Limited Observability as a Constraint in Contract Design*

Stefan Krasa and Steven R. Williams
Department of Economics
University of Illinois
1206 South Sixth Street
Champaign IL 61820
http://www.staff.uiuc.edu/~skrasa
swillia3@uiuc.edu

April 18, 2003

Abstract

Limited observability is the assumption that economic agents can only observe a finite amount of information. Given this constraint, contracts among agents are necessarily finite and incomplete in comparison to the ideal complete contract that we model as infinite in detail. We consider the extent that finite contracts can approximate the idealized complete contracts. The objectives of the paper are: (i) to identify properties of agents' preferences that determine whether or not finiteness of contracts causes significant inefficiency; (ii) to evaluate the performance of finite contracts against the ideal optimal contract in a bilateral bargaining model.

Keywords: limited observability, bounded rationality, contractual incompleteness

JEL Classification Numbers: D82, C78

1 Introduction

People are limited in their ability to observe and to process information. Contracts among individuals and institutions are therefore limited in length and may fail to address all of the different contingencies that may arise. This incompleteness can lead to inefficiency in the contractual outcome, as evidenced by legal disputes or costly renegotiation. We develop in this paper a model that encompasses both the limited contracts that are used in practice and the ideal contracts that address all contingencies. The goal

*This paper originated in In-Koo Cho's seminar on bounded rationality that we attended in the fall of 1998. We thank In-Koo for inviting us to participate in his seminar and for his comments concerning this paper. We also thank Steven Matthews, Nabil Al-Najjar, Jim Peck, Stan Reiter, and a referee of this journal for their suggestions.

of this paper is to identify properties of agents' preferences that determine whether or not limits on contractual length cause significant inefficiency in comparison to an ideal complete contract.

To illustrate the issues that we address, consider a service provider and a system owner who negotiate a service contract for a computer system. A *contingency* is a "yes" or "no" answer that indicates the occurrence of some aspect of the state of the world. Contingencies of interest to the owner include all aspects of his business that determine the value he would receive from a properly maintained system, including future decisions by current and prospective customers and employees. For the service provider, the contingencies include everything that affects his cost of providing the service, including service demands from other customers, the difficulty of repairing different components of the system, and whether or not the various components require service. To model the complexity of the world, we assume that the number of contingencies is countably infinite. The provider and the owner, however, are assumed capable of only observing a finite number of contingencies and hence can address only a finite number in their contract. At issue is whether or not the finite contracts that they may write can approximate the ideal infinite contracts.

It is quite possible that beyond a certain number of contingencies the aggregate effect of all other contingencies upon the welfare of the contracting parties is small. We state this formally as a continuity condition that is assumed to hold throughout the paper. Not surprisingly, continuity of this kind is useful for proving that an infinite contract can be approximated arbitrarily closely with finite contracts. The optimal service contract in this example might thus be approximated with a finite contract that does not address an infinite number of contingencies that are insignificant in their cumulative effect upon the owner and the provider.

We also find, however, that certain contingencies cannot be addressed in a contract because of the constraint of incentive compatibility. Suppose that the owner or the provider observes privately the realization of a particular contingency. If the contract is to address this contingency, then the agent who observes its realization must be induced to reveal his observation. Whether or not this can be accomplished depends upon the relationship between the contingency in question and other contingencies in their effect upon the reporting agent's welfare. Roughly, a contingency is *reversible* if other contingencies may be realized in a number of different ways so as to undo (or reverse) the effect upon the agent's welfare of the given contingency, regardless of its realization. Conversely, the contingency is *strongly irreversible* if its effect upon welfare cannot be masked by other contingencies. Each of these properties concern the relationship between a contingency and other contingencies in their effect upon welfare; neither is simply an issue of the absolute magnitude of the contingency's effect upon welfare, and hence these properties are distinct from the continuity condition discussed above. A reversible contingency that is observed privately by an agent presents an opportunity for misrepresentation that may be unsolvable in the sense that an incentive compatible and finite contract cannot address it, even if its impact upon welfare is large. Conversely, if a contingency is strongly irreversible, then the agent who observes it can be compensated so as to induce him to reveal its realization, which facilitates addressing it in a contract.

We refer to an agent's measure of well-being as *preferences* in our general model, though it may have a more precise interpretation (e.g., as the profit of the system owner) in examples such as the one above. We thus characterize two properties (reversibility and strong irreversibility) that concern how a contingency affects an agent's preferences in relation to other contingencies. Strong irreversibility or reversibility affect whether an incentive compatible finite contract either can or can not address the given contingency. Assuming a finite ability to observe, our paper thus approaches through the properties of agents' preferences the problems of (i) which contingencies may and which contingencies may not be addressed in an incentive compatible and finite contract, and (ii) when a substantial welfare loss necessarily occurs because of contractual finiteness in comparison to the ideal of infinite contracts.

Model and Results. The model is presented in section 2. There are two agents. For each agent j and for any $n \in \mathbb{N}$, *contingency n* is a random variable whose value is either 0 or 1. The *type* α_j of agent j is a sequence $\alpha_j = (a_{j,n})_{n\in\mathbb{N}}$ in which $a_{j,n}$ is a realization of the *n*th contingency. The type set A_j of agent j is $\{0, 1\}^{\mathbb{N}}$. The *set of states of the world* is $A = A_1 \times A_2$. The utility $u_j(\alpha_j, c)$ of each agent is a function of his type α_j and the choice c. A *contract* is a function $f : A \to C$ from A into a set C of possible collective choices for the agents. A contract is *finite* if it only depends on a finite number of contingencies. Otherwise, it is *infinite*.

We deviate from Harsanyi's (1967-68) theory of games with incomplete information by assuming that agent j does not fully observe his type once it is realized. *Limited observability* is the assumption that agent j can choose to costlessly observe the realization of any finite number of the contingencies that define his type but not the entire type itself. His type can thus be regarded as information that he can access to the extent of his bounded ability. Except for this assumption, the agents are otherwise perfectly rational in the sense that they maximize their expected return conditional upon their finite observations.[1] Limited observability is formalized in section 3.[2]

We prove in section 3 that a contract between agents who are constrained by limited observability is necessarily finite. A finite contract in our paper is thus feasible and an infinite contract is an ideal. This is analogous to using an infinite number of traders to model perfect competition. Markets in reality never have an infinite number of traders and we are not proposing that the contingencies in an actual contracting problem are truly infinite. Rather, in the same way in which an infinity of traders avoids the quantification of how rapidly market power diminishes as market size increases, an infinity

[1]While it may not capture all aspects of a human being's capacity for observation, our model of the relationship between an agent and the world is motivated by interpreting the agent as an empirical researcher. The selection of a finite number of contingencies to observe is the design of an experiment and the realization of those contingencies is a data set. Data must necessarily be expressible as a sequence of *bits* (or binary digits) for scientific analysis (e.g., using statistics or a computer). Our assumption that an agent's type is a sequence of bits thus simply assumes that data about the state of the world is collected in the reduced form in which it must ultimately be presentable. This interpretation was suggested to us by Nabil Al-Najjar.

[2]The term "limited observability" is drawn from Radner (2000), which classifies the various approaches to bounded rationality in economic theory. Our approach is closest in spirit to his category of "costly rationality models" in that an agent in our model is rational in his acquisition and use of a finite number of bits of information, given that (i) there is no cost to observing any finite number of contingencies, and (ii) it is infinitely costly to fully observe his type. We thus follow the Savage approach of modeling bounded rationality as rationality subject to the costs of collecting and processing information. This cost function for observing contingencies, however, is not explicitly studied within the paper.

of contingencies models the unattainable complexity of addressing every contingency without the need for assumptions about the feasible length of a contract or the marginal costs of lengthening it.[3]

In the same sense that a perfectly competitive market is meaningful as an abstraction to the extent that it can be approximated by finite markets, an infinite contract is meaningful only if it is approached in the limit by a sequence of finite contracts. Our interest here is primarily in whether or not there are costs associated with contractual finiteness. A contract is *recordable* if its efficiency is matched or surpassed by the limiting performance of some sequence of finite contracts. We have selected the term "recordable" to suggest that the gist of the infinite contract can be written in a finite contract, with only details omitted, to whatever degree of accuracy is sought. If a contract is not recordable, then finite contracts are bounded away from the contract in the performance measure of the problem. Recordability is a rather weak requirement to impose on an infinite contract as a way of modeling bounds on contractual length. It is far less severe than the alternative of simply restricting attention to finite contracts.

We investigate recordability in section 5 in a generalization of the Chatterjee-Samuelson (1983) bilateral bargaining model. Reflecting the presence of incomplete information, we consider the recordability of those contracts that are optimal in the sense of maximizing the ex ante expected gains from trade subject to the constraints of incentive compatibility and interim individual rationality. The issue is whether or not such an optimal contract is recordable with the additional requirement that the finite contracts in the sequence converging to the optimal contract must also be incentive compatible and interim individual rational. Recordability is shown to depend crucially upon reversibility of contingencies. The following results are derived:[4]

1. An optimal contract is infinite.

2. If all but a finite number of contingencies are reversible, then an optimal contract is not recordable.

3. Conversely, if all contingencies are strongly irreversible, then an optimal contract is recordable.

Contractual incompleteness describes a situation in which a meaningful welfare loss occurs because a contract fails to address some contingencies. Result 2. is the most provocative because it describes a case in which contractual incompleteness arises endogenously in a contracting problem because of limited observability and properties of the agents' preferences.

Four examples are worked in the paper that concern the case in which agent j's type $\alpha_j = \left(a_{j,q}\right)_{q\in\mathbb{N}}$ affects his utility through a real value $v_j\left(\alpha_j\right)$ given by the formula $v_j\left(\alpha_j\right) = \Sigma_{q=1}^{\infty} a_{j,q}\delta^q$ for some $\delta \in (0, 1)$. All of the ideas of this paper are illustrated with this simple family of examples by varying the common ratio δ. These examples

[3]As emphasized in Radner (2000), determining the cost of acquiring information is a difficult empirical problem whose solution may depend greatly upon the context.

[4]We have proven similar results in Krasa and Williams (2001) concerning a principal-agent model. Though we have omitted this second example in the interests of brevity, it does suggests some generality to points 1-3.

are instructive because they illustrate some rather difficult ideas using little more than the formula for the sum a geometric series together with some arithmetic. We find it encouraging that our ideas arise in such a common family of formulas and do not require odd examples for the sake of illustration. None of the theorems in the paper, however, depend upon this special form of payoff function, and the issues that we raise are clearly not restricted to this particular family of examples.

Related Work. Our paper originates most directly in the work of Anderlini and Felli (1994, 1998), who grounded the theory of contractual incompleteness in the theory of computational complexity. The most fundamental idea that we draw from their paper is finiteness as a characteristic of real contracts and infiniteness as an ideal of contracting. There are three significant differences between our approach and the Anderlini-Felli model. The first is a matter of emphasis: incomplete information and incentive compatibility are of central interest in our paper but not in Anderlini and Felli's work. The second is a difference in modeling. We postulate a complex world outside of each agent that he cannot fully *observe*. Anderlini and Felli instead postulate a complex inner state of the agent (his valuation) that he knows but cannot fully *describe*.[5] Third, Anderlini and Felli consider only a fixed language for their finite contracts (the digits of the binary expansion of each agent's valuation). Our agents in contrast can write contracts in whatever language they wish; no contract is a priori excluded.[6] Instead, the constraints of incentive compatibility and limited observability are applied in this paper to deduce restrictions on contracts.

We assume limited observability in part to address a gap in the Anderlini-Felli approach that arises in addressing incomplete information and the consequent issue of revelation by agents. An agent in their model is incapable of answering a simple question about his valuation if the answer requires examining an infinite number of the digits that determine the binary expansion of his valuation. This is true even though (i) answering the question may only require a "yes" or "no" response, which is hardly complex, and (ii) the agent may know the answer by virtue of knowing his valuation.[7] We resolve this difficulty first by assuming that a trader does not know his valuation and second by allowing contracts to ask arbitrary questions of the agents in the sense that the language in which an agent is to respond is not fixed a priori in our paper. While operationally we come back to the constraint of finiteness of contracts as originally posed in Anderlini and Felli (1994), the issue here is the coherence of the story that supports this constraint.[8]

Similar to our approach, Segal (1999) models contracting within a complex environment by agents who are limited in their ability to describe the world. Our paper

[5] See Anderlini and Felli ((1994, p. 1115), (1998, sec. 3.2)).

[6] We do prove a result similar to the revelation principle (Theorem 2) that reduces the study of arbitrary contracts to a canonical form.

[7] For instance, an agent cannot respond to the question "Is your valuation irrational?" even though he knows the answer.

[8] Anderlini and Felli cite the familiar story of Justice Potter Stewart's assessment of obscenity as evidence of indescribility of a complex inner state: "He could not define obscenity, he had written, but 'I know it when I see it"' (Woodward and Armstrong (1979, p. 15-16). Our point here is that Stewart's inability to codify his definition of obscenity in legal terms did not prevent him from deciding whether or not a work was obscene and then reporting his decision.

and Segal's both use a complex environment as a key ingredient in modeling boundedly rational behavior. Segal (1999, p. 74) explains this point as follows:

> While much has been said about the role of bounded rationality in explaining contractual incompleteness, existing models have not been able to explain how people could be irrational enough not to be able to describe all of the possible contingencies ex ante, yet rational enough to foresee their payoffs ex ante and to describe any given contingency ex post (see e.g. Maskin and Tirole 1999). In our view, any attempt to model bounded rationality in a simple environment is doomed to fall into the trap of describing decision makers as either "completely dumb" or "perfectly rational". Neither is an attractive alternative for modeling "transaction costs". It is only in environments reflecting the real world's complexities that an intermediate region of "bounded rationality" emerges.

We believe that our use of $\{0, 1\}^{\mathbb{N}}$ as the set of states of the world together with the constraint of limited observability creates a model of human behavior that successfully lands within Segal's intermediate region.[9]

2 The Model

We consider two probability spaces $(A_1, \mathcal{A}_1, \pi_1)$ and $(A_2, \mathcal{A}_2, \pi_2)$ together with their product space $(A, \mathcal{A}, \pi)$. As in the Introduction, $A_j = \{0, 1\}^{\mathbb{N}}$ is agent j's *type space*. A *type* $\alpha_j \in A_j$ of agent j is written as $\alpha_j = \left(a_{j,q}\right)_{q\in\mathbb{N}}$. The set $A = A_1 \times A_2$ is the set of *states of the world* and π is the common prior of the two agents.

The σ-algebra $\mathcal{A}_j$ of measurable sets is defined with limited observability in mind. For $n \in \mathbb{N}$, the *initial string* $\alpha_{j,n-}$ and the *tail* $\alpha_{j,n+}$ determined by α_j and n are

$$\alpha_{j,n-} = \left(a_{j,q}\right)_{1\leq q<n} \textit{ and } \alpha_{j,n+} = \left(a_{j,q}\right)_{q>n},$$

respectively. Notice that the initial string and the tail determined by α_j and n omit the realization $a_{j,n}$ of the nth contingency. Let $A_{j,n-}$ denote the set of all initial strings of length $n-1$ and $A_{j,n+}$ the set of all tails from the $(n+1)$st contingency to infinity. It is essential for our purposes that cylinder sets of the form

$$\left\{a_{j,n-}\right\} \times A_{j,n-1+} \tag{1}$$

are measurable with respect to π_j so that probabilities are well-defined conditional on the observation by agent j of any initial string. We thus define $\mathcal{A}_j$ as the σ-algebra generated by all sets of the form (1).

Utility. Let $(C, \mathcal{C})$ be a measurable space. The set C is the *choice set*. Each agent j's utility $u_j(\alpha_j, c)$ is *quasilinear* in the sense that

$$u_j(\alpha_j, c) = h_j(c)v_j(\alpha_j) + t_j(c), \tag{2}$$

[9] A different tact to complexifying the state space can be found in Al-Najjar (2000). While his set of states is a subinterval of the real line (as in Anderlini and Felli (1994)), Al-Najjar allows measures on this interval that are only finitely additive. This permits functions that are not computable in the sense that they cannot be approximated by simple functions.

where $h_j(c)$, $v_j(\alpha_j)$, and $t_j(c)$ are real-valued functions. The function $v_j(\alpha_j)$ is agent j's *valuation function*, which is the part of his utility that is determined directly by his type. We thus consider an independent private value model in this paper. The function $t_j(c)$ is a monetary transfer to agent j and $h_j(c)$ is a level or portion of $v_j(\alpha)$ that agent j receives as a result of the choice c. It is assumed throughout the paper that $h_j(C)$ is bounded. The functions h_j and t_j are measurable on $(C, \mathcal{C})$. Let $v(\alpha)$ denote the *valuation mapping*

$$v(\alpha) = (v_1(\alpha_1), v_2(\alpha_2)).$$

Each agent j's valuation function v_j is assumed throughout the paper to satisfy the following condition:

$$\begin{aligned} \textit{for every } \varepsilon &> 0\textit{, there exists an } n \in \mathbb{N} \textit{ such that} \\ \left| v_j(\alpha_j) - v_j(\alpha'_j) \right| &< \varepsilon \textit{ for all } \alpha_j,\, \alpha'_j \in A_j \textit{ with } \alpha_{j,n+1-} = \alpha'_{j,n+1-}. \end{aligned} \tag{3}$$

All contingencies in a tail $\alpha_{j,n+}$ beyond the nth contingency are thus *details* in the sense that together they have only a minor effect on the value of v_j. As explained below, (3) is equivalent to continuity of v_j relative to a particular topology on A_j.[10] Our use of this condition is motivated below in subsection 2.1 as part of our discussion of recordability.

Topology. The product topology on $A_j = \{0, 1\}^{\mathbb{N}}$ in which $\{0, 1\}$ is assigned the discrete topology is useful at several points in this paper. Several of its properties are now noted:

1. The cylinder sets of the form (1) are a base for this topology.
2. The Tychonoff Theorem implies that A_j is compact relative to this topology.
3. Condition (3) on a valuation function v_j is equivalent to continuity of this function on A_j relative to this topology.[11]

Given property 3., we refer to condition (3) as *continuity* of v_j. Properties 1.-3. have two implications for a valuation function v_j. First, 1. and 3. together with the definition of $\mathcal{A}_j$ imply that v_j is measurable with respect to $(A_j, \mathcal{A}_j, \pi_j)$. Second, 2. and

[10] It also resembles continuity at infinity of payoffs in an infinitely repeated game when each $\alpha_{j,n-}$ is identified with a history in the game of a particular length (see, for instance, Fudenberg and Tirole (1991, Def. 4.1, p. 110)).

[11] Taking property 1. into account, v_j is continuous at α_j if for every $\varepsilon > 0$ there exists an $n(\alpha_j) \in \mathbb{N}$ such that $\left| v_j(\alpha_j) - v_j(\alpha'_j) \right| < \varepsilon$ for all $\alpha'_j \in A$ with $\alpha_{j,n(\alpha_j)+1-} = \alpha'_{j,n(\alpha_j)+1-}$. Condition (3) thus implies continuity at every $\alpha_j \in A_j$. The converse requires that a *single* value of $n \in \mathbb{N}$ exist with the above property for all $\alpha_j \in A_j$. Given property 2., the existence of such an n follows because a continuous function on a compact metrizable topological space is necessarily uniformly continouous. The metric

$$\left\| \alpha_j - \alpha'_j \right\| = \sum_{q=1}^{\infty} \left| a_{j,q} - a'_{j,q} \right| 2^{-q}$$

is an alternative way to define the product topology on A_j, and the desired value of n can be inferred from this formula.

3. imply that

$$v_j\left(A_j\right) \subset \left[\underline{v}_j, \overline{v}_j\right]$$

for some $\underline{v}_j$ $\overline{v}_j \in \mathbb{R}$. Let μ_j denote the induced probability distribution on $\left[\underline{v}_j, \overline{v}_j\right]$ defined by π_j and v_j.

Example 1 *Consider* $v_j\left(\alpha_j\right) = \sum_{q=1}^{\infty} a_{j,q}\delta^q$ *for* $\delta \in (0,1)$. *If* $\alpha_{j,n-} = \alpha'_{j,n-}$, *then*

$$\left|v_j(\alpha_j) - v_j(\alpha'_j)\right| = \left|\sum_{q=n}^{\infty}\left(a_{j,q} - a'_{j,q}\right)\delta^q\right| \leq \sum_{q=n}^{\infty}\delta^q = \frac{\delta^n}{1-\delta}, \tag{4}$$

from which it is clear that v_j *is continuous.*

Contracts. A *contract* is a measurable mapping $f : A \to C$. A contract f is *finite* if and only if there exists $n_1, n_2 \in \mathbb{N}$ such that $f(\alpha) = f(\alpha')$ for all $\alpha = (a_{1,q}, a_{2,q})_{q\in\mathbb{N}}$ and $\alpha' = (a'_{1,q}, a'_{2,q})_{q\in\mathbb{N}}$ with $\alpha_{i,n_i+1-} = \alpha'_{i,n_i+1-}$ for $i = 1, 2$. Otherwise, f is *infinite.* A finite contract $f(\alpha)$ is sometimes written as $f(\alpha_{1,n_1-},\alpha_{2,n_2-})$, reflecting the fact that $(\alpha_{1,n_1-},\alpha_{2,n_2-})$ determines $f(\alpha)$.

Letting α_j denote an observed type and α_j^* a reported type, a contract f is *incentive compatible* if and only if

$$E_{A_i}\left[u_j(\alpha_j, f(\alpha))\right] \geq E_{A_i}\left[u_j(\alpha_j, f(\alpha_j^*, \alpha_i))\right] \tag{5}$$

for $j = 1, 2$, $i \neq j$, and all $\alpha_j, \alpha_j^* \in A_j$. Define

$$H_j(\alpha_j^*) = E\left[h_j(f(\alpha)) \middle| \alpha_j = \alpha_j^*\right], \textit{ and} \tag{6}$$

$$T_j(\alpha_j^*) = E\left[t_j(f(\alpha)) \middle| \alpha_j = \alpha_j^*\right]. \tag{7}$$

Independence of types insures that $H_j(\alpha_j^*)$ and $T_j(\alpha_j^*)$ depend only upon the reported type α_j^* of agent j and not upon his observed type α_j. A contract f is thus incentive compatible if

$$H_j(\alpha_j)v_j(\alpha_j) + T_j(\alpha_j) \geq H_j\left(\alpha_j^*\right)v_j(\alpha_j) + T_j(\alpha_j^*) \tag{IC}$$

for $j = 1, 2$ and all $\alpha_j^*, \alpha_j \in A_j$. Let r_j denote agent j's reservation utility. The contract f is *interim individually rational* for agent j if

$$H_j\left(\alpha_j\right)v_j(\alpha_j) + T_j(\alpha_j) \geq r_j \tag{IIR}$$

for all $\alpha_j \in A_j$.

2.1 Recordability

Recordability indicates that an infinite contract reflects the potential of contracting in the sense that its performance can be approached or surpassed by a sequence of finite contracts. A precise definition is given in the context of each problem considered. Intuitively, a contract f is *recordable* if there exists a sequence of finite contracts $(f_q)_{q\in\mathbb{N}}$ such that

$$\lim_{q\to\infty} \left\| f_q \right\| \geq \| f \| , \tag{8}$$

where "$\|\cdot\|$" denotes a performance measure that is appropriate in the particular contracting problem. In the bilateral trade problem of section 5, for instance, $\| f \|$ is the ex ante expected gains from trade in the contract f. Each contract f_q in the sequence is also required to have properties appropriate to the problem, such as incentive compatibility and interim individually rationality. If f is not recordable, then all finite contracts are bounded below f according to the performance measure for the problem, i.e.,

$$\left\| f^* \right\| < \| f \| - \kappa ,$$

for some constant $\kappa > 0$ and any finite contract f^*. A contract f that is not recordable therefore overstates the performance potential of contracting. If f is an optimal contract, then *contractual incompleteness* necessarily occurs if f is not recordable.

To the extent that the perfomance measure addresses the welfare of the agents, continuity of each valuation function v_j is an obvious ingredient in proving recordability. If v_j is instead discontinuous, then the tail $\alpha_{j,n+}$ always has a nonnegligible effect upon agent j's valuation and hence also upon the performance measure, regardless of the value of n. Recordability is thus unlikely without continuity of the valuation functions and so it is assumed throughout the paper.[12] Conversely, the results in the paper that demonstrate contractual incompleteness are especially interesting because they are proven despite the continuity of the valuation functions.

2.2 Contracts and Mechanisms

A goal of this paper is to investigate how preferences determine whether or not contractual incompleteness necessarily occurs in the sense that an optimal contract is infinite but not recordable. Optimization has not been previously studied for contracts in which an agent's private information is an element of $\{0, 1\}^{\mathbb{N}}$. We address this issue in this subsection by connecting the theory of contracts as developed in this paper to the rich literature on optimal mechanisms. The main conclusion here is that standard results in mechanism design characterize the optimal contract in the bilateral trade model that we address in section 5.

[12] Continuity of the valuation functions can in fact be sufficient to insure recordability of optimal contracts in models with complete information. Theorem 1 of Krasa and Williams (2000) is a result of this kind. Anderlini and Felli (1998) presents continuity conditions that are sufficient for approximating an optimal principal-agent contract with a computable contract together with examples that illustrate how various kinds of discontinuities can prevent such an approximation. These conditions are distinct from our definition of continuity because they do not concern the function v_j.

A contract is a function $f : A \to C$ that determines a choice $f(\alpha)$ for each state α. Consistent with the terminology of mechanism design, a *mechanism* is a function $\hat{f} : v(A) \to C$ that determines a choice $f(v_1, v_2)$ for each pair of valuations $(v_1, v_2) \in v(A)$.[13] A mechanism $\hat{f}$ defines a contract f through composition with the valuation mapping v; a contract f defines a mechanism $\hat{f}$, however, only if it selects the same choice for all states that determine the same pair of valuations. The set of contracts is larger than the set of mechanisms, and an optimal contract may thus in principle surpass the performance of an optimal mechanism. Our objective in this subsection is to state conditions under which this does not happen, so that the optimal mechanism derived by standard methods characterizes the performance of the optimal contract.

Let

$$\Phi : C \times \left[\underline{v}_1, \overline{v}_1\right] \times \left[\underline{v}_2, \overline{v}_2\right] \to \mathbb{R}$$

be the *objective* of the contracting problem. In a principal-agent model, for instance, Φ is the principal's ex post payoff, and in a bilateral trading problem Φ is the ex post gains from trade. The *optimal contract problem* is

$$\max_{f} E_A \left[\Phi\left(f\left(\alpha\right), v(\alpha)\right)\right] \quad \textit{s.t. IC and IIR,}$$

and the *optimal mechanism problem* is

$$\max_{\hat{f}} E_{\left[\underline{v}_1, \overline{v}_1\right] \times \left[\underline{v}_2, \overline{v}_2\right]} \left[\Phi\left(\hat{f}\left(v\right), v\right)\right] \quad \textit{s.t. IC and IIR,}$$

where *IC* and *IIR* should be interpreted appropriately in the case of the optimal mechanism problem.

Theorem 9 states that if C is a convex subset of $\mathbb{R}^m$ and $\Phi\left(c, v_1, v_2\right)$ is concave in c for each v_1 and v_2, then an optimal mechanism $\hat{f}$ defines an optimal contract f through composition with the valuation mapping v. Conversely, given these conditions on C and Φ, an optimal contract f defines an optimal mechanism $\hat{f}$ through the formula $\hat{f} \circ v = g$, where the contract $g(\alpha)$ is defined by averaging $f(\alpha^*)$ over all states α^* for which $v(\alpha^*) = v(\alpha)$.[14] A formal statement and proof of Theorem 9 is in the Appendix. The bilateral trade model considered in this paper satisfies the hypotheses of Theorem 9, which provides us with the information we need concerning optimal contracts in this model.

3 Limited Observability

The purpose of this section is to justify our focus in the remainder of the paper upon incentive compatible finite contracts. This is accomplished by grounding these contracts

[13] In the broader literature, "contracts" and "mechanisms" are not distinguished as they are here by their domains; the words are used almost interchangeably, depending upon the subject of the model. The distinction we make here is purely for our expositional purposes.

[14] This converse reflects the common result of mechanism design that there are no gains in ex ante performance from introducing lotteries into the operation of a mechanism. The lottery in this case is the dependence of the choice $f(\alpha^*)$ upon α^* given that $v(\alpha^*) = v(\alpha)$. Given the assumptions of the theorem, ex ante expected performance can only improve by replacing each lottery over choices $\left\{f(\alpha^*) \,\middle|\, v(\alpha^*) = v(\alpha)\right\}$ with its certainty equivalent, which corresponds to a mechanism.

in an assumption concerning the limited abilities of the agents. *Limited observability by agent j of his type* is the assumption that agent j can choose to costlessly observe any initial string $\alpha_{j,n-}$ determined by his type α_j of arbitrary but finite length. There is no a priori bound on the number $n-1$ of contingencies that he may observe, and he may choose to observe different numbers of contingencies for different realizations of his type. The practical implication of this constraint is that any action or report that the agent takes conditional upon his type α_j must be determined by $\alpha_{j,n-}$ for some sufficiently large n.

Consider a game in which M_j is agent j's action set and $\eta : M_1 \times M_2 \to C$ is the outcome mapping. The strategy $\gamma_j : A_j \to M_j$ of agent j is *finite* if there exists $n \in \mathbb{N}$ such that

$$\text{if } \alpha_{j,n-} = \alpha^*_{j,n-}, \text{ then } \gamma_j\left(\alpha_j\right) = \gamma_j\left(\alpha^*_j\right) \tag{9}$$

for all $\alpha_j, \alpha^*_j \in A_j$. The first $n-1$ contingencies of α_j thus determines $\gamma_j\left(\alpha_j\right)$ for every $\alpha_j \in A_j$. We write $\gamma_j\left(\alpha_j\right) = \gamma_j\left(\alpha_{j,n-}\right)$ in this case.

An agent who is constrained by limited observability is capable of using any finite strategy. The following theorem goes further by proving that finite strategies are the only strategies that such an agent can use.

Theorem 1 *An agent who is constrained by limited observability can use a strategy if and only if the strategy is finite.*

The proofs of Theorem 1 and all other theorems are in the Appendix. The proof demonstrates that the constraint of using only initial strings to select actions implies the existence of a uniform length $n-1$ of initial string that is sufficient for selecting the action $\gamma_j\left(\alpha_j\right)$ for all α_j. Theorem 1 thus shows that limited observability is more severe as an ex ante constraint than as an interim constraint in the following sense: while an agent can base his choice of an action on as large of a finite number of contingencies of his realized type as he wishes, a well-defined strategy ex ante defines a single upper bound on the number of contingencies that affect the value of this strategy. Theorem 1 thus reveals the limitations on an agent's actions at the interim that are imposed by the requirement of coherently specifying those actions ex ante in a strategy.

If the action $\gamma_j\left(\alpha_j\right)$ is interpreted a signal of the agent's type α_j, then a finite strategy uses only a finite set of signals. Theorem 1 thus proves that limited observability implies *limited communication* in our model, i.e., an agent's language is finite in the sense that he uses only a finite number of messages.[15] The agent also conveys at most the finite number of contingencies given in the initial string $\alpha_{j,n-}$ that determines the value of $\gamma_j\left(\alpha_j\right)$. As Example 2 at the end of this subsection reveals, limited observability is a more restrictive constraint on an agent's abilities than limited communication. The theorem also implies that a contract implemented through a game is necessarily finite if the agents are constrained by limited observability.

Limited observability also influences how we define an equilibrium in a game.

[15] This behavioral constraint has been considered by Dow (1991), Meyer (1991), and Rubinstein ((1993),(1998, Chapter 5)).

Definition *A pair of strategies (γ_1, γ_2) in the game $(M_1 \times M_2, \eta)$ is a* ***Bayesian-Nash equilibrium with limited observability*** *if:*

1. *each agent's strategy is finite in the sense of (9);*
2. *for $j = 1, 2$ and assuming that γ_j depends only upon the first $n-1$ contingencies in α_j, the inequality*

$$E_A\left[u_j\left(\alpha_j, \eta\left(\gamma_j\left(\alpha^*_{j,n-}\right), \gamma_i(\alpha_i)\right)\right)\middle|\alpha_{j,q-} = \alpha^*_{j,q-}\right] \geq \qquad (10)$$
$$E_A\left[u_j\left(\alpha_j, \eta(m_j, \gamma_i(\alpha_i))\right)\middle|\alpha_{j,q-} = \alpha^*_{j,q-}\right]$$

holds for every $\alpha^*_j \in A_j$, $q \geq n$ and $m_j \in M_j$.

Condition 2 requires that $\gamma_j(\alpha^*_{j,n-})$ is a best response to γ_i conditional upon agent j knowing *any* initial string $\alpha^*_{j,q-}$ of length $q \geq n$ that agrees with $\alpha^*_{j,n-}$ on its first $n-1$ contingencies. There are two equilibrium conditions in (10), the first reflecting agent j's ability to freely choose his action in the game and the second reflecting his ability to condition that choice upon the observation of as many contingencies as he wishes. The first is the standard condition that agent j cannot profit by deviating from $\gamma_j(\alpha^*_{j,n-})$ if observes at least enough information $\alpha^*_{j,q-}$ as may be required to compute $\gamma_j(\alpha^*_{j,n-})$. The second is that agent j does not profit from observing additional contingencies $(a^*_{j,k})_{n \leq k < q}$ beyond $\alpha^*_{j,n-}$, for doing so never provides him with grounds for profitably deviating from $\gamma_j(\alpha^*_{j,n-})$. The second condition means that an equilibrium of this kind is not simply a Bayesian-Nash equilibrium for the case in which each agent observes only $n-1$ contingencies.[16] Unlike the standard definition of a Bayesian-Nash equilibrium, the expected values in (10) are computed not just with respect to the unknown value of the other agent's type α_i but also with respect to the unknown tail $\alpha_{j,q-1+}$ of contingencies that are not observed by agent j.

Restricting attention to incentive compatible finite contracts is grounded in the constraint of limited observability through the following theorem.

Theorem 2 *Suppose that each agent's valuation function is continuous. A contract f is implemented by a Bayesian-Nash equilibrium with limited observability in some game if and only if f is finite and incentive compatible in the classical sense of (5).*

Given continuity, incentive compatible finite contracts are thus exactly those that result when the agents are constrained by limited observability.

Limited observability also alters the constraint of interim individual rationality, though in a way that is easily seen to be inconsequential when each v_j is continuous. Suppose the finite contract f depends only upon the initial string $\alpha_{j,n-}$ of length $n-1$ observed by each agent. The contract f is *interim individually rational given limited observability* for agent j if

$$E_A\left[u_j\left(\alpha_j, f\left(\alpha^*_{j,n-}, \alpha_{i,n-}\right)\right)\middle|\alpha_{j,q-} = \alpha^*_{j,q-}\right] \geq r_j \qquad (11)$$

[16]Case 2 of Example 4 in section 5 illustrates how a Bayesian-Nash equilibrium may not be a Bayesian-Nash equilibrium with limited observability precisely because of this second equilibrium condition.

for all $q \geq n$ and $\alpha_j^* \in A_j$. In words, agent j's expected return conditional upon observing $\alpha_{j,q-}^*$ is at least his reservation utility r_j if he reports $\alpha_{j,n-}^*$, regardless of how many additional contingencies $(a_{j,k}^*)_{n\leq k<q}$ he may choose to observe beyond the $(n-1)$st. A contract that is interim individually rational clearly satisfies (11); conversely, if v_j is continuous, then (11) implies interim individual rationality. Imposing interim individual rationality on finite contracts is therefore consistent with limited observability.

We conclude this subsection with an example that emphasizes the impact of limited observability as a constraint on an agent's knowledge of his valuation. The reader should pay particular attention to cases 2 and 3 in this example, which show that an agent may not be able to place the most elementary of bounds upon his valuation if he is constrained by limited observability.

Example 2 *Agent j's valuation function is*

$$v_j(\alpha_j) = \frac{1-\delta}{\delta} \sum_{q=1}^{\infty} \delta^q a_{j,q}, \tag{12}$$

where δ will be selected below. The agent is constrained by limited observability. The formula for the sum of a geometric series along with some elementary analysis implies $v_j(A_j) = [0, 1]$ for $\delta \in [0.5, 1)$.[17] *Reflecting limited communication, suppose that agent j wishes only to announce whether his valuation is high (h) in the case of $v_j(\alpha_j) \geq 0.5$ or low (l) in the case of $v_j(\alpha_j) \leq 0.5$. The formula*

$$\frac{1-\delta}{\delta} \sum_{q=2}^{\infty} \delta^q a_{j,q} \leq \frac{1-\delta}{\delta} \sum_{q=2}^{\infty} \delta^q = \frac{1-\delta}{\delta} \cdot \frac{\delta^2}{1-\delta} = \delta \tag{13}$$

is helpful in this discussion.

Case 1: $\delta = 0.5$. *It is clear from (13) that $v_j(\alpha_j) \geq 0.5$ if $a_{j,1} = 1$ and $v(\alpha_j) \leq 0.5$ if $a_{j,1} = 0$. The strategy*

$$\gamma(\alpha_j) = \begin{cases} h \text{ if } a_{j,1} = 1 \\ l \text{ if } a_{j,1} = 0 \end{cases}$$

thus accurately communicates whether the valuation is high or low. This strategy is finite and is therefore compatible with limited observability.

Case 2: $\delta > 0.5$. *Formula (13) applies to show that*

$$\begin{aligned} v_j\left(\left\{a_{j,1} = 0\right\} \times A_{j,1+}\right) &= [0, \delta], \text{ and} \\ v_j\left(\left\{a_{j,1} = 1\right\} \times A_{j,1+}\right) &= [1-\delta, 1]. \end{aligned}$$

Both of these intervals contain 0.5 in their interiors, and so agent j cannot identify his valuation as high or low with certainty based upon $a_{j,1}$.

[17] The case of $\delta < 0.5$ is not considered in this example because $v_j(A_j) \subsetneq [0, 1]$ in this case, which complicates the discussion. This case is also omitted from Example 4 of section 5 because results that we invoke in this section to characterize optimal mechanisms require that $v_j(A_j)$ is a closed interval.

Case 3. *The problem observed in case 2 for* $\delta > 0.5$ *is also true of longer initial strings and alternatives to the above definitions of* h *and* l *as the language. For any* $k \geq 2$ *and any* $x \in (0, 1)$ *, agent* j *cannot determine with certainty whether* $v_j(\alpha_j) \leq x$ *or* $v_j(\alpha_j) \geq x$ *by observing* $\alpha_{j,k+1-}$ *for every choice of* $\alpha_{j,k+1-} \in A_{j,k+1-}$*. Formally, this means that for any* $x \in (0, 1)$ *there exists an* $\alpha^*_{j,k+1-} \in A_{j,k+1-}$ *such that*

$$x \in Int\left(v_j\left(\left\{\alpha^*_{j,k+1-}\right\} \times A_{j,k+}\right)\right),$$

where $v_j(\{\alpha^*_{j,k+1-}\} \times A_{j,k+})$ *is a closed interval. The proof is in the Appendix.*

4 Reversibility and Strong Irreversibility

Reversibility of contingency n is a property of agent j's valuation function v_j that can make incentive compatible revelation of $a_{j,n}$ bind as a constraint in the design of a finite contract f. Conversely, strong irreversibility of contingency n is a property of v_j that can insure that incentive compatible revelation of $a_{j,n}$ does not constrain the design of f. Reversibility and strong irreversibility are thus useful in determining which contingencies can be addressed by incentive compatible finite contracts. These properties will be shown in subsequent sections to be significant in determining whether or not an optimal contract is recordable.

Definition *Contingency* n *is* ***reversible*** *for agent* j *if for any initial string* $\alpha_{j,n-}$ *one can select at least two pairs of tails* $(\underline{\alpha}^1_{j,n+}, \overline{\alpha}^1_{j,n+})$, $(\underline{\alpha}^2_{j,n+}, \overline{\alpha}^2_{j,n+})$ *so that the following two properties hold.*

1. *Each pair of tails* $(\underline{\alpha}^k_{j,n+}, \overline{\alpha}^k_{j,n+})$ *perfectly reverses the effect upon agent* j*'s valuation of contingency* n*: for* $k = 1, 2$*,*

$$v_j(\alpha_{j,n-}, 0, \underline{\alpha}^k_{j,n+}) = v_j(\alpha_{j,n-}, 1, \overline{\alpha}^k_{j,n+}). \tag{14}$$

2. *The pairs of tails differ in their effects upon agent* j*'s valuation:*

$$v_j(\alpha_{j,n-}, 0, \underline{\alpha}^1_{j,n+}) \neq v_j(\alpha_{j,n-}, 0, \underline{\alpha}^2_{j,n+}). \tag{15}$$

Reversibility of contingency n is depicted in Figure 1 and the possibility that multiple types might determine the same valuation is illustrated in Example 2 of the last section. This definition pinpoints a problem in incentive compatibility for finite contracts. To illustrate this point, consider an incentive compatible and finite contract f that is determined by the initial strings of length n. Suppose that the nth contingency observed by agent j is reversible. Agent j's interim expected utility given his type α_j is

$$H_j(\alpha_{j,n+1-})v_j(\alpha_j) + T_j(\alpha_{j,n+1-}),$$

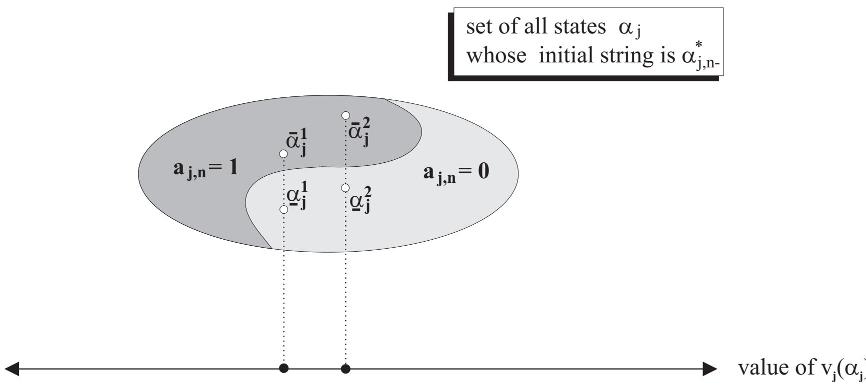

Figure 1: Figure 1: Reversibility of Contingency n for Agent j.

where the tail $\alpha_{j,n+}$ beyond the nth contingency is omitted from $H_j(\alpha_j)$ and $T_j(\alpha_j)$ because f does not depend upon it. Because $v_j(\alpha_{j,n-}, 0, \underline{\alpha}^k_{j,n+}) = v_j(\alpha_{j,n-}, 1, \overline{\alpha}^k_{j,n+})$, it must be true that

$$H_j(\alpha_{j,n-}, 0) v_j(\alpha_{j,n-}, 0, \underline{\alpha}^k_{j,n+}) + T_j(\alpha_{j,n-}, 0) = \\ H_j(\alpha_{j,n-}, 1) v_j(\alpha_{j,n-}, 1, \overline{\alpha}^k_{j,n+}) + T_j(\alpha_{j,n-}, 1), \tag{16}$$

or else agent j with type equal to either $(\alpha_{j,n-}, 0, \underline{\alpha}^k_{j,n+})$ or $(\alpha_{j,n-}, 1, \underline{\alpha}^k_{j,n+})$ would report whichever of these two types produced the larger of the two sides of this equation. This would contradict incentive compatibility for one of these two types. Now let $v \in \mathbb{R}$ denote a variable and consider the equation

$$\left[H_j(\alpha_{j,n-}, 0) - H_j(\alpha_{j,n-}, 1)\right] \cdot v = T_j(\alpha_{j,n-}, 1) - T_j(\alpha_{j,n-}, 0). \tag{17}$$

Statements (15) and (16) imply that (17) holds for distinct values of v, from which we conclude that

$$H_j(\alpha_{j,n-}, 0) = H_j(\alpha_{j,n-}, 1) \;\; and \;\; T_j(\alpha_{j,n-}, 1) = T_j(\alpha_{j,n-}, 0)$$

for all $\alpha_{j,n-}$. Given our assumptions here on f, the two conditions (14) and (15) that define reversibility place conflicting incentive constraints upon an agent's interim expected utility function. The result is that the two functions H_j and T_j that capture the effect of the contract f upon interim expected utility cannot depend upon the reversible contingency. The reversible contingency can thus affect interim expected utility through agent j's valuation v_j but not through its effect upon the contract f.

The above discussion shows that an incentive compatible finite contract f may be constrained in how it depends upon a reversible contingency. As will be shown, this constraint upon f can cause inefficiency. The following theorem presents a slightly different scenario in which the same conclusion holds. With an eye towards our later results, the assumption that f does not depend upon $\alpha_{j,n+}$ is replaced in this theorem with the assumption that each contingency in the tail is reversible.

Theorem 3 *For some $n \in \mathbb{N}$, suppose that every contingency observed by agent j after the nth is reversible. If the contract f is incentive compatible and finite, then the functions $H_j(\alpha_j)$ and $T_j(\alpha_j)$ determined by f depend only upon the first n contingencies. Consequently, agent j's interim expected utility in the contract f depends upon the tail $\alpha_{j,n+}$ only through its effect upon his valuation function $v_j(\alpha_j)$.*

The proof is straightforward. Suppose that f does not depend on any contingency observed by agent j after the qth for some $q > n$. The argument above shows that $H_j(\alpha_j)$ and $T_j(\alpha_j)$ cannot depend upon $a_{j,q}$. The theorem then follows by backwards induction.

Theorem 3 begins our efforts to show how reversibility of a contingency constrains contract design, an idea that we will explore more deeply in the context of the bilateral trade model that follows. It is common in mechanism design to show that an incentive compatible mechanism can be constructed in a particular problem only by appropriately building inefficiency into the collective choice. As suggested by the above discussion, we will show that reversibility in these models can force the agents to design inefficiency into their contract by making it insensitive to certain contingencies as a means of achieving incentive compatibility.

Definition *Contingency n is* ***strongly irreversible*** *for agent j if for every initial string $\alpha_{j,n-}$ there do not exist two pairs of tails $(\underline{\alpha}^1_{j,n+}, \overline{\alpha}^1_{j,n+})$, $(\underline{\alpha}^2_{j,n+}, \overline{\alpha}^2_{j,n+})$ that satisfy (14) and (15).*

"Strongly irreversible" is more demanding than "irreversible": contingency n is irreversible if there is *at least one* initial string $\alpha_{j,n-}$ for which no pair of tails exists satisfying (14) and (15), while contingency n is strongly irreversible if no such pair exists for *any* initial string $\alpha_{j,n-}$. We define "strongly irreversible" as above because it is useful in this form as a sufficient condition for proving that an optimal contract is recordable. Strong irreversibility is interpreted after the following example.

Example 3 *Let $v_j(\alpha) = \sum_{q=1}^{\infty} \delta^q a_{j,q}$ for $\delta \in (0, 1)$. We show in this example that each of the properties of strong irreversibility and reversibility of contingency n holds for an interval of values of $\delta \in (0, 1)$. This supports the hypothesis that neither of these two properties of v_j is degenerate in our general model of contracting. Let $\alpha_{j,n-}$ be any initial string. The inequality*

$$\begin{aligned} & v_j(\alpha_{j,n-}, 0, \underline{\alpha}_{j,n+}) - v_j(\alpha_{j,n-}, 1, \overline{\alpha}_{j,n+}) \\ & \leq v_j(\alpha_{j,n-}, 0, 1, 1, \dots) - v_j(\alpha_{j,n-}, 1, 0, 0, \dots) \qquad (18) \\ & = \delta^n \left(\frac{2\delta - 1}{1 - \delta} \right) \qquad (19) \end{aligned}$$

holds for any two tails $\underline{\alpha}_{j,n+}, \overline{\alpha}_{j,n+}$. Reversibility or strong irreversibility of contingency n depends upon whether δ is at most or exceeds 0.5.

Case 1: $\delta \in (0, 0.5]$. *Each contingency is strongly irreversible. For arbitrary $\alpha_{j,n-}$ and $\delta = 0.5$, the right side of (19) equals* 0 *and (18) is strict except when $\underline{\alpha}_{j,n+} = (1, 1, \dots)$ and $\overline{\alpha}_{j,n+} = (0, 0, \dots)$. Condition (14) in the definition of reversibility thus* *holds only for this one choice of* $\underline{\alpha}_{j,n+}$ *and* $\overline{\alpha}_{j,n+}$. *For* $\delta \in (0, 0.5)$ *the right side of (19) is negative, which means that (14) can never hold.*

Case 2: $\delta \in (0.5, 1)$. *Each contingency is reversible. The right side of (19) is positive in this case. It is also true that*

$$v_j(\alpha_{j,n-}, 0, \underline{\alpha}_{j,n+}) - v_j(\alpha_{j,n-}, 1, \underline{\alpha}_{j,n+}) = -\delta^n.$$

Holding the initial string $\alpha_{j,n-}$ *constant,* $v_j(\alpha_{j,n-}, 0, \underline{\alpha}_{j,n+}) - v_j(\alpha_{j,n-}, 1, \overline{\alpha}_{j,n+})$ *covers an interval on the real line that contains* 0 *in its interior as* $\overline{\alpha}_{j,n+}$ *is varied. Choose* $\underline{\alpha}^1_{j,n+}$ *and* $\underline{\alpha}^2_{j,n+}$ *so that (15) holds. For each of these two tails* $\underline{\alpha}^k_{j,n+}$*, there exists a corresponding* $\overline{\alpha}^k_{j,n+}$ *so that (14) is satisfied, which verifies the reversibility of contingency* n.

Theorem 4 shows that strong irreversibility of each contingency implies an "orderliness" of an agent's valuation function with respect to initial strings. As explained below, this property facilitates the proof of recordability. Suppose that $A = \{0, 1\}^{\mathbb{N}}$ and $v_j(A_j) = [\underline{v}_j, \overline{v}_j]$. If every contingency is strongly irreversible and v_j is continuous, then Theorem 4 asserts the existence of a set of points

$$\underline{v}_j = x_{j,1} \leq x_{j,2} \leq \ldots \leq x_{j,2^{n-1}+1} = \overline{v}_j \tag{20}$$

such that the set of valuations determined by all types that share a particular initial string $\alpha_{j,n-}$ satisfies

$$v_j\left(\left\{\alpha_{j,n-}\right\} \times A_{j,n-1+}\right) = \left[x_{j,k}, x_{j,k+1}\right]$$

for some $1 \leq k \leq 2^{n-1}$. This is illustrated for $n = 2$ in case 1 of Example 2, where $0 = \underline{v}_j = x_{j,1}$, $x_{j,2} = 0.5$, and $x_{j,3} = 1$. As illustrated by cases 2 and 3 of Example 2, $[\underline{v}_j, \overline{v}_j]$ is not partitioned according to the initial strings of a given length if contingencies are reversible: reversibility instead implies that the sets of valuations determined by distinct initial strings of the same length intersect nontrivially.

Theorem 4 *Assume that every contingency observed by agent* j *is strongly irreversible, that his valuation function* v_j *is continuous, and that* $v_j(A_j) = [\underline{v}_j, \overline{v}_j]$. *For any initial string* $\alpha_{j,n-}$*, let*

$$D_{\alpha_{j,n-}} = v_j\left(\left\{\alpha_{j,n-}\right\} \times A_{j,n-1+}\right).$$

Then the following statements hold.

1. *Each set* $D_{\alpha_{j,n-}}$ *is a closed interval.*
2. *If* $\alpha_{j,n-} \neq \alpha'_{j,n-}$*, then* $D_{\alpha_{j,n-}} \cap D_{\alpha'_{j,n-}}$ *contains at most one point.*

Theorem 4 plays the following role in proving the recordability of a contract. Applying this theorem, strong irreversibility of contingencies implies a one-to-one correspondence between (i) finite contracts f that are determined by the initial strings of length $n-1$ observed by the two agents and (ii) mechanisms $\hat{f}$ that are constant on the rectangles of the form $\left[x_{1,k}, x_{1,k+1}\right] \times \left[x_{2,q}, x_{2,q+1}\right]$ as defined in (20). The mechanism $\hat{f}$ is defined by step functions whose discontinuities occur on the edges of the rectangles and the correspondence is through the composition $f = \hat{f} \circ v$. Continuity of v_j implies that the intervals in (20) define an increasingly fine partition of $[\underline{v}_j, \overline{v}_j]$ as n increases. Any given mechanism on $[\underline{v}_1, \overline{v}_1] \times [\underline{v}_2, \overline{v}_2]$ that satisfies certain regularity properties can thus be approximated arbitrarily closely by a step function mechanism $\hat{f}$ of this kind. It is therefore approximated by the corresponding finite contract f. This is the insight that underlies our construction of sequences of finite contracts to prove the recordability of optimal contracts in Theorem 6 below.

5 Reversibility in a Model of Bilateral Trade

A seller can provide a service to a buyer. The state of the world specifies every detail that affects the value of the service to the buyer and the cost of provision to the seller. The buyer's and the seller's types are $\alpha_B = (a_{B,q})_{q \in \mathbb{N}} \in A_B$ and $\alpha_S = (a_{S,q})_{q \in \mathbb{N}} \in A_S$, respectively, with π_B and π_S denoting the distributions of these types. A *contract* is a pair (p, t) that specifies for each α_B and α_S a probability $p(\alpha_B, \alpha_S)$ that the seller provides the service to the buyer and a transfer $t(\alpha_B, \alpha_S)$ from the buyer to the seller. Contracts are thus assumed to be ex post budget balanced throughout this discussion. The buyer's utility is

$$u_B(\alpha_B, p, t) = p \cdot v_B(\alpha_B) - t$$

and the seller's utility is

$$u_S(\alpha_S, p, t) = t - p \cdot v_S(\alpha_S),$$

where $v_B : A_B \to [\underline{v}_B, \overline{v}_B] \subset \mathbb{R}^+$ and $v_S : A_S \to [\underline{v}_S, \overline{v}_S] \subset \mathbb{R}^+$. A contract is required to be incentive compatible and interim individually rational given that each trader's default utility is 0. It is assumed throughout this section that any initial string observable by either trader occurs with positive probability: for $j = B, S$,

$$\pi_j\left(\alpha_{j,n-}\right) = \pi_j\left(\left\{\alpha_{j,n-}\right\} \times A_{j,n-1+}\right) > 0 \tag{21}$$

for any $n \in \mathbb{N}$ and $\alpha_{j,n-} \in A_{j,n-}$.

Let μ_B denote the distribution of the buyer's valuation defined by v_B and π_B, and let μ_S denote the distribution of the seller's valuation defined by v_S and π_S. The model of bilateral trade of Chatterjee and Samuelson (1983) is a special case of our model in which the densities μ'_B and μ'_S are continuous and nonzero on $[\underline{v}_B, \overline{v}_B]$ and $[\underline{v}_S, \overline{v}_S]$, respectively. Our approach extends their model by modeling as states of the world those aspects of the service or good that determine the payoffs from trading. We thus addresses a weakness of noncooperative bargaining theory, which is that bargainers negotiate only price and perhaps quantity in most models, whereas real bargaining problems typically concern a multitude of issues.

The *optimal contract* (p^*, t^*) maximizes the expected gains from trade subject to *IC* and *IIR*: (p^*, t^*) solves

$$\max_{(p,t)} \int_A [v_B(\alpha_B) - v_S(\alpha_S)]\, p(\alpha) d\pi\,(\alpha) \quad \textit{s.t. IC and IIR.} \tag{22}$$

As discussed in section 2.2, Theorem 9 reduces the problem of designing optimal contracts to the problem of designing optimal mechanisms. The optimal mechanism is characterized in Myerson and Satterthwaite (1983).

Reflecting the objective in (22) of maximizing the expected gains from trade, a contract (p, t) is *recordable* if there exists a sequence of finite contracts $(p_m, t_m)_{m\in\mathbb{N}}$ satisfying *IC* and *IIR* such that

$$\begin{aligned} \lim_{m\to\infty} \int_A [v_B(\alpha_B) - v_S(\alpha_S)]\, p_m(\alpha) d\pi\,(\alpha) \geq \\ \int_A [v_B(\alpha_B) - v_S(\alpha_S)]\, p(\alpha) d\pi\,(\alpha)\,. \end{aligned} \tag{23}$$

Suppose that all but a finite number of contingencies observed by each trader are reversible. Given a recordable contract, it is shown in the proof of Theorem 5 that an incentive compatible, interim individually rational finite contract exists that achieves at least as much of the potential gains from trade as the given recordable contract. It follows that optimal contracts are not recordable in problems in which only infinite contracts can be optimal. The Myerson and Satterthwaite (1983) characterization suggests that optimal contracts are typically infinite. Theorem 5 in this sense provides sufficient conditions for contractual incompleteness.

Theorem 5 *For some $b, s \in \mathbb{N}$, suppose that all contingencies $a_{B,n}$ with $n \geq b$ and $a_{S,n}$ with $n \geq s$ are reversible for the trader who observes its realization. If only infinite contracts can be optimal in the given problem, then an optimal contract is not recordable.*

The next theorem concerns a case in which each contingency is strongly irreversible and the optimal contract is both infinite and recordable. This theorem applies the Myerson-Satterthwaite characterization of the optimal mechanism, which requires that μ_B and μ_S are *regular* in the sense that: (i) the densities μ'_B and μ'_S exist, are continuous, and have [0, 1] as their common support; (ii) the *virtual valuation functions*

$$V_B(v_B, k) = v_B + k\frac{\mu_B(v_B) - 1}{\mu'_B(v_B)}, \textit{ and} \tag{24}$$

$$V_S(v_S, k) = v_S + k\frac{\mu_S(v_S)}{\mu'_S(v_S)} \tag{25}$$

are increasing on [0, 1] for each $k \in [0, 1]$.

Theorem 6 *Assume that:*

1. *each contingency observed by a trader is strongly irreversible;*

2. *the valuation functions v_B and v_S are continuous;*

3. *the distributions of valuations μ_B and μ_S are regular in the sense defined above.*

Then an optimal contract (p^*, t^*) is recordable.

The sequence of contracts $((p_n, t_n))_{n\in\mathbb{N}}$ that demonstrates the recordability of the optimal contract (p^*, t^*) in the proof of Theorem 6 satisfies $\lim_{n\to\infty} p_n(\alpha) = p^*(\alpha)$ except on a set of states of π-measure zero. Finite contracts are thus shown to not only approximate the expected gains from trade of the optimal contract (p^*, t^*) but also to approximate the trades of p^*.

Our final theorem applies Theorem 5 to reveal the severe contractual incompleteness when (i) every contingency is reversible, and (ii) it is not common knowledge ex ante that trade should occur. It is shown that gains from trade cannot be contracted ex ante in a recordable contract that is incentive compatible and interim individually rational.

Theorem 7 *Assume that:*

1. *each contingency observed by a trader is reversible;*

2. *there exists types α_B^* and α_S^* for the buyer and the seller such that $v_B(\alpha_B^*) < v_S(\alpha_S^*)$, so that trade should not occur for these types.*

Then the ex ante expected gains from trade are zero in any recordable, incentive compatible, and interim individually rational contract.

Two points should be emphasized about this theorem. First, it concerns a special case of our model. In contrast, Theorem 6 identifies an alternative case in which the loss in ex ante performance can be made arbitrarily small. The second point concerns the way in which Theorem 7 appears in familiar bargaining games. Recall that Theorem 2 identifies Bayesian-Nash equilibria with limited observability of arbitrary games with finite incentive compatible contracts. Given reversible contingencies, Theorem 7 implies that there are no ex ante gains from trade in any Bayesian-Nash equilibrium with limited observability in any game. The practical implication of this in most games is nonexistence of equilibrium, or existence only of a trivial no-trade equilibrium. This is illustrated in the example that follows. Nonexistence of equilibrium models the impossibility of coherently specifying equilibrium behavior ex ante across all states that may arise at the interim, which is the essential task of a contract in our approach. We thus take nonexistence in this case as affirming Theorem 7's assertion that gains from trade cannot be contracted ex ante when contingencies are reversible.

Example 4 *Assume that $\left[\underline{v}_B, \overline{v}_B\right] = \left[\underline{v}_S, \overline{v}_S\right] = [0, 1]$ and fix a price $\rho \in (0, 1)$.*[18] *The game operates as follows: with access to his type, each trader announces "yes" or "no", with trade occurring at the price of ρ if and only if each trader announces "yes". This is inspired by the fixed price game of Hagerty and Rogerson (1985),*

[18]This example was suggested by a question from Jim Peck.

with the minor change here that each trader announces "yes" or "no" rather than his valuation so that communication is limited. If each trader knows his valuation at the interim, then it is a dominant strategy for each trader to announce "yes" if and only if he can profitably trade at the price of ρ. *Trade occurs in this equilibrium for* (v_B, v_S) *satisfying* $v_S \leq \rho \leq v_B$ *and so there are gains from trade.*

Let

$$v_B = \frac{1-\delta}{\delta}\sum_{q=1}^{\infty}\delta^q a_{B,q} \text{ and } v_S = \frac{1-\delta}{\delta}\sum_{q=1}^{\infty}\delta^q a_{S,q}$$

for $\delta \in [0.5, 1)$. *Case 1 below assumes that all contingencies are strongly irreversible for each trader* ($\delta = 0.5$). *It is shown that the dominant strategy equilibrium for the fixed price of* $\rho = 0.5$ *defines a Bayesian-Nash equilibrium with limited observability. Case 2 assumes that all contingencies are reversible for each trader* ($\delta \in (0.5, 1)$). *It is shown in this case that the dominant strategy equilibrium determined by* ρ *does not define a Bayesian-Nash equilibrium with limited observability, regardless of the choice of* ρ. *As an illustration of the no-trade result of Theorem 7, it is then shown that the outcome any Bayesian-Nash equilibrium with limited observability in the case of* $\delta \in (0.5, 1)$ *is the same as a particular no-trade equilibrium in almost all states* α.

Case 1: $\delta = 0.5$. *Fix the price at* $\rho = 0.5$. *The buyer knows whether or not* v_B *is as big or as small as* ρ *based solely upon observing his first contingency* $a_{B,1}$. *Formally,*

$$\begin{aligned} a_{B,1} &= 0 \Rightarrow v_B \leq 0.5, \text{ and} \\ a_{B,1} &= 1 \Rightarrow v_B \geq 0.5. \end{aligned}$$

The strategy

$$\gamma_B(\alpha_B) = \begin{cases} \text{no if } a_{B,1} = 0 \\ \text{yes if } a_{B,1} = 1 \end{cases}$$

is finite and maximizes the buyer's expected payoff against any strategy of the seller, regardless of how long of an initial string the buyer may observe. Similar remarks apply to the seller. The trading outcome

$$p(v_B, v_S) = \begin{cases} 1 \text{ if } v_S \leq 0.5 \leq v_B \\ 0 \text{ otherwise} \end{cases}$$

with a transfer of 0.5 *if and only if trade occurs is therefore sustainable as a Bayesian-Nash equilibrium with limited observability. The corresponding contract*

$$\begin{aligned} p^*(\alpha_B, \alpha_S) &= \begin{cases} 1 \text{ if } a_{S,1} = 0 \text{ and } a_{B,1} = 1 \\ 0 \text{ otherwise} \end{cases}, \text{ and} \qquad (26) \\ t^*(\alpha_B, \alpha_S) &= \begin{cases} 0.5 \text{ if } a_{S,1} = 0 \text{ and } a_{B,1} = 1 \\ 0 \text{ otherwise} \end{cases} \end{aligned}$$

is incentive compatible, interim individually rational and finite.

Case 2: $\delta \in (0.5, 1)$. *Let the price ρ be any fixed value in $(0, 1)$. Case 3 of Example 2 implies that each contingency observed by a trader is reversible, and hence Theorem 7 together with Theorem 2 imply that the expected gains from trade are zero in any Bayesian-Nash equilibrium with limited observability. A trader can of course determine conditional upon a finite string of length $n-1$ whether his expected valuation is no more than or at least ρ. At first glance, the finite strategies*

$$\begin{aligned}\gamma_B(\alpha_B) &= \begin{cases} \text{no if } E\left[v_B \left|\alpha_{B,n-}\right.\right] < \rho \\ \text{yes if } E\left[v_B \left|\alpha_{B,n-}\right.\right] \geq \rho \end{cases}, \text{ and} \\ \gamma_S(\alpha_S) &= \begin{cases} \text{yes if } E\left[v_S \left|\alpha_{S,n-}\right.\right] \leq \rho \\ \text{no if } E\left[v_S \left|\alpha_{S,n-}\right.\right] > \rho \end{cases}\end{aligned} \tag{27}$$

may seem to sustain positive expected gains from trade in equilibrium. These strategies, however, do not define a Bayesian-Nash equilibrium with limited observability because a trader may wish to deviate from his specified message after observing a finite number of additional contingencies beyond the $(n-1)$st. This can be formalized as follows. It is shown in the analysis of case 3 of Example 2 that there exists $\alpha^*_{B,n-}$ such that $v_B(\{\alpha^*_{B,n-}\} \times A_{B,n+})$ is a closed interval with ρ in its interior. Consequently, there exists $q > n$ and $\alpha'_B, \alpha''_B \in A_B$ such that

$$\alpha'_{B,n-} = \alpha''_{B,n-} = \alpha^*_{B,n-},$$

but

$$\begin{aligned}E\left[v_B(\alpha_B) \left|\alpha_{B,q-} = \alpha'_{B,q-}\right.\right] &< \rho, \text{ and} \\ E\left[v_B(\alpha_B) \left|\alpha_{B,q-} = \alpha''_{B,q-}\right.\right] &> \rho.\end{aligned}$$

Given the seller's use of γ_S, the buyer's unique best response is to report "no" if he observes $\alpha'_{B,q-}$ and "yes" if he observes $\alpha''_{B,q-}$. The single report $\gamma_B(\alpha'_B) = \gamma_B(\alpha''_B) = \gamma_B(\alpha^*_B)$ is not a best response in each of these two instances, which means that (γ_B, γ_S) is not a Bayesian-Nash equilibrium with limited observability.

The strategies γ^*_B, γ^*_S in which each trader reports "no" regardless of this type define a Bayesian-Nash equilibrium with limited observability in which trade never occurs. An equilibrium thus exists for every $\delta \in (0.5, 1)$ and $\rho \in (0, 1)$. We now show that (γ^*_B, γ^*_S) is the only equilibrium, except for trivial variations of γ^*_B and γ^*_S over sets of measure zero. The argument is by contradiction. Suppose that (γ'_B, γ'_S) is a Bayesian-Nash equilibrium with limited observability that differs from (γ^*_B, γ^*_S) on some set of states of positive π-measure. This means that at least one of the traders reports "yes" with positive probability in the equilibrium (γ'_B, γ'_S). Suppose for notational convenience that this is true of the seller; the contradiction is derived by considering the buyer's best response to γ'_S. The finiteness of the strategies in a Bayesian-Nash equilibrium with limited observability implies that there exists $n \in \mathbb{N}$ such that $\gamma'_B(\alpha_B)$ is determined by $\alpha_{B,n-}$ for all $\alpha_{B,n-} \in A_{B,n-}$. Because the seller's use of γ'_S presents the buyer with the positive probability of profitable trade, $\gamma'_B(\alpha_B)$ must *equal $\gamma_B(\alpha_B)$ as defined in (27). The argument made above now provides the desired contradiction: for the initial strings $\alpha'_{B,q-}$, $\alpha''_{B,q-}$ and $\alpha^*_{B,n-}$ defined above, the buyer would deviate from $\gamma'_B\left(\alpha^*_{Bn-}\right)$ upon observing one of $\alpha'_{B,q-}$ or $\alpha''_{B,q-}$. The pair (γ'_B, γ'_S) is thus not a Bayesian-Nash equilibrium with limited observability, which completes the contradiction.*

Concluding Remarks on Bilateral Trade. We conclude by connecting our analysis to the broader aims of models of incomplete contracts. Results that show that agents may accomplish more in a later stage than they can achieve ex ante in a contract are desirable in a model of contractual incompleteness, first because they demonstrate the nonnegligible losses that make contractual incompleteness compelling as an issue, and second because they demonstrate the activity in the interim and ex post stages through which the agents try to accomplish what they failed to achieve ex ante. Renegotiation exemplifies this later activity. That such activity occurs in reality and is inconsistent with complete contracts is the principle issue that motivates the theory of contractual incompleteness. We now discuss the possibility that the buyer and the seller may trade at an interim state despite the absence of an ex ante contract that organizes their trading. Consider the fixed price game of Example 4 with the price of ρ and assume that specific types α^*_B and α^*_S of the traders are realized. Unless $v_B\left(\alpha^*_B\right)$ or $v_S\left(\alpha^*_S\right)$ equals ρ, continuity of each trader's valuation function insures that each trader can determine after observing a sufficiently large but finite number of contingencies whether his valuation is more or less than ρ. If each trader uses his dominant strategy of price-taking, then trade occurs if $v_B\left(\alpha^*_B\right) > \rho > v_S\left(\alpha^*_S\right)$.

It is thus arguable that profitable trade may occur in an interim state (α^*_B, α^*_S) even in the case of reversible contingencies in which gains from trade can not be contracted ex ante.[19] This stands in stark contrast to the "no information-based trading in equilibrium" result of Milgrom and Stokey (1982) or, more generally, the result that ex ante efficiency implies interim efficiency of Holmström and Myerson (1983, p. 1806). Both of these classic results depend crucially upon the assumption that complete contracts are available to the agents ex ante. The primary difference between our results and these results is our constraint of limited observability and its implication that contracts are necessarily finite. As captured by Theorem 1, the fact that limited observability is more severe as a constraint ex ante than at the interim explains why ex ante performance may be worse in our model than interim performance.

[19] Two cautionary points, however, should be noted about the prospect of trade at the interim. First, statements about the possibility of trading across all interim states are problematic in the case of reversible contingencies for precisely the same reasons that an ex ante contract can not arrange gains from trade in this case. Second, unlike the fixed-price game above, a trader may not have a dominant strategy in an arbitrary game in a particular state (α^*_B, α^*_S); his choice at the interim would therefore depend upon the strategy of his opponent. The specification of such a strategy leads back through Theorem 1 to the inefficiency result of Theorem 7 in the case of reversible contingencies.

6 Conclusion

Contractual incompleteness is inefficiency of a nonnegligible magnitude that occurs because a contract fails to address some contingencies. We develop a model in which all contingencies are foreseen by the agents in the sense that they know the structure of the model and yet contractual incompleteness may necessarily occur. The causes of contractual incompleteness in our model are: (i) limited observability, which is the inability of agents to observe more than a finite number of contingencies; (ii) a set of states of the world that is complex in the sense that there are an infinite number of possible contingencies; (iii) incomplete information and the consequent problem of incentive compatibility; (iv) reversibility of contingencies, which is a property of an agent's preferences over states of the world. Conversely, we identify strong irreversibility as a property of preferences under which contractual incompleteness need not occur in the context of (i)-(iii) in the sense that the optimal contract is recordable.

We emphasize in this paper the properties of preferences over states of the world that determine whether or not contractual incompleteness must occur. Our results identify attributes of contingencies that determine whether or not those contingencies can be successfully addressed in a contract. The task of identifying aspects of the state of the world that either can or can not be successfully contracted upon is a promising problem that merits further study. It is important because it identifies the potential content of contracts.

7 Appendix: Proofs of Results

7.1 Contracts and Mechanisms

The proof of the following lemma is straightforward.

Lemma 8 *If the mechanism $\hat{f}$ and the contract f satisfy $f = \hat{f} \circ v$, then:*

1. *$\hat{f}$ is incentive compatible with respect to the revelation of valuations if and only if f is incentive compatible with respect to revelation of types;*

2. *$\hat{f}$ is interim individual rational for each valuation of each agent if and only if f is interim individually rational for each type of each agent.*

Theorem 9 *Suppose that:*

1. *C is a convex subset of $\mathbb{R}^m$ and the objective $\Phi(c, v_1, v_2)$ is concave in c for each v_1 and v_2.*

2. *The functions h_j and t_j are affine[20] in c for each agent j.*

Then the following statements are true:

[20] That is, $h_j(\beta c + (1-\beta)c) = \beta h_j(c) + (1-\beta)h_j(c)$ for all $c \in C$ and $\beta \in [0, 1]$, and similarly for t_j.

1. *For every incentive compatible and interim individually rational contract f, there exists an incentive compatible and interim individually rational mechanism $\hat{f}$ that ex ante weakly dominates f:*

$$E_{v(A)}\left[\Phi\left(\hat{f}(v), v\right)\right] \geq E_A\left[\Phi\left(f(\alpha), v(\alpha)\right)\right]. \tag{28}$$

2. *As a consequence of statement 1, if a given mechanism $\hat{f}$ solves the optimal mechanism problem, then the induced contract $f = \hat{f} \circ v$ solves the optimal contract problem. Conversely, if a contract f solves the optimal contract problem, then the mechanism $\hat{f}$ defined in the proof of statement 1. solves the optimal mechanism problem.*

Proof. Define the contract g by averaging f over all states that determine the same valuations: for $\alpha^* \in A$,

$$g(\alpha^*) = E_A[f(\alpha)|v(\alpha) = v(\alpha^*)].$$

The mechanism $\hat{f}$ that is sought is defined as $\hat{f} \circ v = g$. Inequality (28) follows from the concavity of Φ together with an application of Jensen's Inequality:

$$\begin{aligned}
E_{v(A)}\left[\Phi\left(\hat{f}(v), v\right)\right] &= E_A\left[\Phi\left(g(\alpha), v(\alpha)\right)\right] \\
&= E_A\left[\Phi\left(E_A\left[f\left(\alpha^*\right)\left|v(\alpha^*) = v(\alpha)\right.\right], v(\alpha)\right)\right] \\
&\geq E_A\left[E_A\left[\Phi\left(f\left(\alpha^*\right), v(\alpha^*)\right)\left|v(\alpha^*) = v(\alpha)\right.\right]\right] \\
&= E_A\left[\Phi\left(f(\alpha), v(\alpha)\right)\right].
\end{aligned}$$

Applying Lemma 8, the mechanism $\hat{f}$ is shown to satisfy *IC* and *IIR* by showing that the contract g has these properties. For notational simplicity, we do this for $j = 1$. For $\alpha_1^* \in A_1$, define

$$\begin{aligned}
H_1'(\alpha_1^*) &= E_{A_2}\left[h_1(g(\alpha))\left|\alpha_1 = \alpha_1^*\right.\right], \textit{ and} \\
T_1'(\alpha_1^*) &= E_{A_2}\left[t_1(g(\alpha))\left|\alpha_1 = \alpha_1^*\right.\right].
\end{aligned}$$

The fact that h_1 is affine implies

$$\begin{aligned}
E_{A_1}[H_1(\alpha_1)\left|v_1(\alpha_1) = v_1(\alpha_1^*)\right.] &= E_{A_1}\left[E_{A_2}[h_1(f(\alpha))]\left|v_1(\alpha_1) = v_1(\alpha_1^*)\right.\right] \\
&= E_{A_2}[E_{A_1}[h_1(f(\alpha))|v_1(\alpha_1) = v_1(\alpha_1^*)]] \\
&= E_{A_2}[E_A[h_1(f(\alpha))|v(\alpha) = v(\alpha^*)]] \\
&= E_{A_2}[h_1(E_A[f(\alpha)|v(\alpha) = v(\alpha^*)])] \\
&= E_{A_2}[h_1(g(\alpha^*))] = H_1'(\alpha_1^*).
\end{aligned}$$

A similar argument shows that

$$E_{A_1}[T_1(\alpha_1)\left|v_1(\alpha_1) = v_1(\alpha_1^*)\right.] = T_1'(\alpha_1^*).$$

These equalities are now applied to demonstrate that g satisfies *IC* and *IIR* for agent 1. Incentive compatibility of f implies

$$H_1(\alpha_1^*)v_1(\alpha_1^*) + T_1(\alpha_1^*) \geq H_1(\alpha_1)v_1(\alpha_1^*) + T_1(\alpha_1) \tag{29}$$

for all $\alpha_1, \alpha_1^* \in A_1$. Because (29) holds for all $\alpha_1 \in A_1$, it follows that for all $\alpha_1^{**} \in A_1$,

$$\begin{aligned} & H_1\left(\alpha_1^*\right) v_1(\alpha_1^*) + T_1(\alpha_1^*) \\ \geq\ & E_{A_1}\left[H_1\left(\alpha_1\right) v_1(\alpha_1^*) + T_1(\alpha_1) \left| v_1\left(\alpha_1\right) = v_1\left(\alpha_1^{**}\right)\right.\right] \\ =\ & H_1'(\alpha_1^{**})v_1(\alpha_1^*) + T_1'(\alpha_1^{**}). \end{aligned} \tag{30}$$

Because (30) holds for all $\alpha_1^* \in A_1$, it follows that

$$\begin{aligned} & H_1'\left(\alpha_1^*\right) v_1(\alpha_1^*) + T_1'(\alpha_1^*) \\ =\ & E_{A_1}\left[H_1\left(\alpha_1\right) v_1(\alpha_1^*) + T_1(\alpha_1) \left| v_1\left(\alpha_1\right) = v_1\left(\alpha_1^*\right)\right.\right] \\ =\ & E_{A_1}\left[H_1\left(\alpha_1\right) v_1(\alpha_1) + T_1(\alpha_1) \left| v_1\left(\alpha_1\right) = v_1\left(\alpha_1^*\right)\right.\right] \\ \geq\ & E_{A_1}\left[H_1'\left(\alpha_1^{**}\right) v_1(\alpha_1) + T_1'(\alpha_1^{**}) \left| v_1\left(\alpha_1\right) = v_1\left(\alpha_1^*\right)\right.\right] \\ =\ & H_1'(\alpha_1^{**})v_1(\alpha_1^*) + T_1'(\alpha_1^{**}), \end{aligned}$$

and so g satisfies *IC*. Turning to *IIR*, we have

$$\begin{aligned} & H_1'(\alpha_1^*)v_1(\alpha_1^*) + T_1'(\alpha_1^*) \\ =\ & E_{A_1}[H_1(\alpha_1)v_1(\alpha_1^*) + T_1(\alpha_1)|v_1(\alpha_1) = v_1(\alpha_1^*)] \geq r. \end{aligned}$$

■

7.2 Limited Observability

Proof of Theorem 1. The subscript j denoting an agent is omitted for notational simplicity in this proof. Consider a strategy γ used by an agent who is constrained by limited observability. This constraint means that there exists for every $\alpha \in A = \{0, 1\}^{\mathbb{N}}$ a number $n \in \mathbb{N}$ such that $\gamma(\alpha)$ is determined by α_{n+1-}, i.e., γ is constant on the set $\{\alpha_{n+1-}\} \times A_{n+}$. Define $\mathcal{N} : A \to \mathbb{N}$ as the function whose value at α is the smallest natural number $\mathcal{N}(\alpha)$ with this property. As in section 2, consider the product topology on $A = \{0, 1\}^{\mathbb{N}}$ when $\{0, 1\}$ is assigned the discrete topology. For each $\alpha \in A$, the equality

$$\mathcal{N}\left(\{\alpha_{\eta(\alpha)+1-}\} \times A_{\eta(\alpha)+}\right) = \mathcal{N}(\alpha) \tag{31}$$

holds by definition of $\mathcal{N}$. Because $\{\alpha_{\mathcal{N}(\alpha)+1-}\} \times A_{\mathcal{N}(\alpha)+}$ is open, (31) shows that there exists a neighborhood of α whose image under the function $\mathcal{N}$ is $\mathcal{N}(\alpha)$. The function $\mathcal{N}$ is therefore continuous on A, regardless of the topology on its range $\mathbb{N}$. As noted in section 2, the Tychonoff Theorem implies that A is compact. Consequently, the function $\mathcal{N}$ has a maximum n^* on A. For all $\alpha \in A$, $\gamma(\alpha)$ is determined by α_{n^*+1-} and so it is finite.[21] ■

[21] This proof was suggested by a referee of this journal.

Proof of Theorem 2. Let (γ_1, γ_2) be a Bayesian-Nash equilibrium with limited observability in the game $(M_1 \times M_2, \eta)$ that implements f. Let $n \in \mathbb{N}$ be sufficiently large that each strategy $\gamma_j\left(\alpha_j\right)$ is determined by $\alpha_{j,n-}$. Because $f = \eta \circ (\gamma_1, \gamma_2)$, the contract f is finite. It is now shown by contradiction that f is incentive compatible. If not, then there exists $\alpha_j^*, \alpha_j^{**} \in A_j$ and $\varepsilon > 0$ such that

$$E_{A_i}\left[u_j(\alpha_j^*, f(\alpha_j^{**}, \alpha_i))\right] - E_{A_i}\left[u_j(\alpha_j^*, f(\alpha_j^*, \alpha_i)))\right] > \varepsilon > 0. \tag{32}$$

Because of our assumption that $h_j(C)$ is bounded, continuity of v_j implies the existence of $k \in \mathbb{N}$ such that

$$\left|u_j(\alpha_j'', c) - u_j(\alpha_j', c)\right| < \frac{\varepsilon}{2} \tag{33}$$

for all $c \in C$ and for all $\alpha_j', \alpha_j'' \in A$ such that $\alpha_{j,k-}'' = \alpha_{j,k-}'$. For $q \geq k$, it follows that

$$\left|E_A\left[u_j\left(\alpha_j, f(\alpha_j^{**}, \alpha_i)\right)\middle|\alpha_{j,q-} = \alpha_{j,q-}^*\right] - E_{A_i}\left[u_j(\alpha_j^*, f(\alpha_j^{**}, \alpha_i))\right]\right| < \frac{\varepsilon}{2}, \textit{ and}$$

$$\left|E_A\left[u_j\left(\alpha_j, f(\alpha_j^*, \alpha_i)\right)\middle|\alpha_{j,q-} = \alpha_{j,q-}^*\right] - E_{A_i}\left[u_j(\alpha_j^*, f(\alpha_j^*, \alpha_i))\right]\right| < \frac{\varepsilon}{2}.$$

Substitution into (32) implies

$$\begin{aligned} &E_A\left[u_j\left(\alpha_j, f(\alpha_j^{**}, \alpha_i)\right)\middle|\alpha_{j,q-} = \alpha_{j,q-}^*\right] - \\ &E_A\left[u_j\left(\alpha_j, f(\alpha_j^*, \alpha_i)\right)\middle|\alpha_{j,q-} = \alpha_{j,q-}^*\right] > 0 \end{aligned} \tag{34}$$

for all $q \geq k$. Replacing f with $\eta \circ (\gamma_1, \gamma_2)$ in the case of $q \geq n$ produces an inequality that contradicts the assumption that (γ_1, γ_2) is a Bayesian-Nash equilibrium with limited observability.

Conversely, suppose that f is incentive compatible and finite. Finiteness implies the existence of $n \in \mathbb{N}$ such that $f(\alpha_1, \alpha_2)$ depends only upon $\alpha_{1,n-}$ and $\alpha_{2,n-}$, which allows us to write

$$f(\alpha_1, \alpha_2) = f\left(\alpha_{1,n-}, \alpha_{2,n-}\right). \tag{35}$$

Consider the game in which $M_j = A_{j,n-}$ for $j = 1, 2$ and $\eta(m_1, m_2) = f(m_1, m_2)$. Let $\gamma_1(\alpha_1) = \alpha_{1,n-}$ and $\gamma_2(\alpha_2) = \alpha_{2,n-}$. Because $f = \eta \circ (\gamma_1, \gamma_2)$, it is sufficient to show that these finite strategies define a Bayesian-Nash equilibrium with limited observability. This is also proven by contradiction. Suppose that

$$\begin{aligned} &E_A\left[u_j\left(\alpha_j, \eta\left(\gamma_j\left(\alpha_j^{**}\right), \gamma_i(\alpha_i)\right)\right)\middle|\alpha_{j,q-} = \alpha_{j,q-}^*\right] > \\ &E_A\left[u_j\left(\alpha_j, \eta\left(\gamma_j\left(\alpha_j^*\right), \gamma_i(\alpha_i)\right)\right)\middle|\alpha_{j,q-} = \alpha_{j,q-}^*\right] \end{aligned} \tag{36}$$

for agent j, some $\alpha_j^*, \alpha_j^{**} \in A_j$ and some $q \geq n$. This is equivalent to

$$E_A\left[u_j\left(\alpha_j, f(\alpha_j^{**}, \alpha_i)\right)\Big|\alpha_{j,q-} = \alpha_{j,q-}^*\right] > \\ E_A\left[u_j\left(\alpha_j, f(\alpha_j^{*}, \alpha_i)\right)\Big|\alpha_{j,q-} = \alpha_{j,q-}^*\right]. \tag{37}$$

Inequality (37) can hold only if

$$E_{A_i}\left[u_j\left(\alpha_j, f(\alpha_j^{**}, \alpha_i)\right)\right] > E_{A_i}\left[u_j\left(\alpha_j, f(\alpha_j^{*}, \alpha_i)\right)\right] \tag{38}$$

for some $\alpha_j \in A_j$ such that $\alpha_{j,q-} = \alpha_{j,q-}^*$. Equation (35) implies that $f(\alpha_j^*, \alpha_i) = f(\alpha_j, \alpha_i)$ and so (38) contradicts the incentive compatibility of f. ■

Analysis of Case 3 in Example 2. We again omit the subscript j for notational simplicity. Let $S(\alpha_{n-})$ denote the sum

$$S(\alpha_{n-}) = \frac{1-\delta}{\delta}\sum_{q=1}^{n-1}\delta^q a_q$$

for any $\alpha_{n-} = \left(a_q\right)_{1\leq q\leq n-1}$. The formula

$$v\left(\{\alpha_{n-}\}\times A_{n-1+}\right) = \left[S(\alpha_{n-}), S(\alpha_{n-}) + \delta^{n-1}\right], \tag{39}$$

which follows from properties of the geometric series along with the assumption that $\delta > 0.5$, is needed below.

The argument is by contradiction. Given $x \in (0,1)$, let k be the smallest element of $\mathbb{N}$ such that $v(\alpha) \geq x$ or $v(\alpha) \leq x$ can be decided based upon α_{k+1-} for any choice of $\alpha_{k+1-} \in A_{k+1-}$. Case 2 implies that $k > 1$. The minimality of k and formula (39) together imply that there exists an initial string α_{k-}^* such that

$$S\left(\alpha_{k-}^*\right) < x < S\left(\alpha_{k-}^*\right) + \delta^{k-1},$$

so that $v(\alpha) \leq x$ or $v(\alpha) \geq x$ cannot be decided with certainty based upon α_{k-}^*. Define $\overline{\alpha} = \left(\overline{a}_q\right)_{q\in\mathbb{N}}$ and $\underline{\alpha} = \left(\underline{a}_q\right)_{q\in\mathbb{N}}$ as follows:

$$\overline{a}_q = \begin{cases} a_q^* & \text{if } q \leq k-1 \\ 1 & \text{if } q \geq k \end{cases};$$

$$\underline{a}_q = \begin{cases} a_q^* & \text{if } q \leq k-1 \\ 0 & \text{if } q \geq k \end{cases}.$$

Because $v(\overline{\alpha}) = S\left(\alpha_{k-}^*\right) + \delta^{k-1} > x$, it must be the case that $S(\overline{\alpha}_{k+1-}) \geq x$ else $v(\alpha) \leq x$ or $v(\alpha) \geq x$ could not be decided based upon $\overline{\alpha}_{k+1-}$. Because $v(\underline{\alpha}) = S\left(\alpha_{k-}^*\right) < x$, similar reasoning using formula (39) in the case of $\alpha_{n-} = \underline{\alpha}_{k+1-}$ implies that $S(\underline{\alpha}_{k+1-}) + \delta^k \leq x$. Combining these inequalities produces

$$\frac{1-\delta}{\delta}\cdot\delta^k = S(\overline{\alpha}_{k+1-}) - S(\underline{\alpha}_{k+1-}) \geq x - \left(x - \delta^k\right) = \delta^k,$$

where the first equality follows directly from the definitions of S, $\overline{\alpha}$ and $\underline{\alpha}$. This inequality cannot hold because $(1-\delta)/\delta < 1$ for $\delta > 0.5$. ■

7.3 Reversibility and Strong Irreversibility

Proof of Theorem 4. We again omit the subscript j for notational simplicity. The proof is by induction on the length $n-1$ of the initial string α_{n-}. The case of $n-1=0$ is obvious, given the assumption that $v(A)=[\underline{v},\overline{v}]$. Assume that statements 1. and 2. hold for initial strings of length $n-1$. We first show that $D_{(\alpha,n-,0)}$ and $D_{(\alpha_{n-},1)}$ are closed intervals whose intersection consists of a single point. It is then shown that statement 2. holds for all pairs of distinct initial strings of length n.

We begin by noting that $D_{(\alpha_{n-},0)}$ and $D_{(\alpha_{n-},1)}$ are closed sets. The Tychonoff Theorem implies that A is compact in the product topology when each set $\{0,1\}$ is assigned the discrete topology. The sets $\{\alpha'|\alpha'_{n-}=\alpha_{n-},a'_n=0\}$ and $\{\alpha'|\alpha'_{n-}=\alpha_{n-},a'_n=1\}$ are closed subsets of A in this topology and are therefore also compact. Continuity of $v(\cdot)$ implies that $D_{(\alpha_{n-},0)}$ and $D_{(\alpha_{n-},1)}$ are compact and therefore closed.

Because $D_{(\alpha_{n-},1)}\cup D_{(\alpha_{n-},0)}=D_{\alpha_{n-}}$ and $D_{\alpha_{n-}}$ is a closed interval by the induction hypothesis, the connectedness of $D_{\alpha_{n-}}$ implies that $D_{(\alpha_{n-},1)}\cap D_{(\alpha_{n-},0)}\neq\emptyset$. Assume by way of contradiction that $D_{(\alpha_{n-},0)}$ and $D_{(\alpha_{n-},1)}$ either fail to be closed intervals or else they are closed intervals whose interiors overlap. Either of these statements can be true only if there exists $y_1,y_2\in D_{(\alpha_{n-},a)}$ and $z\in D_{(\alpha_{n-},b)}\setminus D_{(\alpha_{n-},a)}$ with $y_1<z<y_2$ and $a\neq b\in\{0,1\}$. We derive a contradiction in the case of $a=0$ and $b=1$; the case of $a=1$ and $b=0$ is similar. Let

$$\overline{y}=\sup\{y\in D_{(\alpha_{n-},0)}|y\leq z\}\ \ and\ \ \underline{y}=\inf\{y\in D_{(\alpha_{n-},0)}|y\geq z\}.$$

The points $\underline{y}$, $\overline{y}$ exist and are elements of $D_{(\alpha_{n-},0)}$ because this set is compact, and $\underline{y}<\overline{y}$ because otherwise $z=\underline{y}=\overline{y}\in D_{(\alpha_{n-},0)}$. It is also the case that

$$\overline{y}=\inf\{y\in D_{(\alpha_{n-},1)}|y\geq\overline{y}\}\ \ and\ \ \underline{y}=\sup\{y\in D_{(\alpha,n-,1)}|y\leq\underline{y}\},$$

and consequently $\underline{y},\overline{y}\in D_{(\alpha_{n-},1)}$. For $k=1,2$, $\underline{\alpha}^k$ and $\overline{\alpha}^k$ therefore exist such that $\underline{\alpha}^k_{n+1-}=(\alpha_{n-},0)$, $\overline{\alpha}^k_{n+1-}=(\alpha_{n-},1)$, $v(\underline{\alpha}^1)=v(\overline{\alpha}^1)=\underline{y}$, and $v(\underline{\alpha}^2)=v(\overline{\alpha}^2)=\overline{y}$. This contradicts the assumption that every contingency is strongly irreversible, and so $D_{(\alpha_{n-},0)}$ and $D_{(\alpha_{n-},1)}$ satisfy the conclusion.

Turning to statement 2 for $\alpha_{n+1-}\neq\alpha'_{n+1-}$, either $\alpha_{n-}=\alpha'_{n-}$ or $\alpha_{n-}\neq\alpha'_{n-}$. We have just shown in the first case that $D_{\alpha_{n+1-}}\cap D_{\alpha'_{n+1-}}$ contains one element. If $\alpha_{n-}\neq\alpha'_{n-}$, then the induction hypothesis together with the fact that $D_{\alpha_{n+1-}}\subset D_{\alpha_{n-}}$ and $D_{\alpha'_{n+1-}}\subset D_{\alpha'_{n-}}$ imply that $D_{\alpha_{n+1-}}\cap D_{\alpha'_{n+1-}}$ contains at most one element. ■

7.4 The Model of Bilateral Trade

Proof of Theorem 5. The theorem is proven by showing that if the contract (p,t) is recordable, then there exists an *IC* and *IIR* finite contract (p^*,t^*) such that:

1. $(p^*(\alpha),t^*(\alpha))$ does not depend upon $a_{B,n}$ for $n\geq b$ and $a_{S,n}$ for $n\geq s$;
2. the ex ante gains from trade in (p^*,t^*) are at least as large as in (p,t), i.e.,

$$\int_A[v_B(\alpha_B)-v_S(\alpha_S)]\,p^*(\alpha)d\pi(\alpha)\geq\int_A[v_B(\alpha_B)-v_S(\alpha_S)]\,p(\alpha)d\pi(\alpha)\,.$$

The contract (p, t) thus cannot be optimal.

Let $((p_m, t_m))_{m\in\mathbb{N}}$ be a sequence of contracts that demonstrates the recordability of (p, t). The first step is to construct for each $m \in \mathbb{N}$ an *IC* and *IIR* contract (p_m^*, t_m^*) that is interim payoff equivalent to (p_m, t_m) and whose value at α is determined by $\alpha_{B,b-}$ and $\alpha_{S,s-}$. Define $(p_m^*(\alpha'), t_m^*(\alpha'))$ by averaging (p_m, t_m) over all states $\alpha = (\alpha_B, \alpha_S)$ such that $\alpha_{B,b-} = \alpha'_{B,b-}$ and $\alpha_{S,s-} = \alpha'_{S,s}$:

$$p_m^*(\alpha') = \int_{A_{B,b-1+}\times A_{S,s-1+}} p_m(\alpha) d\pi(\alpha \left| \alpha_{B,b-} = \alpha'_{B,b-}, \alpha_{S,s-} = \alpha'_{S,s} \right.),$$

and

$$t_m^*(\alpha') = \int_{A_{B,b-1+}\times A_{S,s-1+}} t_m(\alpha) d\pi(\alpha \left| \alpha_{B,b-} = \alpha'_{B,b-}, \alpha_{S,s-} = \alpha'_{S,s} \right.).$$

For any $\alpha'_B \in A_B$, Theorem 3 implies that $\int_{A_S} p_m(\alpha_B, \alpha_S) d\pi_S(\alpha_S)$ is constant over the set of all α_B such that $\alpha_{B,b-} = \alpha'_{B,b-}$. It follows that:

$$\begin{aligned}
&\int_{A_S} p_m(\alpha'_B, \alpha_S) d\pi_S(\alpha_S) = \\
&\int_{A_{B,b-1+}} \int_{A_S} p_m(\alpha_B, \alpha_S) d\pi_S(\alpha_S) d\pi_B \left(\alpha_B \left| \alpha_{B,b-} = \alpha'_{B,b-} \right.\right) \\
&= \int_{A_{S,s-}} \int_{A_{B,b-1+}\times A_{S,s-1+}} p_m(\alpha_B, \alpha'_S) \\
&\cdot d\pi \left(\alpha \left| \alpha_{B,b-} = \alpha'_{B,b-}, \alpha'_{S,s} = \alpha_{S,s-} \right.\right) d\pi_S(\alpha_{S,s-}) \\
&= \int_{A_{S,s-}} p_m^*(\alpha'_B, \alpha_S) d\pi_S(\alpha_{S,s-}) \\
&= \int_{A_S} p_m^*(\alpha'_B, \alpha_S) d\pi_S(\alpha_S).
\end{aligned}$$

The second equality is a change in the order of integration, the third applies the definition of $p_m^*(\alpha'_B, \alpha_S)$, and the last is true because $p_m^*(\alpha'_B, \alpha_S)$ is determined by $\alpha_{S,s-}$. Similar arguments prove that for all $\alpha'_B \in A_B$ and $\alpha'_S \in A_S$,

$$\begin{aligned}
\int_{A_S} t_m(\alpha'_B, \alpha_S) d\pi_S(\alpha_S) &= \int_{A_S} t_m^*(\alpha'_B, \alpha_S) d\pi_S(\alpha_S), \\
\int_{A_B} p_m(\alpha_B, \alpha'_S) d\pi_B(\alpha_B) &= \int_{A_B} p_m^*(\alpha_B, \alpha'_S) d\pi_B(\alpha_B), \textit{and} \\
\int_{A_B} t_m(\alpha_B, \alpha'_S) d\pi_B(\alpha_B) &= \int_{A_B} t_m^*(\alpha_B, \alpha'_S) d\pi_B(\alpha_B).
\end{aligned}$$

The constraints of *IC* and *IIR* thus follow for (p_m^*, t_m^*) from the corresponding properties of (p_m, t_m). The interim expected utility function of each trader and the ex ante expected gains from trade (p_m^*, t_m^*) are the same as in (p_m, t_m).

The value of (p_m^*, t_m^*) at α is determined by $\alpha_{B,b-}$ and $\alpha_{S,s-}$, and the set $A_{B,b-} \times A_{S,s-}$ of all pairs $\left(\alpha_{B,b-}, \alpha_{S,s-}\right)$ of such initial strings is finite. As a probability, $p_m^*(\alpha) \in [0, 1]$. The boundedness of the sets $v_B(A_B)$ and $v_S(A_S)$, assumption (21) of section 5, and *IIR* together imply that $(t_m^*(\alpha_{B,b-}, \alpha_{S,s-}))_{m\in\mathbb{N}}$ is bounded. By taking a subsequence (if necessary), it can thus be assumed without loss of generality that $(p_m^*(\alpha), t_m^*(\alpha))_{m\in\mathbb{N}}$ converges for all $\alpha \in A$. The contract (p^*, t^*) defined by

$$(p^*(\alpha), t^*(\alpha)) = \lim_{m\to\infty} (p_m^*(\alpha), t_m^*(\alpha))$$

for $\alpha \in A$ is thus well-defined. The contract (p^*, t^*) inherits *IC*, *IIR*, and independence of $\alpha_{B,b-1+}$ and $\alpha_{S,s-1+}$ from the contracts in the sequence. To verify that the expected gains from trade in (p^*, t^*) is as large as in (p, t), Lebesgue's Convergence Theorem implies

$$\begin{aligned}
&\int_A (v_B(\alpha_B) - v_S(\alpha_S))\, p^*(\alpha) d\pi\,(\alpha) \\
&= \lim_{m\to\infty} \int_{A_B\times A_S} (v_B(\alpha_B) - v_S(\alpha_S))\, p_m^*(\alpha) d\pi\,(\alpha) \\
&= \lim_{m\to\infty} \int_{A_B\times A_S} (v_B(\alpha_B) - v_S(\alpha_S))\, p_m(\alpha) d\pi\,(\alpha) \\
&\geq \int_A (v_B(\alpha_B) - v_S(\alpha_S))\, p(\alpha) d\pi\,(\alpha)\,,
\end{aligned}$$

where the last two lines are the recordability inequality. ■

Proof of Theorem 6. Theorem 4 implies that the sets $D_{\alpha_j,n-}$ defined in the theorem are intervals $\left[x_{j,q}^n, x_{j,q+1}^n\right]$ for $j = B, S$ and $1 \leq q \leq m_n$. Assumption (21) of section 5 along with assumption 3 of the theorem imply that each of these intervals has a nonempty interior. Define $\xi_B^n(v_B)$ and $\xi_S^n(v_S)$ as follows:

$$\begin{aligned}
\xi_B^n(v_B) &= \sup\left\{x_{B,q}^n \,\middle|\, x_{B,q}^n \leq v_B, 1 \leq q \leq m_n\right\}, \\
\xi_S^n(v_S) &= \inf\left\{x_{S,q}^n \,\middle|\, v_S \leq x_{S,q}^n, 1 \leq q \leq m_n\right\}.
\end{aligned}$$

The function ξ_B^n rounds a buyer's valuation downward while ξ_S^n rounds a seller's valuation upward, in each case to the nearest boundary of one of the intervals $D_{\alpha_B,n-}$ or $D_{\alpha_S,n-}$, respectively. Continuity of v_B and v_S implies that $\lim_{n\to\infty} \xi_B^n(v_B) = v_B$ and $\lim_{n\to\infty} \xi_S^n(v_S) = v_S$ for all $v_B \in [\underline{v}_B, \overline{v}_B]$ and $v_S \in [\underline{v}_S, \overline{v}_S]$.

Theorem 9 implies that an optimal contract has the form $\left(\hat{p}^*(v(\alpha)), \hat{t}^*(v(\alpha))\right)$, where p^* and v^* solve the optimal mechanism problem

$$\max_{(p,t)} \iint (v_B - v_S)\, p(v_B, v_S) d\mu_B d\mu_S \;\; \textit{s.t. IC and IIR.} \tag{40}$$

Given the regularity of μ_B and μ_S, Theorem 2 of Myerson and Satterthwaite (1983) characterizes a constant $k^* \in [0, 1]$ such that $\hat{p}^*(v_B, v_S)$ has the form

$$\hat{p}^*(v_B, v_S) = \begin{cases} 1 & \textit{if } V_B(v_B, k^*) \geq V_S(v_S, k^*); \\ 0 & \textit{otherwise.} \end{cases}$$

Define the probability function $\hat{p}_n(v_B, v_S)$ as

$$\hat{p}_n(v_B, v_S) = \begin{cases} 1 & \textit{if } V_B(\xi_B^n(v_B), k^*) \geq V_S(\xi_S^n(v_S), k^*); \\ 0 & \textit{otherwise.} \end{cases}$$

The sequence $(\hat{p}_n)_{n \in \mathbb{N}}$ converges pointwise to $\hat{p}^*$. A comparison of $\hat{p}_n(v_B, v_S)$ with $\hat{p}^*(v_B, v_S)$ shows that $\hat{p}_n(v_B, v_S)$ satisfies inequality (2) of Myerson and Satterthwaite (1983) because $\hat{p}^*(v_B, v_S)$ satisfies it. It also inherits from $\hat{p}^*(v_B, v_S)$ the monotonicity properties required by Theorem 1 in their paper because ξ_B^n and ξ_S^n are nondecreasing. Formula (6) of their paper thus defines a transfer function $\hat{t}_n(v_B, v_S)$ such that the revelation mechanism $(\hat{p}_n, \hat{t}_n)$ satisfies *IC* and *IIR*. Like $\hat{p}_n(v_B, v_S)$, $\hat{t}_n(v_B, v_S)$ is constant on the interior of each set of the form $D_{\alpha_{B,n-}} \times D_{\alpha_{S,n-}}$.

The sequence $(p_n(\alpha), t_n(\alpha))_{n \in \mathbb{N}}$ that demonstrates the recordability of the optimal contract $(\hat{p}^*(v(\alpha)), \hat{t}^*(v(\alpha)))$ is defined as follows: for $\alpha = (\alpha_B, \alpha_S) \in A$, $(p_n(\alpha), t_n(\alpha))$ equals the value of $(\hat{p}_n(v), \hat{t}_n(v))$ in the interior of $D_{\alpha_{B,n-}} \times D_{\alpha_{S,n-}}$. It is straightforward to show that $(p_n(\alpha), t_n(\alpha))$ satisfies *IC* and *IIR* because $(\hat{p}_n(v), \hat{t}_n(v))$ has these properties. It is clear that

$$\lim_{n \to \infty} p_n(\alpha) = \lim_{n \to \infty} \hat{p}_n(v(\alpha)) = \hat{p}^*(v(\alpha))$$

except at those states $\alpha = (\alpha_B, \alpha_S)$ for which either $v_B(\alpha_B)$ or $v_S(\alpha_S)$ is an endpoint of one of the intervals $\left[x_{j,q}^n, x_{j,q+1}^n\right]$ for $j = B$ or S, respectively, some $n \in \mathbb{N}$ and $1 \leq q \leq m_n$. Assumption 3 implies that this set of states has π-measure zero.

We conclude the proof by showing that the recordability inequality holds:

$$\begin{aligned} & \int (v_B(\alpha_B) - v_S(\alpha_S)) \, \hat{p}^*(v_B(\alpha_B), v_S(\alpha_S)) \, d\pi(\alpha) \\ = & \iint (v_B - v_S) \left(\lim_{m \to \infty} \hat{p}_n(v_B, v_S) \right) d\mu_B(v_B) \, d\mu_S(v_S) \\ \leq & \liminf \iint (v_B - v_S) \, \hat{p}_n(v_B, v_S) \, d\mu_B(v_B) \, d\mu_S(v_S) \\ = & \liminf \int (v_B(\alpha_B) - v_S(\alpha_S)) \, p_n(\alpha_B, \alpha_S) \, d\pi(\alpha). \end{aligned}$$

Because $(v_B - v_S) \, \hat{p}_n(v_B, v_S) \geq -\left|\underline{v}_B - \overline{v}_S\right|$, Fatou's Lemma implies the inequality. The last equality follows because $p_n(\alpha) = \hat{p}_n(v(\alpha))$ for π-a.e. α. By selecting a subsequence of $(p_n, t_n)_{n \in \mathbb{N}}$, the "lim inf" in the last line can be replaced with "lim", which completes the proof of recordability. ■

Proof of Theorem 7. For the recordable contract (p, t), the proof of Theorem 5 demonstrates the existence of an *IC* and *IIR* contract (p^*, t^*) such that: (i) (p^*, t^*) is state independent and hence constant; (ii) the ex ante gains from trade in (p^*, t^*) are as least as large as in (p, t). Suppose $p^* > 0$. Interim individual rationality implies

$$v_S(\alpha_S) \leq \frac{t^*}{p^*} \leq v_B(\alpha_B)$$

for all $\alpha_B \in A_B$ and $\alpha_S \in A_S$. It follows that $v_S(\alpha_S) \leq v_B(\alpha_B)$ for all $\alpha_B \in A_B$ and $\alpha_S \in A_S$, which contradicts assumption 2 in the theorem. The contradiction implies that $p^* = 0$ and so the ex ante expected gains from trade in (p^*, t^*) are zero. Property (ii) above together with interim individual rationality therefore imply that they are also zero in (p, t). ■

References

[1] Al-Najjar, N.: Complexity as a barrier to competitive imitation. Northwestern University working paper (2000).

[2] Anderlini, L., and Felli, L.: Incomplete written contracts: undescribable states of nature. Quarterly Journal of Economics **109**, 1085-1124 (1994).

[3] Anderlini, L., and Felli, L.: Describability and agency problems. European Economic Review **42**, 35–59 (1998).

[4] Chatterjee, K., and Samuelson, W.: Bargaining under incomplete information. Operations Research **31**, 835–851 (1983).

[5] Dow, J.: Search decisions with limited memory. Review of Economic Studies **58**, 1–14 (1991).

[6] Fudenberg, D., and Tirole, J.: Game theory. Cambridge: MIT Press 1991.

[7] Hagerty, K., and Rogerson, W.: Robust trading mechanisms. Journal of Economic Theory **42**, 94-107 (1985).

[8] Harsanyi, J.: Games with incomplete information played by bayesian players, parts I, II, and III. Management Science **14**, 159-182, 320-334, 486-502 (1967-68).

[9] Holmström, B., and Myerson, R.: Efficient and durable decision rules with incomplete information. Econometrica **51**, 1799-1819 (1983).

[10] Krasa, S., and Williams, S. R.: Incompleteness as a constraint in contract design. University of Illinois working paper (2000).

[11] Krasa, S., and Williams, S. R.: Limited observability as a constraint in contract design. University of Illinois working paper (2001).

[12] Maskin, E., and Tirole, J.: Unforeseen contingencies and incomplete contracts. Review of Economic Studies **66**, 83-114 (1999).

[13] Meyer, M.: Learning from coarse information: biased contests and career profiles. Review of Economic Studies **58**, 15–41 (1991).

[14] Milgrom, P., and Stokey, N.: Information, trade and common knowledge. Journal of Economic Theory **26**, 17-27 (1982).

[15] Myerson, R., and Satterthwaite, M. A.: Efficient mechanisms for bilateral trade. Journal of Economic Theory **29**, 265–281 (1983).

[16] Radner, R.: Costly and bounded rationality in individual and team decision-making. Industrial and Corporate Change **9**, 623-658 (2000).

[17] Rubinstein, A.: On price recognition and computational complexity in a monopolistic model. Journal of Political Economy **101**, 473–484 (1993).

[18] Rubinstein, A.: Modeling bounded rationality. Cambridge: MIT Press 1998.

[19] Segal, I.: Complexity and renegotiation: a foundation for incomplete contracts. Review of Economic Studies **66**, 57-82 (1999).

[20] Woodward, B., and Armstrong, S.: The brethren. New York: Simon and Schuster 1979.