Learning Across Subgames: An Application to the Hold-up Problem

Julian Kolm*

Vienna Graduate School of Economics, University of Vienna, Mariatheresienstraße 3/18, 1010 Vienna, Austria

August 2011

Abstract

This paper studies learning in an evolutionary model of the hold-up problem. One player decides upon the investment level and the resulting surplus is then divided with another player in a bargaining game. Players learn about their opponents' bargaining behavior from past situations that differ with respect to the available surplus. They rely on inferences about either absolute or relative bargaining demands. The interplay of these two learning heuristics will overcome the hold-up problem. The payoff for the investor is lower than in the existing literature, where players only learn from past instances of identical bargaining situations.

JEL classification: C78, D83.

Keywords: Evolution, Learning, Reasoning by Similarity/Analogy, Game Theory, Hold-up Problem.

1. Introduction

The hold-up problem has been studied extensively in the literature on vertical integration (e.g., Grout, 1984; Grossman and Hart, 1986; Tirole, 1986). It describes how inefficiency might emerge in situations that feature relation-specific investment. The inefficiency arises if it is impossible to write a contract on the investment level and on the compensation of the investor. Because the investment is relation-specific, the concerned parties will bargain over the terms of the transaction after the investment costs are sunk. If the bargaining parties obtain shares according to some fixed rule, like the Nash bargaining solution (Nash, 1950), then there is necessarily underinvestment, as the marginal return on investment will not be fully captured by the investing party.

* *Phone:* +43 1 4277 37449; *Fax:* +43 1 4277 9374
Email address: julian.kolm@univie.ac.at (Julian Kolm)

If the bargaining game exhibits multiple Nash equilibria, as in the case of the Nash demand game[1] (Nash, 1953), then multiple equilibria with efficient and inefficient investment exist. The evolutionary literature has focused on the question of whether evolutionary learning will select an equilibrium with efficient investment. Tröger (2002) and Ellingsen and Robles (2002), TER henceforth[2], find a unique stochastically stable convention where the most efficient investment level is chosen. The investing player obtains a payoff that is larger than the whole net surplus generated by every other investment level. If investment levels are very similar this implies that the investing player will obtain almost all the surplus.[3]

Models of learning in extensive form games traditionally assume that players learn independently for each information set (e.g., Fudenberg and Levine, 1998). Learning is thus restricted to identical situations which neglects the fact that players can learn from "similar", albeit different, situations. Learning across "similar" situations, however, is clearly a feature of real world learning.[4]

Also TER consider players who learn only from identical situations. In Ellingsen and Robles (2002), players update their strategies independently for each bargaining subgame, based on the strategy distribution of the population. In Tröger (2002), players' beliefs about a bargaining subgame are based exclusively on past bargaining behavior that followed the same investment level.

Contrary to the literature, I consider players who learn across "similar", albeit different, bargaining games. Learning across different bargaining situations seems to be natural because they do indeed have a very similar structure. Players treat different bargaining subgames as "similar" in the sense that they believe that their opponents will make similar bargaining demands independently of the investment levels. Players' beliefs about bargaining behavior following a specific investment level are based on the observed bargaining behavior following all investment levels. This notion of similarity is consistent with the concept of "analogous" situations proposed by Jehiel (2005).

Different investment levels, however, imply that players bargain about surpluses of different sizes. In such a context, it is not clear what constitutes similar demand in different bargaining subgames. Two possibilities seem to be natural: First, two bargaining demands are similar if they are of the same absolute size.

[1] This is also the case for the Rubinstein-Ståhl bargaining game, if there exists a smallest money unit (Van Damme et al., 1990).

[2] These two papers consider exactly the same game and derive the same results. The modeling details of the evolutionary learning process, however, differ because Tröger (2002) extends the model of Young (1993b) while Ellingsen and Robles (2002) apply the framework of Nöldecke and Samuelson (1993).

[3] Dawid and MacLeod (2001, 2008) show that if investment is two sided, then the Hold-up problem persists in evolutionary models.

[4] Moreover, the assumption commonly made in evolutionary learning models, that the different past instances of the situation players find themselves in are exactly the same, is, of course, a fiction employed to model the idea that players face similar, albeit slightly different, situations.

Second, two bargaining demands are similar if they claim the same relative share of the available surplus.

I allow for both interpretations to coexist in the model: Some players form beliefs about absolute demands, while others form beliefs about relative demand shares. This allows me to account for the fact that different players might indeed employ different learning heuristics.[5] The number of players who form their beliefs according to either one of this heuristics has no impact on the results as long as each heuristic is used by some players.

I find that evolution will overcome the hold-up problem. I also show that the distribution of the stochastically stable conventions will be such that the investing players obtain a payoff that is (weakly) larger than the equal split and (weakly) lower than in TER. If the investment levels are very similar, the difference to TER can amount to almost half the gross surplus of the efficient investment level.

In fact, efficient investment-bargaining norms emerge directly from the best response behavior of the players. Thus, evolution solves the hold-up problem whether the players make errors or not.

My result is driven by the interplay of the two learning heuristics. Players who believe that their opponents will make a constant absolute demand essentially consider themselves to be residual claimants. Thus, these players will choose efficient investment if they expect to obtain any positive payoff, accounting for their investment cost. Players who believe that their opponents will demand the same relative share in all bargaining subgames will choose high investment if their share of the gross surplus is large enough. If this is the case, then choosing efficient investment maximizes their payoff accounting for their investment costs.

Because some players consider themselves to be residual claimants, no convention with inefficient investment will be stable. The evolutionary bargaining process will ensure that players coordinate on a stable convention with the efficient investment level. These conventions are stable because the distribution of the surplus will adapt, such that all players have an incentive to choose efficient investment.

In the stochastically stable states, the investing players get a share of the gross surplus that is high enough that both types of players choose high investment. The number of errors necessary to displace one convention for another convention closer to the Nash bargaining solution is lower than for a displacement in the reverse direction.[6] Thus it follows that the distribution of the gross surplus is as close to the Nash bargaining solution as possible under the constraint that this distribution can be a stable convention i.e., the share of the investing players is not to low.

The remainder of the paper is structured as follows. Section 2 discusses the formal model. Section 3 introduces the concepts used to analyze the long-run

[5] In an earlier paper (Kolm, 2010), I considered a similar model where all players only considered the absolute size of past bargaining demands.

[6] This is also the case in Young (1993b).

properties of the players' learning process. Section 4 presents and discusses the results. Section 5 concludes. Formal proofs are relegated to the Appendix.

2. The Model

I consider a model with two finite populations. Every period $t = 1, 2, ...$ a player, α, is randomly selected from the first population and another player, β, is randomly selected from the second. These two players are matched to play the following two stage investment-bargaining game.

2.1. The Game

At the investment stage, a player α chooses an investment level I, which is associated with a private cost C_I. The investment of the player creates a gross surplus of the size V_I.

At the bargaining stage, the players divide the gross surplus by playing a Nash demand game. Players α and β simultaneously make demands x and y, respectively. Players receive their respective demand if, and only if, the sum of their demands is smaller than or equal to the gross surplus that is available. The surplus of α and β in the investment bargaining game is then given by

$$(\pi_\alpha(I,x,y), \pi_\beta(I,x,y)) = \begin{cases} (x - C_I, y) & \text{if } x + y \leq V_I \\ (-C_I, 0) & \text{otherwise} \end{cases}$$

Two demands x and y are said to be compatible if given I , $x + y \leq V_I$.

Because every division is a Nash equilibrium of the Nash demand game (Nash, 1953), every investment level that yields a positive net surplus $V_I - C_I$ can be supported in a subgame perfect Nash equilibrium. To see this, consider the following strategy combination: At the investment stage, the player α chooses an investment any I. In the bargaining subgame following this investment level, the players choose bargaining demands x_I and y_I such that $x_I + y_I = V_I$ and $x_I - C_I$ and y_I are greater than zero. In the bargaining subgames following all other investment levels I', the strategies of players are given by $x_{I'} = 0$ and $y_{I'} = V_{I'}$. It is obvious that, given these strategies, players α do not have an incentive to deviate from their investment level because this would yield a negative payoff of $-C_{I'}$. Because every division of the surplus that satisfies $x + y = V$ is a Nash equilibrium of the Nash bargaining game, the above strategies constitute a subgame perfect Nash equilibrium. The above game thus allows for subgame perfect Nash equilibria with an inefficient investment level, but it also allows for equilibria where efficient investment is chosen. The aim of the following evolutionary analysis is to identify those equilibria that are most likely to occur in the long run.

For technical reasons it is assumed that the action sets of players are finite. The investment levels I are chosen from the finite set $\mathbb{I} = \{I_0, I_1, \ldots, I_N\}$. Let δ be a base unit that is chosen such that for all possible investment levels, $\frac{V_I}{\delta}$, $\frac{C_I}{\delta}$, and $\frac{\max_{I \in \mathbb{I}} V_I}{2\delta}$ are integers. The demands of the players are assumed to be elements of the set of possible demands $X = \{\delta, 2\delta, \ldots \max_{I\in\mathbb{I}}(V_I) - \delta\}$. The lower bound of this set is motivated by the fact that a demand of zero is weakly dominated by any other demand that a player could make. The upper bound of this set is justified by the fact that, given the lower bound of X, the demand $\max_{I\in\mathbb{I}}(V_I)$ is dominated by all other demands in X. The action sets of players α and β are then given by the finite sets $\mathbb{I} \times X$ and X, respectively.

To make the analysis more tractable, I assume that there only exist two investment levels, high investment, H and low investment, L. The number of investment levels is not crucial, but must be finite. I assume further that high investment is efficient, which is the case if

$$V_H - C_H > V_L - C_L \tag{1}$$

This clearly is the relevant case for the discussion of the hold-up problem.

2.2. Players' Behavior

Players consider all bargaining situations analogous in the sense of Jehiel (2005), i.e., they expect their opponents to make bargaining demands similar to those in previous bargaining situations. As mentioned in the introduction, the different sizes of the available surplus allow for two intuitive notions of similarity: First, they can be considered similar if they claim the same absolute value. Second, they can be considered similar if they claim the same share of the surplus.[7]

Let us assume that both populations of players consist of players of both types.[8] Players who interpret past bargaining demands as absolute values are denoted by α_a and β_a, and players who interpret past bargaining demands as relative shares are denoted by α_r and β_r. In what follows, relative demand shares x_t/V_{I_t} and y_t/V_{I_t} will be denoted by χ_t and υ_t.

The beliefs of the players about their opponents' strategies are based on partial observations of the history of plays. Every period the active player α draws a sample A, which consists of k^α entries from the last m records of play. Similarly, player β draws a sample B consisting of k^β entries from the last m records of play.

The beliefs of players α_r regarding the bargaining behavior of players β are given by the observed cumulative distribution of relative demand shares $F_r(\upsilon_t|A)$. The beliefs of players α_a are given by the observed cumulative absolute bargaining

[7] Heterogeneity of agents with respect to their belief formation has been explored in Saez-Marti and Weibull (1999) and Matros (2003). In their models a fraction of "clever" players does not play a best response to the history of their opponents' past actions, but rather a best response to their opponents' best response to their own past actions. Matros (2003), however, shows that this modification does not change the minimal curb set in generic finite two player games as long as the share of "clever" players is below one in all populations.

[8] Alternatively, it could be assumed that whenever a player is chosen to play the investment bargaining game, he randomly selects whether to interpret past bargaining demands as absolute demands or relative share. It is only important that in every period, there exists a positive probability for both types of players to play.

demands $F_a(y_t|A)$. Analogously the beliefs of β_r and β_a can be characterized by $F_r(\chi_t|B)$ and $F_a(x_t|B)$, respectively.

Thus, given an investment level I, the payoff expected by players α_r demanding x is given by

$$E_r(\pi_\alpha(x)|I,A) = xF_r(1-\frac{x}{V_I}|A) - C_I$$

The payoff expected by players α_a is given by

$$E_a(\pi_\alpha(x)|I,A) = xF_a(V_I - x|A) - C_I$$

The beliefs of players β_r and β_a are defined analogously.

2.2.1. Best response behavior

Given their beliefs, the players choose a myopic best response, that is, players α choose $\arg\max_{x\in X, I\in\mathbb{I}} E(\pi_\alpha(x)|I,A)$ and players β choose $\arg\max_{x\in X} E(\pi_\beta(y)|I,B)$. If multiple action profiles yield the highest expected payoff for some player, then this player plays a probability distribution over these action profiles. The support of this distribution is given by the set of action profiles that yield the maximum expected payoff and for which there exists a compatible demand from their opponent, given the investment level I. That is, any action profile (I,x,y) that is played with positive probability satisfies $x < V_I$ and $y < V_I$, respectively. This is motivated by the fact that a bargaining demand $x \geq V_I$ or $y \geq V_I$ is weakly dominated by every other bargaining demand in X given I.

It is possible that some past play (I,x,y) was such that the expected bargaining demands (given this past play) are not compatible with any feasible bargaining demand. For a player α_r, this is the case when given a different investment level I', $(1-\frac{y}{V_I})V_{I'} < \delta$, e.g., when $y = V_I - \delta$ and $V_I > V_{I'}$. For a player α_a, this is the case when given a different investment level I', $y \geq V_{I'}$. Analogous conditions apply for players β_r and β_a. This can be interpreted as a player's insecurity about the distribution of an opponent's demand, in which case the player does not regard this instance of past play as informative and expects that the worst case will happen, i.e., that no coordination will take place. If for all demands in the sample there does not exist a compatible demand (given the chosen investment level), then the expected payoff of all demands is zero and, consequently, the player will play a probability distribution with support $\{z \in X : z < V_I\}$.

In addition, players make occasional mistakes. With probability ϵ, a player α makes a mistake at the bargaining stage and with the same probability ϵ, a player α or β makes a mistake a the bargaining stage. The joint distribution of making an error at the investment stage and the bargaining stage does not matter for the results of this paper. If a player makes an error at some stage of the game, he or she chooses a random action from a distribution that has full support over the action set of this stage.

3. Conventions

The subsequent analysis will largely focus on states where players behave uniformly over time. Such states will be called conventions.

The play in period t is denoted by (I_t, x_t, y_t), which is abbreviated by $(I, x, y)_t$. Let $\mathbb{S}$ be the set of all truncated histories of plays of length m. Let a state $s_t \in \mathbb{S}$ denote the play in the last m periods in at period t, i.e.,

$$s_t = ((I, x, y)_{t-m}, (I, x, y)_{t-m+1}, ...(I, x, y)_t).$$

As the players draw their samples from the last m periods of play, their behavior at time t only depends on s_t. Thus, the behavior of players determines a Markov chain P^ϵ on $\mathbb{S}$. This Markov chain is called the perturbed investment-bargaining process. Because the strategy spaces of players $\mathbb{I} \times X$ and X are finite, the state space $\mathbb{S}$ is finite as well.

3.1. Stability

One concept to describe the long run behavior of the investment-bargaining process is provided by the notion of *stability*.[9] Stability describes those sets of states that will only be left by the investment-bargaining process if players make errors.

Let P^0 denote the Markov chain that is determined exclusively by the best response behavior of players, i.e., when $\epsilon = 0$. This Markov chain is called the unperturbed investment bargaining process.

Definition. A set of states S_i is *stable* if for all states $s, s' \in S_i$ there exists a positive transition probability from s to s' under P^0 and for all states $s^- \notin S_i$ there does not exist a positive transition probability from s to s^- under P^0.

3.2. Conventions

States where players behave uniformly are important because for these states, the beliefs of players are the same for all samples that can be drawn from the history of plays.

Definition. A state $s \in \mathbb{S}$ is a *convention* σ if it consists of m identical records, and the demands of the players are compatible. A convention is *stable* if it is a *stable* set.

Stable conventions are the only states where all possible beliefs of players coincide with all possible best response actions of their opponents. Thus, as long as the process is in a stable convention, beliefs of players will constantly be confirmed.

[9] These sets are called "absorbing" by Ellingsen and Robles (2002) and "recurrent communication classes" by Young (1993a)

3.3. Stochastic stability

The concept of s*tochastic stability* was introduced by Foster and Young (1990); it describes the long run probability that a state occurs. It provides a stronger notion than *stability* and draws on constant perturbations of the best response dynamic.

The perturbed investment-bargaining process P^ϵ is irreducible and aperiodic.[10] From the theory of Markov Processes, it follows that there exists a unique stationary distribution μ^ϵ of the Markov process P^ϵ where $\mu^\epsilon(s)$ is the cumulative relative frequency with which a state s will occur if the process P^ϵ runs for a very long time.

Definition. A state $s \in \mathbb{S}$ is *stochastically stable* if $\lim_{\epsilon \to 0} \mu^\epsilon(s) > 0$.

Let S^* denote the set of stochastically stable states. For very small error probabilities ϵ, the process P^ϵ will be in S^* most of the time. Young (1993a) has shown that the set of stochastically stable states must be a subset of the union of all stable states.

4. Results

In this section, it will be shown that the set of stable states consists of stable conventions with efficient investment. The distribution of the surplus will be such that the investing players who believe that they get the same relative share of the gross surplus independently of the investment level will have an incentive to choose the efficient investment level. The surplus distribution of the *stochastically* stable conventions is such that the investing players get either exactly as much as necessary to make high investment profitable for them (given the above belief), or the investing players get a share of the gross surplus that corresponds to the (generalized) Nash-bargaining solution.

As a point of reference, briefly consider the case where the surplus V_I is split among both players according to the Nash-bargaining rule that attributes $V_I/2$ to each player. Clearly, in this case, a player α will choose high investment if $\frac{V_H}{2} - C_H > \frac{V_L}{2} - C_L$. It follows then that the hold-up problem would lead to an inefficient, low level of investment if

$$V_H - V_L \in (C_H - C_L, 2C_H - 2C_L).$$

The following evolutionary analysis will analyze the case where $V_H > V_L$ and $C_H > C_L$, which is a necessary condition for the hold-up problem to arise.

[10] To see this, suppose that P^ϵ is in state s_t at time t. Suppose further that both players make errors at the investment and the bargaining stage for m consecutive periods. After this m periods, the state s_{t+m} consists entirely of random demands. Consequently, the state s_{t+m} can be every state in $\mathbb{S}$, which implies irreducibility. Clearly, the same state could also be reached after $m+1$ periods of consecutive errors, which implies aperiodicity.

4.1. Stable High Investment

The first result concerns the set of stable conventions.

Lemma 1. *Assume that* $(C_H - C_L)/(V_H - V_L)V_H$ and $(C_H - C_L)/(V_H - V_L)V_L$ are multiples of δ. *If the shares of players* α_r, α_a, β_r, *and* β_a *are strictly positive in the populations of players* α *and* β, *respectively, then the set of stable conventions* Σ *is given by the set of conventions* $(H, \xi, V_H - \xi)$ *with* $\xi > (C_H - C_L)/(V_H - V_L)V_H$.

Proof. see Appendix. □

To discuss the idea behind the proof, I introduce the following notation: Denote the lower bound of the gross surplus that investing players may receive in a stable convention by $\underline{\xi} := (C_H - C_L)/(V_H - V_L)V_H$ and denote the lower bound of the relative shares of the gross surplus that investing players may receive in a stable convention by $\underline{\chi} := (C_H - C_L)/(V_H - V_L)$.

The fact that the set of stable conventions contains only conventions with high investment is driven by the presence of players α_a in the population. Those players believe that all players β will make the same absolute bargaining demands, irrespective of the investment level chosen. This implies that players α_a believe that they are the residual claimants. Consequently, they will maximize the total surplus in order to maximize their own net surplus. Because high investment is efficient, this implies that no stable convention can be associated with low investment.

The fact that the share of the investing players must be larger than $\underline{\xi}$ is driven by the presence of players α_r in the population. Thus, players believe that players β will demand the same share υ at all investment levels. This implies that high investment is profitable if the following condition is met: $(1 - \upsilon)V_H - C_H > (1 - \upsilon)V_L - C_L$. This is the case if and only if $(1 - \upsilon) > \underline{\chi}$, which, given high investment, implies that $\xi > \underline{\xi}$.

This result holds as long as the population of players α contains at least one player of each type. The reason is that, as long as one player of each type is present in the population, the probability for a player of each type to play in any given period is positive. Hence, no convention can be stable that is not stable given the best response correspondence of both types. Thus, if a population consists of several types of players, then the set of stable conventions consists of the intersection of the sets of conventions that are stable for a single type.

This line of reasoning also implies the following result:

Lemma 2. *Let the population of players* α *consist of arbitrary player types and let it contain a positive share of players* α_a. *If the set of stable conventions is non empty, it contains only conventions with high investment.*

Proof. Follows from inspection of the Proof of Lemma 1 and the fact that every stable convention must be a fixed point of the best response correspondence of all players in the populations. □

The mechanism driving this result is that players α_a will always choose high investment. Hence, no convention with low investment can be stable if the probability of a player α_a to play in any given period is positive. Thus, the presence of players who consider themselves to be the residual claimants is sufficient to prevent the emergence of inefficient stable conventions.

The next result concerns the convergence of the evolutionary learning process to the set of stable conventions with populations consisting of players α_r, α_a, β_r, and β_a.

Proposition 1. *Assume investment to be efficient and the sample lengths of both players to be smaller or equal to $m/3$. Then, from any initial state s, the unperturbed bargaining process P^0 converges almost surely to a stable convention.*

The proof of this proposition shows that if players draw the "right" samples, it is possible for them to coordinate on a stable convention. The following example demonstrates how such coordination takes place.

Example 1. Let $k := \max\{k^\alpha, k^\beta\}$ and let A_t denote the last k records in s_t. Abusing language, say that a player samples A_t if he draws a sample contained in A_t.

Suppose that in period $t+1$ to $t+k$, inclusive, players α_a and β_a sample A_t. Thus, the best response behavior of players implies that they will play the same actions in period t to $t+k$, inclusive. Hence, A_{t+k} consists of identical records (I, x, y). Suppose that in period $t+k+1$ to $t+2k$, inclusive, players α_a sample A_t and players β_a sample A_{t+k}. With positive probability, player α_a will take the same actions as in the last k periods. Players β_a will make bargaining demands $V_I - x$. Thus, A_{t+2k} consists of identical records $(I, x, V_I - x)$.

If $(I, x, V_I - x)$ satisfies the conditions of Lemma 1, then if players sample A_{t+2k} in period $t+2k+1$ to $t+k+m$, inclusive, then s_{t+k+m} will be a stable convention.

If $(I, x, V_I - x)$ does not satisfy the conditions of Lemma 1, then the proof of Proposition 1 establishes that it is still possible to reach a stable convention in finitely many periods. Indeed, as long as $\xi \leq \underline{\xi}$, it is possible to increase the share players α get by a sequence of plays where investment levels alternate between high and low. Such alternations are possible because for $\xi \leq \underline{\xi}$, players α_a choose high investment while players α_r choose low investment. Players' α share can increase because when investment changes from low to high, the absolute share of players β remains constant with positive probability, thus decreasing their relative share. Conversely, when investment changes from high to low, the relative share of players β remains constant with positive probability, thus decreasing their absolute share. This gives players α the possibility to increase their share above $\underline{\xi}$ in finitely many periods. Hence, a stable convention can be reached.

Because investment is high in all stable conventions, as stated in Lemma 1, Proposition 1 implies that the unperturbed investment bargaining process will almost surely converge to a convention with high investment. This result is consistent with the results of TER.

My result, however, is driven by the interplay of two different learning heuristics. A convention with high investment will emerge directly from the best response behavior of players. Thus, it does not depend on whether players make errors. In contrast to this, TER rely on the concept of stochastic stability to show efficiency and, thus, use a limit result as the probability of players making errors goes to zero. My result, therefore, adds support for the claim that evolution can resolve the hold-up problem.

4.2. Surplus Distribution

The following theorem characterizes the distribution of the stochastically stable convention in the limit as $\delta \to 0$.

In what follows, let ξ^N denote the asymmetric Nash bargaining solution

$$\arg\max_x ((x)^{k^\alpha/m}(V_H - x)^{k^\beta/m}).$$

If both players draw samples of the same length, then $\xi^N = V_H/2$.

Theorem 1. *Assume δ to be sufficiently small and k and m to be sufficiently large. As $\delta \to 0$ the distribution of the surplus ξ of all stable conventions $(H, \xi, V_H - \xi)$ converges to*

$$\xi^* = \max\left\{\frac{C_H - C_L}{V_H - V_L} V_H, \xi^N\right\} \tag{2}$$

Proof. see Appendix. □

As stated in the Theorem, the investing players will at least get a share according to the (asymmetric) Nash bargaining solution. If, however, this share would not be sufficient to give players α_r an incentive to choose high investment, then the surplus will be higher. It will be just high enough to give players α_r an incentive to choose high investment. Thus the investing players will be compensated for their investment only to the extent that this gives them an incentive to invest, while the non-investing players will receive the remaining surplus that is generated by the investment.

Essentially, the reason for this is that the number of errors that is necessary to displace a convention in favor of a convention that is closer to the Nash bargaining solution is lower than for a displacement in the reverse direction. This, of course, only holds for stable conventions, which explains why $\xi^* > \frac{C_H - C_L}{V_H - V_L} V_H$. Thus, the distribution ξ^* is as close to the Nash bargaining solution as possible under the constraint that the distribution provides an incentive for players α_r to choose high investment.

The proof of this Theorem combines an adaption of the ideas provided in Young (1993b) with techniques developed in Ellison (2000).

It is useful to compare expression 2 to Proposition 4.1 in Tröger (2002) and Proposition 4 in Ellingsen and Robles (2002), which state that as $\delta \to 0$ the distribution of the stochastically stable convention converges to

$$\xi^{TER} = \max\left\{V_L - C_L + C_H, \xi^N\right\}$$

In TER, $\xi \geq V_L - C_L + C_H$ is a necessary condition for a stable convention because the beliefs about bargaining behavior following investment levels that deviate from "conventional" play can "drift", i.e., the beliefs can easily take a random form over time. Thus, players can easily believe that they will obtain the whole surplus after some "unconventional" investment level. Such extreme beliefs can destabilize any investment-bargaining convention where $\xi < V_L - C_L + C_H$. Thus, the only stable conventions are those where the efficient investment level is chosen and the investing players obtain a payoff that is higher than the whole net surplus generated by any other investment level.

In the present model the investing players obtains a weakly smaller surplus than in TER. This follows from simple rearrangement of terms in $\xi^* \leq \xi^{TER}$, which yields $V_L((V_H - V_L) - (C_H - C_L)) \geq 0$. This inequality holds true for efficient investment (equation 1).

If the investment levels are very similar, then the difference between ξ^* and ξ^{TER} can be very large. For instance, suppose that $V_H = V_L + 4\delta$ and $C_H = C_L + 2\delta$. If both players draw samples of the same length, then this implies that $\xi^* = \frac{1}{2}V_H$ and $\xi^{TER} = V_H - 2\delta$. Thus, the difference amounts to almost half the gross surplus available under efficient investment.

The intuition behind this is that in the present model, the stochastically stable convention must only be robust to the investing players believing that they will get the same relative share of the gross surplus if they choose a different investment level. In TER, however, the stochastically stable convention must be robust to the investing players believing that they get the whole surplus if they choose another investment level. However, it seems very unlikely that players, in fact, believe that they will get the whole surplus if they choose another investment level.

5. Conclusion

The presented model of the hold-up problem provides an example of a game where different subgames clearly have a very similar structure. In the context of evolutionary learning where players always learn from past situations, it seems natural that players learn across different bargaining subgames.

Two different learning heuristics are plausible because the surplus available in different bargaining subgames differs and, hence, it is not unambiguous when two bargaining demands are similar. The flexibility of the evolutionary model allows me to consider the interplay of two different learning heuristics. This is an advantage of the evolutionary approach over the strategic model of analogy based expectation equilibrium by Jehiel (2005).

Coordination on efficient conventions arises directly from the best response behavior of the players. The reason for this is that some players consider themselves to be residual claimants, which makes them very reluctant to choose low investment.

Learning across subgames imposes some consistency on the beliefs that players can hold about different bargaining subgames. This consistency of beliefs greatly alters the stability of investment-bargaining conventions. As a result, the investing players obtain a (weakly) lower payoff than in Tröger (2002) and Ellingsen and Robles (2002).

The present model shows how learning across similar situations can influence the results of evolutionary learning. This is important because learning across similar situations is clearly a feature of real-world learning.

6. Acknowledgments

The author thanks Maarten Janssen, Karl Schlag, Larry Blume, Michael Greinecker and Katarzyna Rymuza for helpful comments and discussion.

7. References

Dawid, H., MacLeod, W. B., 2001. Holdup and the evolution of bargaining conventions. Eur. J. Econ. Soc. Syst. 15, 153–169.

Dawid, H., MacLeod, W. B., 2008. Hold-up and the evolution of investment and bargaining norms. Games Econ. Behav. 62, 26–52.

Ellingsen, T., Robles, J., 2002. Does evolution solve the hold-up problem? Games Econ. Behav. 39, 28–53.

Ellison, G., 2000. Basins of attraction. long-run stochastic stability, and the speed of step-by-step evolution. Rev. Econ. Stud. 67, 17–45.

Foster, D., Young, H. P., 1990. Stochastic evolutionary game dynamics. Theor. Popul. Biol. 38, 219–232.

Freidlin, M., Wentzel, A., 1984. Random Perturbations of Dynamical Systems. Springer, New York.

Fudenberg, D., Levine, D. K., 1998. The Theory of Learning in Games. MIT Press Series on Economic Learning and Social Evolution. The MIT Press, Cambridge, Massachusetts and London, England.

Grossman, S. J., Hart, O. D., 1986. The costs and benefits of ownership: A theory of vertical and lateral integration. J. Polit. Economy 94, 691–719.

Grout, P., 1984. Investment and wages in the absence of a binding contract. Econometrica 52, 449–460.

Jehiel, P., 2005. Analogy-based expectation equilibrium. J. Econ. Theory 123, 81–104.

Kolm, J., 2010. An evolutionary model of the hold-up problem. Master's thesis, University of Vienna.

Matros, A., 2003. Clever agents in adaptive learning. J. Econ. Theory 111, 110–124.

Nash, J., 1950. The bargaining problem. Econometrica 18, 155–162.

Nash, J., 1953. Two-person cooperative games. Econometrica 21, 128–140.

Nöldecke, G., Samuelson, L., 1993. An evolutionary analysis of backward and forward induction. Games Econ. Behav. 5, 425–454.

Saez-Marti, M., Weibull, J., 1999. Clever agents in young's evolutionary bargaining model. J. Econ. Theory 86, 268–279.

Tirole, J., 1986. Procurement and renegotiation. J. Polit. Economy 94, 25–259.

Tröger, T., 2002. Why sunk costs matter for bargaining outcomes: An evolutionary approach. J. Econ. Theory 102, 375–402.

van Damme, E., Selten, R., Winter, E., 1990. Alternating bid bargaining with a smallest money unit. Games Econ. Behav. 2, 188–201.

Young, H. P., 1993a. The evolution of conventions. Econometrica 61, 57–84.

Young, H. P., 1993b. An evolutionary model of bargaining. J. Econ. Theory 59, 145–168.

AppendixA. Proofs

AppendixA.1. Necessary Concepts

Techniques to characterize the set of stochastically stable states were pioneered by Freidlin and Wentzel (1984) and further developed by Young (1993a,b) and Ellison (2000). This subsection introduces several concepts and results from this literature as well as the notation that will be used in the subsequent analysis.

Let a path ρ from s to s' be a sequence of one period transitions that leads from s to s' where the probability of each one period transition in ρ under P^ϵ is larger than zero. The resistance of a path $r(\rho)$ is the sum of the number of errors that is necessary for each one period transition in ρ. For any two sets of states $S, S' \subseteq \mathbb{S}$ let $r(S, S')$ denote the least resistance of all paths that start in S and end in S'.

AppendixA.1.1. Radius and Modified Coradius

This subsection is a slight adaption of the definitions in Ellison (2000):

Definition. For every union of stable sets $\cup S_i$, let the radius $R(\cup S_i)$ denote the least resistance of all paths that start in $\cup S_i$ and end in some stable set $S_j \nsubseteq \cup S_i$

The radius describes the minimal number of mistakes that is necessary to enter the basin of attraction of a stable set outside of $\cup S_i$ and thus provides a measure of its persistence.

Let $(S_1, S_2, ...S_r)$ be the stable sets through which the path ρ passes. The modified resistance of a path $r^*(\rho)$ is given by $r(\rho) - \sum_{i=2}^{r-1} R(S_r)$. For any two sets of states $S, S' \subseteq \mathbb{S}$ let $r*(S, S')$ denote the least modified resistance of all paths that start in S and end in S'.

Definition. For every union of stable sets $\cup S_i$, let the modified coradius $CR^*(\cup S_i)$ denote

$$\min_{s' \in \cup S_i} \max_{s \in S \setminus \cup S_i,} r^*(s, s').$$

Theorem 2 in Ellison (2000) states that if $R(\cup S_i) > CR^*(\cup S_i)$, then the set of stochastically stable states must be a subset of $\cup S_i$.

AppendixA.1.2. Stochastic potential

This subsection is a slight adaption of the definitions in Young (1993b, p. 154):

Define a graph G as follows. There is a vertex for every set S_i and a directed edge from every vertex to every other vertex. The "weight" or resistance of any directed edge $S_i \to S_j$ is given by $r(S_i, S_j)$.

An i-Tree is a collection of edges in G, such that from every Vertex $S_{j \neq i}$, there is a unique directed path to its rootS_i, and there are no cycles. Let $\mathbb{T}_i$ be the set of all i-Trees.

Definition. The stochastic potential of a stable set S_i is the least resistance among all i-Trees with root S_i:

$$\gamma_i = \min_{T \in \mathbb{T}_i} \sum_{(S,S') \in T} r(S, S').$$

Theorem 2 in Young (1993a) states that the elements of a stable set S_i are stochastically stable if S_i has the least stochastic potential among all stable sets $S_j \in S$.

AppendixA.1.3. Notation

The following additional Notation will be used:

For any number z and every investment level I, let $[z]_x$ denote $\max\{x \in X_I : x \leq z\}$ and let $[z]_y$ denote $\max\{y \in X_I : V_I - z \leq V_I - y\}$. If given an investment level I, players α_r face a distribution of identical relative demands $y/V_{I'}$ such that, given the current investment level I, $V_I - (y/V_{I'})V_I \notin X_\delta$, these players will choose the bargaining demand $[V_I - (y/V_{I'})V_I]_x$. Similarly, if players β_r face a distribution of identical relative demands $x/V_{I'}$ such that, given the current investment level I, $V_I - (x/V_{I'})V_I \notin X_\delta$, these players will choose the bargaining demand $[V_I - (x/V_{I'})V_I]_y$.

Write $(I, x, y)^l$ for a sequence of length l that consists of identical records (I, x, y). Also, let A_t denote the last k records in s_t where $k := \max\{k^\alpha, k^\beta\}$.

Abusing language, say that a player samples A_t if he draws a sample contained in A_t.

Also remember that $\underline{\xi} := (C_H - C_L)/(V_H - V_L)V_H$ and $\underline{\chi} := (C_H - C_L)/(V_H - V_L)$.

AppendixA.2. Proofs of Lemma 1 and Proposition 1

Proof of Lemma 1. Suppose that in period t the state is the convention $\sigma = (I, x, y)^m$.

First, observe why no convention with low investment $(L, x, y)^m$ can be stable. This follows from the fact that players who interpret their opponents' demands as absolute demands believe that their opponents will demand y at all investment levels. But if investment is efficient, then $V_H - C_H - y > V_L - C_L - y$. Thus, for all $y < V_H - C_H + C_L$, players α_a will choose high investment. Because for every convention with low investment $y \leq V_L < V_H - C_H + C_L$, no such convention can be stable.

Second, observe why no convention with high investment $(H, x, y)^m$ and $\chi \leq \underline{\chi}$ can be stable. This follows from the fact that players who interpret their opponents' demands as relative shares believe that their opponents will demand υV_I at all investment levels. Hence, the expected payoff for players α_r who choose low investment is given by $[(1-\upsilon)V_L]_x - C_L$. Because $(C_H - C_L)/(V_H - V_L)V_H$ and $(C_H - C_L)/(V_H - V_L)V_L$ are multiples of δ, $(1-\upsilon)V_H - C_H > [(1-\upsilon)V_L]_x - C_L \Leftrightarrow (1-\upsilon) > \frac{C_H - C_L}{V_H - V_L} = \underline{\chi}$. Thus, if $\chi \leq \underline{\chi}$ players α_r will choose low investment with positive probability. Hence, no such convention can be stable.

Consider a convention with high investment $(H, x, y)^m$ and $\chi > \underline{\chi}$. By the argument used in the previous paragraph, it follows that players α_r will choose high investment. Note that $(1-\upsilon)V_H - C_H > (1-\upsilon)V_L > 0$ implies that $x - C_H > 0$, which implies $y < V_H - C_H$. Thus, players α_a will choose high investment as well. At the bargaining stage, players α and β will clearly make demands x and y, respectively, because this maximizes their expected payoffs given their beliefs. □

Proof of Proposition 1. Suppose that in period $t+1$ to $t+k$, inclusive, players α_a and β_a sample A_t. Thus, with positive probability, they will play the same actions in period t to $t+k$, inclusive, which implies that A_{t+k} consists of identical records (I, x, y). Suppose that in period $t+k+1$ to $t+2k$, inclusive, players α_a sample A_t and players β_a sample A_{t+k}. With positive probability, players α_a will take the same actions as in the last k periods. Players β_a will make bargaining demands $V_I - x$. Hence, A_{t+2k} consists of identical records $(I, x, V_I - x)$.

Let $t' := t + 2k$ and $A_{t'} = (H, \chi_{t'} V_H, \upsilon_{t'} V_H)^k := A_{t+2k}$. Distinguish three cases:

Case 1. A_{t+2k} consists of identical records $(H, x, V_H - x)$ and $(1 - \upsilon_{t'}) > \underline{\chi}$. Suppose that in period $t'+1$ to period $m-k$, inclusive, players α_r and β_r sample $A_{t'}$. From the condition defining this case, it follows that players α_r will choose high investment At the bargaining stage, players α_r and β_r will demand

$\chi_{t'} V_H$ and $\upsilon_{t'} V_H$. Thus $A_{t'+m-k} = (H, \chi_{t'} V_H, \upsilon_{t'} V_H)^k \equiv (H, x', V_H - x')^k$. Thus, $s_{t'+3k+m} = (H, x', V_H - x')^m$ is a stable convention because it consists of identical records, $I = H$ and $(1 - \upsilon_{t'}) > \underline{\chi}$

Case 2. A_{t+2k} consists of identical records $(H, x, V_H - x)$ and$(1 - \upsilon_{t'}) \leq \underline{\chi}$. Consider the following sequence of events:

1. Suppose that in period $t' + 1$ to period $t' + k$, inclusive, players α_r and β_r sample $A_{t'}$. From the condition defining this case, it follows that players α_r will choose low investment. At the bargaining stage, because $1 - \upsilon = \chi$, players α_r and β_r will demand $[\chi_{t'} V_L]_x$ and $[\upsilon_{t'} V_L]_y$, which are smaller or equal to $\chi_{t'} V_L$ and $\upsilon_{t'} V_L$. Thus $A_{t'+k} = (L, [\chi_{t'} V_L]_x, [\upsilon_{t'} V_L]_y)^k$.
2. Suppose that in period $t' + k + 1$ to period $t' + 2k$, inclusive, players α_a and β_a sample $A_{t'+k}$. Because $\upsilon_{t'} V_L \leq V_L$, it follows that players α_a will choose high investment. At the bargaining stage, players α_a and β_a will demand $V_H - [\upsilon_{t'} V_L]_y$ and $V_H - [\chi_{t'} V_L]_x$, respectively. Thus, $A_{t'+2k} = (H, V_H - [\upsilon_{t'} V_L]_y, V_H - [\chi_{t'} V_L]_x)^k$.
3. Suppose that in period $t' + 2k + 1$ to period $t' + 3k$, inclusive, players α_a sample $A_{t'+k}$ and players β_a sample $A_{t'+2k}$. Players α_a take the same actions as in the previous k periods. Players β_a will demand $\upsilon_{t'} V_L$ at the bargaining stage. Thus, $A_{t'+3k} = (H, V_H - [\upsilon_{t'} V_L]_y, [\upsilon_{t'} V_L]_y)^k$.

Clearly, $\upsilon_{t'+3k} \leq \upsilon_{t'} V_L / V_H$. Thus, repeating the sequence described by (1)-(3), n times starting in period t' yields $\upsilon_{t'+n3k} \leq \upsilon_t (V_L/V_H)^n$. Because $V_L/V_H < 1$, $\upsilon_t (V_L/V_H)^n \underset{n\to\infty}{\longrightarrow} 0$. Thus, there exists a finite $\underline{n}$ for which $1 - \upsilon_{t'+\underline{n}3k} > \underline{\chi}$, and hence, there exists a positive probability of reaching a state $s_{t'+\underline{n}3k}$ where $A_{t'+\underline{n}3k} = (H, x', y')$ fulfills the condition of case 1 above.

Case 3. A_{t+2k} consists of identical records $(L, x, V_L - x)$.
Suppose that in period $t + 2k + 1$ to period $t + 3k$, inclusive, players α_a and β_a sample A_{t+2k}, which has positive probability. Because $V_L - x \leq V_L < V_H - C_H$, players α_a will choose high investment. At the bargaining stage, players α_a and β_a will make demands $V_H - V_L + x$ and $V_H - x$, respectively. Thus, $A_{t+3k} = (H, V_H - (V_L - x), V_H - x)^k$.
Suppose that in period $t + 3k + 1$ to period $t + 4k$, inclusive, players α_a A_{t+k} and β_a sample A_{t+2k}. Players α_a will take the same actions as in the previous k periods. Players β_a will make bargaining demands of $V_L - x$. Thus, $A_{t+4k} = (H, V_H - (V_L - x), V_L - x)^k$ fulfills the conditions of case 1 or 2 above.

Together the above cases show that from any initial state there exists a positive probability to reach a stable convention within finite time, which establishes the Proposition. □

AppendixA.3. Proof of Theorem 1

In this section, a slightly more general version of Theorem 1 is proved. The proof uses several Lemmata: Lemma 3 is an adaption Lemma 1 in Young (1993b), Lemma 4 and 5 are concerned with the existence of two specific paths of least resistance between two stable sets.

Lemma 3. Let $R_\delta(\xi)$ denote the least integer that is larger or equal to

$$r_\delta(\xi) = \min\left\{k^\alpha(1 - \frac{\xi-\delta}{\xi}), k^\beta(1 - \frac{V_H-\xi-\delta}{V_H-\xi}), k^\beta\frac{V_H-\xi}{V_H-\delta}\right\} \quad \text{(A.1)}$$

The minimum number of errors that is necessary to displace a stable convention

1. is given by R_δ where $1-\upsilon > \frac{C_H-C_L+\delta}{V_H-V_L}$.
2. is smaller or equal than R_δ where $1-\upsilon \leq \frac{C_H-C_L+\delta}{V_H-V_L}$.

Proof. It will first be shown that $r_\delta(\xi)$ is given by

$$\min\left\{k^\alpha(1 - \frac{\xi-\delta}{\xi}), k^\beta(1 - \frac{V_H-\xi-\delta}{V_H-\xi}), k^\beta\frac{V_H-\xi}{V_H-\delta}, k^\alpha\frac{\xi}{V_H-\delta}\right\} \quad \text{(A.2)}$$

It will then be argued why the rightmost term can be omitted.

Suppose first that the player who changes his or her best response interprets demand shares as absolute demand shares. Such players take into account only the absolute bargaining demands that their opponents have made, and not the investment levels that have preceded those demands.

1. If players α make mistakes ξ' that are *higher* than their conventional demands ξ, then the minimal number of mistakes that is necessary to make a player β choose $V_H-\xi'$ is given by $k^\beta(1-\frac{V_H-\xi'}{V_H-\xi})$. This expression is minimized by $V_H-\xi' = V_H-\xi-\delta$, which yields the second term of r_δ.
2. If players β make mistakes $(V_H-\xi')$ that are *lower* than their conventional demands $(V_H-\xi)$, then the minimal number of mistakes that is necessary to make a player α choose ξ' is given by $k^\alpha\frac{\xi}{\xi'}$. This expression is minimized if $\xi' = V_H-\delta$, which yields the fourth term of r_δ.
3. If players β make mistakes $(V_H-\xi')$ that are *higher* than their conventional demands $(V_H-\xi)$, then the minimal number of mistakes that is necessary to make a player α choose ξ' is given by $k^\alpha(1-\frac{\xi'}{\xi})$. This expression is minimized by $\xi' = \xi-\delta$, which yields the first term of r_δ.
4. If players α make mistakes ξ' that are *lower* than their conventional demands ξ, then the minimal number of mistakes that is necessary to make a player β choose $V_H-\xi'$ is given by $k^\beta\frac{V_H-\xi}{V_H-\xi'}$. This expression is minimized by $V_H-\xi' = V_H-\delta$, which yields the third term of r_δ.

Suppose now that the player who changes his best response interprets demand shares as relative demand shares. In this case, the opponents' relative demands are derived from the chosen investment levels and the absolute demands following them.

5. If players α and/or β make mistakes such that players β demand a *larger* share $(V_I - \xi')/V_I$ than their conventional share $(V_H - \xi)/V_H$, it might become profitable for players α to choose low investment. This is because for some $\upsilon' = (V_I - \xi')/V_I$ it can be the case that $(1-\upsilon')V_L \in X_\delta$ but $(1-\upsilon')V_H \notin X_\delta$, which might imply that $[(1-\upsilon')V_H]_x - C_H < (1-\upsilon')V_L - C_L$ and would make players α_r choose low investment.
   In the most extreme case $[(1-\upsilon)V_H]_x$ can be almost as small as $(1-\upsilon)V_H - \delta$. Because
$$(1-\upsilon)V_H - \delta - C_H \le (1-\upsilon)V_L - C_L \Leftrightarrow 1-\upsilon \le \frac{C_H - C_L + \delta}{V_H - V_L}$$
   the set of stable conventions where $[(1-\upsilon')V_H]_x - C_H \le (1-\upsilon')V_L - C_L$ is possible is bounded by the above expression.
   The minimal number of different demands that is necessary make a player α_r choose $(H, [(1-\upsilon')V_H]_x)$ is given by $k^\alpha(1 - \frac{[(1-\upsilon')V_H]_x}{\xi})$. This expression is minimized by $[(1-\upsilon)V_H]_x = \xi - \delta$, which is fulfilled if players α do not make an error, i.e., choose high investment, and players β make errors $V_H - \xi' = V_H - \xi + \delta$. Thus, for m$1 - \upsilon > \frac{C_H - C_L + \delta}{V_H - V_L}$ the number of errors is given by the first term of r_δ and for $1 - \upsilon \le \frac{C_H - C_L + \delta}{V_H - V_L}$ it must be smaller than or equal to the first term of r_δ.
6. If players α and/or β make mistakes such that players β demand a *smaller* share $(V_I - \xi')/V_I$ than their conventional share $(V_H - \xi)/V_H$, then, by the same arguments as in num. 5, a player α_r might either choose high investment and $[(1-(V_I - \xi')/V_I)V_H]_x$ or low investment and $[(1-(V_I - \xi')/V_I)V_L]_x$ in response. The minimal number of errors that is necessary for this is given by $k^\alpha \frac{\xi}{[(1-(V_I-\xi')/V_I)V_H]_x}$ and $k^\alpha \frac{\xi + C_L - C_H}{[(1-(V_I-\xi')/V_I)V_L]_x}$, respectively. Clearly, both expressions are minimized at $\min_{I\in\mathbb{I}, \xi'\in X_\delta}((V_I - \xi')/V_I) = -\delta/V_H$, which can be obtained if players β make errors $V_H - \xi' = \delta$ and players α do not make errors, i.e., choose high investment. Because $k^\alpha \frac{\xi}{[V_H - \delta]_x} = k^\alpha \frac{\xi}{V_H - \delta} < k^\alpha \frac{\xi + C_L - C_H}{[((V_H - \delta)/V_H)V_L]_x}$ for $\xi < (C_H - C_L)V_H/(V_H - V_L)$, this yields the fourth term of r_δ.
7. If players α make mistakes ξ'/V_I that claim a *larger* share than their conventional share ξ/V_H, then the minimal number of mistakes that is necessary to make a player β choose $[(1-\xi'/V_I)V_H]_y$ is given by $k^\beta(1 - \frac{[(1-\xi'/V_I)V_H]_y}{V_H - \xi})$. This expression is minimized by $[(1-\xi'/V_I)V_H]_y = V_H - \xi - \delta$, which is fulfilled if players α choose high investment and only make errors $\xi' = \xi + \delta$ at the bargaining stage. This yields the second term of r_δ.
8. If players α make mistakes ξ'/V_I that demand a *smaller* share than their conventional share ξ/V_H, then the minimal number of mistakes that is necessary to make a player β choose $[(1-\xi'/V_I)V_H]_y$ is given by $k^\beta \frac{V_H - \xi}{[(1-\xi'/V_I)V_H]_y}$. This expression is minimized by $[(1-\xi'/V_I)V_H]_y = V_H - \delta$, which is fulfilled if players α make no errors at the investment stage and make errors $\xi' = \delta$. This yields the third term of r_δ.

Thus, 1-8 together imply that for $1-\upsilon > \frac{C_H - C_L + \delta}{V_H - V_L}$ the minimal number of errors that is necessary to displace a convention is given by expression (A.2) and that it constitutes an upper bound for $1-\upsilon \leq \frac{C_H - C_L + \delta}{V_H - V_L}$.

Because $k^\beta\left(1 - \frac{V_H - \xi - \delta}{V_H - \xi}\right)$ and $k^\alpha \frac{\xi}{V_H - \delta}$ are strictly increasing in ξ and $k^\alpha(1 - \frac{\xi - \delta}{\xi})$ and $k^\beta \frac{V_H - \xi}{V_H - \delta}$ are strictly decreasing in ξ and X is a grid, r_δ is maximized by a unique value ξ^*or two adjunct values ξ^* and $\xi^* - \delta$.

Because of this, the minimal number of mistakes necessary to displace a convention is only given by $\left\lceil k^\alpha \frac{\xi}{V_H - \delta} \right\rceil$ if $\xi < \xi^*$ and $k^\beta\left(1 - \frac{V_H - \xi - \delta}{V_H - \xi}\right) \geq k^\alpha \frac{\xi}{V_H - \delta}$. From this, it follows that $\xi \leq \check{\xi} := \frac{V_H}{2} - \sqrt{\frac{V_H^2}{4} - \frac{k^\beta}{k^\alpha}\delta(V_H - \delta)}$ because for δ small enough, the determinant clearly is positive and $\frac{V_H}{2} + \sqrt{\frac{V_H^2}{4} - \frac{k^\beta}{k^\alpha}\delta(V_H - \delta)} > \xi^*$. Because, clearly, $\frac{k^\beta}{k^\alpha}\delta(V_H - \delta) \underset{\delta \to 0}{\longrightarrow} 0$, $\check{\xi} \underset{\delta \to 0}{\longrightarrow} 0$. Thus, for δ small enough $\check{\xi} < \underline{\xi}$. But Lemma 1 implies that for all $s \in \Sigma_H$, $\xi > \underline{\xi}$. Hence, the rightmost term of expression (A.2) can be omitted and r_δ is given by expression (A.1). □

In what follows, let $\overline{\xi}$ denote $[(C_H - C_L + \delta)V_H/(V_H - V_L) + \delta]_x$ and let $\hat{\xi}$ denote $\frac{V_H}{2} + \sqrt{\frac{V_H^2}{4} - \frac{k^\alpha}{k^\beta}\delta(V_H - \delta)}$.

Lemma 4. *If the minimal number of mistakes that is necessary to displace a convention* $(H, \xi, V_H - \xi)^m$ *is given by* $\left\lceil k^\beta \frac{V_H - \xi}{V_H - \delta} \right\rceil$ *and* δ *is small enough, then there exists a path with resistance* $\left\lceil k^\beta \frac{V_H - \xi}{V_H - \delta} \right\rceil$ *from* $(H, \xi, V_H - \xi)^m$ to a stable convention $(H, \xi', V_H - \xi')^m$ with $R_\delta(\xi') < \left\lceil k^\beta \frac{V_H - \xi'}{V_H - \delta} \right\rceil$ *and* $\xi' > \overline{\xi}$.

Proof. Because $k^\beta\left(1 - \frac{V_H - \xi - \delta}{V_H - \xi}\right)$ and $k^\alpha \frac{\xi}{V_H - \delta}$ are strictly increasing in ξ and $k^\alpha(1 - \frac{\xi - \delta}{\xi})$ and $k^\beta \frac{V_H - \xi}{V_H - \delta}$ are strictly decreasing in ξ, the minimal number of mistakes necessary to displace a convention is only given by $\left\lceil k^\beta \frac{V_H - \xi}{V_H - \delta} \right\rceil$ if $\xi > \xi^*$ and $k^\alpha(1 - \frac{\xi - \delta}{\xi}) \geq k^\beta \frac{V_H - \xi}{V_H - \delta}$. From this, it follows that $\xi \geq \hat{\xi} = \frac{V_H}{2} + \sqrt{\frac{V_H^2}{4} - \frac{k^\alpha}{k^\beta}\delta(V_H - \delta)}$ because for δ small enough the determinant clearly is positive and it can be shown that $\frac{V_H}{2} - \sqrt{\frac{V_H^2}{4} - \frac{k^\alpha}{k^\beta}\delta(V_H - \delta)} < \xi^*$. Observe also that $\hat{\xi} \underset{\delta \to 0}{\longrightarrow} V_H$ as clearly $\frac{k^\alpha}{k^\beta}\delta(V_H - \delta) \underset{\delta \to 0}{\longrightarrow} 0$.

In the reminder of this proof, let e denote $\left\lceil k^\beta \frac{V_H - \xi}{V_H - \delta} \right\rceil$. Consider the following sequence of events:

1. Suppose that the process is in state $s_t = (H, \xi, V_H - \xi)^m$ at time t and that $\xi \geq \hat{\xi}$ and $\chi_t > \underline{\chi}$. Suppose then that in period $t+1$ to $t+e$ inclusive, players α play (H, δ) by mistake and players β do not make any mistakes and thus demand $V_H - \xi$. Suppose that in period $t+e+1$ to $t+e+k$, inclusive, players α and β sample A_{t+e}. Thus, players α will play (H, ξ) and players β will demand $V_H - \delta$. Thus, $A_{t+e+k} = (H, \xi, V_H - \delta)^k$.

2. Suppose that in period $t+e+k+1$ to $t+e+2k$, inclusive, players α_r and β_r sample A_{t+e+k}. Because $1-\frac{V_H-\delta}{V_H} < \underline{\chi}$, players α_r will choose low investment. Because $\delta(V_L/V_H) < \delta$, there does not exist a compatible demand in X_δ and thus $[\delta(V_L/V_H)]_x = \{\}$. Thus, players α_r will make a random demand with support X_δ. Suppose that in period $t+e+k+1$ to $t+e+2k$ inclusive, players α_r make demands $[\frac{1}{2}(1+\underline{\chi})V_L]_x$, which has positive probability. Players β_r will demand $[(1-\chi_t)V_L]_y$. Thus, $A_{t+e+2k} = (L, [\frac{1}{2}(1+\underline{\chi})V_L]_x, [(1-\chi_t)V_L]_y)^k$.
3. Suppose that in period $t+e+2k+1$ to $t+e+3k$, inclusive, players α_r and β_a sample A_{t+e+k} and A_{t+e+2k}, respectively. Suppose that players α_r take the same actions as in the previous k periods. Then players β_a will demand $V_L-[\frac{1}{2}(1+\underline{\chi})V_L]_x$. Thus, $A_{t+e+3k} = (L, [\frac{1}{2}(1+\underline{\chi})V_L]_x, V_L-[\frac{1}{2}(1+\underline{\chi})V_L]_x)^k$.
4. Suppose that in period $t+e+3k+1$ to $t+e+4k$, inclusive, players α_r and β_r sample A_{t+e+3k}. Because $\underline{\chi} < 1$, $V_L - [\frac{1}{2}(1+\underline{\chi})V_L]_x < (1-\underline{\chi})V_L$ for δ small enough. Thus, players α_r will choose high investment and will make demands $[[\frac{1}{2}(1+\underline{\chi})V_L]_x(V_H/V_L)]_x$. Players β_r will demand $[(V_L-[\frac{1}{2}(1+\underline{\chi})V_L]_x)(V_H/V_L)]_y$. Thus, $A_{t+e+4k} = (H, [[\frac{1}{2}(1+\underline{\chi})V_L]_x(V_H/V_L)]_x, [(V_L-[\frac{1}{2}(1+\underline{\chi})V_L]_x)(V_H/V_L)]_y)^k$.
5. Suppose that in period $t+e+4k+1$ to $t+e+5k$, inclusive, players α_r sample A_{t+e+4k} and players β_a sample A_{t+e+3k}. Thus, players α_r will choose the same actions as in the previous k periods. Players β_a will choose $V_H - [[\frac{1}{2}(1+\underline{\chi})V_L]_x(V_H/V_L)]_x$. Thus, $A_{t+e+4k} = (H, [[\frac{1}{2}(1+\underline{\chi})V_L]_x(V_H/V_L)]_x, V_H - [[\frac{1}{2}(1+\underline{\chi})V_L]_x(V_H/V_L)]_x)^k$.

Let $\xi' = [[\frac{1}{2}(1+\underline{\chi})V_L]_x(V_H/V_L)]_x$, which can be reached with positive probability as argued above. To see that $\xi' \in (\overline{\xi}, \hat{\xi})$ observe $[[\frac{1}{2}(1+\underline{\chi})V_L]_x(V_H/V_L)]_x \in [(\frac{1}{2}(1+\underline{\chi})V_L-\delta)(V_H/V_L)-\delta, \frac{1}{2}(1+\underline{\chi})V_L(V_H/V_L)]$. Because $\underline{\chi} < 1$, clearly $\frac{1}{2}(1+\underline{\chi})V_L(V_H/V_L) < V_H$ and thus smaller than $\hat{\xi}$ for δ small enough. Also note that $\overline{\xi}$ can be written as $[\underline{\chi}V_H + \delta(1+\frac{V_H}{V_H-V_L})]_x$ and that $(\frac{1}{2}(1+\underline{\chi})V_L - \delta)(V_H/V_L) - \delta < \delta'$ can be written as $\frac{1}{2}(1+\underline{\chi})V_H - \delta(1-\frac{V_H}{V_L})$. Becasue $\underline{\chi} < 1$ it clearly follows that $\xi' > \overline{\xi}$ for δ small enough.

Thus, if in period $t+e+5k+1$ to $t+e+4k+m$, inclusive, players α and β sample A_{t+5k}, a stable convention $(H, \xi', V_H-\xi')^m$ will be reached which proves the Lemma. □

Lemma 5. *If a convention $\sigma = (H, \xi, V_H - \xi)$ with $\xi < \hat{\xi}$ is displaced by $R_\delta(\xi)$ or fewer errors in a way such that players α_a or α_r choose low investment, then for δ small enough there exists a path with resistance smaller or equal to $R_\delta(\xi)$ to a stable convention $\sigma' = (H, \xi', V_H - \xi')$ with $\hat{\xi} > \xi' > \overline{\xi}$.*

Proof. Num. 3 & 5 in the proof of Lemma 3 describe how a convention with $\xi < \hat{\xi}$ can be displaced by $r_\delta(\xi)$ or fewer errors such that players α prefer to choose low investment.

It is easy to see that if a convention is displaced in this way, players α_r will choose actions (L, x) with $x \in [(C_H-C_L)V_L/(V_H-V_L) - 2\delta, (C_H - C_L + \delta)V_L/(V_H-V_L)]$.

Suppose that in period t, A_t is such that if a player α_r samples A_t, it will choose actions (L, x). Suppose that in period $t+1$ to $t+k$, inclusive players α_r and β sample A_t. Thus $A_{t+k} = (L, x, y)^k$.

Suppose that in period $t+k+1$ to $t+2k$, inclusive, players α_r sample A_t and players β_a sample A_{t+k}. Then $A_{t+2k} = (L, x, V_L - x)$.

Suppose that in period $t+2k+1$ to $t+3k$, inclusive, players α_a and β sample A_{t+2k}. Then players α_a will choose $(H, V_H - V_L + x)$ and thus $A_{t+3k} = (H, V_H - V_L + x, y')^k$.

Suppose that in period $t+3k+1$ to $t+4k$, inclusive, players α_a sample A_{t+2k} and players β sample A_{t+3k}. Then $A_{t+4k} = (H, V_H - V_L + x, V_L - x)^k$.

To see that $V_H - V_L + x > \frac{C_H - C_L + \delta}{V_H - V_L} V_H + \delta \geq \overline{\xi}$ note that because $(C_H - C_L)V_L/(V_H - V_L) - 2\delta \leq x$, and $\frac{C_H - C_L}{V_H - V_L} < 1$,

$$\begin{aligned}
\frac{C_H - C_L + \delta}{V_H - V_L} V_H + \delta - x &\leq \frac{C_H - C_L + \delta}{V_H - V_L} V_H - \frac{C_H - C_L}{V_H - V_L} V_L + 2\delta \\
&= \frac{C_H - C_L}{V_H - V_L}(V_H - V_L) + \delta\left(\frac{V_H}{V_H - V_L} + 2\right) \\
&< V_H - V_L
\end{aligned}$$

for δ small enough.

To see that $V_H - V_L + x < \hat{\xi}$ note that because $x \leq (C_H - C_L + \delta)V_L/(V_H - V_L)$, $V_H - V_L + x \leq V_H - V_L + \frac{C_H - C_L + \delta}{V_H - V_L} V_L < V_H$. Thus, because $\hat{\xi} \underset{\delta \to 0}{\longrightarrow} V_H$,$V_H - V_L + x < \hat{\xi}$ for δ small enough.

Thus, if in period $t+4k+1$ to $t+3k+m$, inclusive, players α and β sample A_{t+4k} a stable convention $(H, V_H - V_L + x, V_L - x)^m$ is reached which proves the Lemma. □

In what follows let ξ^* denote $\min\{\arg\max_{\xi \in \Sigma} r_\delta(\xi)\}$.

Theorem (1'). *Assume δ to be sufficiently small and k^α, k^β and m to be sufficiently large. Then*

1. *if $\xi^* > \overline{\xi}$ the set of stochastically stable states is either given by a unique convention $(H, \xi^*, V_H - \xi^*)$ or, at most, two adjunct conventions $(H, \xi^*, V_H - \xi^*)$ and $(H, \xi^* + \delta, V_H - \xi^* - \delta)$.*
2. *if $\xi^* \leq \overline{\xi}$ the set of stochastically stable states is a subset of $\{\sigma \in \Sigma : \xi \leq \overline{\xi}\}$.*

As $\delta \to 0$ the distribution of the surplus ξ of all stable conventions converges to

$$\max\left\{\frac{C_H - C_L}{V_H - V_L} V_H, \xi^N\right\}$$

The following Proof relies on the ideas developed in Young (1993b) and combines them with techniques from Ellison (2000).

Proof. As already pointed out, $k^\beta\left(1 - \frac{V_H - \xi - \delta}{V_H - \xi}\right)$ and $k^\alpha \frac{\xi}{V_H - \delta}$ are strictly increasing in ξ and $k^\beta(1 - \frac{\xi - \delta}{\xi})$ and $k^\beta \frac{V_H - \xi}{V_H - \delta}$ are strictly decreasing in ξ. Becasue

X is a grid, r_δ is maximized by a unique value ξ^*or two adjunct values ξ^* and $\xi^* - \delta$ on X. For sufficiently large k, this implies that R_δ is also uniquely maximized by one or two adjunct values.

Distinguish two cases:

Case 1. ξ^* is smaller or equal to $\overline{\xi}$.

Let $\mathbb{S}^*$ be the set of stable conventions $\sigma = (H, \xi, V_H - \xi)$ with $\xi \leq \overline{\xi}$ and $R(\sigma) > \overline{r} := r_\delta(\overline{\xi} + \delta)$. Let $\overline{\sigma} := (H, \overline{\xi}, V_H - \overline{\xi})^m$, $\tilde{\xi} := \min\{\xi \in X : (H, \xi, V_H - \xi) \in \mathbb{S}^*\}$ and $\tilde{\sigma} := (H, \tilde{\xi}, V_H - \tilde{\xi})$.

To see that $\mathbb{S}^*$ is non-empty note that by construction $\overline{\xi}/V_H > (C_H - C_L + \delta)V_H/(V_H - V_L)$. Thus, Lemma 3 implies that $R(\overline{\sigma})$ is given by $R_\delta(\overline{\xi})$. Because r_δ is decreasing in ξ for $\xi > \xi*$, $r_\delta(\overline{\xi}) > \overline{r}$, which implies $R_\delta(\overline{\xi}) > \lceil \overline{r} \rceil$ for the sample size k^α high enough.

It will now be shown that $CR^*(\mathbb{S}^*) < R(\mathbb{S}^*)$. By Theorem 2 in Ellison (2000), this implies that the set of stochastically stable states must be contained in $\mathbb{S}^*$

By construction, the Radius of $\mathbb{S}^*$, $R(\mathbb{S}^*)$ is larger than $\lceil \overline{r} \rceil$.

To compute the $CR^*(\mathbb{S}^*)$, construct the following graph G_δ on $(\Sigma \setminus \mathbb{S}^*) \cup \overline{\sigma} \cup \tilde{\sigma}$:

1. For every stable convention $(H, \xi, V_H - \xi)^m$ with $\xi > \overline{\xi}$ and $r_\delta = k^\alpha(1 - \frac{\xi - \delta}{\xi})$, put the directed edge $((H, \xi, V_H - \xi)^m, (H, \xi - \delta, V_H - \xi + \delta)^m)$ into G_δ. From Lemma 3 num. 3 and 5, it follows that the resistance of these edges is given by R_δ
2. For every stable convention $(H, \xi, V_H - \xi)^m$ with $\xi > \overline{\xi}$ and $r_\delta = k^\beta \frac{V_H - \xi}{V_H - \delta}$, put the directed edge $((H, \xi, V_H - \xi)^m, (H, \xi', V_H - \xi')^m)$ with $\xi' = [[\frac{1}{2}(1 + \underline{\chi})V_L]_x (V_H/V_L)]_x$ into G_δ. From Lemma 4, it follows that the resistance of these edges is given by R_δ.
3. For every stable convention $(H, \xi, V_H - \xi)^m$ with $\xi < \overline{\xi}$ and $R(\sigma) < r_\delta(\xi)$, put a directed edge $((H, \xi, V_H - \xi)^m, (H, \xi\text{'}, V_H - \xi')^m)$ with $\xi' \in \left(\overline{\xi}, \hat{\xi}\right)$and resistance $R(\sigma)$ into G_δ. It follows from Lemma 5 that this is possible.
4. For every stable convention $(H, \xi, V_H - \xi)^m$ with $\xi < \overline{\xi}$ and $R(\sigma) = r_\delta(\xi)$, put a directed edge $((H, \xi, V_H - \xi)^m, (H, \xi + \delta, V_H - \xi - \delta)^m)$ into G_δ. From Lemma 3 num. 1 and 7, it follows that the resistance of these edges is given by R_δ.

This graph consists of two disjoint subgraphs, one i-tree $\overline{T}$ with root $\overline{\sigma}$, and another possibly degenerate i-tree $\tilde{T}$ with root $\tilde{\sigma}$.

For all $\sigma \in \Sigma \setminus \mathbb{S}^*$, let $cr(\sigma)$ denote $\overline{\sigma}$ if σ is a vertex in $\overline{T}$ and $\tilde{\sigma}$ if σ is a vertex in $\tilde{T}$. Also, for all $\sigma \in \Sigma \setminus \mathbb{S}^*$, let τ_σ^{CR} denote the set of vertices that are contained in the subgraph of G_δ that connects σ with $cr(\sigma)$.

From Lemma 3 and the construction of the graph, it follows that for every edge $(\sigma, \sigma') \in G_\delta$ the resistance $r(\sigma, \sigma') = R(\sigma)$. Thus, for $\sigma \in \Sigma \setminus \mathbb{S}^*$, $r(\sigma, cr(\sigma)) \leq \sum_{\sigma' \in \tau_\sigma^{CR}} R(\sigma')$ and the modified resistance

$$r^*(\sigma, cr(\sigma)) = r(\sigma, cr(\sigma)) - \sum_{\sigma' \in \tau_\sigma^{CR} \setminus \sigma} R(\sigma') \leq R(\sigma).$$

Because by construction $r_\delta(\xi) \leq \overline{r}$ for all $\sigma \in \Sigma \setminus \mathbb{S}^*$ with $\xi < \overline{\xi} + \delta$ and r_δ is decreasing in ξ for $\xi > \xi*$, $\max_{\sigma \in \Sigma \setminus \mathbb{S}^*} R(\sigma) = \lceil \overline{r} \rceil$. It follows that $CR^*(\mathbb{S}^*) = \max_{\sigma \in \Sigma \setminus \mathbb{S}^*} r^*(\sigma, cr(\sigma)) \leq \lceil \overline{r} \rceil < R(\mathbb{S}^*)$. Hence, it follows from Theorem 2 in Ellison (2000) that the set of stochastically stable states must be contained in $\mathbb{S}^*$.

Case 2. ξ^* is larger than $\overline{\xi}$.

Construct the following i-tree T^* on Σ with root ξ^*.

1. For every stable convention $(H, \xi, V_H - \xi)^m$ with $\xi > \xi^*$ and $r_\delta = k^\alpha(1 - \frac{\xi - \delta}{\xi})$, put the directed edge $((H, \xi, V_H - \xi)^m, (H, \xi - \delta, V_H - \xi + \delta)^m)$ into T^*.
2. For every stable convention $(H, \xi, V_H - \xi)^m$ with $\xi > \xi^*$ and $r_\delta = k^\beta \frac{V_H - \xi}{V_H - \delta}$, put the directed edge $((H, \xi, V_H - \xi)^m, (H, \xi', V_H - \xi')^m)$ with $\xi' = [[\frac{1}{2}(1 + \underline{\chi})V_L]_x (V_H/V_L)]_x$ into T^*.
3. For every stable convention $(H, \xi, V_H - \xi)^m$ with $\xi < \xi^*$and $R(\sigma) < R_\delta$, put a directed edge $((H, \xi, V_H - \xi)^m, (H, \xi\text{'}, V_H - \xi')^m)$ with $\xi' \in \left(\overline{\xi}, \hat{\xi}\right)$ and resistance $R(\sigma)$ into T^*.
4. For every stable convention $(H, \xi, V_H - \xi)^m$ with $\xi < \xi^*$and $R(\sigma) = R_\delta = \left\lceil k^\beta (1 - \frac{V_H - \xi'}{V_H - \xi}) \right\rceil$, put the directed edge $((H, \xi, V_H - \xi)^m, (H, \xi + \delta, V_H - \xi - \delta)^m)$ into T^*.

Lemmata 3-5 together imply that for every edge $(\sigma, \sigma') \in T^*$, $r(\sigma, \sigma') = R(\sigma)$.

If there exists a second maximizer of r_δ on X with radius equal to r_δ, then construct an analogous tree T^{**} with root $\xi^* - \delta$.

Because ξ^* and potentially $\xi^* - \delta$ are the unique maximizers of r_δ, T^* and T^{**} are the unique minimal i-Trees for k^α and k^β large enough. Together with Theorem 2 in Young (1993a), this implies that for $\xi^* > \overline{\xi}$, $(H, \xi^*, V_H - \xi^*)$ and possibly $(H, \xi^* + \delta, V_H - \xi^* - \delta)$ are the unique stochastically stable conventions.

Together with Lemma 3 in Young (1993b), which shows that $\lim_{\delta \to \infty} \arg\max_{\xi \in X} r_\delta(\xi)$ equals the Nash bargaining solution and the fact that $\overline{\xi} \underset{\delta \to 0}{\longrightarrow} (C_H - C_L)V_H/(V_H - V_L)$, this completes the proof of the Theorem. □