Contracting with Imperfect Commitment and Noisy Communication

Helmut Bester and Roland Strausz*

Version 2.0, January 3, 2006

Abstract

This paper provides new analytical tools for studying principal–agent problems with adverse selection and limited commitment. By allowing the principal to use general communication devices we overcome the literature's common, but overly restrictive focus on one–shot, direct communication. In addition, general communication devices solve two fundamental problems of contracting with imperfect commitment: First, they allow us to identify the 'local downward' incentive constraints as the relevant ones if the agent's preferences satisfy a single–crossing property. Second, we show how one may restrict the cardinality of the message spaces of the communication device. An example illustrates our arguments and the suboptimality of one–shot, direct communication.

Keywords: contract theory, communication, imperfect commitment, adverse selection; *JEL Classification No.: D82, C72*

*Free University Berlin, Dept. of Economics, Boltzmannstr. 20, D-14195 Berlin (Germany); Email addresses: *hbester@wiwiss.fu-berlin.de* and *strausz@wiwiss.fu-berlin.de*, respectively. Support by the German Science Foundation (DFG) through SFB/TR 15 is gratefully acknowledged. We wish to thank Daniel Krähmer, Jean–Jacques Laffont, Timofiy Mylovanov, an associate editor and a referee for helpful comments and suggestions.

1 Introduction

This paper provides a new analytical framework for studying contracting problems with adverse selection and limited commitment. We consider a principal–agent setup in which the principal is imperfectly informed about the agent's type. Further, he cannot contractually commit himself to some actions. The principal may, however, extract information from the agent by employing a general *communication device*. This device uses as input a report submitted by the agent and generates as output a publicly observable message. The set of possible inputs and outputs and the rules for transforming inputs into outputs are part of the contract.

General communication devices are well–known from game theory (e.g. Myerson (1982), Forges (1986)), but are not generally used in contract theory.[1] This is surprising because it is well–known that the set of implementable allocations may be strictly larger when players use indirect rather than direct communication. Yet, contract theory normally restricts attention to one–shot, direct communication, where the agent simply sends a single message directly to the principal. This restriction is unproblematic when the principal has full commitment; but for settings with imperfect commitment it is overly restrictive in two respects. First, already Forges (1990b) has demonstrated that multiple communication stages between one privately informed (the agent) and one uninformed party (the principal) enlarges the set of outcomes. More recently, Krishna and Morgan (2004) show that in the cheap talk game of Crawford and Sobel (1982) two rounds of communication typically generate Pareto improvements.[2] Second, as the communication literature (e.g. Myerson (1982), Forges (1986)) shows, even mechanisms with multiple stages of direct communication are restrictive in comparison to mechanisms that allow for indirect communication.[3] For these reasons, the use of one–shot, direct communication in contracting problems with imperfect commitment appears questionable. Consequently, we adopt general communication devices to study such contracting problems.

Moreover, general communication devices also allow us to solve two analytical

[1] In a rather different context Laffont and Martimort (1997) consider a communication device to study the bargaining behavior between a supervisor and an agent under two–sided asymmetric information.

[2] See also Aumann and Hart (2003) for analytical tools to study optimal communication in cheap talk games.

[3] See also Krishna and Morgan (2004) and Mitusch and Strausz (2005).

problems that prevent a tractable analysis of contracting problems with imperfect commitment: First, with general communication devices, we are able to identify the relevant incentive restrictions. As a result, even under imperfect commitment the principal's contracting problem can be solved by following a routine that is familiar from the theory of contracting with full commitment. Indeed, we first show that the global incentive compatibility constraints can be replaced by the usual *local downward* constraints, if the agent's preferences satisfy a familiar single–crossing property. This information about the structure of the optimal contract allows us to concentrate on a relaxed problem that is much easier to solve than the original contracting problem. Second, for the relaxed problem it is sufficient to consider message spaces that have the same cardinality as the agent's type space: Not only the set of ingoing reports but also the set of outgoing messages can be taken as a copy of the set of the agent's types. Thus the agent reveals his type truthfully to the communication system, which then generates a message suggesting a type of the agent. Yet, typically the principal will remain imperfectly informed because the optimal communication system is noisy; it converts the agent's report into a message according to non–degenerate transition probabilities.

We therefore stress that communication devices not only generalize but also simplify the analysis of contracting problems with imperfect commitment. Indeed, the literature so far has been rather unsuccessful in developing a manageable framework for these types of contracting problems, which inevitably arise when addressing issues such as ex post renegotiation (e.g. Dewatripont (1989)), repeated short–term contracting (e.g. Laffont and Tirole (1986)), or environments in which the principal takes some non–verifiable, and hence non–contractible, decision (e.g. Khalil (1997)). Laffont and Tirole (1986) illustrate the difficulties of extending the standard contracting framework to problems of imperfect commitment. It leads them to conclude that "the lack of commitment in repeated adverse–selection situations leads to substantial difficulties for contract theory" (Laffont and Tirole (1993), p. 377). This paper shows that these difficulties are in fact related to the restriction to one–shot, direct communication. Indeed, for this type of communication Bester and Strausz (2001) establish a variation of the Revelation Principle[4] which allows stating the contracting problem as a well–defined maximization program with the usual incentive compatibility constraints. Yet, even when the agent's preferences satisfy a single–crossing condition, it re-

[4]See Gibbard (1973), Green and Laffont (1977), Dasgupta *et. al.* (1979) and Myerson (1979).

mains unclear which of these constraints are binding. Solving the contracting problem requires a laborious checking of all combinations of incentive constraints (e.g. Laffont and Tirole (1987)).

Our first step in showing that general communication devices improve tractability is to adopt the Revelation Principle for Bayesian games (see: Myerson (1982), Forges (1986)). This principle states that, without loss of generality, the principal may use an *incentive compatible* contract. Under such a contract, the set of the agent's reports is simply the set of his types and his optimal strategy is to announce his type truthfully to the communication system. Therefore the contracting problem may be formulated as a maximization problem with incentive and individual rationality constraints. We show that this problem may be drastically simplified if the agent's preferences satisfy a single–crossing condition. Similarly to contracting problems with full commitment, this condition implies a monotonicity property of implementable allocations. Conversely, the monotonicity property together with the *downward* incentive constraints guarantees that all of the incentive compatibility and individual rationality conditions are satisfied. If there are $|T|$ types of the agent, this means that the $|T| \times |T-1| + |T|$ incentive compatibility and individual rationality constraints can be replaced by $2\,|T| - 1$ downward incentive and monotonicity constraints.

Replacing the global incentive constraints by local constraints is, however, only the first step towards making the contracting problem tractable. According to the Revelation Principle for Bayesian games, the communication device involves a set of outgoing messages with the same dimensionality as the set of the principal's actions. This creates a difficulty for computing an optimal communication device when the principal's action set is large. We address the problem of the cardinality of the message space by noting that monetary transfers and the vectors of transition probabilities of the communication device enter the principal's problem linearly. Therefore, we can apply a result from the theory of linear semi-infinite programming to show that a message set of cardinality $|T|$ is sufficient to support an optimal contract as long as the monotonicity constraints are not binding. When some monotonicity constraints are binding, the cardinality of the message set increases by the number of binding monotonicity constraints. Since there are exactly $|T| - 1$ monotonicity constraints, the cardinality of the optimal message set is at most $2|T| - 1$.

In summary, our findings allow us to derive the solution of the contracting problem by using the same methodology as for contracting with full commitment:

For a communication system with $|T|$ outgoing messages, we consider a maximization program where the global incentive and individual rationality constraints are replaced by the downward incentive compatibility constraints. If the solution of this program satisfies the monotonicity constraints, then it represents an optimal contract. If not, then one may use the repetitive procedure of including a binding monotonicity constraint in the maximization program and raising the number of outgoing messages by one, until one obtains a solution under which all the omitted monotonicity constraints are not binding. We illustrate this methodology in an example with $|T| = 2$, where monotonicity is automatically satisfied. Also, in this example noisy communication is superior to direct communication.

In Section 2 we describe the contracting environment. The concept of a communication device is explained in Section 3. In Section 4 we use the single–crossing assumption to show that the global incentive constraints can be replaced by local constraints. We address the cardinality of the message space under an optimal communication device in Section 5. Section 6 contains concluding remarks. All proofs are relegated to an appendix.

2 The Environment

Consider a two–stage contracting problem between a principal and an agent. In the first stage, the principal can contractually commit himself to a decision $x \in X$ and a monetary transfer $w \in \mathbb{R}$. In the second stage, he selects a decision $y \in Y(x) \subset Y$. The latter decision is not contractible in the first stage and so the principal will choose y at his own discretion. We allow the contractible decision x to limit the feasible set $Y(x)$ of non–contractible decisions.

In addition to the restriction of imperfect commitment, the contracting parties face a problem of one–sided asymmetric information: In stage 1 the agent has private information about his type $t_i \in T = \{t_1, \ldots, t_{|T|}\}$, where $|T| < \infty$. The principal only knows the probability distribution $\gamma = (\gamma_1, \ldots, \gamma_{|T|})$ of the agent's type, with $\gamma_i > 0$ and $\sum_i \gamma_i = 1$. The payoffs of the two parties depend on the agent's type t_i, the decisions (x, y) and the monetary transfer w. We denote the principal's payoff by

$$V_i(x, y, w) = v_i(x, y) - w \tag{1}$$

and the agent's payoff by

$$U_i(x, y, w) = u_i(x, y) + w. \tag{2}$$

The agent has the option to refuse to contract with the principal; his reservation payoff equals zero, independently of his type. Of course, under an optimal contract the decisions (x, y) and the transfer w will depend on the information that the principal can extract from the agent. In the subsequent sections, we investigate the properties of optimal communication devices.

Contract theory makes extensive use of single–crossing conditions on the agent's preferences. These conditions imply an ordering of the agent's types, which is natural in many economic environments. In our context we impose a single–crossing property by the following assumption:

Assumption 1 (Single–Crossing Property) There exists a $z = (z_1, \ldots, z_{|T|}) \in \mathbb{R}^{|T|}_{++}$ and a pair of functions $\varphi\colon X \times Y \to \mathbb{R}$, $\psi\colon X \times Y \to \mathbb{R}$ such that

$$u_i(x, y) = z_i\,\varphi(x, y) + \psi(x, y). \tag{3}$$

Moreover, either $\varphi(x, y) < 0$ and $z_i > z_{i+1}$, or $\varphi(x, y) > 0$ and $z_i < z_{i+1}$, for all $(x, y) \in X \times Y$ and $t_i \in T \backslash \{t_{|T|}\}$.

Our formulation of the single–crossing property is an extension of the familiar multiplicative form that is used in many contracting problems with full commitment (e.g. Baron and Myerson (1982), Maskin and Riley (1984), and Mussa and Rosen (1978)). Indeed, if the set $Y = \{y_0\}$ is a singleton, there is no commitment problem and only the contractible decision x is available for screening purposes. In this case, we may redefine $\tilde{x} = \varphi(x, y_0)$ so that Assumption 1 implies that the agent's marginal rate of substitution between $\tilde{x}$ and the transfer w

$$\frac{\partial U_i / \partial \tilde{x}}{\partial U_i / \partial w} \tag{4}$$

equals z_i and is monotone in type.

Whenever, for some $x \in X$, the set $Y(x)$ contains more than one element, the principal faces a problem with imperfect commitment. An example is Khalil (1997), where the principal has the possibility to audit the agent after the contracting stage. Since the principal is unable to commit contractually to an auditing strategy, he selects ex–post the probability y of an audit from the set $Y = [0, 1]$. In Khalil's model, the principal pays the agent the wage w for producing some output q. The agent of type t_i has the cost $z_i\,q$ for producing q units of output. After output has been produced, the principal can perform a non–contractible audit at a cost c to verifiably reveal the agent's type. Therefore, in addition to a quantity q, a contract specifies a transfer $\bar{w}_0$ if the principal does not audit the agent and a transfer $\bar{w}_i$ if the principal audits and discovers that the agent is of type t_i. The transfers may be negative, but Khalil assumes that they are bounded to rule out unlimited penalties. Hence, the contractible decision is $x = (q, \bar{w}_0, \bar{w}_1, \ldots, \bar{w}_{|T|}) \in X \subset \mathbb{R}_+ \times \mathbb{R}^{|T|+1}$ and the non–contractible decision is the probability of monitoring $y \in [0,1]$. Thus, an agent of type t_i receives the payoff

$$U_i(x, y, w) = w + y\bar{w}_i + (1-y)\bar{w}_0 - z_i q, \tag{5}$$

which satisfies our Assumption 1 for $\varphi(x,y) = -q$ and $\psi(x,y) = y\bar{w}_i + (1-y)\bar{w}_0$.

We want to stress, however, that the decision variables x and y may be more general than simple actions. For instance, the decision y may represent a continuation contract in a framework of repeated contracting without commitment. Consider for example a two–period model of short–term procurement contracts, in which the agent produces some output q in each period. The t_i–agent's production cost is $z_i\, q$. In the first period, the principal offers the payment w for the output q. In the second period, there is no commitment problem and the principal optimally offers an incentive compatible menu $y = ((\bar{q}_1, \bar{w}_1), ..., (\bar{q}_{|T|}, \bar{w}_{|T|}))$ so that the agent of type t_i selects the output $\bar{q}_i$ to receive the payment $\bar{w}_i$. Let Y denote the set of incentive compatible and individually rational second period contract menus.[5] In our terminology, the contractible decision is the first period output $x = q \in X$ and $y \in Y$ describes the non-contractible second period decision. The agent's payoff in this environment is

$$U_i(x, y, w) = w + \delta\bar{w}_i - z_i(q + \delta\bar{q}_i), \tag{6}$$

where $\delta \in (0,1)$ represents his discount factor. This payoff is consistent with Assumption 1 for $\varphi(x,y) = -(q + \delta\bar{q}_i)$ and $\psi(x,y) = \delta\bar{w}_i$. The specification in (6) is immediately applicable to a repeated version of the Baron and Myerson (1982) model. But (6) can also be adapted to the regulatory framework of Laffont and Tirole (1986, 1987, 1993), where the agent produces a single unit and incurs the observable cost c. This cost depends on the agent's type $t_i \in \mathbb{R}_{++}$ and his unobservable effort e according to

$$c = t_i - e. \tag{7}$$

[5] That is, each $y \in Y$ satisfies for all t_i the individual rationality condition $\bar{w}_i - z_i\,\bar{q}_i \geq 0$ and the incentive compatibility condition $\bar{w}_i - z_i\,\bar{q}_i \geq \bar{w}_j - z_i\,\bar{q}_j$ for all t_j

The agent's effort cost is $k(e)$. In the first period, the principal requires a cost level c in exchange for a transfer w. In the second period, he offers an incentive compatible menu $y = ((\bar{c}_1, \bar{w}_1), ..., (\bar{c}_{|T|}, \bar{w}_{|T|}))$ that induces the t_i–agent to select the cost level $\bar{c}_i$ in combination with the transfer $\bar{w}_i$. Therefore, the agent's payoff is

$$w - k(t_i - c) + \delta[\bar{w}_i - k(t_i - \bar{c}_i)]. \tag{8}$$

If the agent's effort cost is exponential so that $k(e) = \exp(e) - 1$, we may define $z_i \equiv \exp(t_i)$ and $q \equiv \exp(-c)$ to rewrite the agent's utility as

$$w + \delta\bar{w}_i - z_i(q + \delta\bar{q}_i) + (1 + \delta), \tag{9}$$

which is of the same form as the r.h.s. of (6) and thus fits our framework.

Since we allow the decision x to restrict the set Y, we may also use our environment to study problems of renegotiation (e.g. Laffont and Tirole (1990), Rey and Salanie (1996)), where the first period decisions restrict the feasible set in the second period. By employing the technique of dynamic programming, our environment can further be extended to multi–stage contracting problems with limited commitment (e.g. Hart and Tirole (1988)).[6] Hence, our framework is applicable to the kind of contracting problems with imperfect commitment that have been addressed in the literature.

Under our Assumption 1 private information is essentially one–dimensional. Consequently, our model addresses the class of one–dimensional screening problem with imperfect commitment. Indeed, as emphasized in Matthews and Moore (1987), the single–crossing condition only makes sense in problems with one–dimensional private information. We do not study multi–dimensional screening problems, which generate a number of analytical difficulties already within the framework of full commitment (see e.g. Rochet and Chone (1998)).

In the following sections we illustrate the role of noisy communication by an example which is based on Miyazaki (1977). The same example is used in Bester and Strausz (2001), where the analysis is restricted to single–stage face–to–face communication.

Example: There are two types of agents; each type is equally likely. The principal chooses the agent's speed of work $y \in \mathbb{R}_+$ and pays him a wage w. When the agent's type is t_i, the principal's and the agent's payoffs are $v_i(y) - w$ and

[6] For an exact demonstration of this approach see Bester and Strausz (2001).

$u_i(y) + w$, respectively, with

$$v_1(y) = 10y - y^2, v_2(y) = 10y - y^2/4, \; u_1(y) = -y^2/5, u_2(y) = -y^2/6. \qquad (10)$$

The agent's utility satisfies Assumption 1 for the specification $\varphi(x, y) = -y^2$, $\psi(x, y) = 0$, $z_1 = 1/5$, and $z_2 = 1/6$. $\diamond$

3 Communication

To address the problem of asymmetric information, the principal selects a *communication device*, which is a convenient, technical description of information transmission in most general form.[7] The device allows the agent to send a report upon which the principal receives a message. More specifically, a communication device $D = (R, M, \mathcal{B})$ specifies a set of reports $R = \{r_1, \ldots, r_k, \ldots, r_{|R|}\}$ with $|R| \leq \infty$, a set of messages $M = \{m_1, \ldots, m_h, \ldots, m_{|M|}\}$ with $|M| \leq \infty$, and a mapping $\mathcal{B}\colon R \to \Delta(M)$, where $\Delta(M)$ denotes the set of probability distributions over M.[8] In what follows, we use the notation $\mathcal{B}(r_k) = \beta_k = (\beta_{k1}, \ldots, \beta_{kh}, \ldots, \beta_{k|M|})$. The interpretation of a communication system is that the principal receives message m_h with probability β_{kh} after the agent has chosen the report r_k. Note that the principal cannot directly observe the agent's report.

A communication device $D = (R, M, \mathcal{B})$ is *deterministic* if $R = M$ and $\beta_{kk} = 1$. In this case, the principal receives the message m_k with probability one, when the agent sends the report m_k. Deterministic communication devices describe standard face–to–face communication, because the agent's report is directly transmitted to the principal without noise. Due to his lack of commitment, however, the principal may prefer not to receive too much information. By using a non–deterministic communication device, he is able to fine–tune the amount of information that is actually transferred to him.

The message received by the principal is publicly verifiable. For a given communication system D, a *contract* specifies a first–stage decision $x = (x_1, \ldots, x_h, \ldots, x_{|M|})$ and a monetary transfer $w = (w_1, \ldots, w_h, \ldots, w_{|M|})$ contingent upon

[7] The communication literature (e.g. Barany (1992), Ben-Porath (1998, 2003), Forges (1988, 1990a)) has investigated to what extent the somewhat mechanical concept of a communication device is equivalent to other, more conventional forms of communication in settings with complete and incomplete information.

[8] To simplify the exposition and to avoid measure–theoretic complications, we restrict both R and M to be countable. Our analysis can be extended to the case where R and M are metric spaces.

the message received by the principal. Of course, the principal will also use this message to update his beliefs about the agent's type before selecting his second–stage decision. Let $p = (p_{11}, \ldots, p_{ih}, \ldots, p_{|T||M|})$ denote the principal's beliefs. Thus, upon observing message m_h, the principal believes that the agent is of type t_i with probability $p_{ih} \geq 0$, where $\sum_i p_{ih} = 1$. For each message m_h, the principal's strategy $y = (y_1, \ldots, y_h, \ldots, y_{|M|})$ specifies a second–stage decision y_h.[9]

Given a communication system D and a contract (x, w), the principal and the agent are involved in the following game: First the agent chooses a report, which results in a message to the principal. After receiving a message the principal selects a second–stage decision based upon his beliefs over the agent's type. The contracting parties are constrained to the outcomes that can be realized as a Perfect Bayesian equilibrium of this game.

We allow the agent to employ a mixed reporting strategy and denote by $q_i = (q_{i1}, \ldots, q_{ik}, \ldots, q_{i|R|})$ the strategy of type t_i. Thus, the t_i–agent selects report r_k with probability $q_{ik} \geq 0$, where $\sum_k q_{ik} = 1$. When selecting a report, the agent anticipates the principal's decision in the second stage. Therefore, the agent's reporting strategy $q = (q_1, \ldots, q_i, \ldots, q_{|T|})$ is optimal if

$$q_i \in \text{argmax}_{q_i'} \sum\nolimits_{k,h} q_{ik}' \, \beta_{kh} \, [u_i(x_h, y_h) + w_h], \tag{11}$$

for all $t_i \in T$. Given the belief p, the principal's behavior in the second stage has to satisfy

$$y_h \in \text{argmax}_{y_h'} \sum\nolimits_i p_{ih} \, [v_i(x_h, y_h') - w_h], \tag{12}$$

for all $m_h \in M$. Finally, the principal's belief is consistent with Bayesian updating if

$$p_{ih} = \frac{\gamma_i \sum_k q_{ik} \, \beta_{kh}}{\sum_j \gamma_j \sum_k q_{jk} \, \beta_{kh}}, \tag{13}$$

for all $m_h \in M$ such that $q_{jk} \, \beta_{kh} > 0$ for some $(t_j, r_k) \in T \times R$. In summary, (q, y, p) constitutes a Perfect Bayesian equilibrium if conditions (11)–(13) are satisfied.

The message m_h is not used if $\sum_k q_{jk} \, \beta_{kh} = 0$ for all types $t_j \in T$. The solution concept of Perfect Bayesian Equilibrium does not put any restrictions

[9] Note that Y may contain the set of probability distributions over some underlying set of deterministic decisions. Therefore, we do not rule out random decisions.

on the principal's beliefs in (13) for such an out–of–equilibrium message and so $p_{ih} \geq 0$ with $\sum_i p_{ih} = 1$ is arbitrary. Actually, all messages that are not used in equilibrium can simply be deleted from the message set M without changing the equilibrium outcome. We use this insight in Section 5 to derive restrictions on the cardinality of the message set M.

Part of the principal's problem is finding an optimal communication system D. The following result provides a first step in this direction by applying the Revelation Principle for Bayesian games (see: Myerson (1982), Forges (1986)). It shows that, without loss of generality, one can assume that the set of the agent's reports is a copy of his types and that the agent reveals his type truthfully to the communication system.

Lemma 1 *Consider a given contract (x, w). Suppose (q, y, p) is a Perfect Bayesian equilibrium under the communication system $D = (R, M, \mathcal{B})$. Then there exists a communication system $\hat{D} = (\hat{R}, \hat{M}, \hat{\mathcal{B}})$ with $\hat{R} = T$ and $\hat{M} = M$ such that $(\hat{q}, y, p)$ with $\hat{q}_{ii} = 1$, for all $t_i \in T$, is a Perfect Bayesian equilibrium under $\hat{D}$. Moreover, $\sum_k \hat{q}_{ik}\,\hat{\beta}_{kh} = \sum_k q_{ik}\,\beta_{kh}$ for all $(t_i, m_h) \in T \times M$, i.e. $\hat{q}_i$ and q_i induce the same the probability distribution over M.*

By Lemma 1, the principal may restrict himself to an incentive compatible communication system, under which the agent reports his type honestly and the communication device garbles this information when sending a message to the principal.

Lemma 1 leaves open which restrictions can be imposed on the cardinality of M. The Revelation Principle for Bayesian games actually goes beyond the lemma by showing that, without loss of generality, the principal may set $M = Y$ so that he will select y after receiving the message $m = y$. The 'canonical' communication device $(T, Y, \mathcal{B})$ may therefore be interpreted as a mediator who first asks the agent for his private information and subsequently recommends some action y to the principal. This insight may be helpful for solving the contracting problem if the set Y contains only a few elements. In most applications, however, Y will be large because it includes continuous action choices or the set of continuation contracts in a multi–stage environment. Therefore, we will establish a more suitable restriction on M below in Section 5.

EXAMPLE: In Bester and Strausz (2001) it is shown that in our example the principal's maximum payoff is 55/2 when he is restricted to deterministic communication. The principal attains this payoff with the contract $w^* = (5, 20)$ and the deterministic communication device $D = (T, T, \mathcal{B})$ with $\beta_{11} = \beta_{22} = 1$ and $\beta_{12} = \beta_{21} = 0$. The payoff is supported by the Perfect Bayesian Equilibrium (q^*, y^*, p^*) with $q_1^* = (1/2, 1/2)$, $q_2^* = (0, 1)$, $y^* = (5, 10)$, $p_1^* = (1, 0)$ and $p_2^* = (1/3, 2/3)$. Thus, type t_1 randomizes between messages and so the principal remains imperfectly informed after receiving the message t_2. In line with Lemma 1, this outcome under the deterministic device D can be replicated by a noisy communication device $\hat{D}$ so that the agent reports his type truthfully. Indeed, it is easy to see that under $\hat{D} = (\hat{R}, \hat{M}, \hat{\mathcal{B}})$ with $\hat{R} = \hat{M} = T$ and $\hat{\beta}_{11} = \hat{\beta}_{12} = 1/2$, $\hat{\beta}_{21} = 0$ and $\hat{\beta}_{22} = 1$ the outcome $(\hat{q}, y^*, p^*)$ with $\hat{q}_1 = (1, 0)$ and $\hat{q}_2 = (0, 1)$ constitutes a Perfect Bayesian equilibrium. Yet, we will show below in Section 5 that $\hat{D}$ is not optimal within the class of noisy communication devices. Through a noisy communication channel the principal may be able to achieve a higher payoff than through deterministic communication. $\diamond$

4 Optimal Contracts

For a given set of messages M, Lemma 1 allows us to state the principal's problem as a programming problem in which the agent's reporting behavior has to satisfy standard incentive compatibility restrictions. Let $\beta_i = (\beta_{i1}, \ldots, \beta_{ih}, \ldots, \beta_{i|M|})$ and $\beta = (\beta_1, \ldots, \beta_i, \ldots, \beta_{|T|})$. Then the principal's objective is to maximize his expected payoff

$$\max_{x,y,w,\beta,p} V(x, y, w, \beta, p) \equiv \sum_{i,h} \gamma_i \beta_{ih} [v_i(x_h, y_h) - w_h] \tag{14}$$

subject to the incentive compatibility constraints

$$\sum_h \beta_{ih} [u_i(x_h, y_h) + w_h] \geq \sum_h \beta_{jh} [u_i(x_h, y_h) + w_h], \tag{15}$$

for all $t_i, t_j \in T \times T$; the agent's individual rationality constraints

$$\sum_h \beta_{ih} [u_i(x_h, y_h) + w_h] \geq 0, \tag{16}$$

for all $t_i \in T$; the no–commitment constraint

$$y_h \in \operatorname{argmax}_{y'_h} \sum_i p_{ih} v_i(x_h, y'_h); \tag{17}$$

and the Bayesian consistency constraint

$$p_{ih} = \frac{\beta_{ih} \gamma_i}{\sum_j \beta_{jh} \gamma_j}, \tag{18}$$

for all m_h such that $\beta_{jh}\gamma_j > 0$ for some $t_j \in T$. In what follows, we refer to (14)–(18) as the *principal's contracting problem* for a given message set M. Let $\mathcal{V}(M)$ denote the principal's expected payoff from a solution to this problem.

There remain two difficulties to derive a tractable procedure for solving the principal's problem. First, problem (14)–(18) is stated for a given message set M. Therefore, a characterization of the *optimal* message set is required. Second, it is unclear which of the incentive compatibility and individual rationality constraints are actually binding. The first difficulty is a fundamental one, while the latter is more of a computational nature. It turns out that the two problems are nevertheless related. By finding an answer to the second problem in this section, we are able to handle the more fundamental one in the following section.

In order to identify the binding constraints, we proceed by relaxing the principal's contracting problem in two directions. First, we follow the standard approach and focus on local rather than global constraints. In problems with full commitment this approach is valid if the agent's preferences satisfy a natural *single–crossing* condition. Effectively, the single–crossing condition reduces the complexity of the contracting problem because it identifies which of the global incentive–constraints are binding. It is well–known (e.g. Bester and Strausz (2001)), however, that this assumption fails to simplify the principal's problem under imperfect commitment if he is restricted to a deterministic communication device. Indeed, with this type of communication and lack of commitment "any incentive constraint could turn out to be binding at the optimum" (Laffont and Tirole (1993), p. 377). It is an important insight of our analysis below that this is no longer the case if the principal is able to employ a noisy communication system. In this situation, the standard approach can be used to study contracting problems for which a single–crossing condition such as Assumption 1 holds.

In addition to considering only local incentive constraints, we relax the problem in a second direction. Rather than considering message–dependent transfers w, we introduce type–dependent transfers $\omega = (\omega_1, \ldots, \omega_i, \ldots, \omega_{|T|})$. As long as $|T| < |M|$, replacing $w \in \mathbb{R}^{|M|}$ by $\omega \in \mathbb{R}^{|T|}$, reduces the number of variables in the principal's problem. Since we can, for a given β, transform any message–dependent transfer w into a type–dependent transfer ω by specifying $\omega_i = \sum_h \beta_{ih} w_h$, allowing the principal to use type–dependent transfers relaxes his contracting problem.

Specifically, we relax the principal's contracting problem by replacing the global constraints (15) and (16) in the principal's contracting problem by *down-*

ward incentive and *monotonicity* constraints. The downward incentive constraints require that

$$\sum_h \beta_{ih} u_i(x_h, y_h) + \omega_i \geq \sum_h \beta_{i-1,h} u_i(x_h, y_h) + \omega_{i-1}, \tag{19}$$

for all $t_i \in T$, where $\omega_0 \equiv 0$ and $\beta_{0h} \equiv 0$ for all $m_h \in M$.[10] The monotonicity constraints are satisfied if

$$\sum_h \beta_{ih} |\varphi(x_h, y_h)| \geq \sum_h \beta_{i-1,h} |\varphi(x_h, y_h)| \tag{20}$$

for all $t_i \in T \backslash \{t_1\}$.

In summary, we analyse the following *relaxed contracting problem* for a given message set M:

$$\begin{array}{cc} \max_{x,y,\omega,\beta,p} & W(x, y, \omega, \beta, p) \equiv \sum_{i,h} \gamma_i [\beta_{ih} v_i(x_h, y_h) - \omega_i] \\ \text{subject to} & (17), (18), (19), \text{ and } (20). \end{array} \tag{21}$$

For a given message set M, let $\mathcal{W}(M)$ denote the principal's expected payoff from a solution of the relaxed contracting problem (21). The following lemma shows that the constraints of the relaxed problem are implied by the constraints of the original problem.

Lemma 2 *If there exists a (x, y, w, β, p) satisfying the constraints of the principal's contracting problem, then there exists a (x, y, ω, β, p) satisfying the constraints of the relaxed contracting problem. Therefore, $\mathcal{W}(M) \geq \mathcal{V}(M)$.*

Obviously, the downward incentive constraints are binding in the relaxed contracting problem, because otherwise the principal could increase his payoff by lowering ω. This in combination with the single–crossing condition on preferences allows us to show that the local constraints (19) and (20) are sufficient to guarantee global incentive compatibility and individual rationality. As a result, we can show that the solution of the principal's contracting problem can be derived from solving the relaxed problem:

Proposition 1 *Let $(x^*, y^*, \omega^*, \beta^*, p^*)$ denote a solution to the relaxed contracting problem for a given message set M. Then, for any w^* such that $\sum_h \beta^*_{ih} w^*_h = \omega^*_i, i = 1, \ldots, |T|$, the tuple $(x^*, y^*, w^*, \beta^*, p^*)$ is a solution to the principal's contracting problem and generates the expected payoff $\mathcal{V}(M) = \mathcal{W}(M)$.*

[10] Thus for type t_1 condition (19) is simply the participation constraint.

The proposition shows that the solution of the principal's contracting problem can be obtained by solving the relaxed problem. Indeed, the relaxed problem (21) is much easier to solve than the original problem (14), because the $|T|\times|T-1|+|T|$ constraints in (15) – (16) are replaced by the $2\,|T|-1$ constraints in (19) – (20). In fact, the usual approach to screening problems with full commitment is to ignore the $|T|-1$ monotonicity constraints in (20) and then to check under what conditions, e.g. on the distribution of the agent's types, they are automatically satisfied by the solution. Proposition 1 shows that the same procedure can be used to solve the principal's problem in the context of imperfect commitment.

5 Optimal Message Spaces

Although Proposition 1 indicates how to simplify the principal's contracting problem for a given message space M, it does not say anything about the optimal cardinality of the message set itself. To investigate this issue, we say that M^* is an *optimal message set* if $\mathcal{W}(M^*) \geq \mathcal{W}(M')$ for any other message set M'. Note that, by Proposition 1, if $(x^*, y^*, \omega^*, \beta^*, p^*)$ solves the relaxed contracting problem under the message set M^*, then $V(x^*, y^*, w^*, \beta^*, p^*) = \mathcal{V}(M^*) = \mathcal{W}(M^*)$ for any w^* such that $\sum_h \beta^*_{ih} w^*_h = \omega^*_i, i = 1, \dots, |T|$. By Lemma 2, therefore, if M^* is an optimal message set, then $\mathcal{V}(M^*) = \mathcal{W}(M^*) \geq \mathcal{W}(M') \geq \mathcal{V}(M')$ for any other message set M'. That is, the principal's expected payoff from solving his contracting problem with an arbitrary message set M' cannot be higher than his expected payoff from $(x^*, y^*, w^*, \beta^*, p^*)$ under the message set M^*.

To determine the cardinality of an optimal message set, an insight from the theory of linear optimization turns out to be useful. To describe a linear program, let A be an $n \times m$–matrix, $c \in \mathbb{R}^m$ and $b \in \mathbb{R}^n$. For finite n and (possibly) infinite m the following program is a linear semi-infinite program with m decision variables $x \in \mathbb{R}^m_+$ and n constraints:[11]

$$\max_x \sum\nolimits_h c_h x_h \quad \text{subject to} \quad Ax = b,\; x \geq 0. \tag{22}$$

By a fundamental result in the theory of linear programming, whenever a finite linear program has a solution, then one can be found among the extreme points of the set of feasible solutions in (22). Therefore the program has a basic solution x^*, i.e. the number of non–zero components of x^* is no greater than the rank of

[11] See Anderson and Nash (1987) for the theory linear of linear programming in infinite–dimensional spaces.

A. The following lemma extends this result to linear problems with an infinite number of decision variables.

Lemma 3 *If there exists a solution $x^* \in \mathbb{R}^m$ to program (22), then there exists a solution x' with at most n non–zero components.*

Of course, the relaxed contracting problem (21) is not a linear programming problem. But, we can apply the above lemma by replacing each vector $\beta_h = (\beta_{1h}, \ldots, \beta_{|T|h})$ by $\lambda_h \beta_h$, with $\lambda_h \geq 0$, and adding the constraints $\sum_h \lambda_h \beta_h = 1$ for all $m_h \in M$. By keeping x, y, β and p fixed, we thus construct a programming problem that is linear in $\lambda = (\lambda_1, \ldots, \lambda_{|M|})$ and ω. Lemma 3 then allows us to show that there is an upper bound on the number of messages m_h for which a solution of the relaxed problem requires that $\lambda_h > 0$. Since all other messages with $\lambda_h = 0$ are redundant, we are able to derive restrictions on the cardinality of the optimal message set.

In what follows, we say that a subset of constraints in the relaxed contracting problem is *not binding* at the solution $(x^*, y^*, \omega^*, \beta^*, p^*)$, if $(x^*, y^*, \omega^*, \beta^*, p^*)$ remains a solution for maximizing W also when this subset of constraints is deleted from the relaxed problem.

Proposition 2 *Let $(x^*, y^*, \omega^*, \beta^*, p^*)$ be a solution of the relaxed contracting problem for the message set $M = T$. Suppose that the monotonicity constraints (20) are not binding. Then $M = T$ is an optimal message set.*

Under the conditions of Proposition 2, an optimal communication system has a rather simple structure: The agent reports his type truthfully and the principal receives a message that indicates a type of the agent. Yet, in general the principal remains imperfectly informed because the communication device is noisy; it thus may suggest a type to the principal that differs from the true type, as reported by the agent.

Effectively, under an optimal contract with $M = T$ the principal offers a menu (x, y, w) that specifies his decisions (x_i, y_i) in combination with a transfer w_i for each type $t_i \in T$ of the agent. Yet, the agent cannot select directly from this menu. Instead, the communication device allows him only to select one out of $|T|$ probability distributions over the $|T|$ elements in (x, y, w). By incentive compatibility, the t_i–type selects the probability distribution β_i so that he receives (x_j, y_j, w_j) with probability β_{ij}.

Proposition 2 reveals how imperfect commitment influences the communication between the contracting parties: The Revelation Principle for contracting games with perfect commitment shows that $M = T$ is an optimal message set. Further, under perfect commitment it is always optimal to set $\beta_{ii} = 1$, i.e. the agent's honest report about his type is transmitted without distortions to the principal. Typically, this form of 'direct' communication is no longer optimal when the principal cannot contractually commit himself to all of his actions. In this situation, he may prefer to become imprecisely informed by setting $\beta_{ii} < 1$.

There remains the obvious question of what happens when some of the monotonicity constraints are binding. In this case the procedure outlined below Lemma 3 is still applicable. Yet, binding monotonicity constraints cannot simply be deleted from the relaxed problem. Because the number of binding constraints increases, also the cardinality of the optimal message set increases.

Proposition 3 *Let $(x^*, y^*, \omega^*, \beta^*, p^*)$ be a solution of the relaxed contracting problem for the message set $M = T \cup \{1, \ldots, K\}$. If for this solution not more than K of the monotonicity constraints are binding, then $M = T \cup \{1, \ldots, K\}$ is an optimal message set.*

Hence, with binding monotonicity we lose the intuitive property that the cardinality of the optimal message space equals the cardinality of the type space and can therefore be interpreted as the type space. Yet, since the number of binding monotonicity constraints is at most $|T| - 1$ the upper bound on the cardinality of an optimal message set is $2|T| - 1$.

The proposition suggests the following algorithm for solving contracting problems with imperfect commitment: One starts with the message set $M = T$ and solves the relaxed problem ignoring the monotonicity constraints. If the solution automatically satisfies these constraints, one has found the optimal message set. If not, one repeatedly increases the cardinality of the message set until one finds a solution for which the number of binding monotonicity constraints matches the number of additional messages.

It would be helpful to have conditions on the primitives of the model which guarantee that $|M| = |T|$ suffices for an optimal message set. To address this question, we first note that monotonicity conditions are also well–known from contracting with full commitment. There incentive compatibility and the single crossing property require that schedules are monotonic in type. Our monotonicity condition (20) is the counterpart of this condition for the case with imperfect commitment. In particular, if the variable y is irrelevant and messages are truthful ($\beta_{ii} = 1$) then the monotonicity condition (20) simplifies to the standard requirement that the schedule $(x_1, \ldots, x_{|T|})$ is monotonic in type. Hence, it is the non–contractible decision y and the need for stochastic messages which make the monotonicity requirement more complicated than under perfect commitment.

Already under perfect commitment the monotonicity condition complicates the analysis because it leads to the problem of bunching. As is well–known, this problem can be avoided by assuming that the probability distribution of types satisfies a monotone hazard rate condition. For problems with imperfect commitment, however, we cannot expect that assumptions on the distribution of types will suffice to circumvent problems of monotonicity. With imperfect commitment, the variable y plays a non–trivial role in the monotonicity condition. Since y is determined by the no–commitment constraint (17), monotonicity depends also on the principal's decision behavior and cannot be guaranteed by conditions on the agent's distribution of types.

We can make this more precise in the context of our example. With two types there is only a single monotonicity condition, which for $|M| = |T| = 2$ can be written as

$$(\beta_{21} - \beta_{11})|\varphi(x_1, y_1)| + (\beta_{22} - \beta_{12})|\varphi(x_2, y_2)| \geq 0. \tag{23}$$

Without loss of generality we may assume that message m_1 is more indicative of type t_1 than message m_2, because messages can always be labelled such that $\beta_{11}/\beta_{21} \geq \beta_{12}/\beta_{22}$. Now using $\beta_{12} = 1 - \beta_{11}$ and $\beta_{22} = 1 - \beta_{21}$, we obtain $\beta_{11} \geq \beta_{21}$ so that (23) simplifies to

$$|\varphi(x_2, y_2)| \geq |\varphi(x_1, y_1)|. \tag{24}$$

In our example this condition is equivalent to $y_2 \geq y_1$ because $\varphi(x, y) = -y^2$. Hence, the monotonicity condition will be satisfied automatically whenever the principal chooses a larger y after he receives a message that is less indicative of type t_1. Since the principal chooses y according to (17), this is exactly the case when

$$\frac{\partial v_2(y)}{\partial y} > \frac{\partial v_1(y)}{\partial y}. \tag{25}$$

As this is satisfied in our example, we can ignore the monotonicity condition and be confident that a message space with $|M| = 2$ is optimal. We close this section by calculating the optimal contract in our example:

EXAMPLE: We will show that $M = T$ is an optimal message set. To simplify notation, let $\beta_1 = (a, 1-a)$ and $\beta_2 = (1-b, b)$. Thus if the agent selects the report t_1, the principal receives the message t_1 with probability a and the message t_2 with probability $1-a$. Similarly, if the agent selects t_2, the principal receives the message t_2 with probability b and the message t_1 with probability $1-b$. We first solve the relaxed problem ignoring the monotonicity constraint.

The (binding) downward constraints (19) are equivalent to

$$\begin{aligned} -\, b\, y_2^2/6 - (1-b)\, y_1^2/6 + \omega_2 &= -a\, y_1^2/6 - (1-a) y_2^2/6 + \omega_1, \\ -a\, y_1^2/5 - (1-a) y_2^2/5 + \omega_1 &= 0. \end{aligned} \tag{26}$$

Solving for (ω_1, ω_2) yields

$$\omega_1 = [a\,(y_1^2 - y_2^2) + y_2^2]/5, \omega_2 = [y_1^2(5 + a - 5b) + y_2^2(1 - a + 5b)]/30. \tag{27}$$

When the principal receives the message t_i, he believes that the agent's type is t_1 with probability p_{1i} and type t_2 with probability $p_{2i} = 1 - p_{1i}$. According to (17) he selects $y_i = 20/[1+3p_{1i}]$. Since $p_{11} = a/(1-b+a)$ and $p_{12} = (1-a)/(1-a+b)$, we have that

$$y_1 = 20(a-b+1)/(4a-b+1), \quad y_2 = 20(a-b-1)/(4a-b-4). \tag{28}$$

The principal's expected payoff is

$$\begin{aligned} W &= \left[a(10y_1 - y_1^2) + (1-a)(10y_2 - y_2^2) - \omega_1\right]/2 + \\ &\quad \left[b(10y_2 - y_2^2/4) + (1-b)(10y_1 - y_1^2/4) - \omega_2\right]/2. \end{aligned} \tag{29}$$

Substitution of (27) and (28) into W and maximizing with respect to a and b yields $a^* = 137/172$ and $b^* = 1$. Hence, $\beta_1^* = (137/172, 35/172)$ and $\beta_2^* = (0, 1)$. The optimal values for y^*, ω^* and p^* are:

$$y_1^* = 5, y_2^* = \frac{345}{26}, \omega_1^* = \frac{30145}{2704}, \omega_2^* = \frac{506245}{16224}, p_{11}^* = 1, p_{12}^* = \frac{35}{207}. \tag{30}$$

Hence $y_1^* < y_2^*$ and so the monotonicity constraint (20) is not binding. Thus $M = T$ is indeed an optimal message set.

Solving $\sum_h \beta^*_{ih} w_h = \omega^*_i, i = 1, 2$, yields

$$w^*_1 = 97745/16224,\ w^*_2 = 506245/16224. \tag{31}$$

By Proposition 1 $(x^*, y^*, w^*, \beta^*, p^*)$ is a solution to the principal's contracting problem; his expected payoff from this solution is $W^* = 71645/2496$. By Proposition 2 this is the highest possible expected payoff that he could achieve by any arbitrary communication device. Also, as we have pointed out at the end of Section 3, by using a deterministic device the principal could get at most a payoff of 55/2, which is less than W^*. This shows that in our example face–to–face communication is dominated by noisy communication. ◇

The example illustrates some important differences between direct and noisy communication. With direct communication it is unclear which of the incentive constraints are binding; in the example this happens to be the incentive constraint of the inefficient type. With noisy communication, however, the efficient type's incentive constraint is binding. As we have shown above, this feature is a general characteristic of optimal general communication devices.[12]

Also, in the example the principal obtains a higher payoff through noisy communication than through direct communication. This can be explained by the importance of stochastic messages in contracting problems with imperfect commitment. With direct communication, such stochastic messages can only be generated if the agent himself mixes between the available reports. This requires him to be indifferent between the outcomes of all reports that he sends with positive probability. In contrast, when the principal uses noisy communication, the agent only has to prefer the overall mixing probabilities associated with his report to the probabilities associated with the other reports. That is, the agent's incentive constraint is only expressed as a weighted average over the induced outcomes rather than for each individual outcome. Consequently, a noisy communication device has more degrees of freedom for generating stochastic messages.

6 Concluding Remarks

This paper presents a framework to study principal–agent problems with adverse selection and limited commitment. It demonstrates that by allowing for gen-

[12] Mitusch and Stausz (2005) obtain a similar qualitative difference in a specific game of cheap talk without transfers.

eral communication devices one can drastically simplify the derivation of optimal contracts. Instead, the literature on such problems has limited itself to direct, one–shot communication between the contracting parties. This not only restricts the parties' communication capabilities, but also makes it difficult to identify the binding incentive constraints of the contracting problem. In contrast, we allow the contracting parties to employ a general communication device. This together with a standard single–crossing assumption enables us to characterize the structure of optimal contracts: In the same way as in screening problems with full commitment, only the local downward incentive constraints turn out to be binding. Further, we derive an upper bound on the cardinality of message sets under an optimal communication device. These insights yield a tractable procedure for solving screening problems with imperfect commitment.

One further advantage of adopting a general communication device is that none of the arguments in this paper depend on there being only one agent.[13] Therefore our methodology is applicable also to the case of multiple agents. This differs from the extension of the Revelation Principle for direct communication derived in Bester and Strausz (2001), which does not hold for multiple agents (see Bester and Strausz (2000)). The reason is that with direct communication the agent has to be indifferent between all messages that he selects with positive probability. Therefore, it may be impossible to reduce the cardinality of the message space for several agents. In contrast, a noisy communication device generates random messages from a pure reporting strategy and so the agent does not have to be kept indifferent between different reports. This facilitates information transmission and gives more flexibility to reduce the number of messages under an optimal contract.

[13] In particular, our basic Lemma 1, the Revelation Principle for Bayesian games, holds for an arbitrary number of agents (see: Myerson (1982), Forges (1986)).

7 Appendix

Proof of Lemma 1: Let (q, y, p) be a Perfect Bayesian equilibrium under the communication system $D = (R, M, \mathcal{B})$, i.e. (q, y, p) satisfies conditions (11)–(13). Define the communication system $\hat{D} = (\hat{R}, \hat{M}, \hat{\mathcal{B}})$ by $\hat{R} = T$, $\hat{M} = M$, and $\hat{\beta}_{ih} = \sum_k q_{ik}\, \beta_{kh}$ for all $(t_i, m_h) \in T \times M$. Further let $\hat{q}$ satisfy $\hat{q}_{ii} = 1$ for all $t_i \in T$, and $\hat{q}_{ij} = 0$ whenever $t_i \neq t_j$.

Then $\sum_k \hat{q}_{ik}\, \hat{\beta}_{kh} = \hat{\beta}_{ih} = \sum_k q_{ik}\, \beta_{kh}$, which proves the second part of Lemma 1. This immediately implies that $(\hat{q}, y, p)$ satisfies conditions (12) and (13) of a Perfect Bayesian equilibrium under the communication system $\hat{D}$. It thus remains to show that $(\hat{q}, y, p)$ satisfies also condition (11). Suppose the contrary, i.e. there exists a q' such that

$$\begin{aligned}\sum\nolimits_{k,h} \hat{q}_{ik}\, \hat{\beta}_{kh}\, [u_i(x_h, y_h) + w_h] &= \sum\nolimits_{h} \hat{\beta}_{ih}\, [u_i(x_h, y_h) + w_h] \\ &< \sum\nolimits_{k,h} q'_{ik}\, \hat{\beta}_{kh}\, [u_i(x_h, y_h) + w_h]\,,\end{aligned} \tag{32}$$

for some q'_i and some $t_i \in T$. This implies that there is a $t_j \in T$ such that

$$\sum\nolimits_{h} \hat{\beta}_{ih}\, [u_i(x_h, y_h) + w_h] < \sum\nolimits_{h} \hat{\beta}_{jh}\, [u_i(x_h, y_h) + w_h]\,. \tag{33}$$

Therefore, by definition of $\hat{\beta}_{ih}$ and $\hat{\beta}_{jh}$,

$$\sum\nolimits_{k,h} q_{ik}\beta_{kh}\, [u_i(x_h, y_h) + w_h] < \sum\nolimits_{k,h} q_{jk}\beta_{kh}\, [u_i(x_h, y_h) + w_h]\,. \tag{34}$$

By this inequality, q_i fails to satisfy (11) and so it is not an optimal reporting strategy for type t_i under the communication system D. Therefore, (q, y, p) is not a Perfect Bayesian equilibrium under the communication system D, a contradiction. Q.E.D.

Proof of Lemma 2: First, we show that the incentive constraints in (15) imply the monotonicity constraints in (20). By Assumption 1, the incentive constraints in (15) imply that

$$\begin{aligned}\sum\nolimits_{h}[\beta_{ih} - \beta_{i-1,h}][z_i\, \varphi(x_h, y_h) + \psi(x_h, y_h) + w_h] &\geq 0, \\ \sum\nolimits_{h}[\beta_{i-1,h} - \beta_{ih}][z_{i-1}\, \varphi(x_h, y_h) + \psi(x_h, y_h) + w_h] &\geq 0,\end{aligned} \tag{35}$$

for all $t_i \in T\backslash\{t_1\}$. Adding these inequalities yields

$$\sum\nolimits_{h}[\beta_{ih} - \beta_{i-1,h}][z_i - z_{i-1}]\varphi(x_h, y_h) \geq 0. \tag{36}$$

Since Assumption 1 implies $(z_i - z_{i-1})\varphi(x, y) > 0$, this shows that (15) implies (20). By defining $\omega_i = \sum_h \beta_{ih} w_h$, it immediately follows that (15) implies (19). This confirms the first part of the lemma.

To prove the second statement, let $(x^*, y^*, w^*, \beta^*, p^*)$ be a solution to the principal's contracting problem leading to the payoff $\mathcal{V}(M)$. Define $\omega_i^* = \sum_h \beta_{ih}^* w_h^*$. Since $(x^*, y^*, w^*, \beta^*, p^*)$ satisfies the constraints (15) and (16), it follows by the argument above that $(x^*, y^*, \omega^*, \beta, p^*)$ satisfies (19) and (20). Therefore, $(x^*, y^*, \omega^*, \beta^*, p^*)$ satisfies all the constraints in (21) and $W(x^*, y^*, \omega^*, \beta^*, p^*) = V(x^*, y^*, w^*, \beta^*, p^*) = \mathcal{V}(M)$. Therefore, the principal's expected payoff from the solution of the relaxed problem cannot be less than $\mathcal{V}(M)$. Q.E.D.

Proof of Proposition 1: Note that for $(x^*, y^*, \omega^*, \beta^*, p^*)$ all the constraints in (19) are binding. Indeed, if the inequality would hold for the downward constraint of some type t_i, then the principal could increase his expected payoff by lowering ω_i without violating any other constraint in (19). Thus for any w^* such that $\sum_h \beta_{ih}^* w_h^* = \omega_i^*, i = 1, \ldots, |T|$, we have

$$\sum\nolimits_h \beta_{ih}^* [u_i(x_h^*, y_h^*) + w_h^*] = \sum\nolimits_h \beta_{i-1,h}^* [u_i(x_h^*, y_h^*) + w_h^*], \tag{37}$$

for all $t_i \in T$.

It remains to show that the combination $(x^*, y^*, w^*, \beta^*, p^*)$ satisfies the incentive compatibility conditions (15) and the individual rationality conditions (16). Define

$$\theta_i \equiv \sum\nolimits_h \beta_{ih}^* \varphi(x_h^*, y_h^*),\ \ \mu_i \equiv \sum\nolimits_h \beta_{ih}^* [\psi(x_h^*, y_h^*) + w_h^*], \tag{38}$$

so that we may rewrite (37) as

$$\mu_i = \mu_{i-1} - z_i\,(\theta_i - \theta_{i-1}), \tag{39}$$

for all $t_i \in T \backslash \{t_1\}$. Applying (39) iteratively yields

$$\mu_i = \mu_1 - \sum\nolimits_{\ell=2}^{i} z_\ell\,(\theta_\ell - \theta_{\ell-1}). \tag{40}$$

Thus, for all $(t_i, t_j) \in T \times T$,

$$\begin{aligned}(z_i\theta_i + \mu_i) \ &- \ (z_i\theta_j + \mu_j) = z_i(\theta_i - \theta_j) \\ &+ \ \sum\nolimits_{\ell=2}^{j} z_\ell\,(\theta_\ell - \theta_{\ell-1}) - \sum\nolimits_{\ell=2}^{i} z_\ell\,(\theta_\ell - \theta_{\ell-1}).\end{aligned} \tag{41}$$

If $i < j$ then

$$\begin{aligned}\sum\nolimits_{\ell=2}^{j} z_\ell\,(\theta_\ell - \theta_{\ell-1}) - \sum\nolimits_{\ell=2}^{i} z_\ell\,(\theta_\ell - \theta_{\ell-1}) &= \sum\nolimits_{\ell=i+1}^{j} z_\ell\,(\theta_\ell - \theta_{\ell-1}) \\ = \sum\nolimits_{\ell=i+1}^{j} (z_\ell - z_i)\,(\theta_\ell - \theta_{\ell-1}) + z_i(\theta_j - \theta_i) &\geq z_i(\theta_j - \theta_i),\end{aligned} \tag{42}$$

where the inequality follows because the monotonicity condition (20) and Assumption 1 imply $\text{Sign}(\theta_\ell - \theta_{\ell-1}) = \text{Sign}(\varphi) = \text{Sign}(z_l - z_i)$ for $l > i$ so that $(z_\ell - z_i)\,(\theta_\ell - \theta_{\ell-1}) \geq 0$.

If $i > j$, then

$$\begin{aligned}\sum\nolimits_{\ell=2}^{j} z_\ell\,(\theta_\ell - \theta_{\ell-1}) - \sum\nolimits_{\ell=2}^{i} z_\ell\,(\theta_\ell - \theta_{\ell-1}) &= -\sum\nolimits_{\ell=j+1}^{i} z_\ell\,(\theta_\ell - \theta_{\ell-1}) \\ = \sum\nolimits_{\ell=j+1}^{i} (z_i - z_\ell)\,(\theta_\ell - \theta_{\ell-1}) + z_i(\theta_j - \theta_i) &\geq z_i(\theta_j - \theta_i),\end{aligned} \tag{43}$$

because $(z_i - z_\ell)\,(\theta_\ell - \theta_{\ell-1}) \geq 0$ as $i > \ell$.

By (42) and (43), we obtain that (41) is non–negative so that $z_i\theta_i + \mu_i \geq z_i\theta_j + \mu_j$ for all $(t_i, t_j) \in T \times T$. This shows that $(x^*, y^*, w^*, \beta^*, p^*)$ satisfies the incentive compatibility conditions in (15).

To show that $(x^*, y^*, w^*, \beta^*, p^*)$ also satisfies the individual rationality conditions (16), note that the equality for t_1 in (37) implies

$$\sum\nolimits_h \beta^*_{1h}[u_1(x^*_h, y^*_h) + w^*_h] = 0. \tag{44}$$

As we have shown that $(x^*, y^*, w^*, \beta^*, p^*)$ satisfies the incentive compatibility conditions (15), we have for all $t_i \in T$

$$\begin{aligned}\sum\nolimits_h \beta^*_{ih}[u_i(x^*_h, y^*_h) + w^*_h] &\geq \sum\nolimits_h \beta^*_{1h}[u_i(x^*_h, y^*_h) + w^*_h] \\ &\geq \sum\nolimits_h \beta^*_{1h}[u_1(x^*_h, y^*_h) + w^*_h] = 0,\end{aligned} \tag{45}$$

where the second inequality follows from the monotonicity condition (20). Thus also the individual rationality conditions in (16) are satisfied.

We conclude that $(x^*, y^*, w^*, \beta^*, p^*)$ satisfies all constraints (15)-(18) of the principal's problem. It yields the principal the payoff $V(x^*, y^*, w^*, \beta^*, p^*) = W(x^*, y^*, \omega^*, \beta^*, p^*) = \mathcal{W}(M)$. By Lemma 2 the original contracting problem cannot yield more than $\mathcal{W}(M)$ and so we have $\mathcal{V}(M) = \mathcal{W}(M)$. Q.E.D.

Proof of Lemma 3: If x^* has finitely many non-zero entries, then a standard argument (e.g. Theorem 2.5 of Anderson and Nash (1987, p.23)) shows that there exists a basic optimal solution to program (22). Since Rank(A) $\leq n$, it follows that a basic solution to (22) has at most n non-zero entries. Now suppose x^* has infinitely many non-zero entries. Following the approach of the proof of Theorem 4.8 of Anderson and Nash (1987, p.76), we show that there exists a solution $\bar{x}$ with at most $n+2$ non-zero entries, which then implies that there is a basic solution that has at most n non-zero entries.

Let $V^* \equiv \sum_h c_h x_h^*$ be the value of program (22). Extend the matrix A by adding the row vector $c = (c_1, c_2, \ldots)$ to $\hat{A}$. That is, $\hat{A} \in \mathbb{R}^{(n+1)\times m}$ consists of the column vectors $\hat{a}_i = (a_{1i}, a_{2i}, \ldots, a_{ni}, c_i)$. Consequently, any $x \geq 0$ which is a solution to

$$\hat{A}x = \hat{b} \equiv (b, V^*) \tag{46}$$

is also a solution to (22), as it satisfies $Ax = b$ and has the value $c.x = V^*$. In particular, x^* is a solution to (46). Define the cone $C = \{\lambda \hat{a}_i | \lambda \geq 0, i = 1, 2, \ldots, \} \subset \mathbb{R}^{n+1}$ generated by the column vectors $(\hat{a}_1, \hat{a}_2, \ldots)$ in $\hat{A}$. Since

$$\hat{b} = \lim_{k\to\infty} \sum\nolimits_{i=1}^{k} \left(\frac{1}{2}\right)^i 2^i \hat{a}_i x_i^*, \tag{47}$$

$\hat{b}$ is an infinite mixture of points in C and, because C lies in the finite Euclidean space $\mathbb{R}^{n+1}$, it follows that $\hat{b}$ lies in the convex hull of C (Rubin and Wesler (1958)). Hence, Caratheodory's theorem (Rockafellar (1970), Theorem 17.1, p. 155) implies that $\hat{b}$ can be written as a convex combination of $k \leq n+2$ elements $(\xi_1, \ldots, \xi_k)$ in C. That is, there exists $(\mu_1, \ldots, \mu_k)$ with $\sum_i \mu_i = 1$ and $\mu_i \geq 0$ such that

$$\hat{b} = \sum\nolimits_{i=1}^{k} \mu_i \xi_i. \tag{48}$$

Since for each $\xi_j \in C$ there exists an $i(j)$ such that ξ_j can be written as $\lambda_j \hat{a}_{i(j)}$, we may rewrite $\hat{b}$ as

$$\hat{b} = \sum\nolimits_{j=1}^{k} \mu_j \xi_j = \sum\nolimits_{j=1}^{k} \mu_j \lambda_j \hat{a}_{i(j)}. \tag{49}$$

Now let $J(i) \equiv \{j | i(j) = i\}$ and define $\bar{x}_i \equiv \sum_{J(i)} \mu_j \lambda_j$. It follows that $\bar{x}$ has at most $k \leq n+2$ non-zero entries and satisfies (46). Q.E.D.

Proof of Proposition 2: Let $(x^*, y^*, \omega^*, \beta^*, p^*)$ be a solution of the relaxed contracting problem for the message set $M = T$ with value $W^* =$

$W(x^*, y^*, \omega^*, \beta^*, p^*) = \mathcal{W}(T)$. If the monotonicity constraints (20) are not binding then $(x^*, y^*, \omega^*, \beta^*, p^*)$ solves the following problem

$$\max_{x,y,\omega,\beta,p} W(x, y, \omega, \beta, p) \tag{50}$$

subject to (17), (18) and

$$\sum_h \beta_{ih} u_i(x_h, y_h) + \omega_i = \sum_h \beta_{i-1,h} u_i(x_h, y_h) + \omega_{i-1} \tag{51}$$

for all $t_i \in T$.

Now suppose the message set T is not optimal. Obviously, the principal cannot get a higher payoff by using a message set $\bar{M}$ with $|\bar{M}| < |T|$. Accordingly, there must exist a message set $\bar{M}$ with $|\bar{M}| > |T|$ such that $\mathcal{W}(\bar{M}) > \mathcal{W}(T) = W^*$. We will show that this yields the contradiction that $(x^*, y^*, \omega^*, \beta^*, p^*)$ does not solve program (50) for the message set $M = T$.

Let the combination $(\bar{x}, \bar{y}, \bar{\omega}, \bar{\beta}, \bar{p})$ represent a solution of the relaxed contracting problem given the message set $\bar{M}$, i.e. $W(\bar{x}, \bar{y}, \bar{\omega}, \bar{\beta}, \bar{p}) = \mathcal{W}(\bar{M})$. Since all downward incentive constraints (19) are binding, $\bar{\omega}$ together with $\lambda_1 = \ldots = \lambda_{|\bar{M}|} = 1$ solves the program

$$\max_{\lambda,\omega} \quad \sum_{i,h} \gamma_i [\lambda_h \bar{\beta}_{ih} v_i(\bar{x}_h, \bar{y}_h) - \omega_i] \tag{52}$$

subject to

$$\sum_h \lambda_h \bar{\beta}_{ih} u_i(\bar{x}_h, \bar{y}_h) + \omega_i = \sum_h \lambda_h \bar{\beta}_{i-1,h} u_i(\bar{x}_h, \bar{y}_h) + \omega_{i-1}, \tag{53}$$

$$\sum_h \lambda_h \bar{\beta}_{ih} = 1, \tag{54}$$

for all $t_i \in T$; and $\lambda_h \geq 0$ for all $m_h \in \bar{M}$. Solving (53) for ω yields

$$\omega_i(\lambda) \equiv -\sum_{j=1}^{i} \sum_h \lambda_h (\bar{\beta}_{jh} - \bar{\beta}_{j-1,h}) u_j(\bar{x}_h, \bar{y}_h) \tag{55}$$

for all $t_i \in T$. By substitution we may therefore rewrite problem (52)–(54) as

$$\max_\lambda \bar{W}(\lambda) \equiv \sum_i \gamma_i \sum_h \left[\lambda_h \bar{\beta}_{ih} v_i(\bar{x}_h, \bar{y}_h) - \omega_i(\lambda) \right] \tag{56}$$

subject to

$$\sum_h \lambda_h \bar{\beta}_{ih} = 1 \text{ for all } t_i \in T; \tag{57}$$

and $\lambda_h \geq 0$ for all $m_h \in \bar{M}$.

The objective function and constraints of problem (56)–(57) are linear in λ. According to Lemma 3 it has therefore a solution λ^* with $k \leq |T|$ strictly positive entries $\lambda_h^* > 0$. Since $\lambda_1 = \ldots = \lambda_{|\bar{M}|} = 1$ is also a solution, $\bar{W}(\lambda^*) = \bar{W}(1, \ldots, 1) = \mathcal{W}(\bar{M})$.

Now consider the combination $(\bar{x}, \bar{y}, \omega(\lambda^*), \beta(\lambda^*), \bar{p})$ with $\beta_{ih}(\lambda^*) \equiv \lambda_h^* \bar{\beta}_{ih}$. By taking from $(\bar{x}, \bar{y}, \omega(\lambda^*), \beta(\lambda^*), \bar{p})$ those k entries for which $\lambda_h^* > 0$ and $|T| - k$ entries for which $\lambda_h^* = 0$, we obtain a combination $(x', y', \omega', \beta', p')$ for a message set M' with $|M'| = |T|$. It follows that $W(x', y', \omega', \beta', p') = W(\bar{x}, \bar{y}, \omega(\lambda^*), \beta(\lambda^*), \bar{p})$ $= \bar{W}(\lambda^*) = \mathcal{W}(\bar{M})$.

By construction, $(x', y', \omega', \beta', p')$ satisfies (51) and therefore also (19). It also satisfies the constraints (18), because

$$p'_{ih} = \frac{\beta'_{ih}\gamma_i}{\sum_j \beta'_{jh}\gamma_j} = \frac{\lambda_h^* \bar{\beta}_{ih}\gamma_i}{\sum_j \lambda_h^* \bar{\beta}_{jh}\gamma_j} = \frac{\bar{\beta}_{ih}\gamma_i}{\sum_j \bar{\beta}_{jh}\gamma_j} = \bar{p}_{ih}, \tag{58}$$

for all m_h such that $\beta'_{jh}\gamma_j > 0$ for some $t_j \in T$. From this property of the beliefs p' it follows immediately that also (17) holds. In summary, $(x', y', \omega', \beta', p')$ satisfies all constraints of program (50). But, as $W(x', y', \omega', \beta', p') > W(x^*, y^*, \omega^*, \beta^*, p^*)$, we obtain the contradiction that $(x^*, y^*, \omega^*, \beta^*, p^*)$ cannot solve program (50). Hence, there cannot exist a message set $\bar{M}$ such that $\mathcal{W}(\bar{M}) > \mathcal{W}(T)$. Q.E.D.

Proof of Proposition 3: Analogous to the proof of Proposition 2 with the exception that program (52) now has in addition the K monotonicity constraints. Hence, program (56) has $|T| + K$ constraints so that according to Lemma 3, it has a solution λ^* with at most $|T| + K$ strictly positive entries. Q.E.D.

8 References

Anderson, E. and P. Nash (1987): "*Linear Programming in Infinite–Dimensional Spaces*," Wiley Interscience Publication – New York.

Aumann, R. and S. Hart (2003): "Long Cheap Talk," *Econometrica* 71, 1619–1660.

Barany, I. (1992): "Fair Distribution Protocols or How the Players Replace Fortune," *Mathematics of Operations Research* 17, 327-340.

Baron D. and R. Myerson (1982): "Regulating a Monopolist with Unknown Costs," *Econometrica* 50, 911-930.

Ben-Porath, E. (1998): "Correlation without Mediation: Expanding the Set of Equilibrium Outcomes by Cheap Pre-Play Procedures," *Journal of Economic Theory* 80, 108-122.

Ben-Porath, E. (2003): "Cheap Talk in Games with Incomplete Information," *Journal of Economic Theory* 108, 45-71.

Bester, H. and R. Strausz (2000): "Imperfect Commitment and the Revelation Principle: The Multi–Agent Case," *Economics Letters* 69, (2000), 165-171.

Bester, H. and R. Strausz (2001): "Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case," *Econometrica* 69, 1077-1098.

Crawford, V. and J. Sobel (1982): "Strategic Information Transmission," *Econometrica* 50, 1431-1451.

Cremer, J. (1995): "Arm's Length Relationships," *Quarterly Journal of Economics* 110, 275-95.

Dasgupta, P., P. Hammond and E. Maskin (1979): "The Implementation of Social Choice Rules," *Review of Economic Studies 46* 185–216.

Dewatripont, M. (1988): "Commitment through Renegotiation–Proof Contracts with Third Parties," *Review of Economic Studies* 55, 377-89.

Dewatripont, M. (1989): "Renegotiation and Information Revelation over Time: The Case of Optimal Labor Contracts," *Quarterly Journal of Economics* 104, 589-619.

Dewatripont, M. and E. Maskin (1995): “Contractual Contingencies and Renegotiation,” *Rand Journal of Economics* 26, 704-19.

Forges, F. (1986): “An Approach to Communication Equilibria,” *Econometrica* 54, 1375-1385.

Forges, F. (1988): “Can Sunspots Replace a Mediator?,” *Journal of Mathematical Economics,* 17, 347-368.

Forges, F. (1990a): “Universal Mechanisms,” *Econometrica* 58, 1341-1364.

Forges, F. (1990b): “Equilibria with Communication in a Job Market Example,” *Quarterly Journal of Economics* 105, 375-398.

Gibbard, A. (1973): “Manipulation for Voting Schemes,” *Econometrica* 41, 587–601.

Hart, O. and J. Tirole (1988): “Contract Renegotiation and Coasian Dynamics,” *Review of Economic Studies* 55, 509-540.

Green, J. and J.–J. Laffont (1977): “Characterization of Satisfactory Mechanisms for the Revelation of Preferences,” *Econometrica* 45, 427–438.

Khalil, F. (1997): “Auditing without Commitment,” *Rand Journal of Economics* 28, 629–640.

Krishna, V. and J. Morgan (2004): ”The Art of Conversation, Eliciting Information from Experts through Multi–Stage Communication” *Journal of Economic Theory* 117, 147–179.

Laffont, J.–J. and D. Martimort (1997): “Collusion under Asymmetric Information,” *Econometrica* 65, 875-911.

Laffont, J.–J. and J. Tirole (1986): “The Dynamics of Incentive Contracts,” *Econometrica* 56, 1153-75.

Laffont, J.–J. and J. Tirole (1987): “Comparative Statics of the Optimal Dynamic Incentive Contract,” *European Economic Review* 31, 901–926.

Laffont, J.–J. and J. Tirole (1990): “Adverse Selection and Renegotiation in Procurement,” *Review of Economic Studies* 57, 597-625.

Laffont, J.–J. and J. Tirole (1993): *"A Theory of Incentives in Procurement and Regulation,"* MIT Press: Cambridge – Massachusetts, London.

Maskin, E. and J. Riley (1984): "Monopoly Selling Strategies with Incomplete Information," *Rand Journal of Economics* 15, 171-196.

Matthews, S. and J. Moore (1987): "Monopoly Provision of Quality and Warranties: An Exploration in the Theory of Multidimensional Screening," *Econometrica* 55, 441-467.

Mitusch, K. and R. Strausz (2005): "Mediation in Situations of Conflict and Limited Commitment", *forthcoming in Journal of Law, Economics and Organization.*

Miyazaki, H. (1977): "The Rat Race and Internal Labor Markets," *Bell Journal of Economics* 8, 394–418.

Mussa, M. and S. Rosen (1978): "Monopoly and Product Quality," *Journal of Economic Theory* 18, 301-317.

Myerson, R. (1979): "Incentive Compatibility and the Bargaining Problem," *Econometrica* 47, 61–73.

Myerson, R. (1982): "Optimal Coordination Mechanisms in Generalized Principal–Agent Problems," *Journal of Mathematical Economics* 10, 67-81.

Rey, P. and B. Salanie (1996): "On the Value of Commitment with Asymmetric Information," *Econometrica* 64, 1395-1414.

Rochet, J.-C. and P. Chone (1998): "Ironing, Sweeping, and Multidimensional Screening," *Econometrica* 66, 783-826.

Rockafellar, R. (1970): *"Convex Analysis,"* Princeton University Press – New Jersey.

Rubin, H. and O. Wesler (1958): "A Note on Convexity in Euclidean n-Space", *Proceedings of the American Mathematical Society* 9, 522–523.