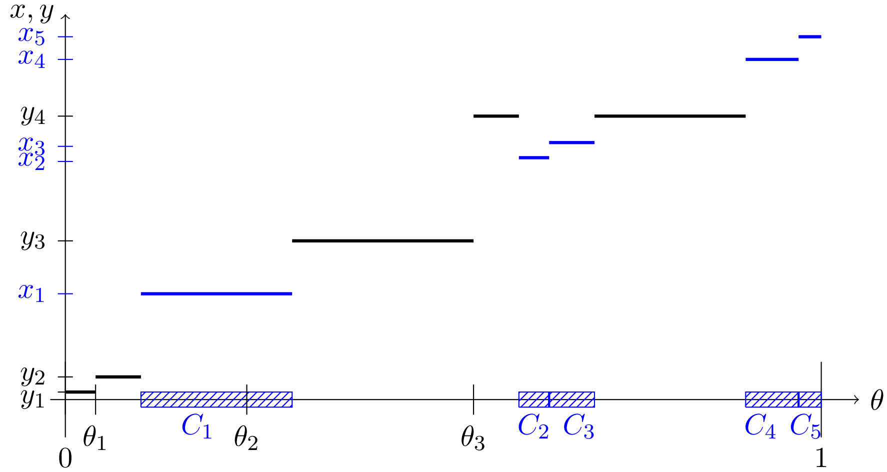

Figure 2: Contract with conditions C_i and instructions x_i, $i = 1, \ldots, 5$, and induced 4-step equilibrium with critical types $\theta_1, \theta_2, \theta_3$ and actions y_j, $j = 1, \ldots, 4$.

4.2 Contract-writing games

Optimal equilibria of contract-writing games have three intuitive properties: (i) They use the maximal number of available clauses. (ii) Keeping the bias fixed, if we let the number of available clauses grow without bound, contracting crowds out communication. (iii) Keeping the number of available clauses fixed, if we let the interests of principal and agent converge toward perfect alignment, communication drives out contracting. The following proposition formalizes these observations. For the third part, we impose the following continuity property: for any sequence of biases $(b_j)_{j=1}^{\infty}$ with $\lim_{j\to\infty} b_j = 0$ and any sequence $(e(b_j))_{j=1}^{\infty}$ of principal-optimal equilibria in the games $(\Gamma^0(b_j))_{j=1}^{\infty}$, the principal's payoffs in those equilibria converge to $\int_{[0,1]} U^P(y^P(\theta, 0), \theta, 0) dF(\theta)$.[15]

Proposition 1 *For every maximal number of clauses $\widehat{K}$ and bias $b > 0$, let $\mathcal{L}_{\widehat{K},b}$ be a gap arising in an optimal equilibrium of the contract-writing game $G(\widehat{K}, b)$ and let $\mathcal{C} = \{(C_k, x_k)\}_{k=1}^{K}$ be an optimal contract. Then,*
(i) $K = \widehat{K}$;
(ii) $\forall b > 0$, $\lim_{\widehat{K}\to\infty} \mathrm{Prob}(\mathcal{L}_{\widehat{K},b}) = 0$; and,
(iii) $\forall \widehat{K} \geq 1$, $\lim_{b\to 0} \mathrm{Prob}(\mathcal{L}_{\widehat{K},b}) = 1$.

[15] Spector (2000), Agastya, Bag and Chakraborty (2015), and Dilmé (2022) provide conditions on primitives that ensure that this continuity property holds.

Property (i) (use of the maximal number of available clauses) follows from two simple observations: In the absence of any clause, the principal can replace one of the steps of any equilibrium in the communication subgame Γ^0 by a condition and replace the corresponding equilibrium action by her favorite action for that condition. If, instead, the candidate optimal contract already includes at least one clause, she can split the corresponding condition and prescribe her favorite actions for each of the newly created conditions. In both cases, the incentive constraints for communication remain satisfied and there is a strict payoff improvement for the principal.[16]

To understand property (ii) (contracting crowding out communication), note that as the number of clauses grows without bound, one can approximate the principal's first-best (full information) payoff with obligationally complete contracts. On the other hand, since there is an upper bound on the number of actions that can be induced with communication, for any set of types that has positive probability the payoff from communication is bounded away from the first-best payoff. Hence, the sizes of the gaps in optimal contracts cannot remain bounded away from zero.

The intuition for property (iii) (communication driving out contracting) is similar: When the bias gets small, one can approximate the principal's first best with equilibria in the communication subgame Γ^0, without any contract clauses. In contrast, with only a finite number of possible contract actions, on any contract region that has positive probability, the payoff is bounded away from the first best.

5 Optimal equilibria in the uniform-quadratic environment

In this section, we examine optimal equilibria under the assumptions that payoff functions are quadratic with the principal having a constant positive bias, that is, $U^P(y, \theta, b) = -(\theta+b-y)^2$ and $U^A(y, \theta) = -(\theta - y)^2$, and that states are uniformly distributed on $[0, 1]$. In the previous section, we obtained limit results on the sizes of the contract and communication regions. Here we focus on the structure of these regions and on how their sizes vary with the bias and the bound on the number of contract clauses.

Given a contract $\mathcal{C} = \{(C_k, x_k)\}_{k=1}^{K}$, we call a union of conditions that forms a connected

[16]It is perhaps worth noting that the literature has found that, sometimes, *prima facie* useful and readily available clauses will not be included in a contract. Allen and Gale (1992) and Spier (1992) have pointed out that in the presence of asymmetric information principals may prefer non-contingent contracts. This is the case when proposing a contingent contract would send an unfavorable signal. Bernheim and Whinston (1998) observe that if some aspects of performance are non-verifiable, it may be advantageous not to include other, verifiable, aspects in a contract. In essence, once a contract needs to be incomplete in some dimensions, the contract will give rise to some form of strategic interaction. In that case, there can be instances in which the quality of that strategic interaction can be improved by not specifying some obligations, even when they are verifiable. In our case, the signaling aspect is absent and while there is strategic interaction for states not covered by the contract, any given contract that does not use all available clauses can be improved upon without impacting the strategic interaction.

set and is not contained in a larger union of conditions that also forms a connected set a *condition cluster*, and denote it by $\boldsymbol{C}$.[17] Hence, there is no communication inside of a condition cluster.

We obtain two "splitting results." On the one hand, we find in Subsection 5.1 that whenever the optimal contract induces influential communication, it splits the communication region: there are at least two communication intervals that are separated by condition clusters. Intuitively, splitting the communication region removes the temptation for small misrepresentations of the state. At the same time, since interests are partially aligned, both principal and agent benefit from communicating large differences between states. Condition clusters can be used as a wedge between two communication regions, so that it becomes in the common interest of principal and agent to indicate whether the state belongs to one or the other. On the other hand, in Subsection 5.2, we find that with a modest lower bound on the number of available clauses, for sufficiently closely aligned incentives, communication splits the contract region. Condition clusters obtained from splitting larger condition clusters can be used to relax incentive constraints for communication.

In Subsection 5.3, we obtain bounds on the size of the contract region and establish monotonicity results for these bounds as well as for the principal's payoff as functions of the bias and the number of available contract clauses.

5.1 Contracts split communication

Whether or not an optimal contract induces communication depends on the bound on the number of clauses, $\widehat{K}$ and the size of the bias, b. The following two observations make this precise: (1) Whenever $\widehat{K} \geq \frac{1}{2b}$, any optimal contract is obligationally complete (implying that there is a single condition cluster), all $\widehat{K}$ conditions are of equal size, and no action is induced through communication. (2) If, on the other hand, $\widehat{K} < \frac{1}{2b}$ then optimality implies that communication is influential, i.e., there are at least two actions induced through communication.[18] Combining these two observations it follows that it is never the case that in an optimal equilibrium only a single action is induced via communication - there is never one-step communication.

The following notions will be useful in characterizing optimal equilibria.

Definition 1 *Consider an equilibrium e^G in the contract-writing game G in which the principal writes a contract $\mathcal{C} = \{(C_k, x_k)\}_{k=1}^K$ that induces an n-step equilibrium $e^{\mathcal{C}}$ in the communication subgame $\Gamma^{\mathcal{C}}$. Then, we call the equilibrium e^G*

1. *partitional – if there is a partition $\mathcal{T} = \{T_1, T_2, \dots, T_{K+n}\}$ of the state space $[0, 1]$ into intervals such that each $T \in \mathcal{T}$ is either a condition of $\mathcal{C}$ or a communication interval; and,*

[17] A subset X of a topological space is connected if it cannot be partitioned into two sets that are open in the relative topology induced on the set X. Here connected sets are intervals.

[18] We verify these two facts in Proposition A.1 in the appendix.

2. *monotonic – if for any two partition elements $T, T' \in \mathcal{T}$, with $T \neq T'$ and* $\inf(T') \geq \sup(T)$, *the actions $a(T')$ and $a(T)$ taken for states in T' and T satisfy $a(T') > a(T)$.*

The next proposition establishes that optimal equilibria are partitional, or, equivalently, that no condition cluster can belong to a communication interval. It also establishes that all actions – induced by communication or by the contract – are increasing in the state, i.e, optimal equilibria are monotonic. Moreover, if there is influential communication, there is at least one condition cluster with communication intervals on either side.

Proposition 2 *Optimal equilibria are partitional and monotonic. Whenever they induce influential communication (i.e., when $\widehat{K} < \frac{1}{2b}$) there is a condition cluster that belongs to the interior of the state space.*

Proposition 2 reveals an interesting interaction between contracts and communication. The principal uses contract clauses to separate events that induce distinct communication actions. This is beneficial because it relaxes incentive constraints in the communication subgame, and the relaxation of incentive constraints makes it possible to equalize the size of communication intervals relative to pure cheap talk.[19] Quadratic loss functions disproportionately penalize actions that are more distant from players' ideal points. Equalizing the sizes of communication intervals amounts to substituting smaller for larger distances from ideal points, and therefore increases expected payoffs. This highlights the dual role of contracting as both substituting for and facilitating communication.

To prove that optimal equilibria are partitional, we start with any equilibrium that is not partitional. We proceed by modifying the corresponding contract in several steps through translations of condition clusters: For any $\delta \in \mathbb{R}$, we refer to the clause $(C_k + \delta, x_k + \delta)$ as the *δ-translation of the clause* (C_k, x_k), to the condition $C_k + \delta$ as the *δ-translation of the condition C_k*, and extend this to condition clusters by applying the same δ-translation to each of the clauses belonging to the cluster.[20] We ensure at each step that the principal's payoff increases: the typical argument is that properly translating a condition cluster enables us to decrease longer communication intervals, while increasing shorter ones. Once all steps are completed, we check that we have obtained an equilibrium.

In the first step, we use the fact that there can be no more than one condition in the interior of any communication interval (see Lemma A.3 in the appendix). We consider a candidate for an optimal contract $\mathcal{C}$ and a corresponding equilibrium $e^{\mathcal{C}}$ with a communication interval containing a single condition in its interior. We then translate that condition to the lower bound of the communication interval. The new contract is $\mathcal{C}_0$. In the second step, we adjust the strategies in the communication game such that, locally (between any two adjacent condition clusters), incentive compatibility is restored. For this step, we rename the contract $\mathcal{C}_1 = \mathcal{C}_0$, and similarly refer to the corresponding communication subgame as $\Gamma^{\mathcal{C}_1}$. We sketch steps one and two in Figure 3.

[19] A similar effect arises in Kolotilin, Li and Li (2013). There the agent has the power to commit to a set of actions from which to chose.

[20] Here, for any set $C \subset \mathbb{R}$ and any $\delta \in \mathbb{R}$, $C + \delta$ denotes the Minkowski sum of the sets C and $\{\delta\}$ – i.e., $C + \delta := \{c' \in \mathbb{R} | \exists c \in C \text{ s.t. } c' = c + \delta\}$.

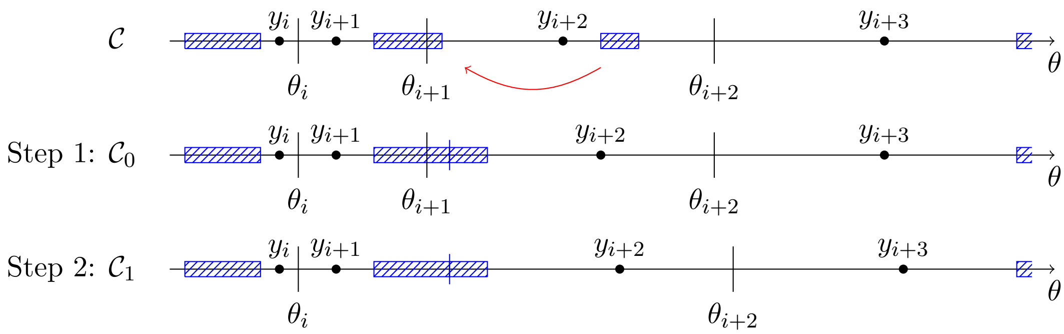

Figure 3: Sketch of the first two steps in the first part of the proof of Proposition 2.

In order to restore incentive compatibility locally, we have to raise the action y_{i+2}. This makes the action less attractive for the type $\tilde{\theta}$ that is at the top of the newly created condition cluster (see Figure 4 for an illustration). In fact, it may make action y_{i+1} more attractive than y_{i+2}. In the third step, we address incentive-compatibility problems of this kind – that is, for types that are separated by condition clusters. To do so, we identify the highest condition cluster such that a type $\tilde{\theta}$ at the upper boundary of that cluster prefers to deviate to a message inducing an action below the cluster. In multiple steps that maintain the local equilibrium conditions, we properly translate the respective condition cluster upwards to restore incentive compatibility for type $\tilde{\theta}$. The resulting contract is $\mathcal{C}_2$. We iterate the third step for all lower condition clusters to obtain a global equilibrium. In conclusion, optimal equilibria must be partitional.

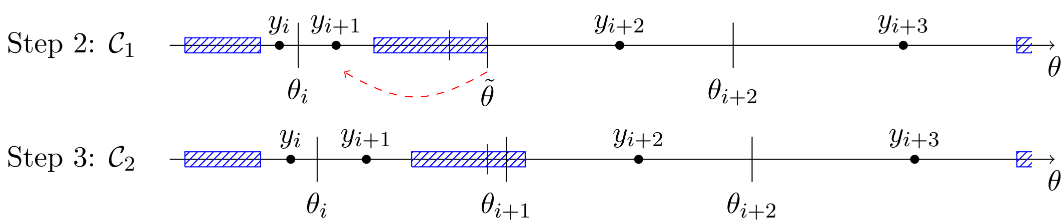

Figure 4: Sketch of the third step in the first part of the proof of Proposition 2.

To prove monotonicity, the relevant case to consider is the one in which a communication interval T' is directly above a condition T. For this case, we show in the appendix, if we had $a(T') \leq a(T)$ (and keeping in mind that $a(T')$ is the agent-optimal action on T' and $a(T)$ is the principal-optimal action on T), we could form a new contract clause $(\overline{T \cup T'}, a(T) + \epsilon)$ that, for sufficiently small ϵ, would result in a payoff improvement for the principal, resulting in a contradiction. This implies that any optimal equilibrium of G is monotonic.

To prove the second statement in Proposition 2, we show that the principal's payoff can be increased when more than one action is induced in equilibrium and all condition clusters are at the extremes. For an illustration of the steps in the argument, see Figure 5.

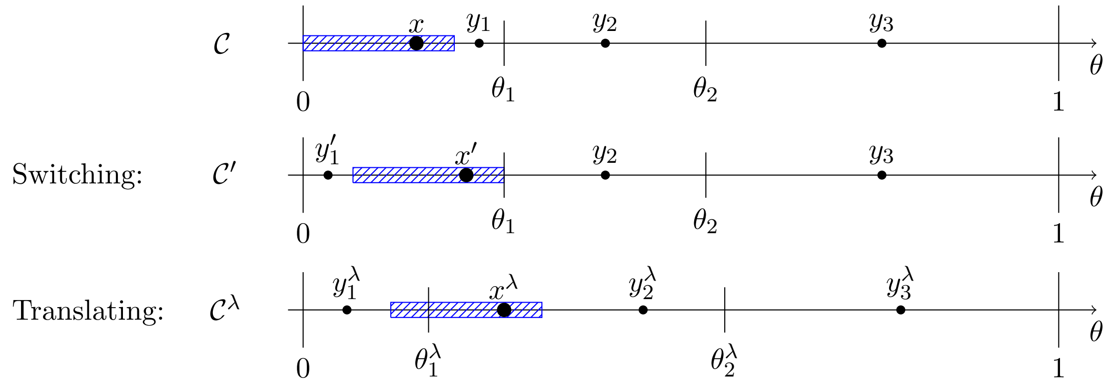

Figure 5: Second part of proof: payoff improvement by translations.

The first panel of Figure 5 shows a contract $\mathcal{C}$ with a single condition located at the left extreme of the state space and a corresponding three-step communication equilibrium. We replace contract $\mathcal{C}$ by a new contract $\mathcal{C}'$ that translates the condition upwards such that the first critical type θ_1 becomes its new upper boundary. Since we do not change the length of any communication interval, payoffs remain the same. However, type θ_1 now strictly prefers action y_2 over y_1'. Moreover, the length of the communication interval inducing action y_1' is smaller than the length of the communication interval inducing y_2. Together, this implies that we can translate the condition further upwards to $\mathcal{C}^\lambda$ while maintaining incentive compatibility and increasing payoffs. This shows that the contract $\mathcal{C}$ that we started with cannot be optimal.

Proposition 2 continues to hold under alternative assumptions about the allocation of bargaining power between the principal and the agent and about the nature of instructions. It remains true if the contract, instead of maximizing the principal's payoff, maximizes either the agent's payoff or weighted social surplus. The same is true if we replace instructions that specify single actions by instructions that specify sets of possible actions or allow for randomness in the specification of actions. To see this, recall that the proposition is proved by translating condition clusters of contracts that violate the properties stated in the proposition. We can decompose the ex ante payoff of whoever writes the contract into a contract payoff, which is derived from states governed by the contract, and a communication payoff, which is derived from the remaining states. Translating condition clusters affects only the communication payoff and ex ante, as long as the size of the communication region remains fixed, principal and agent have the same preferences over communication payoffs. Therefore, the properties of instructions influence the proportion of states allocated to communication and contracting, but not the structure of optimal equilibria.

5.2 Communication splits contracts

The next result shows that with close enough incentive alignment, communication splits the contract region: not all condition clusters can be large, in the sense of being composed of more than three conditions. Hence, when $\widehat{K} > 3$, there are at least two condition clusters that are separated by one or more communication intervals. Intuitively, when a condition cluster is composed of more than three conditions, there is considerable slack in the incentive constraints relevant for the nearest communication actions above and below the cluster. That makes it tempting to break up the cluster and use the conditions made thus available to relax incentive constraints elsewhere, where they were binding before the breakup.

Proposition 3 *For sufficiently small $b > 0$, any optimal contract contains a condition cluster with no more than three conditions.*

The following corollary is an immediate consequence of Proposition 3.

Corollary 1 *With $\widehat{K} > 3$ and sufficiently small $b > 0$, in any optimal contract there are at least two condition clusters.*

To prove the proposition, we start by considering the problem of the principal maximizing her expected payoff subject to the conditions that each cluster is composed of at least four conditions and that incentive constraints for communication hold between but not necessarily across clusters. Since this problem ignores some of the incentive constraints, we call it the "relaxed problem." We show that one can improve on the solution to the relaxed problem by splitting up one of the clusters and that this improvement respects global incentive constraints.

The argument proceeds as follows. Call an interval of states a *communication area*, if it is maximal (with respect to set inclusion) among intervals of states not covered by the contract. This is any set of types bounded by two adjacent condition clusters (or bounded by a condition cluster and either 0 or 1). Note that the union of all communication areas is the communication region. For sufficiently small $b > 0$, for any solution of the relaxed problem there is at least one communication area with two or more communication steps. We 'zoom in' on a cluster adjacent to such a communication area. We then focus on that cluster and its adjacent communication areas. We take the communication area with more than two steps and switch the step that is adjacent to the cluster with the adjacent condition. We show that after the switch all incentive constraints across clusters are slack. This is immediate for the new one-condition cluster. For the remaining cluster it follows from two facts: (1) the cluster is composed of three or more conditions and (2) the solution of the relaxed problem imposes constraints on the lengths of conditions. The incentive constraints across clusters being slack makes it possible to translate clusters in a way that leads to more equal lengths of communication intervals. This improves the principal's expected payoff. For an illustration of these steps, see Figure 6.

To establish the corollary, we only need to show that having a single cluster cannot be optimal. For any solution of the relaxed problem with only a single cluster, we can switch

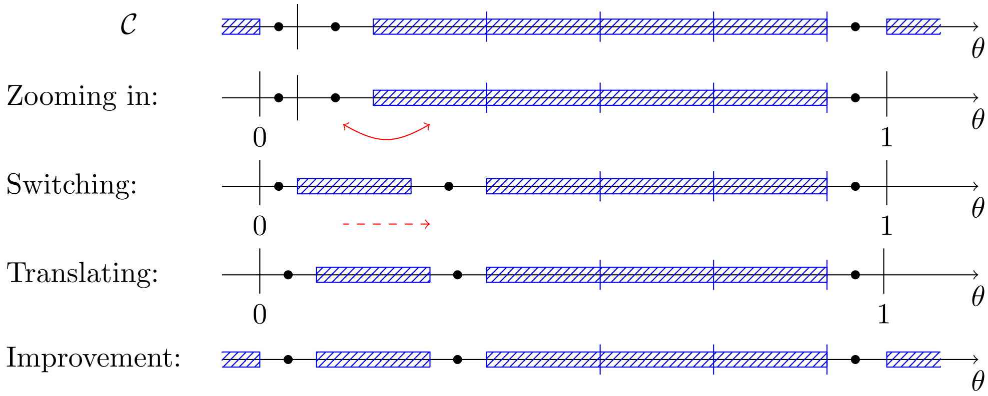

Figure 6: Sketch of the steps in the proof of Proposition 3.

the communication area below and above the cluster and still have a solution. We can therefore assume that for sufficiently small bias there is at least two-step communication below the cluster in a solution to the relaxed problem. In this case, the proof of Proposition 3 goes through, when either the agent writes the contract or the principal only specifies the conditions and not the actions in the contract. Thus, the corollary continues to hold if instead of the principal the agent writes the contract, leaving all else unchanged. The same is true for the case in which the principal only specifies the conditions in the contract, but the agent is free to choose an action for each condition.

5.3 Comparative statics

For conducting comparative statics exercises, it would be helpful to have an explicit expression $Z(b, \widehat{K})$ for the optimal size of the contract region as a (possibly set-valued) function of the bias b and the maximal number of clauses $\widehat{K}$. This, however, would require knowledge of the principal-optimal contracts for all values of these parameters. Short of that, to get a sense of the comparative statics in these parameters, we consider two auxiliary versions of the principal's contracting problem, one that favors communication over contracting and another that does the reverse. Favoring communication tends to deflate the contract region, while favoring contracting tends to inflate it. This provides us with lower and upper bounds on the optimal size of the contract region as functions of the parameters b and $\widehat{K}$.

Denote the size of the contract region obtained from the problem that favors communication by $\underline{Z}(b, \widehat{K})$ and the size of the contract region obtained from the problem that favors contracting by $\overline{Z}(b, \widehat{K})$. Then $\underline{Z}(b, \widehat{K})$ is a lower bound on $Z(b, \widehat{K})$, and, with a qualification discussed below, $\overline{Z}(b, \widehat{K})$ is an upper bound on $Z(b, \widehat{K})$.

We find that the lower bound is nondecreasing in the bias b. The upper bound is nondecreasing in the maximal number of clauses $\widehat{K}$ and (in our numerical illustrations) also in the bias b.

Lower bound In order to obtain a lower bound on the optimal size of the contract region $Z(b,\widehat{K})$, we set up an auxiliary contracting problem in which the principal (1) is restricted to a single condition in the contracting region, thus disfavoring contracting, and (2) achieves the maximum benefit from communication subject only to the constraint that the number of communication actions does not exceed the limit achievable through contracting.

Concerning (1), if we consider only a single condition of size Z and let $a(Z,b,1) = \arg\max_a \int_0^Z -(t+b-a)^2 dt = \frac{2b+Z}{2}$, then the principal's payoff in the contract region equals

$$\underline{\Pi}(Z) = \int_0^Z -(t+b-a(Z,b,1))^2 dt = -\frac{Z^3}{12}.$$

Concerning (2), recall that for any bias b incentive compatibility requires that any two adjacent communication actions have to be separated by a distance of at least $2b$. Therefore, the number of communication actions that are possible in equilibria of communication subgames is bounded by $\left\lceil \frac{1}{2b} \right\rceil$. We can favor communication by postulating that the number of communication actions equals $\left\lceil \frac{1}{2b} \right\rceil$, and, ignoring incentive constraints, letting all communication intervals be of equal length. Letting $\alpha(\ell) = \arg\max_a \int_0^\ell -(t-a)^2 dt = \frac{\ell}{2}$, the payoff from splitting the communication region into $\left\lceil \frac{1}{2b} \right\rceil$ communication intervals of equal length equals

$$\begin{aligned}\underline{\Xi}(b,Z) &= \left\lceil \frac{1}{2b} \right\rceil \int_0^{\frac{1-Z}{\lceil \frac{1}{2b} \rceil}} -\left(t+b-\alpha\left(\frac{1-Z}{\lceil \frac{1}{2b} \rceil}\right)\right)^2 dt \\ &= -\frac{(1-Z)\left((1-Z)^2 + 12b^2 \left\lceil \frac{1}{2b} \right\rceil^2\right)}{12 \left\lceil \frac{1}{2b} \right\rceil^2}.\end{aligned}$$

Let

$$\underline{Z}(b,\widehat{K}) \in \arg\max_{Z\in[0,1]} \left\{\underline{\Xi}(b,Z) + \underline{\Pi}(Z)\right\}.$$

Then $\underline{Z}(b,\widehat{K})$ is a lower bound on $Z(b,\widehat{K})$ for all $\widehat{K}$. Since the function $\underline{\Xi}(b,Z) + \underline{\Pi}(Z)$ is supermodular in (Z,b), it follows from Topkis's Monotonicity Theorem that $\underline{Z}(b,\widehat{K})$ as a (possibly set-valued) function of b is nondecreasing (in the strong set order).

Upper bound With the goal of obtaining an upper bound on $Z(b,\widehat{K})$, we devise an auxiliary contracting problem for the principal in which (1) the principal is free to choose the internal structure of the contracting region and (2) cannot take advantage of contracting to structure the communication region.

Concerning (1), if – for a given size Z of the contracting region – the principal is free to choose the internal structure of the contracting region, without being able to impact incentive compatibility in the communication region, then it is optimal to split the contract region of size Z into $\widehat{K}$ equal-sized conditions. Doing so (at least weakly) increases the direct value of contracting and by itself helps inflate Z relative to $Z(b,\widehat{K})$ – with one qualification: at the optimum of the original problem, which determines $Z(b, \widehat{K})$, the critical types belonging to a condition may need to be in the interior of that condition (call this "interiority").[21]

If we let $a(Z, b, \widehat{K}) = \arg\max_a \int_0^{Z/\widehat{K}} -(t+b-a)^2 dt = \frac{2b\widehat{K}+Z}{2\widehat{K}}$, then the principal's payoff in a contract region of size Z that is split into $\widehat{K}$ equal-size intervals equals

$$\overline{\Pi}(Z, \widehat{K}) = \widehat{K} \int_0^{Z/\widehat{K}} -(t+b-a(Z,b,\widehat{K}))^2 dt = -\frac{Z^3}{12\widehat{K}^2}.$$

Concerning (2), not being able to use contracting to structure communication deflates the value of communication and thus also inflates Z relative to $Z(b, \widehat{K})$. If the principal cannot use contracting to structure communication in the communication region, the maximal achievable payoff from communication is bounded from above by the payoff from an optimal CS equilibrum over an interval of length $1 - Z$. This payoff is given by

$$\overline{\Xi}(b, Z) = (1-Z)\left[-\left(\frac{(1-Z)^2}{12N(b,Z)^2} + \frac{b^2(N(b,Z)^2-1)}{3}\right) - b^2\right],$$

where

$$N(b, Z) = \left\lceil -\frac{1}{2} + \frac{1}{2}\left(1 + \frac{2(1-Z)}{b}\right)^{\frac{1}{2}} \right\rceil.$$

Let

$$\overline{Z}(b, \widehat{K}) \in \arg\max_{Z\in[0,1]} \left\{\overline{\Xi}(b, Z) + \overline{\Pi}(Z, \widehat{K})\right\}.$$

Then $\overline{Z}(b, \widehat{K})$ is an upper bound on $Z(b, \widehat{K})$ whenever interiority holds. Since the function $\overline{\Xi}(b, Z) + \overline{\Pi}(Z, \widehat{K})$ is supermodular in $(Z, \widehat{K})$, it follows from Topkis's Monotonicity Theorem that $\overline{Z}(b, \widehat{K})$ is nondecreasing (according to the strong set order) in $\widehat{K}$. When we calculate $\overline{Z}(b, \widehat{K})$ as a function of b for three different values of $\widehat{K}$, we find it to be nondecreasing in each case.

The dots in Figure 7 indicate the exact optimal sizes of the contract region for selected values of the parameters b (measured along the horizontal axis) and $\widehat{K}$. Small black dots refer to $\widehat{K} = 1$ and large blue dots refer to $\widehat{K} = 2$. In addition, Figure 7 displays the lower bound $\underline{Z}(b, \widehat{K})$ (solid, red) as well as upper bounds $\overline{Z}(b, \widehat{K})$ for $\widehat{K} = 1$ (dotted, black), $\widehat{K} = 2$ (dash-dotted, blue), and $\widehat{K} = 3$ (dashed, green). The overall message that emerges is that increasing either the bias, b, or the limit on the number of contract clauses, $\widehat{K}$, incentivizes the principal to substitute contractual control for cheap-talk communication. This is intuitive: With a larger bias the benefit from controlling decisions rather than delegating them increases, while at the same time the ability to use cheap-talk to convey information diminishes. Likewise, relaxing the constraint on the number of available contract clauses directly increases the value from contracting.

[21] When interiority fails, one can imagine that $Z(b, \widehat{K})$ is kept larger than it would otherwise be by the need to satisfy incentive compatibility for communication. In that case, when we consider the auxiliary problem in which contracting is not used to structure communication, this need is removed, which may lower the size of the contract region. We have found the interiority condition to be satisfied in all cases in which we have obtained explicit expressions for optimal contracts.

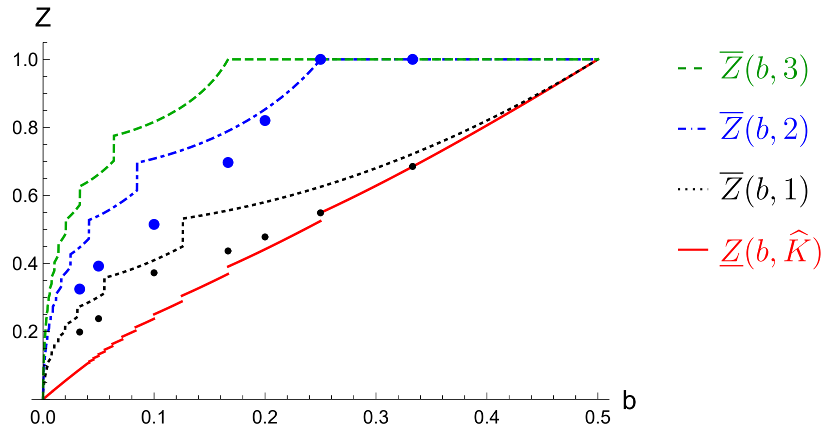

Figure 7: Lower bound $\underline{Z}(b, \widehat{K})$ (solid, red) and upper bounds $\overline{Z}(b, \widehat{K})$ for $\widehat{K} = 1$ (dotted, black), $\widehat{K} = 2$ (dash-dotted, blue), and $\widehat{K} = 3$ (dashed, green).

Principal's payoff Monotonicity of the principal's payoff in terms of $\widehat{K}$: Take an optimal equilibrium for a given b and $\widehat{K}$. With a $\widehat{K}' > \widehat{K}$, the same payoff is achievable by simply not making use of the additional clauses. One can then introduce one of the additional clauses to split one of the existing clauses and impose the principal's optimal actions in the clauses resulting from the split. This does not affect any of the communication incentives, leaves the payoffs from communication unchanged, and strictly increases payoffs in the contract region. Additional benefits are likely to accrue from resizing and repositioning conditions and thus changing the induced communication subgame. Therefore, the principal's payoff is strictly increasing in $\widehat{K}$.

Monotonicity of the principal's payoff in terms of b: Take an optimal equilibrium $e(\widehat{K}, b)$ of the contract-writing game $G(\widehat{K}, b)$ in which the contract is $\mathcal{C}$. Suppose that $\widehat{K} < \frac{1}{2b}$, so that the optimal contract fails to be obligationally complete and therefore induces communication. Fix the contract $\mathcal{C}$, and let $b' < b$. Then, temporarily ignoring communication incentives across condition clusters, within every communication area there exists an equilibrium restricted to that area that has the same number of communication actions as in the original equilibrium. In addition, for the rightmost type in any communication area, the communication action under the new (local) equilibrium is more attractive than before, whereas the next higher communication action, belonging to the next communication area (if there is one) is less attractive. This is the case because the local equilibria in each communication area have communication intervals that are of more equal size than before. The fact that communication intervals are of more equal size also implies that the principal has a strictly higher expected payoff. Therefore, the only possible problem with communication incentives across condition clusters concerns types at the bottom of each (new) communication area. For them the communication action that they induce has become less attractive and the next lower communication action (if there is one) more attractive. This can only be an issue, however, if their associated communication interval is larger than the next lower communication interval. In this case we can translate the intervening communication cluster to the right, restore incentive compatibility across that cluster and further equalize the size of communication intervals, again raising payoffs for the principal. The details of the translation argument match those used in Step 3 of Part I of the proof Proposition 2. Thus whenever there is communication under the original contract, lowering the bias strictly raises the principal's payoff. Equivalently, whenever $\widehat{K} < \frac{1}{2b}$, raising the bias strictly lowers the principal's payoff. When $\widehat{K} < \frac{1}{2b}$, raising the bias has no effect on the principal's payoff because the optimal contract is obligationally complete.

6 Examples and extensions

This section expands the example of Section 2. Suppose that payoff functions are quadratic and the distribution of states is uniform on $[0, 1]$. Consider $G(\widehat{K}, b) = G\left(1, \frac{1}{3}\right)$, the contracting game with maximally one clause and a bias $b = \frac{1}{3}$. Let $\mathcal{C}_n^*(K, b)$ denote a contract that is optimal among contracts that have K conditions and induce n-step communication, and let $\mathcal{C}^*(K, b)$ be an overall optimal contract with K conditions. For biases $b > \frac{1}{4}$ there is no contract in the game $G(1, b)$ that induces an equilibrium with more than two communication actions. The reason is that any two equilibrium actions in an induced cheap-talk game $\Gamma^{\mathcal{C}}$ must be at least a distance $2b$ apart. Therefore, there are four candidates for optimality: a standard cheap-talk game with no contract; an obligationally complete contract with no communication; and contracts with 1-step, or 2-step communication. An optimal contract in this example maximizes the principal's expected payoff among the optima of these four options. We find that the optimal contract with 2-step communication $\mathcal{C}_2^*(1, \frac{1}{3})$ is unique and dominates the two optimal contracts with 1-step communication, which are better than the optimal obligationally complete contract, which improves upon having no contract.

The overall optimal contract $\mathcal{C}^*(1, \frac{1}{3}) = \mathcal{C}_2^*(1, \frac{1}{3})$ (which was presented in Section 2) solves the following maximization problem:

$$\max_{\underline{C}, \overline{C}} \; -\int_0^{\underline{C}} \left(s + \frac{1}{3} - \frac{\underline{C}}{2}\right)^2 ds - \int_{\underline{C}}^{\overline{C}} \left(s + \frac{1}{3} - \left(\frac{\overline{C} + \underline{C}}{2} + \frac{1}{3}\right)\right)^2 ds - \int_{\overline{C}}^{1} \left(s + \frac{1}{3} - \frac{(\overline{C} + 1)}{2}\right)^2 ds$$

$$\text{s.t.} \quad \theta_1 = \frac{\underline{C} + (\overline{C} + 1)}{4} - b \in [\underline{C}, \overline{C}].$$

The first and the third term in the objective function are the principal's expected payoffs conditional on the lower and the upper communication action being taken; the middle term is the principal's expected payoff conditional on the contract action being taken. By Proposition 2, the condition cannot be in the interior of a communication interval and thus $\theta_1 \in [\underline{C}, \overline{C}]$, which explains the constraint.

The solution is unique and given by $\mathcal{C}^*(1, \frac{1}{3}) = \{([0.157, 0.843], 0.833)\}$, where $C = [0.157, 0.843]$ is the condition and $x = 0.833$ the corresponding instruction. See the top left panel of Figure 8, for an illustration. The optimal contract induces two communication actions, y_1 and y_2, while without a contract, the maximal feasible number of communication actions would be one. In this sense, contracting facilitates communication. Notice that the two communication intervals are of equal length. This can be achieved because the contract relaxes incentive constraints for types adjacent to the contract condition. As long as those constraints are slack, it pays to shift the condition in the direction of equalizing the lengths of communication intervals. Note further that the possibility of equalizing the intervals depends on the bias: for $b = \frac{1}{13}$ for example, the optimal contract is such that all communication intervals are of different length. Moreover, as illustrated in the bottom left panel in Figure 8, they are *not* monotonically increasing. This contrasts with non-trivial communication in CS equilibria, where higher communication actions are associated with longer communication intervals.

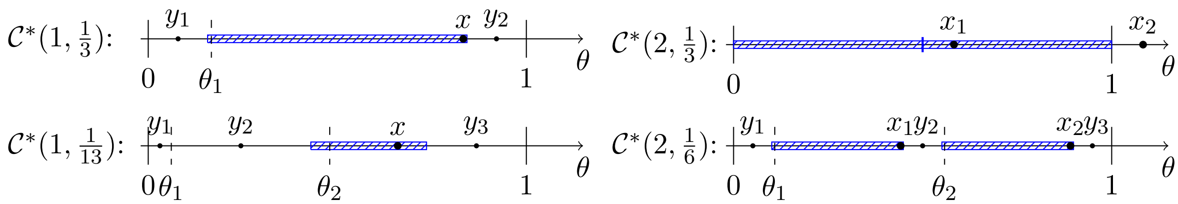

Figure 8: Left: optimal contracts, with $\widehat{K} = 1$ and $b = \frac{1}{3}, \frac{1}{13}$. Right: optimal contracts with $\widehat{K} = 2$ and $b = \frac{1}{3}, \frac{1}{6}$.

We indicate how the optimal contract changes with the parameters b and $\widehat{K}$ in the right panels of Figure 8. In the top panel, we keep $b = \frac{1}{3}$ and relax the constraint on the number of conditions by letting $\widehat{K} = 2$. In that case, since $\widehat{K} \geq \frac{1}{2b}$, the unique optimal contract is obligationally complete with the two conditions dividing the state space into two equal-length intervals, $\mathcal{C}^*(2, \frac{1}{3}) = \{([0, 0.5), 0.583)\,, \ ([0.5, 1]\,, 1.083)\}$.

In the bottom panel, we lower the bias to $b = \frac{1}{6}$ and keep $\widehat{K} = 2$. There is a unique optimal contract $\mathcal{C}^*(2, \frac{1}{6}) = \{([0.101, 0.449]\,, 0.442)\,, ([0.551, 0.899]\,, \ 0.891)\}$, which induces three communication actions. This is, again, more than the maximal number of two actions that can be induced in an equilibrium of the communication game without contracting. Hence, if we keep the bias fixed while increasing the bound on the number of clauses, contracting drives out communication. If, instead, we lower the bias while fixing the upper bound on the number of contract clauses, communication replaces contracting.

6.1 Non-constant Bias

Suppose that instead of the principal's bias being constant at $b = \frac{1}{3}$, it is state dependent and of the form $b(\theta) = \frac{1}{3} + \frac{1}{30}\theta$. Continue to assume that the players' loss functions are quadratic and the state is uniformly distributed on $[0, 1]$.

The structure of the optimal contract is the same as for a constant bias, it allows for two communication actions. In particular, we have $\mathcal{C}_b^*\left(1, \frac{1}{3} + \frac{\theta}{30}\right) = \{([0.183, 0.905], 0.895)\}$.

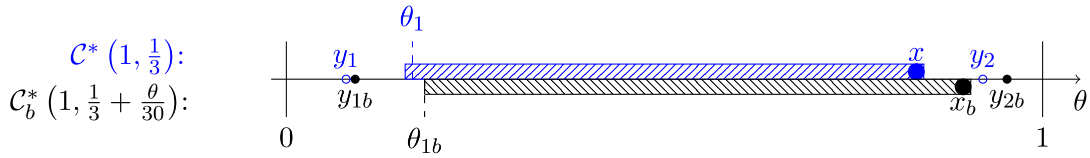

Figure 9: Optimal contracts with constant bias $b = \frac{1}{3}$ on top (blue) and with state-dependent bias $b(\theta) = \frac{1}{3} + \frac{1}{30}\theta$ below (black).

For an illustration see Figure 9. The condition on top of the axis refers to the optimal contract with constant bias while the condition below the axis indicates the optimal contract with state-dependent bias. The figure illustrates the intuitive impact of an increasing bias: the optimal condition shifts upwards and the size of the condition increases. The principal prefers covering states with a higher bias to covering states with a smaller bias, because under communication the agent's action diverges more from the principal's preferred one.

Note that for a larger increase of the bias, the upper communication interval can vanish. For example, with $b(\theta) = \frac{1}{3} + \frac{1}{10}\theta$ the optimal contract $\mathcal{C}_b^*\left(1, \frac{1}{3} + \frac{\theta}{10}\right) = \{([0.284, 1], 1.040)\}$ induces one communication step that is below the condition.

6.2 Nonuniform Distribution

Instead of the state being uniformly distributed, assume now that it is distributed on $[0, 1]$ with density $f(\theta) = \frac{9}{10} + \frac{2}{10}\theta$. Maintain that the players' loss functions are quadratic and that there is a constant bias $b = \frac{1}{3}$.

The optimal contract has one communication action below and one above the condition and is given by $\mathcal{C}_f^*\left(1, \frac{1}{3}\right) = \{([0.184, 0.885], 0.876)\}$. For an illustration, see Figure 10 the contract below the axis.

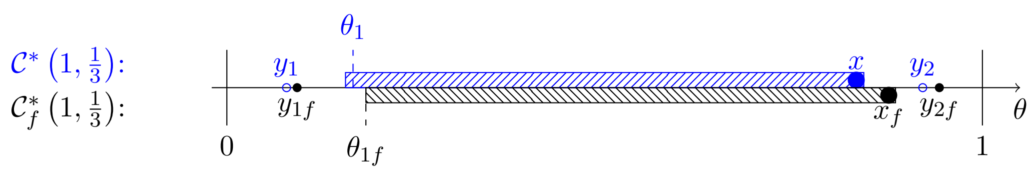

Figure 10: Optimal contracts for a uniform distribution on top and with distribution $f(\theta) = \frac{9}{10} + \frac{2}{10}\theta$ below.

For the distribution with an increasing density compared to a uniform distribution, the optimal instruction as well as the optimal condition shift upwards. The principal prefers covering states that occur more frequently in the contract rather than states that have a lower probability: for states covered by the condition the principal gets her preferred action rather than the agent's preferred action. The contract is used to relax incentive constraints, and thereby makes communication feasible, when otherwise it would not be.

6.3 Transfers

For the main analysis, we abstain from modeling transfers from the principal to the agent. Two common uses of transfers in the literature do not apply to our setup. Under moral hazard, the agent needs to be incentivized to take particular actions; here, however, actions that are governed by the contract are fully under the control of the principal. Under screening, the principal tries to gather information about the agent's private type, whereas in our setup the agent does not have private information. The following example considers the case in which the principal needs to pay a fixed wage to hire the agent. The wage payment is a non-contingent transfer agreed upon as part of the contract offer and acceptance.

If we assume that both the principal's and agent's payoffs are quasi-linear in the wage and that the principal's wage offer needs to meet an individual rationality constraint for the agent, it follows that the principal's problem reduces to maximizing weighted social surplus. That is, the principal maximizes $W(y, \theta, b) = -(1-\alpha)(\theta + b - y)^2 - \alpha(\theta - y)^2$ for some $\alpha \in (0, 1)$. The weight α is a function of the importance that principal and/or agent attach to the wage payment relative to the payoffs U^P and U^A, which derive from the action taken. If, for example, the agent cares primarily about the wage and little about the action utility U^A, then, all else equal, α will be small, and we approximate the setup used in the rest of the paper.

Consider, for example, a bias of $b = \frac{1}{3}$ and a weight $\alpha = 0.01$ on the agent's payoff, which corresponds to a slight departure from the setup without transfers that we consider in our main analysis. The optimal contract with transfers allows for two-step communication and is given by $\mathcal{C}_t^*(1, \frac{1}{3}) = \{([0.160, 0.840], 0.830)\}$. For an illustration see Figure 11. The condition on top of the axis refers to the optimal contract without transfers while the condition below the axis indicates the optimal contract with transfers.

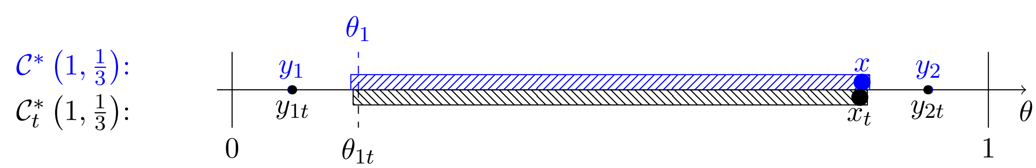

Figure 11: Optimal contract with transfers for $b = \frac{1}{3}$ and $\alpha = 0.01$.

Since the introduction of a transfer changes the principal's problem to one closer to joint-surplus maximization, the principal has to take the agent's payoff into account. It is, therefore, intuitive that the optimal instruction decreases in the direction of the agent-preferred action, which is at the midpoint of the condition. Moreover, the trade-off between communication (agent-optimal action) and contract (principal-optimal instruction) becomes less extreme. As a result, the length of the condition shrinks.[22]

7 Related Literature

Simon (1951) is the first to draw attention to the importance of contractual incompleteness. He notes that many contracts take the form of an "employment contract." An employment contract, in exchange for a fixed wage, transfers authority to the principal rather than providing a detailed specification of the agent's action. In our setting, also, the principal forgoes a detailed specification of the agent's actions, but unlike in Simon (1951), for actions not controlled by the contract, authority resides with the agent, and the principal resorts to communication to influence the agent's action.

Writing costs are sometimes used to rationalize contractual incompleteness. Dye (1985) is the first to make writing and monitoring cost explicit. He notes that contracts with specifications so detailed that they are sensitive to every state are prohibitively expensive to write. The contracts he considers consist of finite lists of clauses, with conditions partitioning the state space. The cost of writing a contract is increasing in the number of clauses.

Battigalli and Maggi (2002) explore the foundations of writing costs by making the language in which contracts are written explicit. A contract specifies a list of clauses and a transfer. Clauses map contingencies into instructions. More elaborate clauses require more "primitive sentences" and are therefore more costly. This results in two types of contractual incompleteness: *rigidity* – insufficient dependence on the state of the world; and *discretion* – insufficient precision in the prescription of behavior. Our environment also gives rise to rigidity and discretion: whenever the optimal contract does not cover all states, the state space splits into a contracting region and a communication region. We have rigidity in the contracting region and discretion in the communication region. Greater alignment of interests, which facilitates communication, favors discretion, and *vice versa*.

Shavell (2006) (see also Schwartz and Watson (2013)) studies the impact of contract interpretation by courts on the writing of contracts. Again, contracts are lists of clauses, each comprised of a condition and an instruction.[23] Because of writing costs, contracts may contain *gaps* – sets of states not covered by any condition. One role of interpretation is to fill gaps, another to replace stated with interpreted clauses. The prospect of interpretation, like the prospect of communication in our setting, shapes how contracts are written.

Since Simon (1951), the interplay of information and authority has played an important role in the study of organizations. Aghion and Tirole (1997) distinguish the right to make a decision (formal authority) from the power to influence a decision (real authority). Either

[22]For the case of joint-surplus maximization, i.e., $\alpha = \frac{1}{2}$, the optimal contract is $\mathcal{C}_t^*(1, \frac{1}{3}) = \{[0.104, 0.646], 0.542\}$. As expected, the condition is shorter and the instruction closer to the midpoint of the condition, compared to a contract with smaller α. Note that the clause is not at the center of the state space; the length of the condition is short relative to the size of the bias, and therefore to maintain incentive compatibility the condition has to be moved towards zero.

[23]Heller and Spiegler (2008) allow for contradictory clauses, in which conditions overlap, but the corresponding instructions differ.

the principal or the agent has formal authority. Real authority requires information that players can acquire at a cost. There is no explicit model of communication.

Dessein (2002) examines the conditions under which an uninformed principal cedes authority to a better-informed agent.[24] He adopts an incomplete contracting approach in which authority, but not actions, can be contracted upon. The principal has a choice between delegating decision rights to the agent and making decisions herself after communicating with the agent. In our setting, the principal has the informational advantage but may cede authority to the agent if sufficiently closely aligned incentives make communication attractive.

In Deimen and Szalay (2019) the principal can choose whether to delegate decision rights to an agent or to rely on communication with the agent. Depending on the principal's choice, the agent decides how much and what kind of information to acquire. In contrast, in our setup there is communication when the decision rights are left with the agent, and it is from the principal to the agent.

Aumann and Hart (2003), Golosov, Skreta, Tsyvinski and Wilson (2014), and Krishna and Morgan (2004) examine different versions of models with repeated cheap talk. One feature that these models have in common with ours is that new communication opportunities may arise as the result of subsets of types having been removed: if at some stage the sender sends a message that is only used by a strict subset of types, at the following stage the receiver can concentrate beliefs on the remaining types. Removing types may facilitate communication for the remaining types since fewer incentive constraints have to be dealt with. In Aumann and Hart (2003) and Krishna and Morgan (2004) types exit because they prefer not to take their chances in a jointly controlled lottery. In Golosov et al. (2014) types are induced to exit by receiver actions that follow each communication round. In our setting, types are removed from the communication game by being covered by a condition in the contract.

In Krishna and Morgan (2008), an uninformed principal incentivizes a privately informed agent to reveal information, using contracts with message contingent payments. In the uniform-quadratic environment, when the principal can contract on transfers but not actions, optimal contracts use payments to induce full revelation for low types and offer no payments to high types. High types convey information through cheap talk and, as in CS, induce a bounded number of actions. Unlike in the present paper, both the contracting region and the communication region are connected; contracting is not used to facilitate communication.

We abstain from modeling transfers explicitly in the main analysis, consistent with Battigalli and Maggi (2002), Shavell (2006), Dessein (2002), and others. In our environment, transfers play no role in providing incentives to supply information or to induce actions.

[24] In the literature on optimal delegation (See, for example, Holmström (1977), Holmström (1984), Melumad and Shibano (1991), Szalay (2005), Alonso and Matouschek (2008), Kováč and Mylovanov (2009), and Amador and Bagwell (2013)), the uninformed principal decides how to optimally constrain the decision rights of the informed agent.

8 Conclusion

Our exploration suggests that when it is difficult to write detailed contracts, there may be a role for non-binding communication. The communication option affects both the scope and the form of contracts that we expect to observe. Greater ease of writing contracts and stronger divergence of incentives favor expanding the scope of optimal contracts. With sufficiently strong divergence of interests, contracts will tend to be obligationally complete. With more closely aligned interests, in contrast, we expect influential communication. Contracts of the appropriate form can help create and enhance a role for communication: The insertion of contract clauses between communication events relaxes incentive constrains for communication. Separating communication events in this manner effectively creates common interest between the contracting parties to have (at least) those events be communicated.

In our formalization of these ideas we have maintained a clean distinction of contracts and communication. For states governed by contracts there is no further cheap-talk communication and for the remaining states there is a well-defined cheap-talk game that can be analyzed in isolation. In future work, one may want to relax this complete decoupling of contracts and communication. With imperfect enforcement of contract clauses, for example, at the interim stage the principal would sometimes prefer to send one of the cheap-talk messages, rather than insist on having the applicable contract clause enforced. One may also want to allow for contract clauses that leave room for communication in states where the clauses apply. We expect the takeaway from our paper, that optimal contracts reflect concerns with communication incentives and sometimes facilitate communication, to go through. Working out the details will be more delicate when contracts no longer induce isolated cheap-talk games.

A Appendix

Lemma A.1 *There exists an $\varepsilon > 0$, uniform over all communication subgames $\Gamma^{\mathcal{C}}$, such that for every equilibrium in $\Gamma^{\mathcal{C}}$ and all actions y and y' induced in that equilibrium, $|y - y'| \geq \varepsilon$. There is an upper bound on the number of actions that are induced in equilibrium that is uniform across all communication subgames*

Proof. This is essentially a restatement of CS's Lemma 1. That ε is uniform over all communication subgames follows from the fact that the type distribution plays no role in the proof. $\square$

Lemma A.2 *For all $b \geq 0$ and all $\eta > 0$, there exists a $\gamma > 0$ such that for all $\Phi \subseteq [0,1]$ with* $\text{Prob}(\Phi) \geq \eta$,

$$\int_{\Phi} U^P(y^P(\theta,b),\theta,b)dF(\theta) - \int_{\Phi} U^P(y^{*P}(\Phi,b),\theta,b)dF(\theta) > \gamma.$$

Proof. By continuity of f and compactness of $[0,1]$, f is bounded. Therefore, for all $\delta > 0$ there is an $\varepsilon_0 > 0$ such that for all $\Phi \subseteq [0,1]$ with $\text{Prob}(\Phi) > \delta$, $\ell(\Phi) > \varepsilon_0$ (where ℓ denotes Lebesgue measure). Hence, for all $\delta > 0$ there is an $\varepsilon_1 > 0$ such that for all $\Phi \subseteq [0,1]$ with $\text{Prob}(\Phi) > \delta$, for all $\theta \in [0,1]$ there exists $\Psi \subseteq \Phi$ such that $|\theta - \theta'| > \varepsilon_1$ for all $\theta' \in \Psi$ and $\ell(\Psi) > \varepsilon_1$. This and the fact that $y^{*P}(\Phi,b)$ is the ideal point of some type $\theta(\Phi) \in [0,1]$ imply that for all $\delta > 0$ there is an $\varepsilon_1 > 0$ such that for all $\Phi \subseteq [0,1]$ with $\text{Prob}(\Phi) > \delta$, there exists $\Psi \subseteq \Phi$ such that $|\theta(\Phi) - \theta'| > \varepsilon_1$ for all $\theta' \in \Psi$ and $\ell(\Psi) > \varepsilon_1$.

Since the derivative of y^P is strictly positive and continuous it has a strictly positive lower bound. Therefore, for all $\varepsilon_1 > 0$ we can find $\varepsilon_2 > 0$ such that for all $\theta, \theta' \in [0,1]$ with $|\theta - \theta'| > \varepsilon_1$, we have $|y^P(\theta,b) - y^P(\theta',b)| > \varepsilon_2$. This and the continuity of U^P imply that for all $\varepsilon_1 > 0$ we can find $\varepsilon_3 > 0$ such that for all $\theta, \theta' \in [0,1]$ with $|\theta - \theta'| > \varepsilon_1$, we have $U^P(y^P(\theta,b),\theta)$ - $U^P(y^P(\theta',b),\theta) > \varepsilon_3$. This, the fact that f is everywhere positive, and the observation at the end of the previous paragraph imply the statement. $\square$

Proof of Proposition 1.
Part (1) Suppose $\mathcal{C}$ is an optimal contract in $G(\widehat{K},b)$. If the contract is empty, $K = 0$, or the union of conditions has probability zero, then $\Gamma^{\mathcal{C}}$ is a CS game. Hence, each equilibrium action in an equilibrium of $\Gamma^{\mathcal{C}}$ is induced by an interval of types. Consider an optimal equilibrium $e^{\mathcal{C}}$ of $\Gamma^{\mathcal{C}}$. Since there are only finitely many equilibrium actions, there is an action $\hat{y}$ that is induced with positive probability. Let $[\underline{\theta},\overline{\theta}]$ be the closure of the set of types who induce action $\hat{y}$ in $e^{\mathcal{C}}$. For every $\varepsilon > 0$ such that $\tau + \varepsilon < \overline{\theta}$, there is a set $[\tau, \tau+\varepsilon] \subset [\underline{\theta},\overline{\theta}]$ with $y^{*A}([\tau,\tau+\varepsilon]) = \hat{y}$. Evidently, also $y^{*A}\left([\underline{\theta},\overline{\theta}] \setminus [\tau,\tau+\varepsilon]\right) = \hat{y}$. Since $y^P(\theta,b) \neq y^A(\theta)$ and both y^P and y^A are continuous and $[0,1]$ is compact, there exists $\varepsilon_0 > 0$ such that $|y^P(\theta,b) - y^A(\theta)| > \varepsilon_0$ for all $\theta \in [0,1]$. Continuity of y^P and y^A and compactness of $[0,1]$ further imply that there exists $\delta > 0$ such that $|y^P(\theta,b) - y^A(\theta+\delta)| > \varepsilon_0$ for all $\theta \in [0,1]$. Hence, if we choose $\varepsilon < \delta$ then $y^{*P}([\tau,\tau+\varepsilon],b) > y^{*A}([\tau,\tau+\varepsilon)] = \hat{y}$. Hence, the

alternative contract $\mathcal{C}' = \{(C_1, x_1)\}$, where $C_1 = [\tau, \tau + \varepsilon)]$ and $x_1 = y^{*P}([\tau, \tau + \varepsilon], b)$ allows an equilibrium $e^{\mathcal{C}'}$ in $\Gamma^{\mathcal{C}'}$ in which types outside of $[\tau, \tau + \varepsilon]$ induce the same actions and receive the same payoffs as in the equilibrium $e^{\mathcal{C}}$ in $\Gamma^{\mathcal{C}}$, while the principal is strictly better off if condition C_1 is realized. It follows that $K \geq 1$, and therefore an optimal contract is never empty.

Consider any contract $\mathcal{C}$ with $K < \widehat{K}$ and an optimal equilibrium in the communication game $\Gamma^{\mathcal{C}}$. Consider replacing the contract $\mathcal{C}$ by a contract $\mathcal{C}'$ that splits the condition $C_K = [\underline{C}_K, \overline{C}_K]$ (taking the condition C_K to be closed is without loss of generality) into two conditions $\tilde{C}_K = [\underline{C}_K, \tilde{C})$ and $\tilde{\tilde{C}}_K = [\tilde{C}, \overline{C}_K)$ with $\underline{C}_K < \tilde{C} < \overline{C}_K$ and leaves all other clauses unchanged. Then $U^P_{11} < 0$ and $U^P_{12} > 0$ imply that $y^{*P}(\tilde{C}_K, b) < y^{*P}(C_K, b) < y^{*P}(\tilde{\tilde{C}}_K, b)$, which implies that the principal is strictly better off under the new contract, conditional on the event C_K being realized, while incentives in the communication games $\Gamma^{\mathcal{C}'}$ and $\Gamma^{\mathcal{C}}$ are identical. This implies that optimal contracts must have $K = \widehat{K}$.

Part (2) Suppose not. Then there is a $b > 0$ and a sequence of gaps $(\mathcal{L}_{\widehat{K},b})_{\widehat{K}=1}^{\infty}$ arising in optimal equilibria $e(\widehat{K}, b)$ of $G(\widehat{K}, b)$ with a subsequence $(\mathcal{L}_{\widehat{K}_i,b})_{i=1}^{\infty}$ and $\kappa > 0$ such that $\text{Prob}(\mathcal{L}_{\widehat{K}_i,b}) > \kappa$ for all i. From Lemma A.1, there is an upper bound $\widehat{N}$ on the number of actions induced in any equilibrium of any communication subgame. Hence for every $\widehat{K}_i$, $i = 1, \ldots,$ there is an action that is induced by a subset $\Phi_{\widehat{K}_i}$ of $\mathcal{L}_{\widehat{K}_i,b}$ that has at least probability $\frac{\kappa}{\widehat{N}}$. Hence, by Lemma A.2 there exists $\varepsilon > 0$ such that

$$\int_{\Phi_{\widehat{K}_i}} U^P(y^P(\theta, b), \theta, b)dF(\theta) - \int_{\Phi_{\widehat{K}_i}} U^P(y^{*P}(\Phi_{\widehat{K}_i}, b), \theta, b)dF(\theta) > \varepsilon$$

for all $i = 1, \ldots.$ This implies that for every $i = 1, \ldots$ the principal's payoffs in $e(\widehat{K}_i, b)$ are bounded from above by

$$\int_{[0,1]} U^P(y^P(\theta, b), \theta, b)dF(\theta) - \varepsilon.$$

Continuity of y^P follows from the maximum theorem and uniform continuity from the fact that $[0, 1]$ is compact. By assumption U^P is continuous. Uniform continuity of U^P follows from compactness of $\left[\min_{\theta\in[0,1]} y^P(\theta, b), \max_{\theta\in[0,1]} y^P(\theta, b)\right] \times [0, 1]$. For any $\widehat{K}$, partition the interval $[0, 1]$ into $\widehat{K}$ equal length intervals $I_1 := [\theta_0, \theta_1]$ and $I_k := (\theta_{k-1}, \theta_k]$, $k = 2, \ldots, \widehat{K}$. For each $\widehat{K} = 1, 2, \ldots$, define the function $U^P_{\widehat{K}} : [0, 1] \to \mathbb{R}$ by the property that $U^P_{\widehat{K}}(\theta) = U^P(y^P(\theta_k, b), \theta, b)$ for all $\theta \in I_k$ and all $k = 1, \ldots, \widehat{K}$. Then $\int_{[0,1]} U^P_{\widehat{K}}(\theta)dF(\theta)$ is the principal's payoff from writing the contract $\mathcal{C}_{\widehat{K}} = \{(C_k, x_k)\}_{k=1}^{\widehat{K}}$ where $C_k = I_k$ and $x_k = y^P(\theta_k, b)$. Uniform continuity of y^P and U^P imply that for any $\tilde{\varepsilon} > 0$ we can choose $\widehat{K}$ sufficiently large (and therefore $\delta := \theta_k - \theta_{k-1}$ appropriately small) such that $0 \leq U^P(y^P(\theta, b), \theta, b) - U^P_{\widehat{K}}(\theta) < \tilde{\varepsilon}$ for all $\theta \in [0, 1]$. Therefore we have

$$\lim_{\widehat{K}\to\infty} \int_{[0,1]} U^P_{\widehat{K}}(\theta)dF(\theta) = \int_{[0,1]} U^P(y^P(\theta, b), \theta, b)dF(\theta),$$

which contradicts the supposition that $e(\widehat{K}_i, b)$ is optimal in $G(\widehat{K}_i, b)$ for all $i = 1, 2, \ldots$.

Part (3) Suppose not. Then there is an $\varepsilon_0 > 0$ and a sequence of gaps $(\mathcal{L}_{\widehat{K},b_j})_{j=1}^{\infty}$ arising in optimal equilibria $e(\widehat{K}, b_j)$ of $G(\widehat{K}, b_j)$ with $b_j \to 0$ and $\text{Prob}(\mathcal{L}_{\widehat{K},b_j}) < 1-\varepsilon_0$ for all j. Hence, for every j there is a condition C^j in the contract $\mathcal{C}^j$ that is part of the optimal equilibrium $e(\widehat{K}, b_j)$ with $\text{Prob}(C^j) \geq \frac{\varepsilon_0}{\widehat{K}}$. By Lemma A.2 there is an $\varepsilon_1 > 0$ such that

$$\int_{C^j} U^A(y^A(\theta), \theta, 0)dF(\theta) - \int_{C^j} U^A(y^{*A}(C^j), \theta, 0)dF(\theta) > \varepsilon_1$$

for all j. The space of intervals of length ℓ, $\frac{\varepsilon_0}{\widehat{K}} \leq \ell \leq 1$ is compact. Hence, the sequence $(C^j)_{j=1}^{\infty}$ has a convergent subsequence. After reindexing, use $(C^j)_{j=1}^{\infty}$ to denote that subsequence in the sequel, and denote the limit by C. By continuity,

$$\int_{C} U^P(y^P(\theta, 0), \theta, 0)dF(\theta) - \int_{C} U^P(y^{*P}(C, 0), \theta, 0)dF(\theta) \geq \varepsilon_1.$$

Hence, appealing to continuity again, for sufficiently large j,

$$\int_{C^j} U^P(y^P(\theta, b_j), \theta, b_j)dF(\theta) - \int_{C^j} U^P(y^{*P}(C^j, b_j), \theta, b_j)dF(\theta) \geq \frac{\varepsilon_1}{2}.$$

This implies that for sufficiently large j in this sequence the principal's payoffs in the equilibria $e(\widehat{K}, b_j)$ are bounded away from $\int_{[0,1]} U^P(y^P(\theta, 0), \theta, 0)dF(\theta)$. This contradicts optimality of the equilibria in the sequence $(e(\widehat{K}, b_j))$, since by the continuity property the communication games $\Gamma^0(b_j)$ have equilibria whose payoffs converge to $\int_{[0,1]} U^P(y^P(\theta, 0), \theta, 0)dF(\theta)$ with $j \to \infty$. $\square$

The remainder of the appendix establishes the results in Section 5. We begin by showing that optimal contracts are either obligationally complete or induce influential communication. We also characterize the conditions for either to be the case in terms of the parameters b and $\widehat{K}$.

Proposition A.1 *(1) If $\widehat{K} \geq \frac{1}{2b}$, then any optimal contract is obligationally complete.*
(2) Optimality requires influential communication if and only if $\widehat{K} < \frac{1}{2b}$.

Proof. (1) In any communication subgame, for any strategy profile, the agent's payoff equals the expected conditional variance and the principal's payoff differs from that by a constant $-\lambda b^2$ when the communication region has size λ. Therefore, the principal's expected payoff from a contract with $\widehat{K}$ conditions that specifies a communication region of size λ is bounded from above by

$$-\lambda b^2 - \widehat{K} \int_0^{\frac{1-\lambda}{\widehat{K}}} \left(x - \frac{1-\lambda}{2\widehat{K}}\right)^2 dx = -\lambda b^2 - \frac{1}{12}\frac{1}{\widehat{K}^2}(1-\lambda)^3.$$

The derivative of this expression with respect to λ, $-b^2 + \frac{(1-\lambda)^2}{4\widehat{K}^2}$, is negative for $b^2 \geq \frac{1}{4\widehat{K}^2}$. Therefore, for $\widehat{K} \geq \frac{1}{2b}$ it is optimal to reduce the size λ of the communication region to zero.

(2) Consider $\widehat{K} < \frac{1}{2b}$. The principal's expected payoff under an obligationally complete contract with $\widehat{K}$ conditions is

$$\widehat{K} \cdot \int_0^{\frac{1}{\widehat{K}}} -\left(\frac{1}{2\widehat{K}} - s\right)^2 ds = -\frac{1}{12\widehat{K}^2}.$$

The principal's expected payoff under one-step communication, where l denotes the length of the communication interval, is given by

$$-\int_0^l \left(\frac{l}{2} - b - s\right)^2 ds - \widehat{K} \cdot \int_0^{\frac{1-l}{\widehat{K}}} \left(\frac{1-l}{2\widehat{K}} - s\right)^2 ds = -lb^2 - \frac{l^3}{12} - \frac{(1-l)^3}{12\widehat{K}^2}.$$

The first-order condition for the optimal length l^* of the communication interval implies that $l^* = \frac{\widehat{K}\sqrt{1-4b^2(\widehat{K}^2-1)}-1}{\widehat{K}^2-1}$ for $\widehat{K} > 1$ and $l^* = \frac{1}{2} - 2b^2$ for $\widehat{K} = 1$. The second derivative is $-\frac{1+l(\widehat{K}^2-1)}{2\widehat{K}^2} < 0$. The optimal length satisfies $l^* > 0$ for $b < \frac{1}{2\widehat{K}}$, the case we are considering.

Inserting the optimal length l^* into the one-step-communication payoff and comparing it to the payoff from the obligationally complete contract, we find that the obligationally complete contract yields a lower payoff when $\widehat{K} = 1$ and $b < \frac{1}{2\widehat{K}}$, and when $\widehat{K} > 1$ and $b \leq \frac{1}{2\sqrt{\widehat{K}^2-1}}$. The latter inequality is satisfied for the case we are considering since $\widehat{K} < \frac{1}{2b}$ is equivalent to $b < \frac{1}{2\widehat{K}}$ and because $\frac{1}{2\widehat{K}} < \frac{1}{2\sqrt{\widehat{K}^2-1}}$.

Evidently, splitting the one-step communication interval in half raises the principal's payoff – as long as incentive compatibility is satisfied. Consider two communication intervals, $\left[0, \frac{l^*}{2}\right]$ and $\left[1 - \frac{l^*}{2}, 1\right]$. The relevant incentive constraint is

$$\frac{l^*}{2} + b - \frac{l^*}{4} \leq 1 - \frac{l^*}{4} - b - \frac{l^*}{2}$$

$$\Leftrightarrow l^* \leq 1 - 2b \Leftrightarrow b \leq \frac{1}{2\sqrt{\widehat{K}^2 - 1}}.$$

Since we are considering the case $b < \frac{1}{2\widehat{K}}$, this condition is satisfied. □

The following two observations record simple facts that we refer to in the proofs. The first of these notes that both players gain from reducing variance.

Observation A.1 *Suppose that, given any distribution over $[\underline{\theta}, \overline{\theta}] \subseteq [0, 1]$, either the principal or the agent takes an optimal action. Then, the principal's and the agent's expected payoffs are decreasing in the variance of that distribution.*

The next observation records the fact that all else equal, the principal gains from equalizing the size of communication intervals.

Observation A.2 *Let* $\underline{\theta}_i < \overline{\theta}_i \leq \underline{\theta}_j < \overline{\theta}_j$ *and* $\overline{\theta}_j - \underline{\theta}_j - \delta > \overline{\theta}_i - \underline{\theta}_i + \delta$. *Suppose that the agent takes action* $y_i^\delta = \frac{\underline{\theta}_i + \overline{\theta}_i + \delta}{2}$ *for types in* $\left(\underline{\theta}_i, \overline{\theta}_i + \delta\right)$ *and action* $y_j^\delta = \frac{\underline{\theta}_j + \overline{\theta}_j + \delta}{2}$ *for types in* $\left(\underline{\theta}_j + \delta, \overline{\theta}_j\right)$ *. Then, the expected payoff of the principal conditional on* $\left(\underline{\theta}_i, \overline{\theta}_i + \delta\right) \cup \left(\underline{\theta}_j + \delta, \overline{\theta}_j\right)$ *is increasing in* δ.

Lemma A.3 notes that inside any communication interval there can be at most a single condition. Intuitively, when there are more conditions inside a communication interval we can translate the outermost conditions further to the extremes in a way that does not affect the agent's optimal action in that interval, and therefore does not upset the principal's communication incentives. Doing so reduces variance and therefore raises the principal's *ex ante* payoff.

Lemma A.3 *For any optimal contract* $\mathcal{C}$ *and any communication interval* $(\underline{\theta}, \overline{\theta})$ *of an optimal equilibrium* $e^{\mathcal{C}}$ *of the communication subgame* $\Gamma^{\mathcal{C}}$ *there is no more than one condition* C *with* $C \subset (\underline{\theta}, \overline{\theta})$.

Proof. Suppose that for an equilibrium $e^{\mathcal{C}}$ there is a communication interval $\left(\underline{\theta}, \overline{\theta}\right)$ for which the conditions C_ℓ, $\ell = 1, \ldots, k$, are the ones satisfying $C_\ell \subset \left(\underline{\theta}, \overline{\theta}\right)$. Then the communication action induced by the types in $\left(\left(\underline{\theta}, \overline{\theta}\right) \cap \mathcal{L}(\mathcal{C})\right)$ solves

$$\max_y - \int_{\underline{\theta}}^{\underline{C}_1} (s-y)^2 \, ds - \sum_{\ell=1}^{k-1} \int_{\overline{C}_\ell}^{\underline{C}_{\ell+1}} (s-y)^2 \, ds - \int_{\overline{C}_k}^{\overline{\theta}} (s-y)^2 \, ds,$$

with the solution given by

$$y^{*A}\left(\left(\underline{\theta}, \overline{\theta}\right) \cap \mathcal{L}(\mathcal{C})\right) = \frac{1}{2} \frac{\overline{\theta}^2 - \sum_{\ell=1}^k \overline{C}_\ell^2 + \sum_{\ell=1}^k \underline{C}_\ell^2 - \underline{\theta}^2}{\overline{\theta} - \sum_{\ell=1}^k \overline{C}_\ell + \sum_{\ell=1}^k \underline{C}_\ell - \underline{\theta}}.$$

Since $\underline{C}_1 > \underline{\theta}$, and $\overline{C}_k < \overline{\theta}$, for sufficiently small ε the $(-\varepsilon)$-translation C_1' of C_1 and the δ-translation C_k' of C_k satisfy $C_1', C_k' \subset \left(\underline{\theta}, \overline{\theta}\right)$. Consider the contract $\mathcal{C}'$ that is obtained from $\mathcal{C}$ by replacing the condition C_1 by C_1', replacing the condition C_k by C_k', and – in case C_1 and/or C_k belong to a condition cluster – forming the closure of the union of conditions thus obtained.

Then, $y^{*A}\left(\left(\underline{\theta}, \overline{\theta}\right) \cap \mathcal{L}(\mathcal{C}')\right) =$

$$\frac{1}{2} \frac{\overline{\theta}^2 - \sum_{l \neq 1,k} \overline{C}_l^2 - \left(\overline{C}_1 - \varepsilon\right)^2 - \left(\overline{C}_k + \delta\right)^2 + \sum_{l \neq 1,k} \underline{C}_l^2 + \left(\underline{C}_1 - \varepsilon\right)^2 + \left(\underline{C}_k + \delta\right)^2 - \underline{\theta}^2}{\overline{\theta} - \sum_{l \neq 1,k} \overline{C}_l - \left(\overline{C}_1 - \varepsilon\right) - \left(\overline{C}_k + \delta\right) + \sum_{l \neq 1,k} \underline{C}_l + \left(\underline{C}_1 - \varepsilon\right) + \left(\underline{C}_k + \delta\right) - \underline{\theta}}.$$

If we require that $y^{*A}\left(\left(\underline{\theta}, \overline{\theta}\right) \cap \mathcal{L}(\mathcal{C}')\right) = y^{*A}\left(\left(\underline{\theta}, \overline{\theta}\right) \cap \mathcal{L}(\mathcal{C})\right)$, we find that this is equivalent to letting $\delta = \varepsilon \frac{\overline{C}_1 - \underline{C}_1}{\overline{C}_k - \underline{C}_k}$.

This implies that the game $\Gamma^{\mathcal{C}'}$ has an equilibrium $e^{\mathcal{C}'}$ in which the agent's strategy is the same as in $e^{\mathcal{C}}$, the principal's strategy is the same for all types in $\mathcal{L}(\mathcal{C}') \setminus (\underline{\theta}, \overline{\theta})$, and types in $(\underline{\theta}, \overline{\theta}) \cap \mathcal{L}(\mathcal{C}')$ send a common message not sent by any of the other types. The change in payoffs from replacing the contract-equilibrium pair $(\mathcal{C}, e^{\mathcal{C}})$ by the pair $(\mathcal{C}', e^{\mathcal{C}'})$ is given by:

$$
-\int_{\overline{C}_1-\varepsilon}^{\overline{C}_1} (s+b-\hat{y})^2\, ds + \int_{\underline{C}_1-\varepsilon}^{\underline{C}_1} (s+b-\hat{y})^2\, ds - \int_{\underline{C}_k}^{\underline{C}_k+\delta} (s+b-\hat{y})^2\, ds + \int_{\overline{C}_k}^{\overline{C}_k+\delta} (s+b-\hat{y})^2\, ds
$$

$$
= \varepsilon \frac{\overline{C}_1 - \underline{C}_1}{\overline{C}_k - \underline{C}_k} \left(\left(\overline{C}_k + \underline{C}_k - \overline{C}_1 - \underline{C}_1\right) \left(\overline{C}_k - \underline{C}_k\right) + \varepsilon \left(\overline{C}_k - \underline{C}_k + \overline{C}_1 - \underline{C}_1\right) \right).
$$

This expression is strictly positive since $\varepsilon > 0$, $\overline{C}_1 - \underline{C}_1 > 0$, $\overline{C}_k - \underline{C}_k > 0$, and $\overline{C}_k + \underline{C}_k > \overline{C}_1 + \underline{C}_1$. □

Proof of Proposition 2. Part I – partitional. To establish that optimal equilibria are partitional, we show that for every condition cluster $\boldsymbol{C}$, there is a critical type $\theta \in \boldsymbol{C}$.

Since for every equilibrium in which the principal mixes there is an outcome equivalent equilibrium in which her strategy is pure, it is without loss of generality to have the principal's strategy be pure in the equilibrium $e^{\mathcal{C}}$. Denote the strategy profile corresponding to the equilibrium $e^{\mathcal{C}}$ by $f^{\mathcal{C}} = \left(\sigma^{\mathcal{C}}, \rho^{\mathcal{C}}\right)$. It follows from Lemma A.3 that it suffices to look at the case where the interior of each communication interval of the equilibrium $e^{\mathcal{C}}$ contains at most one condition. Hence, it suffices to show that for any $k = 1, \ldots, \widehat{K}$, the condition C_k does not belong to the interior of a communication interval for the equilibrium $e^{\mathcal{C}}$.

Suppose otherwise, i.e., for the contract $\mathcal{C}$ and the equilibrium $e^{\mathcal{C}}$ there is at least one communication interval with a condition in its interior. We will gradually replace the contract $\mathcal{C}$ by other contracts and the strategy profile $f^{\mathcal{C}}$ by other strategy profiles. At each iteration, we will ensure that the principal's payoffs strictly increase. At the end, we will verify that the strategy profile we obtain is an equilibrium profile.

Let the equilibrium $e^{\mathcal{C}}$ have n steps, and therefore n communication intervals I_j, $j = 1, \ldots, n$. For each communication interval I_j let the principal send message m_j and denote the action induced by types in I_j by y_j. Denote the critical types from equilibrium $e^{\mathcal{C}}$ by $\theta_j^{\mathcal{C}}$, $j = 0, 1, \ldots, n$. At each replacement of the prevailing contract and strategy profile, the number of steps as well as the number communication intervals remains constant at n. Types in communication interval I_j continue to send message m_j after each replacement and the agent best responds to the replacement of the principal's strategy. After all unsent messages, have the agent use the same response as after message m_1. As the response to m_1 changes with each replacement, change the response to unsent messages in the same way.

Step 1. Replace the contract $\mathcal{C}$ and the strategy profile $f^{\mathcal{C}}$ by a new contract $\mathcal{C}_0$ and a new strategy profile $f^{\mathcal{C}_0}$:

(a) Change the contract as follows: Consider any condition C_k such that there is a communication interval $I_j = (\underline{\theta}_j, \overline{\theta}_j)$ with $C_k \subset (\underline{\theta}_j, \overline{\theta}_j)$. If $\underline{\theta}_j$ does not belong to a condition, replace C_k by its $-(\underline{C}_k - \underline{\theta}_j)$-translation. If $\underline{\theta}_j$ does belong to a condition, replace C_k by the $-(\underline{C}_k - \underline{\theta}_j)$-translation of the left-open interval $C_k \setminus \{\underline{C}_k\}$.

(b) Change the principal's strategy as follows: For any communication interval I_j that was affected by a translation (i.e., there was a condition $C_k \subset (\underline{\theta}_j, \overline{\theta}_j)$), after the translation have the principal send message m_j for types θ with $\underline{\theta}_j + (\overline{C}_k - \underline{C}_k) < \theta < \overline{\theta}_j$. For any communication interval I_j that was not affected by a translation have the principal continue to send message m_j.

(c) Change the agent's strategy as follows: Let the agent best respond to the new strategy of the principal and respond to all unsent messages the same way he responds to message m_1.

We make no claim that the new strategy profile $f^{\mathcal{C}_0}$ is an equilibrium profile of the communication game $\Gamma^{\mathcal{C}_0}$. The question of equilibrium is addressed after the final iteration. By Observation A.1, we have a strict payoff improvement for the principal over the payoff from $e^{\mathcal{C}}$ in $\Gamma^{\mathcal{C}}$ if players adopt the strategy profile $f^{\mathcal{C}_0}$ in the communication game $\Gamma^{\mathcal{C}_0}$.

After the replacement of the contract $\mathcal{C}$ by the contract $\mathcal{C}_0$ there is some number $L \leq \widehat{K}$ of condition clusters $\boldsymbol{C}_\ell$, $\ell = 1, \dots, L$. Denote the minimal (maximal) type in each condition cluster $\boldsymbol{C}_\ell$ by $\underline{\boldsymbol{C}}_\ell$ ($\overline{\boldsymbol{C}}_\ell$). Refer to the communication interval with lower bound $\overline{\boldsymbol{C}}_\ell$ by $I^+(\boldsymbol{C}_\ell, f^{\mathcal{C}_0})$ and let $y^+(\boldsymbol{C}_\ell, f^{\mathcal{C}_0})$ be the agent's best reply to beliefs concentrated on $I^+(\boldsymbol{C}_\ell, f^{\mathcal{C}_0})$. Similarly, let $I^-(\boldsymbol{C}_\ell, f^{\mathcal{C}_0})$ stand for the communication interval with upper bound $\underline{\boldsymbol{C}}_\ell$ and let $y^-(\boldsymbol{C}_\ell, f^{\mathcal{C}_0})$ be the agent's best reply to beliefs concentrated on $I^-(\boldsymbol{C}_\ell, f^{\mathcal{C}_0})$.

Note that type $\underline{\boldsymbol{C}}_\ell$ (weakly) prefers action $y^-(\boldsymbol{C}_\ell, f^{\mathcal{C}_0})$ to action $y^+(\boldsymbol{C}_\ell, f^{\mathcal{C}_0})$: $y^-(\boldsymbol{C}_\ell, f^{\mathcal{C}_0})$ is no further from $\underline{\boldsymbol{C}}_\ell$ than that type's preferred equilibrium action under the original equilibrium $e^{\mathcal{C}}$ and $y^+(\boldsymbol{C}_\ell, f^{\mathcal{C}_0})$ is no closer to $\underline{\boldsymbol{C}}_\ell$ than that type's preferred equilibrium action under $e^{\mathcal{C}}$.

Step 2. As noted before, the strategy profile $f^{\mathcal{C}_0}$ will generally violate incentive compatibility for the principal given the contract $\mathcal{C}_0$ and the agent's strategy. With the ultimate goal of reestablishing equilibrium, we begin by restoring incentive compatibility locally by replacing the strategy profile $f^{\mathcal{C}_0}$ by a new strategy profile $f^{\mathcal{C}_1}$ while leaving the prevailing contract unchanged, i.e., $\mathcal{C}_1 = \mathcal{C}_0$.

Between any two condition clusters $\boldsymbol{C}_\ell$ and $\boldsymbol{C}_{\ell+1}$ with $\ell < L$, and similarly between $\boldsymbol{C}_L$ and 1, restore equilibrium locally. In order to obtain a *local equilibrium* between $\boldsymbol{C}_\ell$ and $\boldsymbol{C}_{\ell+1}$, alter the principal's strategy in that range and the agent's responses to messages sent by types in that range, so that the agent best responds to those messages and the principal's types in that range have no incentive to mimic other types in that range. For now, ignore incentives to mimic types between other condition clusters. We address those incentives later. To this end, modify strategies as follows:

(a) If none of the critical types $\theta^{\mathcal{C}}$ from the equilibrium $e^{\mathcal{C}}$ satisfy $\overline{\boldsymbol{C}}_\ell < \theta^{\mathcal{C}} < \underline{\boldsymbol{C}}_{\ell+1}$, leave the principal's and the agent's strategies unchanged – they already satisfy the local-equilibrium condition. Otherwise, suppose that the critical types $\theta^{\mathcal{C}}$ satisfying $\overline{\boldsymbol{C}}_\ell < \theta^{\mathcal{C}} < \underline{\boldsymbol{C}}_{\ell+1}$ are $\theta^{\mathcal{C}}_i, \dots, \theta^{\mathcal{C}}_{i'}$. Note that given the postulated agent behavior in $f^{\mathcal{C}_0}$, type $\theta^{\mathcal{C}}_i$ is the only critical type in the range $\left(\overline{\boldsymbol{C}}_\ell, \underline{\boldsymbol{C}}_{\ell+1}\right)$ for which incentive compatibility is violated. Define $\lambda^{\mathcal{C}} := \theta^{\mathcal{C}}_i - \overline{\boldsymbol{C}}_\ell$.

(b) In order to restore equilibrium locally between $\boldsymbol{C}_\ell$ and $\boldsymbol{C}_{\ell+1}$, replace $\theta_i^{\mathcal{C}}, \ldots, \theta_{i'}^{\mathcal{C}}$ in the specification of the principal's strategy by $\theta_i, \ldots, \theta_{i'}$, where $\theta_i = \overline{\boldsymbol{C}}_\ell + \lambda$ and $\theta_{j+1} - \theta_j = \theta_{j+1}^{\mathcal{C}} - \theta_j^{\mathcal{C}} - \frac{\lambda - \lambda^{\mathcal{C}}}{i'+1-i}$, $j = i, \ldots, i'-1$, and $\lambda^{\mathcal{C}} \leq \lambda \leq \left(\theta_{i+1}^{\mathcal{C}} - \theta_i^{\mathcal{C}}\right)(i'+1-i) + \lambda^{\mathcal{C}}$. The last condition ensures that the length of the second step $\theta_{i+1} - \theta_i$ (and thus all subsequent steps) remains positive. For types in the range $\left(\overline{\boldsymbol{C}}_\ell, \underline{\boldsymbol{C}}_{\ell+1}\right)$, have the new strategy of the principal prescribe that the principal send message m_i in the interval $(\overline{\boldsymbol{C}}_\ell, \theta_i)$, message m_j in (θ_{j-1}, θ_j), $j = i+1, \ldots, i'$, and message $m_{i'+1}$ for types in $\left(\theta_{i'}, \underline{\boldsymbol{C}}_{\ell+1}\right)$. Otherwise, leave the principal's strategy unchanged. Adjust the agent's strategy so that the agent best responds to messages m_j, $j = i, \ldots, i'+1$, given the new strategy of the principal, leaving all other responses unchanged.

(c) For $\lambda = \lambda^{\mathcal{C}}$, type θ_i (weakly) prefers the action that is induced by types in the interval $(\overline{\boldsymbol{C}}_\ell, \theta_i)$ to the action that is induced by types in the interval (θ_i, θ_{i+1}). If θ_i is indifferent, we are done. Otherwise, it must be the case that the length of the interval (θ_i, θ_{i+1}) exceeds that of $(\overline{\boldsymbol{C}}_\ell, \theta_i)$. Consider increasing λ from $\lambda = \lambda^{\mathcal{C}}$ to the value λ'' at which the lengths of these two intervals become the same. At that point type θ_i strictly prefers the action that is induced by types in the interval (θ_i, θ_{i+1}) to the action that is induced by types in the interval $(\overline{\boldsymbol{C}}_\ell, \theta_i)$. Therefore, existence of a λ' with $\lambda'' \geq \lambda' \geq \theta_i - \overline{\boldsymbol{C}}_\ell$ that restores equilibrium locally between $\boldsymbol{C}_\ell$ and $\boldsymbol{C}_{\ell+1}$ follows from continuity the payoff function, the intermediate value theorem, and the fact that as we vary λ in the manner described, the arbitrage conditions for types θ_j, $j = i+1, \ldots, i'$ continue to be satisfied, since the lengths of adjacent intervals (θ_{j-1}, θ_j), $j = i+1, \ldots, i'$, and $\left(\theta_{i'}, \underline{\boldsymbol{C}}_{\ell+1}\right)$, continue to differ by $4b$.

The total change of behavior required to restore equilibrium locally between $\boldsymbol{C}_\ell$ and $\boldsymbol{C}_{\ell+1}$, as just described, can be decomposed into $i'+1-i$ steps. In the kth step λ is increased by $\frac{\lambda' - \lambda^{\mathcal{C}}}{i'+1-i}$, the intervals $(\theta_{i+(k'-1)}, \theta_{i+k'})$ with $1 \leq k' < k$ are all shifted up by that amount, and the interval $(\theta_{i+k-1}, \theta_{i+k})$ is reduced in size by the same amount by keeping θ_{i+k} fixed while θ_{i+k-1} increases. In the final step the interval whose size is reduced is $\left(\theta_{i'}, \underline{\boldsymbol{C}}_{\ell+1}\right)$. By Observation A.2 we have a payoff improvement at every step. Denote the strategy profile that results from restoring local equilibria in the game $\Gamma^{\mathcal{C}_1}$ between all pairs of adjacent condition clusters by $f^{\mathcal{C}_1}$.

Step 3. We next turn to addressing incentive constraints that involve types that are separated by condition clusters.

Observe that when we replace $f^{\mathcal{C}_0}$ by $f^{\mathcal{C}_1}$ in $\Gamma^{\mathcal{C}_1}$, for any condition cluster $\boldsymbol{C}_\ell$, we have $\left|I^+(\boldsymbol{C}_\ell, f^{\mathcal{C}_1})\right| \geq \left|I^+(\boldsymbol{C}_\ell, f^{\mathcal{C}_0})\right|$ and $\left|I^-(\boldsymbol{C}_\ell, f^{\mathcal{C}_1})\right| \leq \left|I^-(\boldsymbol{C}_\ell, f^{\mathcal{C}_0})\right|$. In combination with type $\underline{\boldsymbol{C}}_\ell$ having preferred action $y^-(\boldsymbol{C}_\ell, f^{\mathcal{C}_0})$ to action $y^+(\boldsymbol{C}_\ell, f^{\mathcal{C}_0})$ prior to the strategy-profile replacement, this implies that none of the types equal to or less than $\underline{\boldsymbol{C}}_\ell$, have an incentive to induce any action greater than $y^-(\boldsymbol{C}_\ell, f^{\mathcal{C}_1})$ available to them given the profile $f^{\mathcal{C}_1}$. Therefore, if none of the types $\overline{\boldsymbol{C}}_\ell$, $\ell = 1, \ldots, L$ have an incentive to induce an action less than $y^+(\boldsymbol{C}_\ell, f^{\mathcal{C}_1})$ available to them given the profile $f^{\mathcal{C}_1}$, the combination of local equilibria forms an equilibrium overall.

If instead there is a type $\overline{\boldsymbol{C}}_\ell$ who prefers inducing an action less than $y^+(\boldsymbol{C}_\ell, f^{\mathcal{C}_1})$ that is available given the profile $f^{\mathcal{C}_1}$, let $\hat{\ell}$ be the maximal ℓ such that this is the case. Consider the set of actions that are induced by types $\theta > \overline{\boldsymbol{C}}_{\hat{\ell}}$. Refer to the types who are indifferent among adjacent actions in this set of actions as $\hat{\ell}$-critical types. Use $\tilde{\ell}$ to denote the minimal $\ell > \hat{\ell}$ such that there is an $\hat{\ell}$-critical type $\tilde{\theta} \in \left[\underline{\boldsymbol{C}}_\ell, \overline{\boldsymbol{C}}_\ell\right)$, if there is such a type. If there is no $\hat{\ell}$-critical type $\tilde{\theta} \in \left[\underline{\boldsymbol{C}}_\ell, \overline{\boldsymbol{C}}_\ell\right)$ for all $\ell > \hat{\ell}$, proceed without introducing $\tilde{\ell}$. Note that in this case we have that either $\overline{\boldsymbol{C}}_\ell$ is an $\hat{\ell}$-critical type for all $\ell > \hat{\ell}$ or $\boldsymbol{C}_{\hat{\ell}}$ is the rightmost condition cluster ($\hat{\ell} = L$).

Note that if θ_{j-1}, θ_j and θ_{j+1} are $\hat{\ell}$-critical types such that $\theta_j = \overline{\boldsymbol{C}}_\ell$ and neither θ_{j-1} nor θ_{j+1} belong to a condition cluster, then we have

$$\theta_j + b - \frac{\theta_{j-1} + (\theta_j - (\overline{\boldsymbol{C}}_\ell - \underline{\boldsymbol{C}}_\ell))}{2} = \frac{\theta_{j+1} + \theta_j}{2} - \theta_j - b,$$

which is equivalent to

$$\theta_{j+1} - \theta_j = \theta_j - \theta_{j-1} + 4b + (\overline{\boldsymbol{C}}_\ell - \underline{\boldsymbol{C}}_\ell). \tag{1}$$

This is the standard arbitrage condition in the CS uniform quadratic example extended to the case where the $\hat{\ell}$-critical type θ_j is the upper endpoint of a condition cluster. If θ_{j-1} belongs to the condition cluster $\boldsymbol{C}_{\ell-1}$, replace θ_{j-1} by $\overline{\boldsymbol{C}}_{\ell-1}$ in the above expression, and if θ_{j+1} belongs to the condition cluster $\boldsymbol{C}_{\ell+1}$, replace θ_{j+1} by $\underline{\boldsymbol{C}}_{\ell+1}$.

Consider replacing the condition cluster $\boldsymbol{C}_{\hat{\ell}}$ by its λ-translation (for notational convenience also denoted by $\boldsymbol{C}_{\hat{\ell}}$) for values $\lambda > 0$ that make it possible to

(a) maintain local equilibrium for types in the range $(\overline{\boldsymbol{C}}_{\hat{\ell}-1}, \underline{\boldsymbol{C}}_{\hat{\ell}})$ (if $\hat{\ell} > 1$, and in the range $(0, \underline{\boldsymbol{C}}_{\hat{\ell}})$ otherwise) (this is achieved by choosing λ sufficiently small and increasing the length of each communication interval in this range by λ divided by the number of communication intervals in this range), and

(b) maintain local equilibrium in the range $(\overline{\boldsymbol{C}}_{\hat{\ell}}, \underline{\boldsymbol{C}}_{\tilde{\ell}})$ and preserve indifference for all types θ such that $\theta = \overline{\boldsymbol{C}}_\ell$ with $\hat{\ell} < \ell < \tilde{\ell}$ (by condition (1), this is achieved by choosing λ sufficiently small and reducing the sizes of communication intervals in the range $(\overline{\boldsymbol{C}}_{\hat{\ell}}, \underline{\boldsymbol{C}}_{\tilde{\ell}})$ all by λ divided by the number of communication intervals in this range).

For each λ, denote the strategy that maintains local equilibrium for types $\theta > \overline{\boldsymbol{C}}_{\hat{\ell}-1}$ by f^λ.

Note that if, prior to the λ-translation of $\boldsymbol{C}_{\hat{\ell}}$, type $\overline{\boldsymbol{C}}_{\hat{\ell}}$ prefers inducing an action less than $y^+(\boldsymbol{C}_{\hat{\ell}}, f^{\mathcal{C}_1})$ that is available given the profile $f^{\mathcal{C}_1}$, as postulated, it has to be the case that $\left|I^+(\boldsymbol{C}_{\hat{\ell}}, f^{\mathcal{C}_1})\right| > \left|I^-(\boldsymbol{C}_{\hat{\ell}}, f^{\mathcal{C}_1})\right|$. As a consequence of replacing $\boldsymbol{C}_{\hat{\ell}}$ by its λ-translation and maintaining local equilibria in the ranges specified above, the lengths of communication intervals in the range $(\overline{\boldsymbol{C}}_{\hat{\ell}-1}, \underline{\boldsymbol{C}}_{\hat{\ell}})$ increase and the lengths of communication intervals in the range $(\overline{\boldsymbol{C}}_{\hat{\ell}}, \underline{\boldsymbol{C}}_{\tilde{\ell}})$ decrease. It is easily checked that for all λ between $\lambda = 0$ and the value of λ that equalizes $\left|I^+(\boldsymbol{C}_\ell, f^\lambda)\right|$ and $\left|I^-(\boldsymbol{C}_\ell, f^\lambda)\right|$ the local equilibria in the ranges $(\overline{\boldsymbol{C}}_{\hat{\ell}-1}, \underline{\boldsymbol{C}}_{\hat{\ell}})$

and $(\overline{\boldsymbol{C}}_{\hat{\ell}}, \underline{\boldsymbol{C}}_{\tilde{\ell}})$ can be preserved, as described above. Hence by payoff continuity and the intermediate value theorem, there exists a value of λ for which we have an equilibrium in the auxiliary game that is obtained by restricting the type space to $(\overline{\boldsymbol{C}}_{\hat{\ell}-1}, \underline{\boldsymbol{C}}_{\tilde{\ell}})$, leaving all condition clusters $\boldsymbol{C}_\ell$ with $\ell \neq \hat{\ell}$ unchanged, and replacing $\boldsymbol{C}_{\hat{\ell}}$ by its λ-translation. Denote this value of λ by λ'. Monotonicity of type $\overline{\boldsymbol{C}}_{\hat{\ell}}$'s payoff differential from actions $y^+(\boldsymbol{C}_{\hat{\ell}}, f^\lambda)$ and $y^-(\boldsymbol{C}_{\hat{\ell}}, f^\lambda)$ implies that λ' is unique. By a similar argument there exists a unique value of λ such that local equilibria in the ranges $(\overline{\boldsymbol{C}}_{\hat{\ell}-1}, \underline{\boldsymbol{C}}_{\hat{\ell}})$ and $(\overline{\boldsymbol{C}}_{\hat{\ell}}, \underline{\boldsymbol{C}}_{\tilde{\ell}})$ are preserved as above and, in addition, we have $\tilde{\theta} = \overline{\boldsymbol{C}}_{\tilde{\ell}}$. Denote this value of λ by λ''.

Define $\lambda_{\min} := \min\{\lambda', \lambda''\}$ and note that with the $\lambda_{\min}$-translation of $\boldsymbol{C}_{\hat{\ell}}$ we have $\tilde{\theta} \in [\underline{\boldsymbol{C}}_\ell, \overline{\boldsymbol{C}}_\ell]$. Let n_1 be the number of communication intervals in the range $(\overline{\boldsymbol{C}}_{\hat{\ell}-1}, \underline{\boldsymbol{C}}_{\hat{\ell}})$ and n_2 the number of communication intervals in the range $(\overline{\boldsymbol{C}}_{\hat{\ell}}, \underline{\boldsymbol{C}}_{\tilde{\ell}})$. If we replace $\boldsymbol{C}_{\hat{\ell}}$ by its $\lambda_{\min}$-translation while preserving local equilibria in the ranges $(\overline{\boldsymbol{C}}_{\hat{\ell}-1}, \underline{\boldsymbol{C}}_{\hat{\ell}})$ and $(\overline{\boldsymbol{C}}_{\hat{\ell}}, \underline{\boldsymbol{C}}_{\tilde{\ell}})$ as indicated above, this increases the length of each communication interval I_j, in the range $(\overline{\boldsymbol{C}}_{\hat{\ell}-1}, \underline{\boldsymbol{C}}_{\hat{\ell}})$ by $\frac{\lambda_{\min}}{n_1}$ and lowers the length of each communication interval I_j in the range $(\overline{\boldsymbol{C}}_{\hat{\ell}}, \underline{\boldsymbol{C}}_{\tilde{\ell}})$ by $\frac{\lambda_{\min}}{n_2}$.

We can decompose the replacement of $\boldsymbol{C}_{\hat{\ell}}$ by its $\lambda_{\min}$-translation and the corresponding preservation of local equilibria in the ranges $(\overline{\boldsymbol{C}}_{\hat{\ell}-1}, \underline{\boldsymbol{C}}_{\hat{\ell}})$ and $(\overline{\boldsymbol{C}}_{\hat{\ell}}, \underline{\boldsymbol{C}}_{\tilde{\ell}})$ into $n_1 \cdot n_2$ steps of size $\frac{\lambda_{\min}}{n_1 \cdot n_2}$. Define $I_j(0) := I_j$. At the rth step, $r = 1, \ldots, n_1 \cdot n_2$,

(1) identify two intervals $I_{j'}(r) \subseteq (\overline{\boldsymbol{C}}_{\hat{\ell}-1}, \underline{\boldsymbol{C}}_{\hat{\ell}})$ and $I_{j''}(r) \subseteq (\overline{\boldsymbol{C}}_{\hat{\ell}}, \underline{\boldsymbol{C}}_{\tilde{\ell}})$ among those that have been established by step $r-1$ and which satisfy $|I_{j'}(r)| < |I_{j'} + \frac{\lambda_{\min}}{n_1}|$ and $|I_{j''}(r)| > |I_{j''} - \frac{\lambda_{\min}}{n_2}|$,

(2) increase the length of the former by $\frac{\lambda_{\min}}{n_1 \cdot n_2}$ by changing its right endpoint,

(3) reduce the length of the latter by the same amount by changing its left endpoint,

(4) replace all intervals $I_j(r) \subseteq (\overline{\boldsymbol{C}}_{\hat{\ell}-1}, \underline{\boldsymbol{C}}_{\hat{\ell}})$ with $j > j'$ by their $\frac{\lambda_{\min}}{n_1 \cdot n_2}$-translation,

(5) replace all intervals $I_j(r) \subseteq (\overline{\boldsymbol{C}}_{\hat{\ell}}, \underline{\boldsymbol{C}}_{\tilde{\ell}})$ with $j < j''$ by their $\frac{\lambda_{\min}}{n_1 \cdot n_2}$-translation,

(6) replace the $\boldsymbol{C}_{\hat{\ell}}$ that resulted from step $r-1$ by its $\frac{\lambda_{\min}}{n_1 \cdot n_2}$-translation,

(7) have the principal send the same message in $I_j(r)$ that she sent in $I_j(r-1)$ for all j,

(8) have the agent best respond to the new strategy of the principal.

By Observation A.2 we have a strict payoff improvement at every step.

Denote the contract that results from replacing $\boldsymbol{C}_{\hat{\ell}}$ by its $\lambda_{\min}$-translation by $\mathcal{C}_2$. Denote the strategy profile that results from preserving local equilibria in the ranges $(\overline{\boldsymbol{C}}_{\hat{\ell}-1}, \underline{\boldsymbol{C}}_{\hat{\ell}})$ and $(\overline{\boldsymbol{C}}_{\hat{\ell}}, \underline{\boldsymbol{C}}_{\tilde{\ell}})$ as described above while otherwise being identical with $f^{\mathcal{C}_1}$ by $f^{\mathcal{C}_2}$.

If $\lambda_{\min} = \lambda'$, identify the maximal ℓ such that type $\overline{\boldsymbol{C}}_\ell$ prefers inducing an action less than $y^+(\boldsymbol{C}_\ell, f^{\mathcal{C}_2})$ that is available given the profile $f^{\mathcal{C}_2}$, if there is such an ℓ. Otherwise we are done. Note that this ℓ necessarily satisfies $\ell < \hat{\ell}$. Make this ℓ the new $\hat{\ell}$ and repeat the construction that, starting with $\mathcal{C}_1$ and the strategy profile $f^{\mathcal{C}_1}$, gave us $\mathcal{C}_2$ and $f^{\mathcal{C}_2}$.

If instead $\lambda_{\min} = \lambda''$, identify the minimal $\ell > \hat{\ell}$ such that there is a critical type $\tilde{\theta}$ in the set $[\underline{\boldsymbol{C}}_\ell, \overline{\boldsymbol{C}}_\ell)$ (note that this ℓ, if it exists, is necessarily larger than $\tilde{\ell}$). If there is no such ℓ we are done. Make this ℓ the new $\tilde{\ell}$ and repeat the construction that, starting with $\mathcal{C}_1$ and the strategy profile $f^{\mathcal{C}_1}$, gave us $\mathcal{C}_2$ and $f^{\mathcal{C}_2}$.

Starting with any $\mathcal{C}_i$ and $f^{\mathcal{C}_i}$ obtained in this manner construct $\mathcal{C}_{i+1}$ and $f^{\mathcal{C}_{i+1}}$ using the same procedure. Since there are finitely many indices ℓ and at each step either $\hat{\ell}$ drops or $\tilde{\ell}$ rises, this process terminates and at that point we have an equilibrium with a strict payoff improvement. This establishes that optimal equilibria are partitional.

Part II – monotonic. To prove monotonicity of optimal equilibria, we need to show that for any two distinct partition elements $T, T' \in \mathcal{T}$ satisfying $\inf(T') \geq \sup(T)$, we have $a(T') > a(T)$. Note that it suffices to show this for T and T' that are adjacent to each other. If T' is a condition of $\mathcal{C}$, or both T and T' are communication intervals, the result is an immediate consequence of our assumptions on the payoff functions U^i, $i = P, A$.

Suppose, therefore, that T' is a communication interval and T a condition of $\mathcal{C}$ with $\inf(T') \geq \sup(T)$. For each $T \in \mathcal{T}$ use $x(T)$ to denote the principal's preferred action given T and $y(T)$ to denote the agent's preferred action given T. In order to derive a contraction, suppose that we have that $a(T') \leq a(T)$ – i.e., $y(T') \leq x(T)$. The cross-partial condition implies that $x(T') > x(T)$. $\int_{T'} U^P(x, \theta) d\theta$ is an integral over strictly concave functions and, therefore, itself strictly concave. This implies that $\int_{T'} U^P(x, \theta) d\theta$ is strictly increasing for all $x < x(T')$. Therefore, since $y(T') \leq x(T) < x(T')$, and

$$\frac{d}{dx} \left(\int_T U^P(x, \theta) d\theta \right) \bigg|_{x = x(T)} = 0,$$

we obtain that for sufficiently small $\varepsilon > 0$,

$$\int_T U^P(x(T), \theta) d\theta + \int_{T'} U^P(y(T'), \theta) d\theta < \int_{T \cup T'} U^P(x(T) + \varepsilon, \theta) d\theta.$$

This, however, implies that there exists $\varepsilon > 0$ such that the principal would strictly prefer to have a single condition $\overline{T \cup T'}$ with instruction $x(T) + \varepsilon$ (of course, a further improvement could be achieved by replacing the instruction $x(T) + \varepsilon$ by the instruction $x\left(\overline{T \cup T'}\right)$ if they differ). This gives us the desired contradiction, thus establishing monotonicity of optimal equilibria.

Part III – interior condition cluster. We want to show that if the equilibrium $e^{\mathcal{C}}$ induces at least two communication actions, then there is a condition cluster $\boldsymbol{C}$ and a critical type $\theta \neq 0, 1$ with $\theta \in \boldsymbol{C}$.

In order to reach a contradiction, suppose that the equilibrium $e^{\mathcal{C}}$ induces at least two communication actions, and that for all critical types $\tilde{\theta} \neq 0, 1$ and all condition clusters $\boldsymbol{C}$, it is the case that $\tilde{\theta} \notin \boldsymbol{C}$. Let $n > 1$ be the number of communication intervals in $e^{\mathcal{C}}$. Then, from Part I, any condition cluster $\boldsymbol{C}$ satisfies either $0 \in \boldsymbol{C}$ or $1 \in \boldsymbol{C}$, and there is a critical type $\theta_1 \in (0, 1)$.

Consider the case where $0 \in \boldsymbol{C}$ for a condition cluster $\boldsymbol{C}$. Let the contract $\mathcal{C}'$ only differ from $\mathcal{C}$ by replacing the condition cluster $\boldsymbol{C}$ by its $(\theta_1 - \overline{\boldsymbol{C}})$-translation, $\boldsymbol{C}'$. Evidently, the game $\Gamma^{\mathcal{C}'}$ has an equilibrium $e^{\mathcal{C}'}$ in which types $\theta \in (0, \theta_1 - \overline{\boldsymbol{C}})$ send the message sent by types in $(\overline{\boldsymbol{C}}, \theta_1)$ in equilibrium $e^{\mathcal{C}}$, and all other types behave as they did before in equilibrium

$e^{\mathcal{C}}$. The principal's expected payoff in the equilibrium $e^{\mathcal{C}'}$ is the same as in $e^{\mathcal{C}}$, type $\theta_1 - \overline{\boldsymbol{C}}$ strictly prefers the action that is induced by types in $(0, \theta_1 - \overline{\boldsymbol{C}})$ to all other equilibrium actions and type θ_1 strictly prefers the action that is induced by types in the communication interval that is bounded below by θ_1 to all other equilibrium actions.

Since the incentive constraints of types $\underline{\boldsymbol{C}}' = \theta_1 - \overline{\boldsymbol{C}}$ and $\overline{\boldsymbol{C}}' = \theta_1$ in the new equilibrium $e^{\mathcal{C}'}$ are slack, for sufficiently small $\lambda > 0$ we can replace the contract $\mathcal{C}'$ by a contract $\mathcal{C}^{\lambda}$ that only differs from $\mathcal{C}'$ by replacing the condition cluster $\boldsymbol{C}'$ by its λ-translation, $\boldsymbol{C}^{\lambda}$, so that the game $\Gamma^{\mathcal{C}^{\lambda}}$ has an equilibrium $e^{\mathcal{C}^{\lambda}}$, in which, relative to $e^{\mathcal{C}'}$, the length of the first communication interval increases by λ and the lengths of all the remaining communication intervals are reduced by $\frac{\lambda}{n-1}$. Combining this with the fact that in $e^{\mathcal{C}}$, and therefore in $e^{\mathcal{C}'}$, the first is the smallest communication interval, repeated application of Observation A.1 implies that for any sufficiently small $\lambda > 0$ the principal's expected payoff from $e^{\mathcal{C}^{\lambda}}$ strictly exceeds that from $e^{\mathcal{C}}$. It follows that $e^{\mathcal{C}}$ cannot have been optimal.

For the case in which $1 \in \boldsymbol{C}$ for a condition cluster $\boldsymbol{C}$, consider the contract $\mathcal{C}''$ that only differs from $\mathcal{C}$ by replacing the condition cluster $\boldsymbol{C}$ by its $-(\underline{\boldsymbol{C}} - \theta_{n-1})$-translation, $\boldsymbol{C}''$. In this case, the game $\Gamma^{\mathcal{C}''}$ has an equilibrium $e^{\mathcal{C}''}$ in which types $\theta \in (1 - (\underline{\boldsymbol{C}} - \theta_{n-1}), 1)$ send the message sent by types in $(\theta_{n-1}, \underline{\boldsymbol{C}})$ in equilibrium $e^{\mathcal{C}}$ and all other types behave as they did before in equilibrium $e^{\mathcal{C}}$. Similar to the previous case, the incentive constraints of types $\underline{\boldsymbol{C}}''$ and $\overline{\boldsymbol{C}}''$ are slack, $\left(\overline{\boldsymbol{C}}'', 1\right] = (1 - (\underline{\boldsymbol{C}} - \theta_{n-1}), 1]$ is the largest communication interval, and therefore for sufficiently small $\lambda > 0$ one can increase equilibrium payoffs by replacing $\boldsymbol{C}''$ by its λ-translation. $\square$

Proof of Proposition 3. Suppose that $\widehat{K} > 3$ (otherwise the proposition holds vacuously). Consider the problem of maximizing the principal's payoff subject to the constraints that each cluster has at least four conditions and that the principal's strategy restricted to any communication area below or above a cluster is an equilibrium strategy for the game restricted to that communication area. Call this the "relaxed problem." The payoff from a solution to the relaxed problem is at least as high as the payoff from a contract that is optimal in the class of contracts in which each cluster has at least four conditions, since the latter has to respect additional incentive constraints. The additional constraints are that types in any communication area do not prefer to mimic types in some other communication area. We will show that one can strictly improve on the solution to the relaxed problem by breaking up one of the condition clusters and that this improvement respects global incentive constraints. This implies that one can strictly improve on the optimal contract in the class of contracts in which each cluster has at least four conditions.

If $b > 0$ converges to zero it is possible to approximate the principal's first-best payoff in the relaxed problem. To see this, note that without using any of the conditions, the optimal communication equilibrium converges to the first best as $b > 0$ converges to zero. Taking advantage of the conditions while respecting the constraints of the relaxed problem cannot lead to a lower payoff (we always have the option to simply replace communication intervals by conditions, in which case the requirement that there are always at least four conditions in a cluster is easily met). If we placed an upper bound on the number of communication steps in the relaxed problem, then as b converges to zero, the principal's payoff in the relaxed problem would remain bounded away from the first best. Therefore, in order to approximate the first best as $b > 0$ converges to zero, the number of communication intervals has to grow without bound. This implies that for small $b > 0$ the solution to the relaxed problem has the property that there is at least one cluster adjacent to a communication area with at least two communication intervals.

For any condition cluster with a communication interval I^+ above the cluster and a communication interval I^- below the cluster, the lengths of these intervals satisfy $|I^-| \geq |I^+|$. This follows from repeated application of Observation A.2. Otherwise we would have all (say m) communication intervals in the communication area below the cluster be shorter than all (say n) communication intervals in the communication area above the cluster. Then we could translate the cluster to the right by some small ϵ, increase the size of each of the communication intervals in the communication area below the cluster by $\frac{\epsilon}{m}$, reduce the size of each of the communication intervals in the communication area above the cluster by $\frac{\epsilon}{n}$, and thereby increase expected payoffs.

Fix a condition cluster $\widetilde{\boldsymbol{C}}$ from the solution to the relaxed problem that is adjacent to a communication area with at least two communication intervals. Use $\widetilde{K}$ to denote the number of conditions in the cluster $\widetilde{\boldsymbol{C}}$. Use $\underset{\sim}{\theta}$ to denote the infimum of the communication area below $\widetilde{\boldsymbol{C}}$ and $\tilde{\theta}$ to denote the supremum of the communication area above $\widetilde{\boldsymbol{C}}$. The solution of the relaxed problem induces a solution of the following "restricted relaxed problem" with the same cluster $\widetilde{\boldsymbol{C}}$ as part of the solution: maximize the principal's payoff with the type distribution restricted to the interval $[\underset{\sim}{\theta}, \tilde{\theta}]$, subject to the constraints that there is a single cluster with $\widetilde{K}$ conditions and that the principal's strategy restricted to any interval in the communication area below or above the cluster is an equilibrium strategy.

Let ι denote the length of the interval $[\underset{\sim}{\theta}, \tilde{\theta}]$. Up to rescaling (i.e., replacing the interval $[\underset{\sim}{\theta}, \tilde{\theta}]$ by $[0, 1]$, replacing the bias b by $\frac{b}{\iota}$ and identifying every type θ in $[\underset{\sim}{\theta}, \tilde{\theta}]$ with a type $\frac{\theta - \underset{\sim}{\theta}}{\iota}$ in $[0, 1]$) the restricted relaxed problem is equivalent to the problem of maximizing the principal's payoff over the original type space $[0, 1]$ subject to the constraints that there is a single cluster with $\widetilde{K}$ conditions and that the principal's strategy restricted to any communication area below or above the cluster is an equilibrium strategy. That is, the rescaled problem has a solution in which each communication interval and each condition corresponds to a communication interval or condition of the original problem up to rescaling by the factor $\frac{1}{\iota}$.

To economize on notation, it is convenient to work with the rescaled version of the restricted relaxed problem. We continue to use θ to denote types and b to denote the bias, keeping in mind that they have been rescaled. Continuing with our slight abuse of notation, we also use $\widetilde{\boldsymbol{C}}$ to denote the cluster that is part of the solution of the the rescaled restricted relaxed problem. We use $\underline{C}_1$ to indicate the lower endpoint of the first condition and $\overline{C}_{\widetilde{K}}$ to indicate the upper endpoint of the $\widetilde{K}$'s condition in the cluster $\widetilde{\boldsymbol{C}}$.

All conditions in the cluster $\widetilde{\boldsymbol{C}}$ will be of equal length. This follows from repeated application of Observation A.2. Denote that length by ℓ.

Suppose the solution to the relaxed problem has m communication intervals below the cluster and n communication intervals above the cluster. In that case, there is a corresponding solution to the restricted relaxed problem that has m communication intervals below the cluster and n communication intervals above the cluster. Either $m > 1$, in which case $n \geq 1$, or $n > 1$, in which case $m \geq 1$. Hence there is a maximal communication interval below the cluster and a least communication interval above the cluster. Denote the maximal communication interval below the cluster by I^- $(= (\theta_{m-1}, \underline{C}_1))$ and the minimal communication interval above the cluster by $I^+ \left(= \left(\overline{C}_{\widetilde{K}}, \theta_{m+1}\right)\right)$. As noted before, the lengths of these communication intervals satisfy $|I^-| \geq |I^+|$.

To simplify notation, define $\phi := \underline{C}_1$ and $L := \phi + \ell$. Fix L. It has to be the case that ϕ solves the problem of optimally dividing the interval $[0, L]$ into a communication area of size ϕ with m communication steps and one condition of length ℓ. That is ϕ solves the problem

$$\max_{\phi,\theta_0,\theta_1,\ldots,\theta_{m-1}} \sum_{i=1}^{m-1} -\int_{\theta_{i-1}}^{\theta_i} \left(s + b - \frac{\theta_{i-1}+\theta_i}{2}\right)^2 ds - \int_{\theta_{m-1}}^{\phi} \left(s + b - \frac{\theta_{m-1}+\phi}{2}\right)^2 ds - \int_{\phi}^{L} \left(s - \frac{\phi+L}{2}\right)^2 ds$$

$$\begin{aligned}
\text{s.t.} \quad & \theta_i + b - \frac{\theta_{i-1}+\theta_i}{2} = \frac{\theta_{i+1}+\theta_i}{2} - \theta_i - b \quad \text{for } i = 1, \ldots, m-2, \\
& \theta_{m-1} + b - \frac{\theta_{m-2}+\theta_{m-1}}{2} = \frac{\phi+\theta_{m-1}}{2} - \theta_{m-1} - b, \quad \text{and} \\
& \theta_0 = 0.
\end{aligned}$$

The FOC for a solution to this problem is

$$-\frac{1}{3}b^2(m^2+2) + \frac{1}{4}\left(L^2 - 2L\phi + \left(1 - \frac{1}{m^2}\right)\phi^2\right) = 0.$$

One checks easily that the SOC is satisfied. Using the fact that $L = \phi + \ell$ and rearranging, we find that

$$9(lm^2 + \phi)^2 = m^2\left(9(l+\phi)^2 + 12b^2(m^4 + m^2 - 2)\right),$$

which is equivalent to

$$\phi = \sqrt{\ell^2 m^2 - \frac{4}{3}b^2(m^2+2)m^2}.$$

First, suppose that $m > 1$. Split the cluster $\widetilde{\boldsymbol{C}}$ by switching the mth communication interval with the lowest condition in the cluster. For the modified strategy of the principal to be an equilibrium strategy given the rearrangement of conditions on the state space of the (rescaled) restricted relaxed problem, it suffices that types in the moved communication interval do not have an incentive to mimic types above the maximal condition and vice versa. The latter requirement is satisfied since $|I^-| \geq |I^+|$. Let $\gamma := |I^-|$. Then the former requirement is satisfied if

$$\phi + \ell + b - \left(\frac{\phi + \ell + (\phi + \ell - \gamma)}{2}\right) < \phi + \ell + (\widetilde{K} - 1)\ell - (\phi + \ell + b). \tag{2}$$

This inequality is equivalent to

$$\gamma < 2(\widetilde{K} - 1)\ell - 4b.$$

Using the standard communication game calculations for the lengths of communication intervals, we obtain that $\phi = m\theta_1 + 2m(m-1)b$ and $\gamma = \theta_1 + (m-1)4b$. Therefore our sufficient condition for incentive compatibility is equivalent to

$$\frac{\phi}{m} + 2(m-1)b < 2(\widetilde{K} - 1)\ell - 4b.$$

Using the fact that $\phi < \ell m$, this is implied by

$$\ell m < 2m((\widetilde{K} - 1)\ell - b(m+1)).$$

Using the fact that $2m(m-1)b < \phi < \ell m$, this inequality holds if

$$\ell < 2\left((\widetilde{K} - 1)\ell - \frac{\ell(m+1)}{2(m-1)}\right),$$

which is equivalent to

$$\widetilde{K} > \frac{1 + \frac{m+1}{m-1}}{2} + 1.$$

The right-hand side of this inequality is no larger than 3. Therefore the sufficient condition for incentive compatibility (2) holds as long as $\widetilde{K} > 3$.

Second, suppose instead that $m = 1$ (and therefore $n > 1$). Then

$$\phi = \sqrt{\ell^2 - 4b^2}. \tag{3}$$

In this case split the cluster $\widetilde{\boldsymbol{C}}$ by switching the second communication interval (i.e., the first above the cluster) with the highest condition in the cluster. For the modified strategy of the principal to be an equilibrium strategy given the rearrangement of conditions on the state space of the (rescaled) restricted relaxed problem, it suffices that types below ϕ do not have an incentive to mimic types in the moved communication interval and vice versa. The latter requirement is satisfied since $|I^-| \geq |I^+|$. Note that $|I^-| = \phi$. The former requirement is satisfied if

$$\phi + b - \frac{\phi}{2} < \phi + (\widetilde{K} - 1)\ell - (\phi + b). \tag{4}$$

Since $\phi < \ell$ and $b < \frac{\ell}{2}$ (both from (3)), this holds as long as $\widetilde{K} \geq 3$.

Returning to the original relaxed problem, to ensure that switching a condition in the cluster $\widetilde{\boldsymbol{C}}$ with a communication interval, as described above, results in an equilibrium on the entire state space, we need to verify in addition that for every other cluster $\boldsymbol{C}'$ with K' conditions types immediately below that cluster have no incentive to mimic types immediately above that cluster, and vice versa. The latter condition, as before, follows from $|I^-| \geq |I^+|$. In case the number of communication intervals m below $\boldsymbol{C}'$ satisfies $m > 1$ the former condition is satisfied if inequality (2) holds with K' replacing $\widetilde{K} - 1$, which is the case if $K' > 2$. If instead $m = 1$, we need inequality (4) to hold, again with K' replacing $\widetilde{K} - 1$, which is the case if $K' \geq 2$. Note that in this case we do not (need to) require that $n > 1$.

Denote the condition that was translated by $\widetilde{C}$. Notice that all the incentive constraints we checked are slack. In the case in which $m > 1$, this implies that we can reduce the length of the (new) communication interval above $\widetilde{C}$ by some small $\epsilon > 0$, increase the lengths of each of the $m-1$ communication intervals below $\widetilde{C}$ by $\frac{\epsilon}{m-1}$ and still have an equilibrium. By Observation A.2 this equilibrium has a strictly higher payoff than the solution of the relaxed problem, and therefore a strictly higher payoff than from any contract in which every cluster has more than three conditions. An analogous argument applies to the case in which $m = 1$.

$\square$

References

Agastya, M., Bag, P. K. and Chakraborty, I. (2015). Proximate preferences and almost full revelation in the Crawford-Sobel game, *Economic Theory Bulletin* **3**(2): 201–212.

Aghion, P. and Tirole, J. (1997). Formal and real authority in organizations, *Journal of Political Economy* **105**(1): 1–29.

Allen, F. and Gale, D. (1992). Measurement distortion and missing contingencies in optimal contracts, *Economic Theory* **2**(1): 1–26.

Alonso, R. and Matouschek, N. (2008). Optimal delegation, *Review of Economic Studies* **75**(1): 259–293.

Amador, M. and Bagwell, K. (2013). The theory of optimal delegation with an application to tariff caps, *Econometrica* **81**(4): 1541–1599.

Aumann, R. J. and Hart, S. (2003). Long cheap talk, *Econometrica* **71**(6): 1619–1660.

Ayres, I. and Gertner, R. (1992). Strategic contractual inefficiency and the optimal choice of legal rules, *Yale Law Journal* **101**(4): 729–773.

Bajari, P. and Tadelis, S. (2001). Incentives versus transaction costs: A theory of procurement contracts, *RAND Journal of Economics* **32**(3): 387–407.

Banerjee, A. V. and Duflo, E. (2000). Reputation effects and the limits of contracting: A study of the Indian software industry, *Quarterly Journal of Economics* **115**(3): 989–1017.

Battigalli, P. and Maggi, G. (2002). Rigidity, discretion, and the costs of writing contracts, *American Economic Review* **92**(4): 798–817.

Bernheim, B. D. and Whinston, M. D. (1998). Incomplete contracts and strategic ambiguity, *American Economic Review* **88**(4): 902–932.

Chakravarty, S. and MacLeod, W. B. (2006). Construction contracts (or "How to get the right building at the right price?"). CESifo Discussion Paper.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission, *Econometrica* **50**(6): 1431–1451.

Crocker, K. J. and Reynolds, K. J. (1993). The efficiency of incomplete contracts: An empirical analysis of air force engine procurement, *RAND Journal of Economics* **24**(1): 126–146.

Deimen, I. and Szalay, D. (2019). Delegated expertise, authority, and communication, *American Economic Review* **109**(4): 1349–74.

Dessein, W. (2002). Authority and communication in organizations, *Review of Economic Studies* **69**(4): 811–838.

Dilmé, F. (2022). Strategic communication with a small conflict of interest, *Games and Economic Behavior* **134**: 1–19.

Dye, R. A. (1985). Costly contract contingencies, *International Economic Review* **26**(1): 233–250.

Eckfeldt, B., Madden, R. and Horowitz, J. (2005). Selling agile: target-cost contracts, *Agile Development Conference (ADC'05)*, IEEE, pp. 160–166.

Golosov, M., Skreta, V., Tsyvinski, A. and Wilson, A. (2014). Dynamic strategic information transmission, *Journal of Economic Theory* **151**: 304–341.

Hart, O. and Moore, J. (1988). Incomplete contracts and renegotiation, *Econometrica* **56**(4): 755–785.

Heller, D. and Spiegler, R. (2008). Contradiction as a form of contractual incompleteness, *Economic Journal* **118**(530): 875–888.

Holmström, B. (1977). *On Incentives and Control in Organizations*, Ph.D. Thesis, Stanford University.

Holmström, B. (1984). *On the Theory of Delegation*, in M. Boyer and R. Kihlstrom (eds.) Bayesian Models in Economic Theory, New York: North-Holland.

Kolotilin, A., Li, H. and Li, W. (2013). Optimal limited authority for principal, *Journal of Economic Theory* **148**(6): 2344–2382.

Kováč, E. and Mylovanov, T. (2009). Stochastic mechanisms in settings without monetary transfers: The regular case, *Journal of Economic Theory* **144**(4): 1373–1395.

Krishna, V. and Morgan, J. (2004). The art of conversation: Eliciting information from experts through multi-stage communication, *Journal of Economic Theory* **117**(2): 147 – 179.

Krishna, V. and Morgan, J. (2008). Contracting for information under imperfect commitment, *RAND Journal of Economics* **39**(4): 905–925.

Kvaløy, O. and Olsen, T. E. (2009). Endogenous verifiability and relational contracting, *American Economic Review* **99**(5): 2193–2208.

Mayer, K. J. and Argyres, N. S. (2004). Learning to contract: Evidence from the personal computer industry, *Organization science* **15**(4): 394–410.

Melumad, N. D. and Shibano, T. (1991). Communication in settings with no transfers, *RAND Journal of Economics* **22**(2): 173–198.

Schwartz, A. and Watson, J. (2013). Conceptualizing contractual interpretation, *Journal of Legal Studies* **42**(1): 1–34.

Shavell, S. (2006). On the writing and the interpretation of contracts, *Journal of Law, Economics, and Organization* **22**(2): 289–314.

Simon, H. A. (1951). A formal theory of the employment relationship, *Econometrica* **19**(3): 293–305.

Spector, D. (2000). Pure communication between agents with close preferences, *Economics Letters* **66**(2): 171–178.

Spier, K. E. (1992). Incomplete contracts and signalling, *RAND Journal of Economics* **23**(3): 432–443.

Szalay, D. (2005). The economics of clear advice and extreme options, *Review of Economic Studies* **72**(4): 1173–1198.