1. Introduction

One of the pricing techniques that is fairly common on the internet is something called click stream pricing. Anyone who shops on the internet has surely encountered this (possibly without recognizing it). For example, a recent request for a price quote for a generic desktop computer offered to bundle the computer with a two year warrant. Below the price quote and its associated “Add to Cart” button was a check box with an offer to buy a 2 year warranty for $79. Checking the box had the effect of adding the warranty to the cart. Leaving the box unchecked then clicking the checkout button, brought up a page with the usual total charges. Below this total charge was a banner saying “People who purchased this Item also liked the Following items:”. There again was the 2 year warranty, but now at a price of $56.

This example illustrates a number of things about the way modern digital markets operate (and how different they are from competitive markets as envisaged by Arrow-Debreu). First, the transaction itself doesn’t involve a seller - just a dumb computer program that comes up with a price offer. This program can’t negotiate or change its mind. However, what it does depends on messages it receives from the buyer. In the click stream pricing case, the message is a stream of urls that the buyer visited before asking for a quote. The actual offer would depend on cookies that had been placed on the buyer’s browser, whether the buyer had 'logged in’ to his or her account with the seller, and so on. The quote might even depend...
on whether they buyer had visited other websites.\footnote{Twitter and Facebook can place cookies on my computer when I load the “Like” or “Follow us on” buttons on other websites into my browser. The dumb program can (presumably for a fee) query Twitter to check whether the same browser has visited a competitors website.} This kind of price discrimination will occur even when there is intense competition between suppliers and goods produce only private values (like warranties).

What makes these markets hard to understand theoretically is that many of the details associated with the mechanisms that sellers use are buried in a complex computer network. It is almost impossible to describe a digital market by listing the mechanisms that players use along with the messages that these mechanisms entail. Moreover, the process by which sellers process messages is continually changing since there is lots of innovation in information processing on the internet. Whatever mechanisms sellers are allowed to use in a model are likely to be obsolete by the time the model produces a theorem.

The purpose of this paper is to show how the theory of competing mechanisms can be used to get around these problems and model these markets. We provide a formal definition of competing mechanism games that is broad enough to encompass any imaginable digital market. We then provide a relatively weak condition on these games, referred to as regularity, that makes it possible to characterize the outcome functions that can be supported as perfect Bayesian equilibrium of games that embed the known parts of the game - referred to here as the default game. We provide a complete characterization of this set of outcome functions. The significance is that equilibrium of games that are incompletely known can be understood using information about the part of the game that is known.

Finally, we provide a single game, called the reciprocal contracting game, that has the property that an outcome function can be supported as a perfect Bayesian equilibrium in a regular competing mechanism game if and only if it can be supported as an equilibrium for the reciprocal contracting game. The logic of equilibrium in the reciprocal contracting game is quite simple, making it a useful stand-in for the true game when this true game is not known.

The reciprocal contracting game itself is a regular competing mechanism game. In particular what that means is that players commit their own actions based on messages they exchange with other players. In this sense, the players are using mechanisms that are conceptually no more complicated or demanding than auctions. The fact that there is a decentralized game like the reciprocal contracting game that supports all incentive compatible and individually rational outcomes at the same time is important.

To see why, recall that the primary practical concern of all this is that very good default games can support very bad equilibrium outcomes when the parts of the game the modeler doesn’t understand allow traders to circumvent the rules of the default game. Myerson (Myerson 1997) has already explained how a coordinator can support outcomes that aren’t Bayesian equilibrium of a default game by having players report types then instructing players about which action they should take. Provided the players are bound somehow to do what the coordinator says, lots of outcomes can be supported by having the coordinator provide different instructions to players. The literature on collusion ((Laffont and Martimort 1997)) often adopts this approach to explain bid manipulation is auctions, for example.
(Kalai, Kalai, Lehrer, and Samet 2010) and (Forges 2012) point out that this can be interpreted in a different way. Once the mapping from types to actions is known, one can imagine defining a game based on this outcome function in which the type reports are to be thought of as commitment devices, while the payoffs associated with these commitment devices are derived using the outcome function constructed by the coordinator. This becomes what Forges calls a commitment game. It follows immediately from this that if an outcome function is incentive compatible and individually rational in Myerson’s sense, there is some commitment game which implements that outcome as a Bayesian Nash equilibrium.

When type reports are interpreted as commitment devices, every incentive compatible and individually rational outcome in Myerson’s sense needs a different commitment game in order to be supported as a Bayesian equilibrium. The default game is the part of the game that is easy to observe, the commitment game that implements a particular individually rational incentive compatible outcome is a game that mechanism designers could be playing. The revelation principle says itself says nothing about what the right commitment game is.

Furthermore, even if we could somehow fix the rules of one of these commitment games to implement some specific outcome, the whole problem in digital markets is that it is likely to be impossible to prevent mechanism designers from dreaming up new ways of communicating with their agents and incorporating this communication into their actions. So a direct mechanism or commitment game is really no more reliable than any other default game in understanding players actions.

Finally, direct mechanisms and their corresponding commitment games are actually quite implausible. What players do when they participate in a direct mechanism or a commitment game is to delegate their action choice to a disinterested coordinator who calculates the appropriate randomization over actions for each player based on the entire profile of type reports and on the outcome function that is supposed to be implemented. Though all games need centralized coordination to enforce rules, it is harder to imagine where this disinterested coordinator comes from, and how he or she diguges out how to make the right recommendations.

The approach taken in this paper goes back to the most basic premise of mechanism design - a mechanism designer can commit only his or her own actions based on messages exchanged with agents. These commitments must be chosen by the mechanism designer on his or her own, and it must be in their interest to offer the contracts they do - no explicit coordination can be present apart from that required to ensure the enforcement of contracts. In fact, this paper effectively assumes that all extensive form games in which principals write contracts that condition their actions on messages received from agents are extensions of the default game that must be considered when evaluating the performance of a default game.

Since it is difficult to imagine exactly what the set of ‘all extensive form games’ looks like in practise, we provide a single game, called the reciprocal contracting game, that supports as a perfect Bayesian equilibrium every outcome function that can be supported as a perfect Bayesian equilibrium in any regular competing mechanism game. This game is immune to contractual innovation in the sense that if mechanism designers find new ways to communicate with agents new contracts to be used with this communication, those innovations won’t change the set of supportable equilibrium outcomes. Though the reciprocal contracting game is abstract, its logic is elementary, so it provides a convenient way of analyzing default games.
and suggesting simpler or more realistic communication protocols that can be used to understand how these markets are working.

The basic idea that messages exchanged by players can be used to convey market information is an old one (for example (Epstein and Peters 1999)). Recent papers have shown how to exploit this to expand the set of supportable outcomes in a default game. For example (Peters and Troncoso-Valverde 2013) show that a modification of the recommendation game described by (Yamashita 2010) supports all individually rational and incentive compatible outcomes (in Myerson’s sense) of the default game as perfect Bayesian equilibrium. The full set of outcome functions supportable in (Peters and Troncoso-Valverde 2013) or (Yamashita 2010) are not known, since they aren’t regular games. In particular, they do not support the property of invariant punishment as described in (Peters and Szentes 2012). Here the main theorem shows that the equilibrium outcomes in the reciprocal contracting game coincide with the set of equilibrium outcomes in regular contracting games. So the paper provides a subset of the set of all mechanism design game for which a complete characterization of equilibrium outcomes is possible.

The concept of regularity is the extensive form equivalent of the invariant punishment property in (Peters and Szentes 2012). The characterization provided in that paper is for a game of computer programs in which players cannot communicate type information except through the programs they offer. As a consequence, the set of outcome functions supportable in (Peters and Szentes 2012) is a subset of the set of outcome functions supportable as equilibrium in reciprocal contracting games.

The contribution here is twofold, relative to that paper. First, it shows how additional additional communication can be added to allow players to communicate their payoff type information directly to each other rather than indirectly, through contract offers, as was done in (Peters and Szentes 2012). Second, by limiting the set of feasible contracts, this paper shows how the issue of self referentiality can be handled without complex higher order logic when contracts condition on one another. Finally, the paper shows how to deal with randomization.

This paper begins with an example. The purpose of the example is twofold. First, it shows how mechanisms based only on messages received from other players can be used by players to work around the equilibrium outcomes associated with a good default game. In particular, the paper shows how private communication can be used to coordinate players behavior without the need for any kind of centralized mediator - the example helps illustrate how this is done.

Second, though the logic of reciprocal contracting is simple, the messages are complex. The example illustrates how the basic idea of reciprocal contracting can be used in an application to suggest more realistic indirect mechanisms.

Finally, the applied message of the paper is that good default games might not support good outcomes. The example illustrates a familiar situation in which this is a real problem.

A reader who believes this already and is only interested in the main theorems should skip to the next section.

2. Example

The example begins with a double auction, which will act as the default game in the arguments that follow. It assigns property rights to the sellers and defines
the messages buyers and sellers can send, along with the payoffs associated with these different messages. In particular, all that buyers and sellers can do in the auction is submit bids. This institution defines the backdrop against which sellers’ mechanisms operate. These institutions also define the outside options for buyers and sellers when they decide not to participate in the mechanisms offered by the sellers.

There are two sellers and two buyers (i.e., four players in all) in the market that is guided by this institution. Each seller has a single unit of output to which he or she assigns a value of 0. Each buyer has a private valuation, either \( v_l \) or \( v_h \), ranked in the obvious way with \( 0 < v_l < v_h \). Payoffs to the seller are equal to the money they receive while payoffs to each buyer are equal to their private valuation when they succeed in trading, less the money they pay. We assume that valuations are correlated. To make life simple suppose that both valuations are the same with probability \( \pi > \frac{1}{2} \) and that they are equally likely to (both be) \( v_h \) or \( v_l \) in that case.

The buyers and sellers submit bids. The two available goods are awarded to the two highest bidders at a price equal to the third highest bid with the proviso that if there are more than two highest bidders, then the good is awarded to buyers whenever possible and randomly otherwise. For the purposes of illustration, focus on pure strategy equilibrium. There is a continuum of (ex post efficient) Bayesian equilibrium outcomes for this game in which all bidders bid \( p \in (0, v_l) \) independent of type. The sellers’ payoff in each of these equilibrium outcomes is \( p \), high value buyers earn \( v_h - p \) and low valuation buyers earn \( v_l - p \). In all of these equilibrium outcomes trade occurs for sure. Evidently, the best equilibrium outcome for sellers occurs when \( p = v_l \).

There may also be inefficient asymmetric equilibrium in which the trading price is above \( v_l \) with positive probability.\(^2\) In these equilibria, one seller bids some price \( p^h \) that lies strictly between \( v_l \) and \( v_h \), while the other bids \( v_l \). Buyers bid \( p^h \) when their value is high and \( v_l \) otherwise. The seller who submits the high bid sells only if both buyers have high values. So his expected revenue is \( \pi \frac{v_h^2}{2} \). He can ensure trade at a price \( v_l \) by cutting his bid to \( v_l \), so a necessary condition for this to be supported in equilibrium is that \( \pi \frac{v_h^2}{2} \geq v_l \). If this inequality is strict, the seller who submits the high bid does better than he does in the efficient equilibrium. From the inequality it is apparent that an inefficient outcome like this can make some seller better off only if \( v_h \) is at least \( \frac{2}{\pi} \) times as large as \( v_l \).

The next step is to illustrate how an efficient set of institutions can come undone, so it will be assumed that \( v_h \) and \( v_l \) are close enough together so that all equilibria in the double auction are efficient. It is now possible to illustrate how sellers’ competition in mechanisms can support a collusive outcome. It is important to note here that the seller’s competition is part of the process of price determination that a modeler may only imperfectly understand. The objective of the example is to create an equilibrium in which sellers collude in a manner that restricts output in the event that both buyers have low valuations. Whatever the sellers do to accomplish this, the ultimate outcome will reduce to three different trading prices, one, \( p_{hh} \), for the case where both buyers have high valuations, one, \( p_{ll} \) for the case where both buyers have low valuations, and one, \( p_{hl} \) for the case where their valuations are different.

\(^2\) I thank Gabor Virag for pointing this out to me.
Assuming that the sellers collude and provide only one unit of output when the
buyers have low valuations, the possible prices are determined by standard incentive
compatibility constraints. In order for the high value buyer to be willing to accept
the lottery between prices \( p_{hh} \) and \( p_{hl} \), he should prefer that lottery to what he
could get by pretending to be a low value buyer. This gives the completely standard
incentive condition

\[
\pi (v_h - p_{hh}) + (1 - \pi) (v_h - p_{hl}) \geq (1 - \pi) \frac{1}{2} (v_h - p_l).
\]  

(2.1)

He trades for sure if his value is high at a price that might depend on the value of
the other buyer. Since one of the sellers submits a high bid, he will fail to trade if
he pretends to be low value and the other buyer has a high value. If the other other
buyer’s value is low, he will have the same chance to trade as the other buyer - \( \frac{1}{2} \).

Similarly, the incentive condition for the low value buyer is

\[
\frac{\pi}{2} (v_l - p_l) \geq \pi (v_l - p_{hl}) + (1 - \pi) (v_l - p_{hh}).
\]  

(2.2)

As for participation, the worst that can happen is that the traders who do
participate manage to drive prices down (if the non-participant is a seller) or up (if a
buyer) so that the non-participating trader can’t earn any surplus by bidding alone.
For the moment, just assume that this is what will happen to a non-participant.
Then the individual rationality constraint simply requires all participants earn non-
negative surplus.

The best possible outcome for both sellers in this set is one that maximizes

\[
\pi \left( \frac{1}{2} p_{hh} + \frac{1}{4} p_l \right) + (1 - \pi) \frac{1}{2} p_{hl}.
\]  

(2.3)

The sellers’ best expected surplus is found by maximizing this expression subject
to the constraints above.

To understand the solution to this problem, observe that the low value bidder
only trades at price \( p_l \). So he won’t be willing to participate if \( p_l > v_l \). It is then
immediate that \( p_l = v_l \) at the solution to this problem. Then most of the solution
can be gleaned from the following Figure 2.1, which represents these two incentive
conditions for the case when \( v_h \) and \( v_l \) aren’t to different, and \( \pi \) is high.

The steeper of the two curves in figure describes the set of \((p_{hh}, p_{hl})\) pairs that
make the high type buyer indifferent between revealing his type and pretending to
be a low value buyer. This presumes that the price when both buyers claim to have
low values is \( v_l \). The high value buyer pays prices \( p_{hh} \) and \( p_{hl} \) when he trades. So
if these are too high, he will be better off pretending to be low value and getting
nothing. The set of price pairs that are incentive compatible for the high value
buyer are those below the curve for this reason.

The flatter of the two curves\(^3\) represents the set of price pairs that make the low
value buyer indifferent between revealing his type truthfully and pretending to be
high value. Reversing the reasoning above, if the prices \((p_{hh}, p_{hl})\) are too low, the
low value buyer will want to pretend to have a high value so he can buy at these
low prices. As a result, the prices that are incentive compatible for the low value

\(^3\)The curves have different slopes because the low and high value buyer have different beliefs
about whether or not the other buyer has a high value.
buyer are those above this curve. The set of prices that are incentive compatible for both are those in the shaded area.

The sellers’ iso-profit function has the same slope as the steeper of the two curves, as is readily seen by comparing (2.1) and (2.3) above. As a result, any pair of prices on the upper right edge of the shaded triangle constitute a solution to the problem defined above.

This diagram has been constructed for a situation where inefficient equilibria don’t exist when the double auction is played without any contracting. If \( v_h \) were increased sufficiently to support such an equilibrium, then the steep line would shift upward and to the right until it no longer crosses the flatter line in the positive orthant. The set of supportable outcomes would become a quadrilateral. The inefficient equilibrium would involve price pairs that lie somewhere on the interior of the line segment from \((0, v_l)\) to \((v_h, v_l)\) within the shaded region. Evidently, such equilibrium must then be less profitable for sellers than the collusive equilibria described her.

Efficient Bayesian equilibria also exist in which both sellers are sure to trade. The most profitable such equilibrium for sellers is the one where all bidders bid \( v_l \). This is less profitable than the most profitable collusive equilibrium described here provided that \( \pi \) (the probability both buyers have the same value) is high enough.

Many of these outcomes along the upper edge of the shaded triangle look quite odd. For example, at the top left of the shaded triangle is an outcome in which the price is very high when the buyers have different values, but falls to zero when both buyers have high values. In the main body of the paper, we will show how all of these outcomes can be supported as perfect Bayesian equilibrium in a reciprocal contracting game. To illustrate the logic, focus on a much more prosaic outcome in which the trading price is \( p^* > v_h \) when both buyers have high values, but falls to \( v_l \) when at least one of the buyers’ values is low.

**The example continued - the competing mechanism game.** This section describes a competing mechanism game and shows that there is an equilibrium of the game that implements the sellers’ favorite outcome, as described above. This game is an extensive game, so even though the example above is stylized, it is hard
to describe strategies and verify that they constitute perfect Bayesian equilibrium. Perhaps this illustrates why the characterization theorem uses an abstract approach, and why a characterization theorem might be useful in the first place.

The double auction described above is the default game. Sellers are going to build on that by writing programs that make price offers. The commitments associated with these programs are going to be readily identifiable, they will simply be fixed price offers, accompanied by a money back guarantee similar to the guarantees made on many airline ticket aggregator sites. Buyers who don’t see the commitments they expect will refuse the initial offer and proceed to bid in the double auction. The fact that all buyers can see the mechanisms acts as a coordinating device that triggers and equilibrium of the double auction that is bad for a deviating seller. When buyers see the right price offer with the right money back guarantee, their strategies will support the outcome described in the previous subsection.

The competing mechanism game begins with sellers writing their computer programs, and with buyers and sellers sending messages to a public website. Think of this as Twitter, or something comparable, where messages are ostensibly related to the product that is being traded, but in which the messages on their own have no meaning. The collection of messages, all voluntarily contributed by buyers and sellers, are used later in the game to correlate the actions of sellers. At the end of this opening stage, all the comments are made public.

Next buyers visit sellers website and receive price offers which they can accept or reject.

Next the double auction opens and sellers and buyer submit asks and bids.

Finally all trades are completed.

The set of feasible programs are all those that commit to sellers based on observables in the game. Since the point of this section is just to show an example, all we need to say about the set of feasible offers for sellers is that it contains all the mechanisms in the following class - the seller offers a fixed price; if any buyer is ever able to purchase at a price no higher than the seller’s from another seller, then the seller promises to refund the buyer money with return of the product.

On the equilibrium path, both sellers offer buyers price \( p^* \) along with a money back guarantee that promises that if buyers are able to purchase the good in the auction at a lower price they will receive a full refund. On the public website where buyers and sellers send their 'tweets', each buyer and seller posts a number between 0 and 1. These tweets will be aggregated in a manner that will determine, through the sellers’ programs, which seller holds out for a high price.

High value buyers accept the sellers initial offer \( p^* \) along with the guarantee. Low value buyers refuse the offer and wait for the double auction.

If both buyers have high values, there are no sellers left to participate in the double auction, so both buyers pay \( p^* \). If one of the buyers has a low value, then one seller will offer his good on the double auction. Participation strategies of the bidders then ensure the high value bidder can bid and win the good in the double auction at price \( v_l \), then have his money refunded on his original purchase. To bring this off, we allow the bidders to observe who participates in the double auction and choose their strategies accordingly. The randomized messages are used

---

4For example, Cheapair.com has something they call a “Price Drop Payback” - literally Buy your ticket, See fares Drop, Get Money Back - only at Cheapair.
to coordinate the participation decisions of buyers and sellers in a manner described below.

As a consequence when there is only a single high value bidder, the high value bidder buys at price $v_l$ while the low value trader is shut out, as required in the mechanism above.

Finally if both bidders have low values, the randomized messages coordinate the entry decisions of the bidders so that only one of them buys at price $v_l$.

If either of the sellers deviates and makes an alternative offer at the beginning of the game, all offers are refused and low bids are submitted to the double auction. This is the reciprocal contracting part that punishes a deviating seller. Buyers see the deviation when they see the sellers’ initial offers.

The strategies are described in detail in the appendix, along with the proof that they constitute a perfect Bayesian equilibrium.

A number of comments are in order. Notice that to understand this market, it isn’t enough simply to analyze the perfect Bayesian equilibrium of the double auction. Sellers have little difficulty in working their way around the rules of the auction. To tell whether the double auction is a good or bad institution, one has to try to understand all the equilibrium outcomes that could be supported when the double auctions acts as the default game. A good institution is one where all the equilibrium outcomes of all the associated competing mechanism games are reasonable. To discover this, one needs to find a method that can be used to characterize supportable outcomes for any default game. The point of the next section is to provide that characterization and method.

Second, a relatively brief proof is given above that the strategies constitute a perfect Bayesian equilibrium. This proof is going to become too cumbersome to follow when there are more buyers and sellers. A general proof is going to require a different approach. The method below is based on a variation on contractible contracts ((Peters and Szentes 2012)) called reciprocal contracting. We now proceed to the general case.

3. Incomplete Information Games and Mechanism Design

The basic approach in what follows is to add the contracting game on top of a basic default game of incomplete information as in the example above. In default game, there are $n$ players. Each player has a finite action set $A_i$ and a finite type set $T_i$. In standard notation $A$, $A_{-i}$ represent cross product spaces representing all players actions and the actions of all the players other than $i$, respectively. Similarly, define $T = \prod_i T_i$, and $T_{-i} = \prod_{j \neq i} T_j$. Types are jointly distributed on $T$ according to some common prior.

Let $q$ be a mixture over the set of action profiles $A$, i.e., $q \in \Delta (A)$. For any action profile $a$, we write $q_a$ to be the probability of $a$ under $q$, and $q_{a_i} = \sum_{a_{-i}} q_{a_i}$. We use notation $q_{A_i}$ to represent the marginal distribution over $A_i$ and $q_{A_{-i}}$ to be the marginal distribution over $A_{-i}$. We assume that players have expected utility preferences over lotteries. Then players preferences are given by $u_i : \Delta (A) \times T \rightarrow \mathbb{R}$ where $u_i$ is linear in $q$. An outcome function is a mapping $\omega : T \rightarrow \Delta (A)$. So player $i$’s payoff from this outcome function is $E \{ u_i (\omega(t), t) | t_i \}$.

Notice that an outcome function here is a description of a set of actions that players use in the default game rather than a set of outcomes in the usual sense. Of course, the default game itself is an indirect mechanism. Ultimately the point
of all this is to evaluate different default games. However, at this stage, this game is taken to be fixed.

An outcome function $\omega$ is implementable (by a mechanism designer) if the usual incentive compatibility and individual rationality conditions hold. Formally, an outcome function $\omega$ is *incentive compatible* if for every $i$, $t_i$ and $t'_i$,

$$E\{u_i(\omega(t), t) | t_i}\geq E\{u_i(\omega(t', t_{-i}), t) | t_i\}. \quad (3.1)$$

This is completely standard so there is no need to discuss it further. What is different in our approach is what happens when a player refuses to participate in the mechanism designer’s scheme. In that case, instead of an exogenous payoff, it will be assumed that the default game takes precedence.

When some player unilaterally refuses to participate, we allow the mechanism designer to implement a ‘punishment’ involving a mixture over profiles of actions of the participating players.

Furthermore, this punishment is written as if the non-participant is involved in his own punishment. Of course, the action of the non-participant in the default game can’t be controlled by the mechanism designer. So from the non-participant’s perspective, the action he is supposed to take should be interpreted as a recommendation from the mechanism designer. The IR constraint should then involve the usual sort of obedience condition that requires that conditional on receiving some recommendation from the mechanism designer, the non-participant actually wants to take that action.

Let $\rho^iT_{-i} \rightarrow \Delta(A_{-i})$ be an outcome function that is implemented when player $i$ chooses not to participate in the mechanism that implements $\omega$. We refer to this outcome function as a *punishment*. The outcome function $\omega$ is *individually rational* if there is a collection of punishments $\{\rho^i\}_{i=1,n}$ such that for every player $i$,

$$E\{u_j(\omega(t), t) | t_i\} \geq$$

$$E\{u_i(\rho^i(t), t) | t_i\} \geq$$

$$\max_{\alpha_i} E\{u_i(a_i, \rho^i(t_{i-1}), t) | t_i\}. \quad (3.2)$$

This is just the constraint used by (Myerson 1997) to describe Bayesian equilibrium outcome functions for a specific 'default game' of incomplete information. What we are going to show is that, provided there are enough players, the same set of outcome functions can be supported as Perfect Bayesian equilibria in some regular competing mechanism game that embeds the default game. Notice a number of things - Bayesian equilibria in the default game is replaced by Perfect Bayesian equilibrium in a regular competing mechanism game. The punishments in Myerson don’t have to be sequentially rational, while they do here.

Secondly, the equivalence only holds for regular competing mechanism games. The next section describes competing mechanism games more generally, and introduces the concept of regularity. Also, it is possible to implement all the outcome functions supportable in some conceivable competing mechanism game with one special, but somewhat abstract, game. This ‘game’ is defined after the next section. The formal statement of the results follow. The proofs of the main theorems are in the appendix.
3.1. **What is a Competing Mechanism Game.** There are probably countably many ways to model competing mechanism games. Examples come from the competing auction literature (for example, (Epstein and Peters 1999), (Yamashita 2010) or (Peters and Troncoso-Valverde 2013)) or the literature on common agency ((Pavan and Calzolari 2001) or (Martimort and Stole 2002) or (Bernheim and Whinston 1986)) in which mechanism designers simultaneously offer mechanisms which make commitments based on a specific group of players called agents. However a useful description should also capture models in which mechanisms are offered sequentially, as in (Pavan and Calzolari 2009) or privately as in (Segal and Whinston 2003) or (Duquietd and Martimort 2012).

We interpret a competing mechanism game as an extensive form game of incomplete information. We interpret the nodes of this game as opportunities for players to send messages. A path through the game is an ordered sequence of messages. Some of these messages convey commitments, some type information, while some are just cheap talk. In order to interpret the messages, we use an outcome function $\lambda$ which assigns a profile of actions to each path through the game tree. The profile of actions indirectly determines each player’s payoff.\(^5\)

The picture that follows shows the reciprocal contracting version of a simple prisoner’s dilemma game (with the cheap talk part left out to make it simpler).

\[\begin{array}{c}
\theta^* \\
\theta_{cd} \\
1 \\
2 \\
\theta_{cd}
\end{array}\]

In this game, players announce public messages representing commitments over two rounds. In the first round, each player can offer a contract that conditions directly on the other player’s contract. This is the contract $\theta^*$ in the picture. The

---

\(^5\)The fact that messages only determine profiles of actions is what distinguishes the approach here from the revelation principle. A standard direct mechanism could be interpreted exactly this way except that any history of messages will generate a probability distribution over profiles of actions.
alternative is a contract $\theta_{cd}$ that allows the player to defer his choice until the second round. The outcome function $\lambda$ is displayed on the far right of the picture. Notice that if player 1 announces the message $\theta^*$ in his first information set, then the outcome function forces him to use action $c$ in every history in which player 2 uses message $\theta^*$, and to use action $d$ in every other history following that choice. So from the outcome function $\lambda$, the interpretation of the message $\theta^*$ is that it is a reciprocal contract that commits player 1 to use action $c$ if player 2 sends signal $\theta^*$, and to use $d$ otherwise.

Any set of behavioral strategies specify a possibly random path consisting of a sequence of messages. The outcome function $\lambda$ converts every sequence of messages into a profile of actions for the players. Player $i$’s payoff in the history in which players send the sequence of messages $m$ is given by $u_i(\lambda(m), t)$. In the extensive form version of the reciprocal contracting game pictured above, the history of messages $\{\theta^*, \theta^*, d, d\}$ supports the profile of actions $\{c, c\}$.

Let $\{\sigma_i, b_i\}_{i=1,...,n}$ be behavioral strategies and beliefs for the players specifying mixtures over messages available to players in each of their information sets, and beliefs about the history of play prior to the information set. Let $\iota$ be an information set for player $i$. The continuation game associated with $\iota$ is the extensive form game of incomplete information in which each player’s type is his payoff type from the original game along with his information about the history of play prior to $\iota$. Beliefs for player $i$ in this continuation game are given by $b_i(\iota)$. For every other player $j$, the player’s type $t_j$ in the continuation game describes (among other things) the most recent information set $\iota_j$ in which he sent a message. So player $j$’s belief in the continuation game coincide with his beliefs in the information set $\iota_j$. Associated with each history $h \in \iota$, there is an outcome function that describes type contingent mixtures over action profiles when all players use the continuation strategies associated with $\{\sigma_i, \sigma_{-i}\}$ from the information set $\iota$ onward. Using $i$’s beliefs in the information set $\iota$, we write $\rho(t_i, t_{-i}|\sigma_i, \sigma_{-i}, \iota)$ as the outcome function conditional on attaining this information set when players are using the continuation strategies associated with $\{\sigma_i, \sigma_{-i}\}$.

Given an array of behavioral strategies $\{\sigma_i, \sigma_{-i}\}$, a collection of information sets $\mathcal{I}$ is attainable with probability $\pi$ by player $i$ in the continuation game associated with $\iota$ if there is a continuation strategy for $i$ at $\iota$ such that an information set in $\mathcal{I}$ is reached with probability at least $\pi$ given $i$’s beliefs $b_i(\iota)$ and the continuation strategies $\sigma_{-i}$.

An information set $\iota$ for player $i$ has the no-commitment property if (i) the outcome function $\rho_{A_{\iota, i}}(t_i, t_{-i}|\theta_{i}, \sigma_{-i}, \iota)$ is independent of $\theta_{i}$, and (ii) for each $a_i \in A_i$, there is a strategy $\sigma_i'$ such that $\rho_{A_{\iota, i}}(t_i, t_{-i}|\sigma_i', \sigma_{-i}, \iota)$ assigns probability 1 to the action $a_i$. In words, a no-commitment information set is one in which $i$ can carry out any action he likes without changing the behavior of the other players. We say that a player $i$ is uncommitted in information set $\iota$ if he has a continuation strategy that attains an information set having the no-commitment property with probability 1.

**Definition 1.** A competing mechanism game is said to be regular if for every profile $\sigma$ of strategies, each player $i$ has a strategy $\sigma'_i$ that attains some no-commitment information set with probability 1.6

---

6This definition is inspired by a similar assumption in (Peters and Szentes 2012).
This restriction is imposed because we are interested in adding contracts that enhance players’ strategy sets, not in contracting games that impose arbitrary restrictions on what players can do. For example, consider the complete information game of matching pennies (with payoffs 1 and $-1$). This game has a unique Nash equilibrium in which both players’ payoff is zero, which makes it pretty predictive by game theoretic standards. We already know we could change the outcome of this game by removing actions, or changing timing. We want to know whether contracts that both players would want to use might change the set of equilibrium outcomes for the game.

Suppose we specify the following contracting game: player 2 is allowed to choose one of two contracts. The first commits him to tails, the second to heads. We now allow player 1, still moving simultaneously with player 2, to commit himself in a manner that depends on the commitment made by player 2. The only equilibrium would then have player 1 committing to match (or mismatch) the commitment of player 2. Payoffs would then be 1 for player 1 and 0 for player 2.

In this example, we would say the contracting game is not regular for player 2. If player 1 is using his equilibrium strategy, then there are no strategies available to player 2 that allow him to change actions without simultaneously changing player 1’s response. The contracting game we build on top of the matching pennies is simply depriving player 2 of the ability to select his action simultaneously with player 1.

It is possible to use the methods we describe below to analyze irregular games. For example, in the matching pennies example, if the asymmetric contract structure seems the right one for some reason, then we could analyze it by changing the original game from matching pennies to sequential matching pennies. The contracting game would then be regular with respect to this sequential game.

The first part of our theorem is then given in the following:

**Theorem 2.** Suppose the outcome function $\omega$ can be supported as a weak Perfect Bayesian equilibrium in some regular contracting game. Then there is a collection of punishments $\{\rho_i\}_{i=1}^n$ such that (3.1), and (3.2) hold.

The formal proof is in the appendix, but the logic is straightforward. If the competing mechanism game is regular, then any player can deviate to the strategy that leads to a no-commitment information set. This deviation induces some continuation play that determines the punishment $\rho_i$. Since this is a deviation from an equilibrium in the game, it must be unprofitable for the deviator no matter what he does in the continuation. This guarantees that (3.2) holds.

3.2. **Reciprocal Contracting.** We turn now to the reverse problem. If there is a collection of punishments $\rho_i$ that make an outcome function $\omega$ implementable, we want to show that there is an equilibrium of some regular competing mechanism game that supports $\omega$ as a perfect Bayesian equilibrium. It is tempting to conclude that since Theorem 2 already shows that outcome functions can only be supported if they satisfy (3.1) and (3.2), then we could just use the generalized revelation principle as stated by (Myerson 1997) to conclude that the outcomes can be implemented with direct mechanisms. With this interpretation, there seems nothing more to prove.

This is unsatisfactory for two reasons. First, it requires an outside mechanism designer to process information then instruct each player individually about which

---

7This is how the meet the competition argument works.
action he or she should take. This is unrealistic in large digital markets in which mechanism designers only make commitments about their own actions, and only communicate with their own 'agents'. The whole question here is whether these outcomes can be implemented with indirect mechanisms that satisfy these two requirements.

The second difficulty with direct mechanisms is that a centralized mechanism designer has to be able to control not just the messages that agents can send to him, but also has to be able to prevent them from trying anything else. The view taken here is that a centralized mechanism designer can do no more than to set the rules of the default game. The reciprocal contracting game is the 'default game'. It plays the same role as the double auction did in the example at the beginning of the paper.

Finally, the reciprocal contracting game uses the approach in (Peters and Szentes 2012) in which contracts can condition directly on other contracts. However, what (Peters and Szentes 2012) assume is that players declare publicly is a verbatim description of their mechanism - for example, they might publicly post the computer code they use to determine their price. This code is no different in principle from any other message they send, so other players should be able to condition on this message. However, in order to simplify the logic slightly, the reciprocal contracting game assumes that players can describe their commitments ‘parametrically’. What this means is that there is a fixed set of messages that players can send at the beginning of the game. Each message in this set describes a commitment. The reciprocal contracting game assumes that players can describe their commitments ‘parametrically’. What this means is that there is a fixed set of messages that players can send at the beginning of the game. Each message in this set describes a commitment. The rules of the default game (which may have been set by some centralized mechanism designer) force the player to obey the commitment associated with any message that he sends.

An analogy is simply the way we typically describe a fixed price mechanism. Literally, the fixed price mechanism names a price then asks the 'agent' to name a quantity. When sent together, the two messages commit the seller to deliver the quantity named by the buyer, and commit the buyer to pay the seller the product of the two messages. Instead of writing all that down, we simply allow the seller to describe his commitment parametrically by naming a price.

The advantage of using this obvious analogy here is that we can write down contracts that are self-referential (in the terminology of (Peters and Szentes 2012)) without having to worry about any infinite regress. As will be apparent below, the parametric approach makes it possible to think through the reciprocal contracting game using logic that is almost as simple as the well know Nash Demand Game.

The reciprocal contracting game takes place over three stages. In the first stage, each player sends a public message, part of which describes his mechanism, part of which describes an encrypted type report; and part of which describes an encrypted number in the interval [0, 1] called a correlating message. At the same time each player sends private messages to each of the other players describing encryption keys which can be used to decode the encrypted part of their public message. In the second stage players tell each other what private messages they received at the end of the first stage. All messages from the first two stages are ultimately verifiable and can be used to make commitments. The third stage of the game is a trivial one in which any player who is uncommitted chooses his action in the default game.

Mechanisms are commitments that are based on the various messages players receive during the game. The mechanisms are described by messages in a set $\Delta$ to
be defined formally momentarily. As in (Peters and Szentes 2012), the commitments associated with the various mechanisms can depend on the public messages that other players send to describe their mechanisms as well as on private messages.

The public messages involve encrypted variables. The treatment here is taken from (Peters and Troncoso-Valverde 2013). Encryption begins with a publicly known function $\kappa : T_1 \times [0, 1] \times K_1 \times K_2$. The sets $K_1$ and $K_2$ contain encryption keys. The function $\kappa$ must have two properties. First for any $k_1$ and $k_2$, the function $\kappa (\cdot, k_1, k_2)$ must be one to one and onto some set $\mathcal{E}$. Secondly, for any $k_i \in K_i$ and $e \in \mathcal{E}$, the set $\{(t_i, x) \in (T_i \times [0, 1]) : \exists k_j \in K_j; \kappa ((t_i, x), k_i, k_j) = e\} = T_i \times [0, 1]$. In words, knowing the encrypted value of some variable and one of the keys conveys no information at all about the unencrypted value of the variable. To make the notation a bit simpler, it is assumed that $T_i$ is the same for each player. Then define $\mathcal{E}$ to be the image of the function $\kappa$ over the domain $T_1 \times [0, 1] \times K_1 \times K_2$.

The private messages are assumed to come from $K_1 \cup K_2 \equiv K$. Then in the first round, each player sends a public message in $\triangle \times \mathcal{E}$, and a private message in $K$ to each of the other players. At the beginning of the second round, each player will have seen each of the public messages describing commitments, and will have received $n - 1$ messages from the other players giving some information about their encryption keys. Each player reports the $(n - 1)$ keys he received at the end of the first round to each of the other players. At the end of the second round, each of the other players will have sent a given player $n$ messages. This means each player has $n \times (n - 1)$ private messages from the other players that he must use in some way to figure out their encryption keys.

At the risk of adding notation, just assume that the set of private messages that a player receives is represented by the conventional notation $M$, where, of course, $M = K^{n(n-1)}$. In our approach, the message space which is processed by a mechanism is $(\triangle^n \times \mathcal{E}^n \times M)$ and a mechanism for player $i$ will be a mapping from each of the points in this message space into his action space $A_i$. For reasons that will become apparent later, there is no need to consider stochastic mechanisms.

To describe the mechanisms that players are going to use in equilibrium, it is necessary to describe how the private messages are used. Each player is going to commit to a method of translating the private messages in $M$ into a pair of encryption keys for each of the players. The encryption keys will be used to decrypt the public messages in $\mathcal{E}^n$. The decrypted versions of these messages will be treated as type reports (and correlating messages) submitted to a direct mechanism.

The exact method of interpreting the private messages will be borrowed from (Peters and Troncoso-Valverde 2013). Discussion of the details of this conversion can be deferred to an appendix, so for the moment define the mapping $\tau_i : M \rightarrow (K_1 \times K_2)^n$ which simply takes a collection of these private messages and uses them to assign a pair of encryption keys to each player. This mapping will represent a part of the commitment that player $i$ makes at the start of the game.

To begin, let $\hat{T} \equiv \prod_j \{T_j \times [0, 1]\}$ be the set containing type declarations and correlating messages of all the players. Let $D_i$ be the set of measurable mappings $d_i : \hat{T} \rightarrow A_i$. We refer to each of these mappings as a direct mechanism even though the message space is bigger than the type space. This should create no confusion below. Players are not allowed to offer direct mechanisms (though they can emulate them) in the reciprocal contracting game. Instead, they specify their commitments by using messages in a special set.
The parameter that defines a mechanism is a list \( \delta^i = \{\delta_1^i, \ldots, \delta_n^i\} \) where \( \delta_j^i = \{d_j^i, \{p_j^i\}_{k \neq j}\} \in D_j^i \). We describe the commitments associated with such a message below. To think about the message that describes a mechanism it might help to think of the analogy with a proposal. For example, \( \delta_j^i \) is a part of the message describing the mechanism used by player \( i \). It could be interpreted to be the direct mechanism that \( i \) proposes that player \( j \) use, along with the punishments that player \( i \) thinks player \( j \) should impose when any player \( k \) 'deviates'. In this sense, each of the messages describing mechanisms could be thought of as a proposal about how the game should be played by each of the players.

The set of messages \( \Delta \) that each player can use to commit himself is then the set of all messages \( \delta^i \). Each of these messages is associated with a commitment that translates all the messages the player observes into the action that he takes. The messages he observes are the commitment messages of the other players, referred to here as \( \delta^{-i} \); the encrypted public messages \( (e^i, e^{-i}) \in \mathcal{E}^n \); and all the private messages \( m^i \in M \) that he receives by the end of the second stage. A mechanism for player \( i \) is a mapping from \( \Delta^i \times \mathcal{E}^n \times M \) into \( A_i \).

As discussed above, fix a mapping \( \tau \) from private messages in \( M \) to pairs of encryption keys. A precise description of \( \tau \) will be given below. The commitment associated with the message \( \delta^i \) is given by

\[
\lambda_i (\delta^i, \delta^{-i}, (e^i, e^{-i}), m) = \begin{cases} d_i (\kappa^{-1} ((e^i, e^{-i}), \tau (m^i))) & \exists \delta^* : \delta^j = \delta^* \forall j \\
p_j^i (\kappa^{-1} ((e^i, e^{-i}), \tau (m^i))) & \exists \delta^* : \exists j : \delta^j \neq \delta^* \land j \neq i \\
 & \text{otherwise.}
\end{cases}
\]

In this notation, \( p_j^i \) and \( d_i \) represent the corresponding elements of \( \delta^i \) and the notation \( \exists ! \) means "there exists a unique". The function \( \kappa^{-1} ((e^i, e^{-i}), \tau (m^i)) \) means that the process \( \tau \) is used by each player to compute a collection of encryption keys from the private messages. The process that each player uses to create the encryption keys is the same, but whether or not players use the same keys depends on what messages are sent.

One aspect of this that may deserve comment is the fact that when some player \( j \) is a unilateral dissenter, the others punish him by using mechanisms that make use of his type and correlating message. Of course, the others can ignore his type and correlating message if they want. If they choose to make use of that information, then the way they use it should be designed so that the dissenter is happy to report his information truthfully.

This formalism now makes it possible to state the main theorem.

**Theorem 3.** If there are seven or more players, any outcome function \( \omega \) satisfying (3.1), and (3.2), can be supported as a perfect Bayesian equilibrium outcome in the reciprocal contracting game. Along the equilibrium path of this game, all players make a common proposal \( \delta^* \), post encrypted values of their true type along with a correlating message chosen using a uniform distribution on \([0, 1]\), 'report' their encryption keys truthfully to the other players, and truthfully report the messages they receive to other players.

The full proof is contained in Section 5.4. This proof collects results from a number of other papers, so the details aren't new. A sketch of the problems that
need to be resolved in the proof, and the method used to resolve them might be helpful.

The first complication stems from the fact that the outcome function being implemented, \( \omega \), is a joint randomization that may include a lot of correlation between the actions of the players. This must be implemented by independent mechanisms. This is what turns the reciprocal contracting game into a regular competing mechanism game.

The device that ties the actions of the players together is the correlating messages that players encrypt along with their type reports. The proof adopts a method from (Kalai, Kalai, Lehrer, and Samet 2010) (and the generalization provided in (Peters and Troncoso-Valverde 2013)) which converts the private correlating messages into something that works like a public randomizing device that the players cannot manipulate. The logic in the proof explains precisely how the mechanisms that players implement along the equilibrium path can be constructed in order to implement the correct joint randomization.

The second complication emerges because players have to have a way to commit to a punishment before they know that some player has deviated. The reason this is important is because once players see that a deviation has occurred, carrying out a punishment may not be sequentially rational. The encrypted types they publish in the first period are there precisely to act as a commitment device. Furthermore, since the players know what deviation has occurred they could provide the others with information that would allow them to carry out a punishment that depends on the deviation. This is what happens in the recommendation game described in (Yamashita 2010). If they can do that, then the competing mechanism game is no longer regular. The set of supportable outcomes should then be larger than the set described here, though there are no formal results about this.

The fine line that the proof has to tread comes from the fact that the first period messages have to indicate that the players have already committed to a punishment without actually sending any information about the types of the players. This is why the messages are encrypted. At the same time, the players have to be able to decrypt these messages in an appropriate way once they learn what the others have offered in the first period. This is where the second round of communication plays a role. The method followed is borrowed from (Peters and Troncoso-Valverde 2013).

This is where the requirement that there be 7 players emerges. The reason for the restriction is that the incentive for accurate communication in the second round is created by the belief that the others are communicating accurate information and that a message that disagrees with the messages being sent by others will be ignored. Each player needs to figure out two encryption keys. A player that learns both keys in the first round of communication can determine another player’s type, and this will ruin incentives. For this reason each player has to convey each of his encryption keys to two different groups of players in the first round.

In the second round, when a player decides whether or not he should accurately convey the key that he received in the first round, he has to believe that there are at least two other players who have received the same key and that they will report this key truthfully to the other players. This means that in the first round each player must send each of his encryption keys to a distinct group of players, each
of which contains at least 3 of the other players. The sender, plus the two distinct groups of 3 players each makes a total of 7 players.

The communication strategies in the second stage depend in a critical way on the outcome of the first stage. If there are more than two distinct proposals made in the first stage, then all players are expected to convey random noise in the second round. Since the message space is a continuum, the probability that two messages agree is zero. The interpretation process $\tau$ ignores the second round messages in this case, so it is sequentially rational for all players to send random noise.

The other special case occurs when there is a unilateral deviation in the first stage. In that event all players are expected to send random noise to the deviator in the second round, but accurate messages to the non-deviators. Once again, this behavior is self enforcing when all the other players are expected to follow it.

4. The Equivalence of Competing Mechanisms and Reciprocal Contracting

Combining this theorem with Theorem 3 gives the following corollary:

**Theorem 4.** An outcome function $\omega$ is supportable as an equilibrium in a regular competing mechanism game with seven or more players, if and only if it is supportable as an equilibrium in the reciprocal contracting game.

Recall that the reciprocal contracting game is intended as a default game. Players could try to devise ways around outcomes that they didn’t like by offering their own contracts outside the game. What Theorem 4 shows is that even if they did this, no new equilibrium outcomes could be supported. In this sense, the reciprocal contracting game is robust to the invention of new trading methods.

From the double auction example described at the beginning of the paper, robustness is not a property that is shared by double auctions. Even though all the equilibrium outcomes of the double auction are good, they don’t tell the whole story. An enlargement of the set of trading mechanisms can increase the set of supportable outcomes.

The reciprocal contracting game is not the only robust indirect mechanism. From Theorem 2, it is apparent that the game described in (Peters and Troncoso-Valverde 2013) is also a robust indirect trading mechanism. The difference between the mechanisms described here and those in (Peters and Troncoso-Valverde 2013) is that the latter implements outcomes without relying on contracts that directly condition on one another.

**Conclusions**

Theorem 3 characterizes the set of outcome functions that can be supported as equilibrium outcomes in regular competing mechanism games. Theorem 4 shows that all these outcome functions can be understood using reciprocal mechanisms. In this sense reciprocal mechanisms play the role of direct mechanisms in competing mechanism games.

The main implication of the Theorem is that for many markets, digital markets being an example, it makes little sense to try to specify a precise model of how the market operates, then analyze its equilibrium outcomes since many of the details of players’ interactions are going to be unknown or misunderstood by the modeler. Instead, the performance of an institution (like the double auction in our example)
should be evaluated relative to the set of outcomes it supports, as described by Theorem 3. The default game will generally support many outcome functions, but for any welfare criteria, the best and worst of them are easy to understand because the characterization takes the form of a set of inequalities that have to be satisfied.

Like direct mechanisms, reciprocal mechanisms make it possible to understand equilibrium outcomes with competition without worrying about the intricacies of particular indirect mechanisms that are used in practice. Apart from the standard logic of incentive constraints, reciprocal mechanisms simply add the logic that if everyone else wants to do something, it is simple to write a contract that commits you to do it too.

It might seem that the reciprocal contracting approach assumes too much commitment. For example, in the simple example given at the start of the paper with a double auction, reciprocal contracting would assume that players commit themselves to make bids in the double auction that are not sequentially rational. Of course, the whole point of mechanism design is that one can benefit by committing themselves ex ante to behave in a way that may not be sequentially rational at some point. In fact the practice is common. The eBay robot, for example, commits to an automatic bidding scheme that binds a bidder to one auction and bids there on a buyer’s behalf. Automatically bidding this way isn’t sequentially rational if there is another auction where the good is available at a lower price.

However, the real issue is not so much whether commitments are plausible within the context of a specific institution. One of the issues that this paper is supposed to address is what to do when most of the details of the trading mechanism either aren’t known, or are incompletely understood by the mechanism designer. So in a sense, the whole issue is about the robustness of trading mechanisms in response to innovation. From a general perspective, the right way to proceed is to assume that any commitment based on a verifiable message is at least potentially enforceable.

The theorem here suggests at least two directions for further research. First, are there other more natural looking indirect mechanisms in which message spaces are simpler than they are here. One example of this sort of approach is (Xiong 2013) who shows how to implement outcomes satisfying (3.1) and (3.2) as Bayesian equilibria using message spaces in which each player reports a number between 0 and 1 to each of the others. The jist of his idea is to break the interval \([0, 1]\) up into as many sub intervals as there are principals in the game plus 1. Each sub interval is isomorphic to \([0, 1]\) and so can be used to convey any required type information. The sub interval within which the report is contained acts as an indicator of which principal has deviated from some putative equilibrium ‘agreement’, with the extra interval included for the case where no deviation occurred.

A related question is whether well known mechanisms are actually robust in the right environments. For example it seems plausible that a double auction might be robust in a large enough market in the sense that very collusive outcomes become less likely with large numbers. One nice property of the theorem here is that this question can be answered without relying on complex equilibrium arguments - the answer can be discerned by using the incentive compatibility and individual rationality inequalities with the double auction as the default game, then increasing the number of participants.
5. Appendix: Proofs

5.1. Strategies and Equilibrium in the Example. The rest of this section describes strategies in detail, and illustrates how the buyers’ and sellers’ tweets on the public website are aggregated to coordinate the sellers pricing strategies.

Buyers make no commitments at all in this story, so it is a bit harder to describe their strategies. They look like this:

- if both sellers offer the mechanism $m^*$, all four players should choose a number in $[0,1]$ uniformly and report it to the feedback website. Once all the feedback results are published on the website, the fractional part of their sum will be referred to as $\tilde{x}$.\footnote{The fraction part of 3.25 is .25.}

  - (b0) in any history in which all four players are participating in the auction, bid 0;
  - if type is high
    * (b1) accept one of the sellers’ offers then bid $p^*$ in the auction if the other buyer chooses to participate in the auction and $v_l$ if the other buyer stays out of the auction.
  - if type is low
    * (b2) bid $v_l$ in an auction with 2 buyers and 1 seller, bid 0 if the only other bidders in the auction are sellers;
    * (b3) reject all offers, then stay out of the auction if one of the sellers’ website offers was accepted.
    * (b4) If both sellers’ offers were rejected, and you are buyer 1 stay out of the auction if $\tilde{x} < \frac{1}{2}$, otherwise enter the auction and bid as above;
    * (b5) if both offers were rejected and you are buyer 2, stay out of the auction if $\tilde{x} \geq \frac{1}{2}$, otherwise enter the auction and bid as above.
- (b6) Finally if either seller offers a mechanism other than $m^*$, reject all offers, enter the auction and bid as above.

The seller’s strategy goes like this

- (s0) in any history in which all four players bid in the auction, bid 0;
- (s1) bid $v_l$ in the auction if the other seller’s offer was accepted and only one buyer participates in the auction, bid $p^*$ in this case if both buyers bid in the auction;
- (s2) if the other seller’s offer was rejected, you are seller 1, and $\tilde{x} < \frac{1}{2}$, give up and stay out of the auction. Seller 2 should do the same thing, except to replace $\tilde{x} < \frac{1}{2}$ with $\tilde{x} \geq \frac{1}{2}$.
- (s3) if the other seller’s offer was rejected, you are seller 1, and $\tilde{x} \geq \frac{1}{2}$, bid $v_l$ in the auction if the other seller chooses not to participate and bid 0 otherwise. As above, use the same rule replacing substituting $\tilde{x} < \frac{1}{2}$ for seller 2.
- (s4) if the other seller offers a mechanism other than $m^*$, bid as described above in the double auction.

These strategies implement the sellers’ preferred outcome as described above. If they both offer the mechanism $m^*$ and the buyers both have high types, then by (b1) above, both buyers will accept the sellers offer $p^*$. Since nothing will be offered
in the auction, both buyers pay $p^*$ in this case. If both the buyers have low types, then by (b4) and (b5), one of the two bidders will stay out of the auction, the other will bid $v_l$. By (s2) and (s3), only one of the sellers will bid in the auction and the bid will be $v_l$. This ensures trade by one of the buyers at price $v_l$. If the buyers have different types, one of them will accept one of the sellers’ offers to trade at $p^*$, then bid $v_l$ in the double auction, which he will win. So the high value bidder pays $v_l$.

Notice in this outcome how the feedback website serves as a correlating device which picks the seller who withholds output. One surprising feature of this technique is that if the other three players are all expected to choose their message uniformly on $[0, 1]$, then the fourth player can’t manipulate the correlating device in the sense that the $x$ is uniformly distributed on $[0, 1]$ independent of the fourth player’s report.\footnote{See (Kalai, Kalai, Lehrer, and Samet 2010) or (Peters and Troncoso-Valverde 2013) for a proof for the case of more than two players.} So ex ante each of the sellers has probability $\frac{1}{2}$ of being the one who forgoes a profitable trade.

The proof that these strategies constitute a perfect Bayesian equilibrium is a bit tedious since there are so many deviations.

The simplest case occurs when one of the sellers deviates and offers a mechanism other than $m^*$, in which case all the players are expected to bid 0 in the double auction without accepting any other offer. This is from the strategies as defined by (b6), (b0), (s4) and (s0). This is a Bayesian equilibrium of this continuation game as we described above. Since sellers’ earn no profits in this continuation equilibrium, it can’t be a profitable deviation.

Otherwise we can start at the end with the auction. If all four players participate, then all bids are expected to be zero. In that case it is a best reply for sellers and both types of buyers to bid 0 as specified by the strategies above. At the other extreme, if one buyer and seller participate in the auction, it is a best reply for both to bid $v_l$ since the seller expects the buyer to bid $v_l$ and loses a profitable sale if he asks for more. The outcome in any history like this is that the buyer purchases at price $v_l$. The only history in which there are two sellers and 1 buyer occurs when a seller who is supposed to stay out of the auction enters anyway. In that case, the buyer and non-deviating seller are both expected to bid 0. This ensures that the buyer trades at price 0 and no trader can benefit by altering their bid. Finally, the history where there are 2 buyers and 1 seller bidding in the auction is one in which a buyer who refused an offer tries to bid in the auction against a buyer who has already agreed to buy at $p^*$. In this case the seller and one of the buyers bids $p^*$. The other buyer cannot win and pay less than $p^*$ when this happens. The buyer who has previously paid $p^*$ may win the auction in this case (depending on what the deviator does), but loses nothing from this since he can have his money refunded by the original seller.

This makes it possible to check deviations that occur before the bidding in the auction starts. Start with the case where both the sellers’ offers are rejected. The seller who is supposed to bid expects to sell at price $v_l$ so it makes no sense for him to deviate and stay out of the auction. The seller who is supposed to stay out expects the other two bidders to bid 0 if he enters, so it is not profitable to enter. The buyer who is supposed to stay out of the auction in this history may or may not win if he enters. If he wins he should expect the price to be $v_l$, which is not
worth it if he is low value. If he is high value he can guarantee a trade at a price $v_1$, but notice that this is the same price he would have paid in this event if he had accepted one of the sellers’ original offers - we come back to this point in a moment.

At the other extreme, if both sellers offers are accepted, both buyers trade at the price $p^*$. A high value buyer could deviate and pretend to be low value, then deviate again and bid in the auction. As described above, he can’t trade at a price below $p^*$ in this case. At the interim stage, the high value buyer doesn’t know whether the other buyer is high or low value, but as described above, in either case he pays the same price by deviating that he would have paid by following his equilibrium strategy. This covers all the potential deviations.

5.2. Proof of Theorem 2.

**Theorem 5.** (Restatement of Theorem) Suppose the outcome function $\omega$ can be supported as a weak Perfect Bayesian equilibrium in some regular contracting game. Then there is a collection of punishments $\{\rho^i\}_{i=1,...,n}$ such that (3.1), and (3.2) hold.

**Proof.** Let $\omega(t)$ be the outcome function supported by some equilibrium of a regular competing mechanism game in which strategies are $\sigma^*$. It satisfies (3.1) by the usual revelation principle.

The approach in the proof is to pick a out of equilibrium behavioral strategy. Deviating to this strategy will be unprofitable. The punishment is going to be the response the other players make to this deviation. The problem is to show that this strategy can be chosen to have the properties described in Theorem 2.

The game is regular, so $i$ has a behavioral strategy $\sigma^i$ that attains a no-commitment information set with probability 1. Since the set of strategies available to players doesn’t depend on their type, we can choose this strategy so that the collection of no-commitment information sets it induces is the same for every one of the player’s types. Write $I(\sigma^i, \sigma^*_{-i})$ be the collection of no-commitment information sets in the support of the strategy pair $(\sigma^i, \sigma^*_{-i})$. By definition, each no-commitment information set has the property that there is a collection of strategies that ensure that no matter what pure action the player wants to play in the continuation, he can play that action and induce the same reaction from his opponents. Since this reaction can’t depend on the deviating player’s type, write it as $\rho_{A_{-i}}(t_{-i}|i)$.

The deviation we want to focus on is the one in which the player adopts some behavioral strategy that attains a no-commitment information set with probability 1. Then whenever the player attains such an information set, he chooses his favorite action conditional on updated beliefs, then implements it using the strategy that induces the common response $\rho_{A_{-i}}(t_{-i}|i)$. With a slight abuse of notation, call this strategy $\sigma^i$.

Of course, deviating to this strategy is unprofitable because $\sigma^i$ is part of a perfect Bayesian equilibrium. The payoff function that prevails when player $i$ uses this behavioral strategy $\sigma^i$ can be written

$$E\left\{u_i\left((\sigma^i, \sigma^*_{-i}), t\right) | t_i\right\} =$$

$$E_{t \in I(\sigma^i, \sigma^*)} \left\{ \max_{-a_i \in A_i} E\left\{u_i\left(a_i, \rho_{A_{-i}}(t_{-i}|i), t\right) | t_i, i\right\} \right\}$$

by the law of iterated expectations.
By Blackwell’s Theorem, this expression is no smaller than
\[
\max_{-a_i \in A_i} \mathbb{E}_{t \in I} \{ u_i(a_i, \rho_{A_{-i}}(t-i|\ell_i), t) | t_i, i \} = \\
\mathbb{E} \left\{ u_i \left( a_i, \mathbb{E}_{t \in I} (\sigma^*_i, \sigma^i) (t-i|\ell_i), t \right) | t_i \right\}
\]
by the linearity of preferences. This verifies that the collection of punishments
\[
\left\{ \mathbb{E}_{t \in I} (\sigma^*_i, \sigma^i) (t-i|\ell_i) =_{i=1,...,n} \right\}
\]
satisfies (3.2). □

5.3. The transformation \( \tau \). The most complex part of the reciprocal contracting game is the transformation \( \tau \) that is used to generate encryption keys from the collection of private messages \( m \) that a player receives. The details of this transformation are described formally in (Peters and Troncoso-Valverde 2013), so we provide a heuristic description here.

Focus on some player \( i \), and imagine that player \( i \) chooses two distinct groups of friends \( F^+_i \) and \( F^-_i \) from among the other players. The groups \( F^+_i \) and \( F^-_i \) are mutually disjoint, and each group consists of at least three of the others. Player \( i \) and his two groups of three friends each explains the restriction in the theorem to 7 or more players.

In the first round, player \( i \) is going to report his first encryption key to his friends in \( F^+_i \) and his second encryption key to his friends in \( F^-_i \). The single encryption keys don’t convey any useful information to his friends about his type - they need both keys to decode his public messages. Player \( i \)’s friends are commonly known by all players. Every other player does the same thing - chooses two distinct groups of friends and reveals one of his encryption keys to each group of friends privately in the first round.

In the second round, every player can try to learn \( i \)’s encryption keys by asking his friends about them privately. There is nothing that compels his friends to reveal the keys accurately, so the other players have to interpret the friends messages in a special way in order for the information to be revealed.

The way this is done is to imagine that when player \( i \) receives reports from \( i \)’s friends, he ignores the reports and uses some arbitrary encryption key, unless at least two of these reports agree. If two or more of the reports agree, then player \( j \) accepts this majority report as \( i \)’s actual encryption key.

The reason this works is that if each of \( i \)’s friends believes that \( i \) reported his encryption key accurately to each of them, and believes that each of \( i \)’s other friends are going to convey those reports accurately to other players, then there is nothing they can do to manipulate the interpretation of the reports since their contrarian messages will be ignored. It is therefore always sequentially rational for them to report the encryption key they heard accurately, no matter how the information is ultimately used.

Player \( i \) could try to manipulate this by sending different encryption keys to his friends, but if the outcome function being implemented is incentive compatible, he cannot profit by doing so.

A formal description of this process is complicated by two small details. First, the transformation \( \tau \) is supposed to be the same for every player. To accomplish this, the reciprocal contracting game assumes that \( i \) commits himself to use the encryption keys his friends report to him to decipher his own public message from
the first round, forcing him to interpret the public message the same way that other players do.

The second complication occurs when player $i$ is a friend of some player $j$ whose encryption key he needs to decrypt. In that case, $i$ commits to use the reports he hears from $j$’s other two friends along with the private message he received from $j$ in the first period.

To provide a formal description of $\tau$, let $m$ be the vector of $n(n-1)$ messages that $i$ receives from the other players, as described above. Let $m^j = \{m^j_1, m^j_2\}$ be the report that $i$ receives from player $j$. The first component of $m^j$ is the private message that $i$ receives from $j$ in the first round. The message $m^j_2$ is a vector of $(n-1)$ messages that player $j$ heard at the end of the first round. The notation $m^j_2$ is the encryption key that was reported to player $j$ by player $k$ at the end of the first stage of the reciprocal contracting game. Define

$$M^+_j(m) = \begin{cases} m' & m^j_2 = m' \forall k \in F^+_j \vee \exists ! m^k_2 \neq m'; k \in F^+_j \\ \emptyset & \text{otherwise.} \end{cases}$$

The notation $M^+_j(m)$ means the majority report of $j$’s friends in $F^+_j$. The idea is that if $j$’s friends in $F^+_j$ (to whom $j$ is supposed to report his first encryption key) all agree about the key they were given, or if only one of them disagrees, then $i$ should use the key they reported to decrypt $j$’s message. To compute the first of $j$’s encryption keys, $i$ would then use the formula

$$\tau^+_j(m) = \begin{cases} R & M^+_j(m) = \emptyset \\ M^+_j(m) & \text{otherwise.} \end{cases}$$

A similar formula describes the encryption keys that $i$ commits to use for each of the other players.

Notice that the transformation $\tau$ is part of player $i$’s mechanism. He commits to use it to decrypt the public messages. As such, the convention below is to adopt this transformation as part of the reciprocal contracting game.

With this last construction, it is possible to state and prove the main theorem of the paper.

5.4. Proof of Theorem 3.

**Theorem.** If the reciprocal contracting game has seven or more players, there is a Perfect Bayesian equilibrium that supports the outcome function $\omega$ if $\omega$ satisfies (3.1), (3.2). Furthermore, along the equilibrium path of this game, all players announce a common mechanism $\delta^*$, post encrypted versions of their true type, along with encrypted versions of a number chosen uniformly in the interval $[0, 1]$. Finally each player truthfully reveals the encryption keys they learned in the first period to every other player in the second period.

**Proof.** Begin with the equilibrium path

Index the action profiles in $A$ in some arbitrary way. Let $\omega^k(t)$ be the probability assigned to action profile $a^k$ by the outcome function $\omega$ when player types are given by the vector $t$. The notation $a^k_i$ means the action taken by player $i$ in action profile $a^k$. For $(t, x) \in \tilde{T}_i$, let $t_i$ and $x_i$ be the type and correlating message declared by
player $i$ in the first round. The following mapping defines a direct mechanism:

$$
(5.2) \quad d_i^\omega(t, x) = \begin{cases} 
  a_k^i : k = \min_{k'} \sum_{\tau=1}^{k'} \omega^\tau(t_i, t_{-i}) \geq \lfloor x_i + \sum_{j \neq i} x_j \rfloor 
\end{cases}
$$

The notation $\lfloor y \rfloor$ means the fractional part of the real number $y$. This function aggregates the correlating messages into a number between 0 and 1, then uses this to choose an action profile in $A$. The mechanism then commits $i$ to carry out his part $a_k^i$ of the corresponding action profile $a^k$.

This will implement outcome $a^k$ with probability $\omega^k(t_i, t_{-i})$ provided each of the $x_j$ are uniformly distributed on $[0, 1]$. The property of this construction that will be especially useful below, is the fact that as long as each of the other players is choosing $x_j$ uniformly, the random variable $|x_i + \sum_{j \neq i} x_j|$ is uniform on $[0, 1]$ for each value of $x_i$. What this means is that the probability distribution over $i$’s actions is independent of $x_i$. As a consequence, it is sequentially rational for $i$ to choose his correlating message uniformly from $[0, 1]$ (no matter what his commitment) provided he thinks the other players are doing the same.

By (3.2), there is a collection of punishments $\{\rho_i\}_{i=1,\ldots,n}$ associated with $\omega$. For every such punishment, define for each of the players other than $i$ the direct mechanism

$$
(5.3) \quad p_j^{\rho_i}(t, x) = \begin{cases} 
  a_k^i : k = \min_{k'} \sum_{\tau=1}^{k'} \rho^\tau(t) \geq \lfloor \sum_i x_i \rfloor 
\end{cases}
$$

As above, this will implement $j$’s part of the punishment providing reports are truthful and correlating messages are uniform.

We are now ready to give the strategies associated with the Perfect Bayesian equilibrium that supports $\omega$. Each player should propose

$$
\delta^* = \{d_i^\omega, \{p_j^{\rho_i}\}_{j \neq i}\}_{i=1,\ldots,n}
$$

then punish an encrypted version of his true type along with an encrypted value of a number chosen uniformly from the interval $[0, 1]$. Each player should then privately give the correct encryption keys to his friends.

In the second stage, if all players offered $\delta^*$, then each player should report the encryption keys he heard at the end of the first round accurately. If a single player $j$ deviated in the first round, then each player should send a completely random set of reports to $j$, but report the keys they heard in the first period accurately to each of the other players. If there are more than two distinct proposals offered in the first round, each player should send noise to each of the other players in the second stage, then play their part of some Bayesian Nash equilibrium in the final stage.

It remains to check deviations. The ‘part of a Bayesian Nash equilibrium’ part described above is straightforward, so attention is restricted to the case where all players offer the same mechanism, and to the case where a single player deviates.

If all players offer the same mechanism, and players otherwise follow the equilibrium strategies, then by the argument following (5.2), each player’s payoff is the
one associated with the outcome function $\omega$. Since $\omega$ is incentive compatible, no player has an incentive to misrepresent his type, or send out false encryption keys since his payoff is at least as large when he conveys his true type as it is when this type is replaced by something else when mechanisms are implemented. In the second stage, each player believes the others have conveyed their types truthfully and reported the same encryption keys to each of their friends. So any player who lies about the encryption keys he learned in the first period believes that his report will be ignored (see (5.1)). In this history, no player has a move to make in the final stage, so the continuation strategies are sequentially rational in any history in which all players offer the same mechanism.

In a history where there is a single deviator, the argument is similar for the non-deviators. They have no incentive to lie to the other non-deviators about the messages they heard in the first stage. The only exception is that each player is supposed to send an uninformative message to the deviator in the second stage. Since the others are believed to be sending uninformative messages, the probability that two messages match is zero. So no player believes his message to the deviator will change any outcome.

As for the deviator, he has no incentive to lie about the encryption keys he learned in the first period since he thinks deviating messages will be ignored. Notice that since the deviator is required to report these messages truthfully to incentivize the others, the deviator participates in his own punishment. The deviator believes the messages he receives about other players encryption keys is uninformative, so the best he can do is to maximize his expected payoff given his interim beliefs against the punishment $\rho_j$ that he expects the others to implement (again following (5.3)). Since the outcome function $\omega$ is individually rational, this deviation is unprofitable.

□

References


