DISCUSSION PAPERS IN ECONOMICS

Working Paper No. 01-11

The Evolution of Contracts and Property Rights

Jack Robles
Department of Economics, University of Colorado at Boulder
Boulder, Colorado

October 2001

Center for Economic Analysis
Department of Economics

University of Colorado at Boulder
Boulder, Colorado 80309

The Evolution of Contracts and Property Rights

Jack Robles

Department of Economics
University of Colorado
Boulder, CO 80309
Campus Box 256
phone: (303) 492-7407
email: roblesj@spot.colorado.edu

First Draft October 8, 2001
Revised October 13, 2001

I thank Peter Norman for comments on an earlier paper which helped to motivate the current project.

The Evolution of Contracts and Property Rights

Abstract

I apply stochastic stability (Kandori, Mailath and Rob, Econometrica 1993, and Young

1 Introduction

Conventions, accepted behavior when multiple agents are involved, are present in many aspects of life, including contracting. Some conventions, in contracting and more generally, are explicit, while other are implicit.[1] Almost by definition, conventions regarding incomplete contracts must be both. They must be explicit in the portion of the relationship which is guided by the written contract, and implicit in the portion of the relationship which is not. Explicit and implicit contractual conventions have been studied, but only in isolation from each other. This paper investigates the simultaneous evolution of explicit and implicit conventions in a fairly standard incomplete contracting problem. The evolutionary process is modeled via stochastic stability.[2]

Young (1998) studies a contracting game in which a pair of agents each name a contract, and enter into it if they name the same contract. Young's main result is that the stochastically stable convention splits the surplus equally.[3] Because these contracts are enforcible, Young's result is a statement about explicit conventions. Ellingsen and Robles (2001) (see also Troger (2000) and Dawid and MacLeod (2000)) study a game in which one agent makes a relationship specific investment, after which he and another agent bargain over the surplus. Ellingsen and Robles (2001) find that evolution leads to

[1]Young (1998) contains examples. For one, employment contracts are generally explicit concerning wages, but not concerning bonuses. Consequently, conventions concerning wages are explicit, while those concerning bonuses are implicit.

[2]Kandori, Mailath and Rob (1993), Young (1993) and No

a convention with an implicit assignment of efficient property rights and, consequently, efficient investment. This paper studies the evolution of conventions in a contracting game which includes relationship specific investment, *implicit* post investment property rights, and *explicit* sharing of surplus.

The contracting game is between a buyer and a seller. It consists of a monetary transfer, a (seller specific) investment by the buyer, and a price at which the good is sold. The only portion of the game which is contractible is the transfer. The game begins with the two agents suggesting a transfer. If they suggest different transfers, then there spend most of the time in a very specific set of conventions. These are the stochastically stable conventions.

A simple composite of the results in Young (1998) and Ellingsen and Robles (2001) would suggest: efficient implicit post investment property rights, efficient investment, and a transfer that shares surplus evenly. The logic of such a composite result requires first letting property rights and investment evolve, and only then applying evolution to the determination of the transfer. A more proper approach must study the simultaneous determination of transfer, investment and post investment bargaining. Interestingly, property rights and investment evolve much more quickly than the division of surplus. Hence, stochastic stability in the contracting game does imply efficient property rights and investment.

However, while stochastic stability precludes hold up, the possibility of hold up drives the selection of the stochastically stable transfer. Stochastic stability is, at least partially, a statement about the number of mutants who must play contrary to a convention in order to change the optimal choices of other agents. With this in mind, consider a convention with efficient investment and property rights. Suppose some sellers mutate to suggest the same transfer, but demand a higher price. This attempted hold up makes the contract less attractive to the buyers, and if enough sellers demand a higher price, then the convention dissolves. Of course, the larger the transfer, the more a buyer has to loose from being held up. Hence, hold up decreases the stability of contracts which give too much of the surplus to the sellers. On the other hand, starting levels of investment: I^0 and I^* with $I^* > I^0$. An investment I creates a surplus of $V(I)$, with $V^* = V(I^*)$ and $V^0 = V(I^0)$. I^* is the efficient investment; $V^* - I^* > V^0 - I^0$. Both the seller's cost of production, and the inefficient investment level, I^0, are normalized to zero. In the third stage, players bargain, via the Nash demand game, over the price P. Let P_i be the price demanded by player i. The buyer's payo **a** *e* N *c* ²M

Assumption 1 *A)* $2\Delta \quad \min\{V^0, I^*, (V^* - I^* - V^0)\}$

$B)\frac{1}{2}V^* - I^* \quad V^0 - \Delta$

Part (A) of the assumption, is an assumption that Δ is small, relative to both efficient investment and the gains from efficient investment. Part (B) is weaker than $\frac{1}{2}V^* - I^* \quad \frac{1}{2}V^0$, which is required for there to a hold up problem, if, for example, P is determined through bargaining a la Rubinstein (1982).

Let $\bar{T}_S = V^* - I^* - \Delta$ and $\bar{T}_B = -(V^* - \Delta)$. If a player i suggests $\bar{T}_i$, then the lowest payoff he can receive is zero. The grid of allowable transfers is $\mathcal{T}(\phi) = \{\bar{T}_B, \bar{T}_B + \phi, ..., \bar{T}_S - \phi, \bar{T}_S\}$. It is assumed here, as in Young (1998), that agents choose only strictly individually rational strategy. That is, if a buyer (resp. seller) plays (T, I, P) (resp. $(T, P(\cdot))$ then $V(I) - I - P_B - T_B > 0$ (resp. $\max_I\{T + P(I)\} > 0$.) Clearly, then a buyer will choose only $T \quad \bar{T}_S$, and a seller will only choose $T > \bar{T}_B$.

Our objective is to predict an outcome in the contracting game. Let us consider then, the outcomes which are possible. If a buyer and seller suggest different transfers T_i, then the outcome is $([T_B, T_S])$. If they suggest $T_B = T_S = T$, demand different prices P_i, and the buyer invests I, then the outcome is denoted $(T, I, [P_B, P_S])$. If the previous case is modified so that $P_B = P_S = P$, then the outcome is denoted (T, I, P). An outcome (T, I, P) is a *convention*, if there are off path beliefs which make it self enforcing.[4] This requires: $P + T \quad 0$, $V(I) - I - P - T > 0$ and $V(I) - I - P \quad \Delta$.

[4]A convention is supportable as the outcome from a subgame perfect equilibrium.

Observe that it is possible for a seller to receive a payoff of zero in f

There are two types of histories after which a buyer must have beliefs: $\emptyset$ (the null history) and (T, I) for some agreed upon transfer T and investment level I. A seller must have beliefs following these histories, and also following T for any agreed upon transfer T. Let $\nu(\cdot|$

based upon this information (beliefs following unreached decision nodes are unchanged)

Υ be the set of equilibria. While useful, the restriction to equilibria is not sufficient; an equilibrium may have multiple outcomes, and so does not necessarily yield a convention.

Local stability is sufficient to restrict attention to conventions. For $\theta \in \Theta$, let $\xi^0(\theta) \subset \Upsilon$ be the set of equilibria, which can be reached from θ through updating alone.[7] For $l > 0$ let $\xi^l(\theta)$ be the elements in Υ which can be reached from some element of $\xi^{l-1}(\theta)$ with updating and no more than a single mutation. If $\theta' \in \xi^l(\theta)$, then a sequence of l transitions between equilibria, each of which required only one mutation, can move the population from θ to θ'. $\xi(\theta) = \cup_{l \quad 0}\xi^l(\theta)$ is the set of equilibria which can be reached from θ with a sequence of single mutation transitions.

Definition 1 *A set of states Y is locally stable if $\forall \theta \in Y$, $\xi(\theta) = Y$.*

A locally stable set is impossible to escape with a single mutation, and does not contain any proper subset with this property. Locally stable sets contain only equilibria.

We now turn to the relationship between conventions and locally stable sets. Let $\rho(\theta)$ denote the unique outcome within θ.[8] For an outcome ρ, the ρ-component is the set $\{\theta \in \Upsilon | \rho(\theta) = \rho\}$.

Clearly, in an efficient convention, $-\Delta \quad T$.

Proposition 2 *Every locally stable set contains the ρ-component for an efficient convention ρ.*

Proposition 2

4 Main Result

From Proposition 2 it is but a small step to know that stochastic stability yields a convention with efficient property rights and investment.[11] The one remaining question then is, what share of the surplus does each player receive? Before answering this question, it is useful to define some expressions.

$$H \equiv V^* + V^0 - 2\Delta. \tag{3}$$

H represent the desirability of hold up for sellers: $V^* - \Delta$ is the highest price a seller can charge in the final stage. The individual rationality of buyers implies that (in the limit as $\phi \to 0$) the highest transfer after which sellers might hope to charge $V^* - \Delta$ is $V^0 - \Delta$.[12] Hence H, the sum of these two terms, is the highest payoff that sellers can hope to receive from holding up a buyer.

It is easiest to characterize the stochastically stable convention in two separate cases. We say that *effi*

Otherwise efficient investment is large. Some feeling for Inequality 4 might be found rewriting H as $(V^0 + I^*) + (V^* - I^*) - 2\Delta$. Written so, we can see that either decreasing V^0, or increasing $V^* - I^*$ makes the inequality easier to satisfy. With this in mind, we can read Inequality 4 as, I^* is small relative to the net benefit of investment $(V^* - I^* - V^0)$ and the net value of the relationship $(V^* - I^*.)$

Results are stated in terms of the share of surplus received by the seller. In the case of small efficient investment, F_1 approximates the seller's share.

$$F_1 \equiv \frac{H}{H + (V^* - I^*)} \tag{5}$$

Note that $\frac{1}{2} \quad F_1 \quad \frac{2}{3}$.

Theorem 1 *Let* $(\tilde{T}, I^*, \Delta)$ *denote a stochastically stable convention. Let* $\Lambda_S \equiv \frac{\tilde{T}+\Delta}{(V^*-I^*)}$ *denote the (non investing) seller's share of net surplus in the stochastically stable outcome. For any level of approximation* $\lambda > 0$, *if* ϕ *is sufficiently small, and* N *is sufficiently large, then:*

If investment is small (Inequality 4 holds,) then $|\Lambda_S - F_1| \quad \lambda$.

For the sake of understanding Theorem 1 let us presume that while agents may choose the transfer to suggest, any agreed upon transfer must be followed by (I^*, Δ). In this case, individual rationality implies that $\pi_S = T + \Delta \quad 0$ and $\pi_B = V^* - I^* -$

$(T + \Delta) \quad 0$ or that $-\Delta \leq T \leq V^* - I^* - \Delta$.[13] To displace a convention (T, I^*, Δ) with a convention (T', I^*, Δ), requires that a sufficient number of buyers or sellers mutate to play (T', I^*, Δ). If it is sellers who mutate, then the proportion that much mutate is r^S such that $r^S(V^* - I^* - (T' + \Delta)) = (1$

symmetric, and the most stable transfer would be $T = \frac{1}{2}(V^* - I^*) - \Delta$ which would split the net surplus equally. However, since H is always greater than $(V^* - I^*)$, the seller gets over half of the net surplus. Further, the seller's share is larger, the greater his hoped for payoff from holding up the seller.

We now turn to the case when investment is *large*, $\frac{I^* - \Delta}{V^* - I^*} > \frac{V^* - I^*}{H + V^* - I^*}$. In this case, the best that can be found is a pair of bounds on the seller's share, which the following two fractions provide.

$$\underline{F}_2 \equiv \frac{H}{H + (V^* - \Delta)} \tag{6}$$

$$\overline{F}_2 \equiv \frac{H}{H + (\frac{1}{2}(V^* + I^*) - \Delta)} \tag{7}$$

Observe that $\frac{1}{2} \quad \underline{F}_2 \quad \overline{F}_2 \quad \frac{2}{3}$.

Theorem 2 *Let $(\tilde{T}, I^*, \Delta)$ denote a stochastically stable convention. Let $\Lambda_S \equiv \frac{\tilde{T} + \Delta}{(V^* - I^*)}$ denote the (non investing) seller's share of net surplus in the stochastically stable outcome. For any level of approximation $\lambda > 0$, if ϕ is sufficiently small, and N is sufficiently large, then:*

If investment is large, *then* $\underline{F}_2 - \lambda \quad \Lambda_S \quad \overline{F}_2 + \lambda$.

Whether investment is large or small, that which moves sellers is their desire for H, the prize for a successful hold up. Hence, H enters $\underline{F}_2$ and $\overline{F}_2$ in the same way it entered F_1. However, when investment is large, there is a force operating on the the buyers which is stronger than their desire to grab the entire surplus, $V^* - I^*$. Rather, what moves buyers in this case, is their desire to avoid being held up. If a seller attempts to hold up an unsuspecting buyer, then there is disagreement at the the price setting stage. This leaves the buyer with a loss of $-(T+I^*)$. When efficient investment is large, avoiding this loss can be much more important than chasing after $V^* - I^*$. Because this loss depends upon the transfer, we can not write down it's exact strength as an incentive. However, since individual rationality implies $T \leq V^* - I^* - \Delta$, we can be sure that the strength of this incentive is less than $V^* - \Delta$, the term which appears in $\underline{F}_2$. On the other hand, the loss from being held up decrease as T decreases. For $T \leq \frac{1}{2}(V^* - I^*) - \Delta$ the incentive to avoid being held up becomes too weak to matter. Adding I^* to this transfer yields the term in $\overline{F}_2$.

In order to better understand how different parameters determine the distribution of surplus, I present two limiting results.

Corollary 1 *As* $I^* \to 0$, $\Lambda_S \to \frac{V^*+V^0}{V^*+2V^0}$.
As $V^* - I^* - V^0 \to 0$ *and* $I^* \to 0$, $\Lambda_S \to \frac{2}{3}$.
As $V^0 \to 0$ *and* $I^* \to 0$, $\Lambda_S \to \frac{1}{2}$.

As I^* becomes vanishingly small, V^0 comes to represent the seller's ability to hold up the buyer. Adding $V^* - I^* - V^0 \to 0$, we might think that the whole issue of investment become irrelevant, so that Young (1998) would suggest an even split. However, we see just the opposite. While investment makes no difference in the surplus from the relationship, it is in this case that the seller has strongest incentive to hold up the buyer.

On the other hand, adding $V^0 \to 0$, all of the surplus is generated from investment. However, the buyer really has no hold up ability, and so receives only half of the surplus.

We turn next to the opposite extreme, when the magnitude of I^* dwarfs the other parameters.

Corollary 2 *Holding $V^* - I^*$ constant, as V^* and $I^* \to \infty$, $\Lambda_S \to \frac{1}{2}$.*

In this case, as I^* dominates, an absolute cap is put on the power gained from the sellers ability to hold up the buyer. While the seller will still attempt hold up, hold up is so costly to the buyers that his attempts to avoid it leave the agents with an even split.

A Proofs of Propositions 1 and 2

Proof of Proposition 1 *All communication classes are singletons.*

From Ellingsen and Robles (2001, Lemma 2) if the transfer were fixed at $T = 0$, then all communication classes would be singletons. The only difference between fixed transfers $T = 0$ and $T' \neq 0$ is a constant offset in payoffs, which leaves incentives unchanged. Hence, a nonsingleton communication classes must involve multiple transfers. Consider a nonsingleton communication class in which different transfers, including $\hat{T}$ are suggested. Since agents are switching between transfers, if there is a buyer who invests I^* following $\hat{T}$, then there must be a state within the communication class in which only one buyer suggests $\hat{T}$ followed by I^* and at least one seller suggests $\hat{T}$

Lemma 1 *Let θ be an equilibrium. In θ:*

1) All agents receive a nonnegative payoff.

2) all agents in the same subpopulation receive the same payoff.

3) If T is sometimes agreed upon, I always follows T, and at least one seller demands P following (T, I) then $T + P \quad 0$.

Further, if payoffs for both populations are strictly greater than zero, then:

4) the same set of transfers are suggested by the two subpopulations.

5) The same set of prices are demanded by both subpopulations following any (T, I) which occurs in θ.

Proof:(1) Otherwise an agent would suggest $\bar{T}_i$ which guarantees him a zero payoff. (2)

population, with payoffs offset by a constant of T, from which the result follows. ♣

For $\theta \in \Theta$, let $H(\theta)$ denote the set of outcomes which occur in θ. Let $\rho^D = ([\bar{T}_S, \bar{T}_B])$ be the disagreement outcome. Let $\bar{\Upsilon} = \{\theta \in \Upsilon | H(\theta) = \{(T, I, P)\}$ or $\{\rho^D\}\}$. Call the elements of $\bar{\Upsilon}$ conventional states.

Lemma 2 *Let θ_1 be an equilibrium, with T, I, P' and T, I, P'' ($P' \neq P''$) both elements of $H(\theta)$. Then $\exists \theta_2 \in \xi(\theta_1) \cap \bar{\Upsilon}$ such that $\rho(\theta_2) = (T, I, P) \in H(\theta_1)$.*

Proof: Denote by P' the lowest price demanded by either population following (T, I) and by P'' the highest. Let one buyer who was suggesting the transfer T and demanding P' following T, I mutate, and change his play only in that he now demands P'' following (T, I). This makes (T, I, P'') the only best response for sellers. Let all sellers update to play T, I, P''. Then let all buyers update; this leaves us at an equilibrium with the unique outcome of (T, I, P''). ♣

Lemma 3 *Let θ_1 be an equilibrium in which the transfer T_1 is sometimes agreed upon. If a unique investment level and price follow the transfer T_1 (i.e. $(T_1, I', P'), (T_1, I'', P'') \in H(\theta_1)$ implies that $I' = I''$ and $P' = P''$,) then $\exists \theta_2 \in \xi(\theta_1) \cap \bar{\Upsilon}$.*

Proof: Of course if only the transfer T is suggested, then θ_1 is a conventional state, and the proof is completed. Assume that this is not so. Denote by (T_1, I_1, P_1) the outcome which occurs when T_1 is agreed upon, and by $(T_2, I_2 . P_2)$ some other outcome which occurs in θ_1. We know that $V(I_1) - I_1 - T_1 - P_1 > 0$, or buyers would never

$\beta = \{(\theta_1 \to \theta_2), (\theta_2 \to \theta_3)...(\theta_{k-1} \to \theta_k)\}$ such that $\theta_1 \in L'$ and $L'' \subset \xi(\theta_k)$. Under these circumstances, we define an $(L, \mathcal{L})$-tree, as a collection of $\mathcal{L}$-edges, such that $\forall L' \in \mathcal{L} \setminus \{L\}$ there is a unique directed path of $\mathcal{L}$-edges from L' to L. We do not define a cost for $\mathcal{L}$-edges, but directly define the cost of $\mathcal{L}$-trees. Let η be an $\mathcal{L}$-tree, and let $E(\eta) = \{E = (\theta' \to \theta'') | \exists \beta \in \eta \text{ with } E \in \beta\}$. The $\mathcal{L}$-cost of an $\mathcal{L}$-tree η is $\sum_{E \in E(\eta)} (C(E) - 1)$. The $\mathcal{L}$-potential of $L \in \mathcal{L}$ is the minimum cost over $(L, \mathcal{L})$-trees.

Theorem 4 *Let $\mathcal{L}$ be a collection of disjoint mutation connected subsets of Υ, such that for every locally stable set L, $\exists L' \in \mathcal{L}$ with $L' \subseteq L$. $\theta \in \Upsilon$ is stochastically stable, if and only if $\theta \in \xi(\theta^*)$ for $\theta^* \in L^*$ and L^* is an element of $\mathcal{L}$ with lowest $\mathcal{L}$-potential.*

Proof: Let L^* have lowest $\mathcal{L}$-potential, $\theta^* \in L^*$, and let η be an $(L^*, \mathcal{L})$-tree which achieves this lowest potential. Let $\mathcal{L}_0 =$ [illegible]

$i > 0$, let $\eta_i = \cup_j \bar{\eta}^j_{i-1}$. Finally, let $\bar{\eta} = \cup_i \eta_i$. We observe that η_0 is a collection of edges which provide escape from every locally stable set but L^*. η_0^0 is empty. Hence, $B(\eta_0^0) = \{\theta^*\}$. η_0^1 consists of all edges $(\theta \rightarrow \theta^*)$ which are single mutation transitions and depart from equilibria from which an edge does not already depart in η_0. η_1 then results from repeatedly adding edges for single mutation transitions which eventual lead to θ^*. Hence η_1 has all of the edges in η_0, plus an edge for a single mutation transition departing from every θ such that $\theta^* \in \xi(\theta)$. Proceeding from here, we see that η_i consist of all the edges in η_{i-1} plus single mutation transition from every θ such that $\theta_1(\beta) \in \xi(\theta)$, where $(\theta_1(\beta) \rightarrow \theta_2(\beta))$ is the first edge in β, the $\mathcal{L}$ edge departing some set $L \in \mathcal{L}_{i-1}$. Further, for every $\theta \in B(\eta_i) \cap \bar{\mathcal{L}}_{i-1}$, there is a sequence of edges which lead from θ to θ^*. Since Θ is finite, this process then eventually yields a θ^*-tree. We observe, that every equilibrium but one must have an edge departing from it for any θ-tree. Each of these edges must have a cost of at least one. Let $\bar{M}$ denote the cardinality of Υ. By construction, $\bar{\eta}$ minimizes $\sum_{E \in \mathcal{G}}(C(E)-1) = (\sum_{E \in \mathcal{G}} C(E)) - (\bar{M}-1)$ for any collection of edges $\mathcal{G}$ which provides an escape from all but one of the locally stable sets. Since any tree must do this, and $\bar{M}$ is a constant, θ^* has lowest stochastic Υ potential, and is stochastically stable. From Samuelson (1994) we know that θ^* is in a locally stable set, and that the stochastically stable set must include that entire locally stable set, $\xi(\theta^*)$.

♣

Theorem 4 and Proposition 2 indicate that it is possible to find the stochastically stable set through trees with edges which include the efficient conventions.

C Resistances between Conventions

To apply Theorem 4 we must determine the number of mutations required to replace an efficient convention with another. What is required is to move the population to a state from enough N, we can work with the proportions r which I term resistances.[15]

There are broadly speaking, two types of transitions between conventions which bear consideration: *direct* and *indirect*. Consider a transition between two efficient conventions (T, I^*, Δ) and (T', I^*, Δ). Let θ_1 be the first equilibrium on the path of this transition such that $\rho(\theta_1) \neq (T, I^*, \Delta)$. If $\exists(I, P)$ such that $\rho(\theta_1) = (T', I, P)$, then it is a direct transition. If $\exists(I, P)$ such that $\rho(\theta_1) = (T, I, P)$ then it is an indirect transition. We focus first on direct transitions.

There are two possible means of affecting a direct transition. A new transfer might be made attractive if agents mutate to suggest it, or the current transfer might be made unattractive by changing behavior following it. Clearly this second is accomplished by changing the price. The worst price that sellers (resp. buyers) can demand is $V(I) - \Delta$ (resp. Δ.) Since the price is already Δ in an efficient convention, there is no point to having buyers mutate in this manner. The buyer's payoff is linear in the number of sellers choosing different strategies. Hence, there is no point in considering transitions which involve both types of mutations: either mutations which decrease the payoff for the old transfer are more effective, or mutations which increase the payoff for the new transfer are more effective. These observations are collected in Lemma 9.

Lemma 9 *Consider a minimal mutation direct transition from an efficient convention with transfer T_1 which results in a convention with transfer $T_2 \neq T_1$.*

---

[15] Th [illegible]

If mutations are to buyers, then all the mutants suggest the transfer T_2.

If the mutation is to sellers, then either all of the sellers suggest the transfer T_2, or all of the sellers demand a price of $V^* - \Delta$ following the current transfer and investment.

Lemma 9 identifies three different types of direct transition. Let us focus first on a transition from (T, I^*, Δ), to a convention with transfer $T' \neq T$, which is affected by mutating buyers who suggest the transfer T'. Clearly the sellers can't k e R g s ea g p g hR g in such a transition of minimum resistance. That is, the minimum resistance of such a transition should be r such that $(1-r)(V^{*}-I^{*}-\Delta-T)=r(u_B(T')-T')$.[16]

Proposition 3 *Consider a minimum resistance transition from an efficient convention* (T, I^{*}, Δ) *to a convention with transfer* $T' \neq T$. *Define* $u_S(T')$, *and* $u_B(T')$ *as above.*

If sellers mutate to suggest T' *then* $u_B(T')=V^{*}-I^{*}-\min\{P \in \mathcal{P}(I^{*})\}$ *such that* $u_B(T')-T' \quad V^{*}-I^{*}$.

If buyers mutate to suggest $T' \quad V^{0}-\Delta$, *then* $u_S(T')=V^{*}-\Delta$.

If buyers mutate to suggest $T' \quad V \quad V^{0}-\Delta$, *then* $u_S(T')$

$(T', P(I^*) = P')$. If a sufficient proportion of sellers have mutated, the buyers update to play (T', I^*, P'). At this point updating by the sellers completes the transition. This demonstrates that the suggested value for $u_S(T)$ is feasible. By assumption $u_S(T) - T' = V^* - I^*$ is not feasible, so we ask is it possible that $u_S(T) - T' > V^* - I^*$. The only way for buyers to receive above $V^* - I^*$ is for sellers to receive a negative payoff. In this case, the non-mutating sellers would not imitate the mutants and the mutants would eventually imitate the nonmutants. Hence if the buyers needed a draw of $u_S(T) - T' > V^* - I^*$ to switch strategies, then they will eventually switch back, and $u_S(T) - T' > V^* - I^*$ is not feasible. Now consider $u_B(T')$. If T' $V^0 - \Delta$, then the mutating buyers play (T', I^0, Δ), while the sellers expect $(I^*, V^* - \Delta)$ following T' and correctly expect Δ following (T', I^0). All the sellers update as soon as they see T'. Then all the buyers

Proposition 4 $R^B(T,T') = \frac{T+\Delta}{T+T'+u_B(T')+\Delta}$.

If $T \neq T^0$, *then* $R^B(T) = R^B(T,T^0) = \frac{T+\Delta}{T+V^*+V^0-\Delta-\delta^0}$.

$R^B(T^0) = R^B(T^0, T^0-\phi) = \frac{V^0-\delta^0}{V^*+2(V^0-\Delta-\delta^0)-\phi}$.

Similarly, let $R^S(T,T')$ denote the resistance for a transition from (T, I^*, Δ) to a convention with transfer T' when *sellers* mutate to suggest T'. $R^S(T,T') = r$ such that $r(-T'+u_B(T')) = (1-r)(V^*-I^*-\Delta-T)$, which solves to $R^S(T,T') = \frac{V^*-I^*-\Delta-T}{V^*-I^*-\Delta-T-T'+u_B(T')}$. By assumption, this expression is minimized when $T' = T^M$ and $u_B(T^M) - T^M = V^* - I^* - \delta^M$. Let $R^S(T) \equiv \min_{T'} R^S(T,T')$.

Proposition 5 $R^S(T,T') = \frac{V^*-I^*-\Delta-T}{V^*-I^*-\Delta-T-T'+u_B(T')}$

If $T \neq T^M$, *then* $R^S(T) = R^S(T,T^M) = \frac{V^*-I^*-\Delta-T}{2(V^*-I^*)-T-\Delta-\delta^M}$.

Ther5 remains one type of direct transition. This last occurs when, starting from

$(1-r)(V^* - \Delta) - I^* - T = 0$. This expression follows because the payoff for suggesting a different transfer is zero, and the buyers only get $V^* - \Delta$ when they are matched with non-mutants, but must pay the investment and transfer in all matchings. Of course, the buyers can guarantee themselves a payoff of $\Delta - T$ by playing (

Proposition 7 *A least resistance indirect transfer can be constructed to end at* ρ^M.

There are nonetheless many different types of indirect transfers. However, it happens that only transitions beginning at conventions with a transfer $T > \frac{1}{2}(V^* - I^*) - \Delta$ are relevant. The following Proposition restricts attention to only one one type of indirect transition.

Proposition 8 *If* $T \quad \frac{1}{2}(V^* - I^*) - \Delta$, *and* $P \neq \Delta$, *then* (T, I, P) *is not strictly individually rational for buyers. The resistance of any transition to such a convention is infinite.*

Proof: By assumption $2\Delta \quad \min\{I^*, (V^* - I^* - V^0)\}$. Hence if $T \quad \frac{1}{2}(V^* - I^*) - \Delta$, and $P \quad \frac{1}{2}V(I)$, then $V(I) - I - P - T \leq \frac{1}{2}V(I) + \Delta - I - \frac{1}{2}(V^* - I^*) \quad 0$ since $\frac{1}{2}V^0 + \Delta \quad \frac{1}{2}(V^* - I^*)$ (if $I = I^0$) and $\Delta \quad \frac{1}{2}I^*$ (if $I = I^*$.) ♣

Let us then consider an indirect transfer which starts at (T, I^*, Δ), passes through (T, I^0, Δ) and then moves on to ρ^M. We presume that buyers correctly expect Δ to follow (T, I^0). In this circumstance, the resistance r, to change (T, I^*, Δ) to (T, I^0, Δ) satisfies $(1 - r)(V^* - \Delta) - I^* - T = V^0 - \Delta - T$. Of course, if $V^0 - \Delta - T \leq 0$, then buyers will not play I^0 following T

from (T, I^*, Δ) to (

Conversely, if the value of the relationship is greater than the cost of hold up, then it is easier for sellers to suggest a new agreement.

Proposition 12 *If* $\min\{V^0, \frac{1}{2}(V^* - I^*)\} \leq T + \Delta$ *then* $R^i(T) \quad \min\{R^S(T), R^d(T)\}$.

Proof: If $V^0 - \Delta \leq T$, then $R^i(T) = \infty$. So assume $V^0 > T + \Delta \quad \frac{1}{2}(V^* - I^*)$. Fix T, and set $X = T + \Delta$. Note that if $I^* + T + \delta^M > V^0$, then $R^i(T) > \frac{(V^*-I^*)-X}{V^*-\Delta}$. Consequently $R^i(T) - R^d(T) > \frac{(V^*-I^*)-X}{V^*-\Delta} - \frac{(V^*-I^*)-X}{V^*-\Delta} = 0$. On the other hand, if $I^* + T + \delta^M \leq V^0$, then $R^i(T) \quad \frac{(V^*-I^*)-X}{V^*-I^*+V^0-X-\delta^M}$. Consequently, $R^i(T) - R^S(T)$ $\frac{(V^*-I^*)-X}{V^*-I^*+V^0-X-\delta^M} - \frac{(V^*-I^*)-X}{2(V^*-I^*)-X-\delta^M} > 0$ since $V^* - I^* > V^0$. ♣

Evidently $R^B(T)$ is increasing in T while both $R^d(T)$ and $R^S(T)$ are decreasing in T. Let $R_0^B(T) = \lim_{\phi\to 0} R^B(T)$, and $R_0^S(T) = \lim_{\phi\to 0} R^S(T)$. Define τ such that $R_0^B(\tau) = \min\{R_0^S(\tau), R^d(\tau)\}$ and $\tilde{\tau} = \arg\max_{T\in\mathcal{T}(\phi)}\{\min\{R^B(T), R^S(T), R^d(T)\}\}$.[17] Clearly $\lim_{\phi\to 0}|\tau - \tilde{\tau}| = 0$. As one might suspect, the transfer $\tilde{\tau}$ is very important. Given Proposition 12, if ϕ is sufficiently small, and $\tau > \frac{1}{2}(V^* - I^*) - \Delta$, then the convention $(\tilde{\tau}, I^*, \Delta)$ is the most difficult convention to upset. Furthermore, it is almost possible to construct a $(\tilde{\tau}, I^*, \Delta)$ tree using only the easiest transition out of every efficient convention. It is now to demonstrate that $\tau > \frac{1}{2}(V^* - I^*) - \Delta$.

Let $X_\tau \equiv \tau + \Delta$. Let X_S be defined such that $R_0^B(X_S - \Delta) = R_0^S(X_S - \Delta)$, and let X_d be defined such that $R_0^B(X_d - \Delta) = R^d(X_d - \Delta)$. Clearly, $X_\tau = \min\{X_S, X_d\}$.

[17] For ease of exposition I presume that there is only one argmax. Observe that if there are two, then they adjacent to each other, and still quite close to τ.

$\mathcal{T}(\phi)|T^2 \leq T \leq T^3\}$, while $\{T \in \mathcal{T}(\phi)|\Delta \leq T \leq \tau\} \setminus \mathcal{T}_i \subseteq \mathcal{T}_B$. First presume that $\tau >$ T^0. Let us construct an $\mathcal{L}$ graph as follows: if $T \in \mathcal{T}_i \cup$

[9] Larry Samuelson. Stochastic stability in games with alternative best replies. *J. Econ. Theory*, 64:35–65, 1994.

[10] Joel Sobel. Evolutionary stability and efficiency. *Econ. Letters*, 42:301–312, 1993.

[11] Thomas Troger. Why sunk costs matter for bargaining outcomes: An evolutionary approach. Mimeo, ELSE, University College London, 2000.

[12] Peyton Young. The evolution of conventions. *Econometrica*, 61:57–84, 1993.

[13] Peyton Young. An evolutionary model of bargaining. *J. Econ. Theory*, 59:145–168, 1993.

[14] Peyton Young. Conventional contracts. *Rev. Econ. Stud.*, 65:773–792, 1998.