A classification of bargaining solutions by evolutionary origin

Sung-Ha Hwang and Jonathan Newton

April 2014

A classification of bargaining solutions by evolutionary origin ☆

Sung-Ha Hwang[a,*], Jonathan Newton[b]

[a]*Sogang University, Seoul, Korea.*
[b]*University of Sydney, Australia*

Abstract

For games of contracting under perturbed best response dynamics, varying the perturbations along two dimensions (uniform vs. logit, directed vs. undirected) gives four possibilities. Three of these select differing major bargaining solutions as stochastically stable. The fourth possibility yields a new bargaining solution which exhibits significant nonmonotonicities and demonstrates the interplay of two key drivers of evolutionary selection: (i) the ease of making errors; (ii) the ease of responding to errors.

Keywords: Evolution, adaptive learning, bargaining.

JEL Classification Numbers: C73, C78.

1. Introduction

For games of contracting characterized by a convex bargaining set, we show that under perturbed best response dynamics, by varying the perturbations along two simple dimensions (uniform vs. logit, directed vs. undirected), any of the major bargaining solutions can emerge as a stochastically stable convention. There are two populations, each corresponding to a position in a contract game: a coordination game with zero payoffs for miscoordination. Most of the time, agents play best responses to the distribution of play of the other population. However, from time to time an agent will make an *error* and play something other than a best response. Error probabilities can be *uniform* – all errors are equally likely, or *logit* – errors which incur a higher payoff loss for the agent making them are less likely to be made. Young (1998a) showed that the best response dynamic with uniform errors leads to the selection of the Kalai and Smorodinsky (1975) bargaining solution. Naidu, Hwang and Bowles (2010) showed that if the support of the perturbation distribution is restricted to be *directed* so that agents' errors only involve demanding more, never less, then the Nash (1950) bargaining solution is selected. The current paper extends the analysis of such models to

---

	Undirected	Directed
Uniform	Kalai-Smorodinsky	Nash bargaining
Logit	$Q4$	Egalitarian

Table 1: Stochastically stable bargaining solutions by error process.

payoff-dependent errors. In doing so, it makes three distinct contributions to the literature: a methodological contribution, an applied contribution, and an analytical contribution.

To combine state-dependent error probabilities with population dynamics, the paper makes a methodological contribution. In contrast to the case of uniform errors, the most probable transitions between conventions of contract games under the logit choice rule can involve errors being made by both populations. It is shown that for contract games played under a popular class of strategy revision rules, for large population size, the cost of such transitions can be well approximated by the cost of the most probable transition which involves errors by only a single population (Theorem 1). A sufficient condition on strategy revision rules for such an approximation is that error probabilities depend log-linearly on payoffs. This condition is satisfied by the class of exponential revision rules, a popular and flexible class of rules which includes the logit choice rule and exponential better reply rules.

Using these results, the paper makes an applied contribution to the literature on the evolution of bargaining solutions. It is shown that if the logit choice rule is used with directed errors, then the Egalitarian bargaining solution (Kalai, 1977) is selected. Moreover, the logit choice rule with undirected errors selects a new solution, which we call the $Q4$ solution as it corresponds to the remaining quadrant in table 1. Although it is developed from the same set of ingredients as the existing solutions, the $Q4$ solution exhibits significantly different properties. For example, holding the bargaining frontier close to the solution fixed, the $Q4$ solution is non-monotonic with respect to maximum obtainable payoffs: an increase in a player's maximum payoff can lead to a decrease in the amount he receives. Moreover, the solution can be non-monotonic with respect to stretches of the bargaining set parallel to the axes. Even Nash's bargaining solution, which is well known to breach the (individual) monotonicity axiom of Kalai and Smorodinsky (1975), is monotonic with respect to such stretches.

The paper makes an analytical contribution by highlighting the importance of interaction between two drivers of selection in evolutionary models: (i) the ease of making errors; (ii) the ease of responding to errors. Factor (i) enters directly through the error distribution. Factor (ii) enters because multiple errors in a population can be required to shift the process away from a convention, with the precise number of errors required depending on both the current

2

convention and on which errors occur. That is, from different conventions, different numbers of errors are required to induce some agent to best respond with an action other than the one chosen at the given convention. Under uniform errors, any mistake is equally likely in any state, so factor (i) plays no role in selection. As noted above, for large and equal population sizes, the Kalai-Smorodinsky bargaining solution is selected in such a setting. Logit errors are payoff dependent and so introduce factor (i). If both populations are of size 1, then factor (ii) plays no role in selection, and logit errors select the Egalitarian bargaining solution. Combining the two effects via logit errors and large populations, we obtain the unexpected result that the new solution does not necessarily lie between the Kalai-Smorodinsky and Egalitarian solutions: the player who receives the greater share under the Kalai-Smorodinsky solution can be the player who receives the lesser share under the $Q4$ solution.

That effects (i) and (ii) can work in opposite directions is by no means obvious from the existing literature. Consider the parallel literature on stochastic stability in Nash demand games. Young (1993b) shows that in two player Nash demand games, the Nash bargaining solution is stochastically stable. Agastya (1999) shows that if a cooperative game is modelled as a generalized Nash demand game, then the stochastically stable states are states in the core at which the maximum payoff over all players is minimized. Newton (2012b) shows that, under some conditions, the addition of joint strategic switching to such models leads to Rawlsian selection within the (interior/strong) core, maximizing the minimum payoff over all players. For the assignment game (Shapley and Shubik, 1971), a cooperative game for which the core has an empty interior so the methods of Newton (2012b) cannot be applied, Nax and Pradelski (2013) have recently shown a maxmin selection result within the core. Interestingly, although both papers attain similar results, these results arise in different ways. Nax and Pradelski (2013) use logit errors: error probabilities depends log-linearly on payoff differences. Selection then comes from (i) how hard it is for a player to make errors. Newton (2012b) uses uniform errors and sampling of opponents' behavior: selection comes from (ii) how hard it is for a player to respond to errors. In the papers cited above, these effects turn out to work towards a similar result. The current paper demonstrates that this is not always the case and that the combination of effects (i) and (ii) can create interesting nonmonotonicities.

This paper adds to the literature on perturbed adaptive dynamics, specifically on best response dynamics under various perturbations. The methodology, that of Freidlin and Wentzell (1984), was introduced to economics by Young (1993a) and Kandori, Mailath and Rob (1993). For any given stable state of an unperturbed dynamic, it is clear that there exists *some* perturbed dynamic such that the stationary distribution of the process gives a

probability close to one of the process being in that state.[1] See Bergin and Lipman (1996); van Damme and Weibull (2002) for more discussion of this kind of result. The point of the current paper is that the most commonly used error processes (uniform and logit), with and without an intuitive restriction on the domain of the errors, suffice to select all three major bargaining solutions: they are all part of the same evolutionary family. In addition, the family has a fourth member which is in some ways alike, but in other ways totally unlike the existing solutions. As well as previous results on processes within this family (Young, 1998a; Naidu et al., 2010) which are included in the classification of the current paper, a related result in Newton (2012a) incorporates coalitional behavior into perturbation structures, showing that if random errors are uniform, but coalitional behavior occurs with higher probability than random errors, then the Nash bargaining solution is selected.

Our results show that similar evolutionary processes to those used to justify common bargaining solutions found in the literature can justify other bargaining norms with unexpected and interesting attributes. The paper is organized as follows. Section 2 introduces the ideas of the paper with a simple example. Section 3 gives the evolutionary model and defines the bargaining solutions. Section 4 gives the approximation result which is the main methodological contribution of the paper. Section 5 uses this result to classify bargaining solutions by the evolutionary perturbations which give rise to them. Section 6 examines the properties of the new bargaining solution. Section 7 concludes.

2. Leading example

Consider the normal form game in figure 1. Note that there are three strict Nash equilibria: (A_α, A_β), (B_α, B_β), (C_α, C_β). We shall consider the evolution of play in three differing dynamic situations.

	A_β	B_β	C_β	D_β
A_α	**7, 5**	0, 0	0, 0	0, 0
B_α	0, 0	**6, 6**	0, 0	0, 0
C_α	0, 0	0, 0	**5, 7**	0, 0
D_α	0, 0	0, 0	0, 0	**0, 10**

Figure 1: A two player normal form game.

---

[1] For example, define a perturbation structure such that there is probability ε of transiting from any state to a given stable state i of the unperturbed dynamic, and let the probability of leaving i be ε^2. Then, for small ε, close to all the probability mass of the stationary distribution will be on state i.

4

2.1. Two players, logit errors.

Consider two players, α and β, who correspond to the positions in the game in figure 1. Each period, one of the players is chosen at random to adjust his strategy. Most of the time, a player so chosen plays a best response to the current action of the opposing player, with each possible best response being chosen with equal probability if there is more than one best response. However, with small probability, the chosen player will make an error and switch to an action which is not a best response. Each possible error occurs with a probability of order ε^l, where ε is some small number and l is the difference between the payoff from playing a best response and the payoff from making the error. For example, if the current actions of the players are (A_α, A_β) and player β is chosen to update his action, then he will play action B_β with a probability of order $\varepsilon^{5-0} = \varepsilon^5$. Following this error, if player α is chosen to update his action, he can best respond with B_α, and the pure Nash profile (B_α, B_β) is reached. In a similar manner, (B_α, B_β) can be reached from (C_α, C_β). However, to leave (B_α, B_β), an error of probability of order ε^6 is required. This implies that for small ε, almost all of the weight of the stationary distribution of this Markov chain will be on the state in which (B_α, B_β) is played. The limiting stationary distribution as $\varepsilon \to 0$ places all weight on this state: it is uniquely stochastically stable. Note that a single error is all that is required to induce a different best response from the opposing player: selection is entirely driven by how easy it is to make errors. This creates a bias towards egalitarianism.

2.2. Two populations of ten agents each, uniform errors.

Now assume that rather than a single player for each position of the game, there exist two populations, each of which comprises 10 agents. One of the populations is associated with position α in the game, the other population with position β. Each period, a single agent from one of the populations is chosen at random to adjust his strategy. Most of the time, an agent so chosen plays a best response to the distribution of current actions of the opposing population. With small probability, the chosen agent will make an error and switch to an action which is not a best response. Each possible error occurs with a probability of order ε. That is, each possible error occurs with similar probability, independent of the payoff loss incurred by making the error. Note that in the population setting, starting from a state in which every agent plays actions corresponding to some Nash equilibrium of the game, multiple errors by agents in one population can be necessary to induce an agent in the opposing population to play a best response that differs from the action corresponding to the original Nash equilibrium. For example, starting from a state in which every agent plays action A, at least 4 errors by α-agents where they switch to D_α are required to induce a β-agent to best respond with D_β. If every β-agent switches to D_β, then when an α-agent

5

is chosen to update his strategy, any action will be a best response. If the α-agents switch to action C_α, then β-agents can best respond with C_β, and the state in which every agent plays C will be reached. A similar transition is possible from the state in which every agent plays B to the state in which every agent plays C. However, from the state in which every agent plays C, at least 5 errors will be necessary to induce a best response other than to play C. Therefore, for small ε, almost all of the weight of the stationary distribution of this Markov chain will be on the state in which every agent plays C. This state is uniquely stochastically stable. Note that in this process every error occurs with similar probability: selection is entirely driven by how easy it is to respond to errors. This creates a bias favouring populations who have some possibility of high payoffs, such as the possibility of a payoff of 10 for β-agents in the game in figure 1.

2.3. Two populations of ten agents each, logit errors.

Now consider the process with two populations of 10 agents each, and with perturbations occuring with probabilities of order ε^l as described above. It can be checked that the easiest way to transition from every agent playing A to any state corresponding to one of the other pure Nash equilibria is as follows. First, 3 of the α-agents make errors and play D_α. These errors occur with probability of order ε^7 each. The payoff loss for a β-agent of playing D_β rather than A_β is then $5 \cdot {}^7/_{10} - 10 \cdot {}^3/_{10} = {}^1/_2$. Next, let all 10 of the β-agents make errors and play D_β. These errors occur with probability of order $\varepsilon^{1/2}$ each. Following this, when an α-agent is chosen to update his strategy, any action will be a best response. Therefore the overall transition probability is of order $(\varepsilon^7)^3 \cdot (\varepsilon^{1/2})^{10} = \varepsilon^{26}$. Note that these lowest cost transitions involve errors by agents in both populations. One contribution of the current paper is to show that for large populations, the calculation of transition costs can be simplified by restricting attention to paths in which errors only occur in a single population.

From the state in which every agent plays B (C), there exists a transition to the state in which every agent plays A in which 4 (5) α-agents make errors and play D_α, giving a transition probability of order ε^{24} (ε^{25}). Therefore, for small ε, almost all of the weight of the stationary distribution of this Markov chain will be on the state in which every agent plays A. This state is uniquely stochastically stable. Selection in this process is driven by both how easy it is to make errors and by how easy it is to respond to errors. By combining these two effects, the outcome is better for α-agents than that selected by either of these effects acting on their own. The remainder of the paper will examine the interplay of these two effects in settings where the range of possible coordination outcomes is given by a convex bargaining set.

3. Evolution and bargaining

3.1. Bargaining Problem

Consider two positions, α and β. Players in these positions bargain over which pair of payoffs is selected from a bargaining set. Let $S \subset \mathbb{R}^2$ be the bargaining set which is convex and compact and $a \in \mathbb{R}^2$ be the disagreement point: the payoffs that players receive when agreement is not reached. We suppose that a is normalized to $(0,0)$ which belongs to the bargaining set. We assume that for each S there exists a decreasing, differentiable, and concave function, f_S, such that $(t, f_S(t))$ is the efficient allocation in which α and β players receive t and $f_S(t)$, respectively. Thus, the maximum payoff that players α and β can obtain from bargaining are

$$\bar{s}_\alpha := \sup \{t : f_S(t) \geq 0\} \ \text{and} \ \bar{s}_\beta := \sup \{f_S(t) : t \geq 0\},$$

respectively. We shall also routinely omit the subscript from $f_S(.)$, writing $f(.)$.

A bargaining solution maps bargaining problems to allocations. A bargaining solution is essentially a rule by which surplus in a bargaining problem is allocated. The three bargaining solutions most commonly used in economics are the Nash bargaining solution (Nash, 1950), the Kalai-Smorodinsky bargaining solution (Kalai and Smorodinsky, 1975), and the Egalitarian bargaining solution (Kalai, 1977).

Definition 1. *Let S be a bargaining set with a bargaining frontier given by $f_S(.)$. Denote by $(t^{.}, f_S(t^{.}))$ the solution associated to the bargaining solution under consideration.*

Kalai-Smorodinsky solution	$\frac{t^{KS}}{\bar{s}_\alpha} = \frac{f_S(t^{KS})}{\bar{s}_\beta}$.
Nash bargaining solution	$t^{NB} \in \arg\max_{0 \leq t \leq \bar{s}_\alpha} t f_S(t)$.
Egalitarian solution	$t^E = f_S(t^E)$.

These solutions each uniquely satisfy distinct sets of intuitively appealing axioms. Such axioms are further discussed in section 6. Furthermore, it has been shown that the Kalai-Smorodinsky and Nash solutions can emerge from plausible models of adaptive behavior (Young, 1998a; Binmore, Samuelson and Young, 2003; Naidu et al., 2010; Newton, 2012a). One contribution of the current paper is to show that the same is true for the Egalitarian solution.

3.2. Evolutionary contracting

Consider two populations of agents − α and β populations − of size N.[2] Two agents, one from each population, are matched to play a coordination game. The set of possible outcomes on which coordination is possible corresponds to a bargaining set as described in section 3.1. Similarly to previous literature on evolution and bargaining, we discretize the bargaining set as follows. Let $n \in \mathbb{Z}_+$, $\delta = \delta_n = n^{-1}\bar{s}_\alpha$, and $\mathcal{I} := \{0, 1, 2, \cdots, n\}$ and suppose that the two agents each play a strategy from the following sets, respectively:

$$\text{player } \alpha: \ \{0, 1\delta, \cdots, n\delta\}, \ \text{player } \beta: \{f(0), f(\delta), f(2\delta), \cdots, f(n\delta)\}.$$

To simplify notation, we will denote by i_α and i_β strategies $i\delta$ and $f(i\delta)$, respectively.[3]

A strategy profile or a state of two populations is described by $x := (x_\alpha, x_\beta)$, where x_α and x_β are vectors giving the number of agents using each strategy. Thus, the state space Ξ is

$$\Xi := \left\{ (x_\alpha, x_\beta) : \sum_{l \in \mathcal{I}} x_\alpha(l) = N, \ x_\alpha(l) \in \mathbb{N}_0, \ \sum_{l \in \mathcal{I}} x_\beta(l) = N, \ x_\beta(l) \in \mathbb{N}_0 \right\}$$

More explicitly, we have $(x_\alpha, x_\beta) = ((x_\alpha(0), x_\alpha(1), \cdots, x_\alpha(n)), (x_\beta(0), x_\beta(1), \cdots, x_\beta(n)))$, where $x_\beta(2)$, for example, denotes the number of β-agents playing strategy $2_\beta = f(2\delta)$.

We consider contract games (Young, 1998a), coordination games in which players who demand the same outcome receive their associated payoffs, and receive nothing otherwise. That is, the payoffs for a contract game are

$$(\pi_\alpha(i_\alpha, j_\beta), \ \pi_\beta(j_\beta, i_\alpha)) = \begin{cases} (i\delta, \ f(i\delta)) & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$

To avoid notational clutter, we shall occasionally denote $\pi_\alpha(i) := \pi_\alpha(i_\alpha, i_\beta)$ and $\pi_\beta(i) := \pi_\beta(i_\beta, i_\alpha)$. Agents from each population are matched to play the contract game and thus, the expected payoff of an α agent who plays strategy i_α is $\pi_\alpha(i_\alpha, x_\beta) := \sum_{l \in \mathcal{I}} \pi_\alpha(i_\alpha, l) x_\beta(l)/N$, given that the fraction of the β population using strategy l is $x_\beta(l)/N$. Similarly, the expected payoff of a β-agent who plays strategy i_β is $\pi_\beta(i_\beta, x_\alpha) := \sum_{l \in \mathcal{I}} \pi_\beta(i_\beta, l) x_\alpha(l)/N$.

---

[2]Exposition is simplified by the assumption that the populations are of the same size. This is always the case when the two populations represent roles played by different agents in the same population. That is, each agent could be considered to appear twice: he will play one strategy when he plays as an α-player, and another strategy when he plays as a β-player.

[3]Note that the discretization is uniform for α, but not for β. This can be reversed without changing results, or, alternatively, the bargaining set can be approximated by the points of a square lattice with nearest neighbor distance δ.

We consider a discrete time strategy updating process defined as follows. At each period, an agent from either the α population or the β population is randomly chosen and matched to play the contract game with an agent from the other population. The chosen agent selects a strategy (which will be used to play the game) based on his evaluation of the expected payoffs of the different strategies. The agents idiosyncratically experiment with non-optimal strategies, or simply make mistakes. The probability of such mistakes will be parameterized by a parameter η, and larger values of η will correspond to higher mistake probabilities.

To study various behavioral rules of strategy revising agents, we suppose that the transition probabilities for the strategy updating process, P^η, admit a real-valued function $V(x, y)$ satisfying

$$\lim_{\eta \to 0} -\eta \ln P^\eta(x, y) = V(x, y) \tag{1}$$

where V is defined over the set of all $x, y \in \Xi$ such that $P^{\hat\eta}(x, y) > 0$ for some $\hat\eta > 0$ (see Beggs, 2005; Sandholm, 2010b). Here, $P^\eta(x, y)$ is the transition probability from state x to state y. The *resistance* of a transition from x to y, $V(x, y)$, measures the rarity of transitions from x to y. To determine the function V for a given P^η, we sometimes use the following fact, called a "largest-exponent wins" principle (See Laplace's method in Dembo and Zeitouni, 1998, p.137 and remark there. See also Freidlin and Wentzell, 1984, pp.71-72).

Lemma 1. *Suppose that f and g are functions on Ξ and g is positive. Then we have*

$$\sum_{x \in \Xi} \exp(\eta^{-1} f(x)) g(x) \asymp \exp(\eta^{-1} \max_{x \in \Xi} f(x))$$

where $a_\eta \asymp b_\eta$ means $\lim_{\eta \to 0} \eta(\log a_\eta - \log b_\eta) = 0$.

To specify transition probabilities more precisely, we first write as $x^{\gamma, l, l'}$ the state induced from x by a γ population ($\gamma = \alpha$ or β) agent's strategy change from l to l'. That is, for $\gamma = \alpha, \beta$, we have:

$$x_\gamma^{\gamma, l, l'}(i) := \begin{cases} x_\gamma(i) & \text{if } i \neq l, l' \\ x_\gamma(i) - 1 & \text{if } i = l \\ x_\gamma(i) + 1 & \text{if } i = l' \end{cases}$$

In the context of population dynamics, the typical form of the transition probability $P^\eta(x, y)$

is as follows:

$$P^\eta(x,y) = \begin{cases} \frac{x_\gamma(l)}{2N} p_\gamma^\eta(l'|l,x) & \text{if } y = x^{\gamma,l,l'} \text{ for some } \gamma, l, l' \\ 1 - \sum_{\gamma,l,l'} \frac{x_\gamma(l)}{2N} p_\gamma^\eta(l'|l,x) & \text{if } y = x \\ 0 & \text{otherwise} \end{cases}, \qquad (2)$$

where $\sum_{l'} p_\gamma^\eta(l'|l,x) \leq 1$. In equation (2), the factor $x_\gamma(l)/2N$ in the first line accounts for the probability of randomly choosing a γ- agent with strategy l. The term $p_\gamma^\eta(l'|l,x)$ gives the conditional probability that a chosen agent from population γ will switch from strategy l to strategy l' given that the state of the populations is x. We will specify this probability shortly. The second line in the transition probability normalizes and the last line means that the only transitions which are possible from x are to states $x^{\gamma,l,l'}$. Specification (2) defines a family of Markov chains parameterized by η. This paper considers processes with perturbations varying in two dimensions: the perturbations can be uniform or logit, and they can be directed or undirected. The definitions of these concepts shall now be given.

3.3. (Generalized) Logit choice rule

Under the generalized logit choice rule, from state x, a strategy-revising agent who is currently playing l, will switch to l' with a probability given by

$$p_\gamma^\eta(l'|l,x) := \frac{q_l \exp(\eta^{-1}\pi_\gamma(l',x))}{\sum_{\tilde{l}} q_{\tilde{l}} \exp(\eta^{-1}\pi_\gamma(\tilde{l},x))} \qquad (3)$$

where q_l, $l \in \mathcal{I}$, are positive constants. When $q_l = 1$ for all $l \in \mathcal{I}$, equation (3) gives the logit choice rule which is well-known in the literature of evolutionary games (see Blume, 1993, 1996).[4] The parameter η measures the degree of perturbation in best response rules and can be interpreted as noise in observing others' strategies and evaluating expected payoffs, the frequency of mistakes or experimentation in strategy revision, and so on. As $\eta \to 0$, the probability of a strategy-revising agent playing anything other than a best response approaches zero. From Lemma 1, it follows that the resistance $V(x, x^{\gamma,l,l'})$ of a transition from x to $x^{\gamma,l,l'}$ equals

$$\lim_{\eta \to 0} -\eta \ln P^\eta(x, x^{\gamma,l,l'}) = \max_{\tilde{l}} \pi_\gamma(\tilde{l},x) - \pi_\gamma(l',x).$$

The interpretation of this is that the probabilities of errors which cause higher payoff losses approach zero faster as $\eta \to 0$. Error probabilities are asymptotically log-linear in payoff

---

[4]Alternatively, it could be the case that $q_{\tilde{l}} = 1$ for some $\tilde{l}$ and $q_{\tilde{l}} \approx 0$ otherwise. Such a behavioral rule might involve a comparison of some target strategy with some fixed strategy (see Weibull, 1995).

loss.

3.4. Uniform mistake rule

When errors are uniform, every error occurs with the same probability. That is, from state x, a strategy-revising agent who is currently playing l, will switch to l' with a probability given by

$$p_\gamma^\eta(l'|l,x) := \begin{cases} \frac{1}{|\arg\max_{\tilde{l}} \pi_\gamma(\tilde{l},x)|}(1-\epsilon) + \frac{1}{n+1}\epsilon & \text{if } l' \in \arg\max_{\tilde{l}} \pi_\gamma(\tilde{l},x) \\ \frac{1}{n+1}\epsilon & \text{otherwise} \end{cases}$$

where $\epsilon = \exp(-\eta^{-1})$. It follows that the resistance $V(x, x^{\gamma,l,l'})$ of a transition from x to $x^{\gamma,l,l'}$ equals

$$\lim_{\eta \to 0} -\eta \ln P^\eta(x, x^{\gamma,l,l'}) = \begin{cases} 0 & \text{if } l' \in \arg\max_{\tilde{l}} \pi_\gamma(\tilde{l},x) \\ 1 & \text{if } l' \notin \arg\max_{\tilde{l}} \pi_\gamma(\tilde{l},x) \end{cases}$$

3.5. Directed & Undirected errors

Let $\Delta_\gamma(x)$ be the set of strategies for an agent of type $\gamma = \alpha, \beta$ which involve demanding at least as much as the agent demands when best responding to the strategy distribution of the other population.

$$\Delta_\gamma(x) := \{l : \pi_\gamma(l,l) \geq \pi_\gamma(l',l') \text{ for some } l' \in \arg\max_{\tilde{l}} \pi_\gamma(\tilde{l},x)\}.$$

Undirected error processes retain the conditional probabilities $p_\gamma^\eta(l'|l,x)$ described above for logit and uniform errors. Directed errors are when agents never demand less than their best response, but can demand more. This fits with an interpretation of the perturbations as idiosyncratic experimentation by agents to see if they can obtain a higher payoff. The conditional probabilities of switching for directed processes are given by:

$$\hat{p}_\gamma^\eta(l'|l,x) := \begin{cases} \frac{p_\gamma^\eta(l'|l,x)}{\sum_{\tilde{l} \in \Delta_\gamma(x)} p_\gamma^\eta(l'|l,x)} & \text{if } l' \in \Delta_\gamma(x) \\ 0 & \text{otherwise} \end{cases}$$

where $p_\gamma^\eta(l'|l,x)$ denotes the conditional probability for the corresponding undirected process.

3.6. Conventions and stochastic stability

The process with $\eta = 0$, or $\epsilon = 0$, is the *unperturbed process*. The recurrent classes of the unperturbed process are the absorbing states in which all α and β agents coordinate on the same strategy, and each agent type receives nonzero payoff. We shall denote by

11

E_i, $i \in \{0, \ldots, n\}$, the state in which all α-agents play δi and all β-agents play $f(\delta i)$, $x_\alpha(i) = N$, $x_\beta(i) = N$. Hence, the absorbing states of the process are precisely those in the set $\Lambda := \{E_1, \ldots, E_{n-1}\}$. Following Young (1993a), we refer to these states as *conventions*. Let $L := \{1, \ldots, n-1\}$ index the states in $\Lambda = \{E_i\}_{i \in L}$.

Stochastic stability analysis selects from amongst conventions by taking a limit of the stationary distributions of perturbed processes as $\eta \to 0$. A preliminary step is to show that such stationary distributions exist and are unique for any given positive value of η.

Lemma 2. *Each process, uniform or logit, undirected or directed, for given $\eta > 0$, has a unique stationary distribution, which we denote μ_η.*

Proof. Note that for all $x \in \Xi$, $n_\alpha \in \Delta_\alpha(x)$, so for all of our processes, from any $x \in \Xi$, unless $x_\alpha(n_\alpha) = N$, we have that $P^\eta(x, x^{\alpha, l, n}) > 0$ for some $l \neq n$. In this way, the process reaches a state with $x_\alpha(n_\alpha) = N$, from which n_β is a best response for any β-agent. Therefore, from any $x \in \Xi$, with positive probability E_n will be reached within $2N$ periods. As the state space is finite, standard results in Markov chain theory[5] imply that for all $\eta > 0$, P^η has a unique recurrent class and μ_η exists and is unique. ∎

By standard arguments (see Young, 1998b), the limit $\mu := \lim_{\eta \to 0} \mu_\eta$ exists, and for any $x \in \Xi$, $\mu(x) > 0$ implies that x is in a recurrent class of the process with $\eta = 0$. In our setting, this implies $x \in \Lambda$.

Definition 2. *A state $x \in \Xi$ is stochastically stable if $\mu(x) > 0$.*

In a similar way that $V(\cdot, \cdot)$ measures the rarity of single steps in the dynamic, we will use a concept, overall cost, that measures the rarity of a transition between any two states over any number of periods. Let $\mathcal{P}(x, x')$ be the set of finite sequences of states $\{x^1, x^2, \ldots, x^T\}$ such that $x^1 = x$, $x^T = x'$ and for some $\hat{\eta} > 0$, $P^{\hat{\eta}}(x^\tau, x^{\tau+1}) > 0$, $\tau = 1, \ldots, T-1$.

Definition 3. *The overall cost of a transition between $x, x' \in \Xi$ is:*

$$c(x, x') := \min_{\{x^1, \ldots, x^T\} \in \mathcal{P}(x, x')} \sum_{\tau=1}^{T-1} V(x^\tau, x^{\tau+1}).$$

If there is no positive probability path between x and x' then let $c(x, x') = \infty$. We shall be interested in the cost of transitions between conventions. In the current setting, this quantity is always finite. Denote the overall cost functions for the undirected-uniform, undirected-logit, directed-uniform and directed-logit processes by c^U, c^L, $\hat{c}^U$, $\hat{c}^L$ respectively.

---

[5]See, for example, "Probability" by Shiryaev (1995, p.586, Theorem 4).

4. Transition costs for exponential family: a boundary problem and a solution

For uniform error processes, any least cost transition from a given convention of the contract game to some other convention is driven by errors within a single population. That is, from some initial convention, errors occur in one of the populations, following which, agents from the other population can best respond in a way which differs from the initial convention. Errors in the population which best responds differently would be superfluous. This logic does not translate to situations in which error costs are state dependent. It is theoretically possible that on a path between conventions, errors in one population could facilitate errors in the other population. Previous work does not explicitly study the implications of such transitions (e.g. Belloc and Bowles, 2013; Staudigl, 2012). It is shown in the following example, that it is in fact possible for least cost transitions to require that both populations make errors. Fortunately, it transpires (Theorem 1) that when population sizes are large, such transitions can be ignored for the purpose of assessing stochastic stability. Define the basin of attraction of a convention, the set of states from which the unperturbed dynamic converges to that convention with probability 1.

Definition 4. *The basin of attraction of E_i is given by*

$$D(E_i) = \{x \in \Xi : \ c(x, E_i) = 0, \ c(x, E_j) > 0 \ \text{for all} \ j \neq i\}$$

For a given convention, E_i, we seek to determine the lowest cost transition path to some state outside of the convention's basin of attraction, $D(E_i)$.

Example 1. *Consider the logit dynamics. Suppose that we have the following game:*

	1_β	2_β
1_α	5,4	0,0
2_α	0,0	7,8

We suppose that $N = 5$. First we compute the minimum escaping cost from E_1 such that $D(E_2)$ is reached by transitions by only one population. We have

$$\left\lceil N \frac{\pi_\alpha(1)}{\pi_\alpha(1) + \pi_\alpha(2)} \right\rceil \pi_\beta(1) = 12, \quad \left\lceil N \frac{\pi_\beta(1)}{\pi_\beta(1) + \pi_\beta(2)} \right\rceil \pi_\alpha(1) = 10.$$

Next, consider the following transitions:

$$\beta \text{ switching from } 1_\beta \text{ to } 2_\beta \ \left\lceil N \frac{\pi_\alpha(1)}{\pi_\alpha(1) + \pi_\alpha(2)} \right\rceil - 1 \text{ times and}$$

$$\alpha \text{ switching from } 1_\alpha \text{ to } 2_\alpha \ \left\lceil N \frac{\pi_\beta(1)}{\pi_\beta(1) + \pi_\beta(2)} \right\rceil \text{ times.}$$

Then this gives a path from E_1 to E_2 and the cost of the path is given by

$$\left(\left\lceil N\frac{\pi_\alpha(1)}{\pi_\alpha(1)+\pi_\alpha(2)}\right\rceil - 1\right)\pi_\beta(1) + \left\lceil N\frac{\pi_\beta(1)}{\pi_\beta(1)+\pi_\beta(2)}\right\rceil$$
$$\times \left[\frac{1}{N}\left(N - \left(\left\lceil N\frac{\pi_\alpha(1)}{\pi_\alpha(1)+\pi_\alpha(2)}\right\rceil - 1\right)\right)\pi_\alpha(1) - \frac{1}{N}\left(\left\lceil N\frac{\pi_\alpha(1)}{\pi_\alpha(1)+\pi_\alpha(2)}\right\rceil - 1\right)\pi_\alpha(2)\right]$$
$$= 2\pi_\beta(1) + 2\left(\frac{3}{5}\pi_\alpha(1) + \frac{2}{5}\pi_\alpha(2)\right) = 8.4$$

which is smaller than the minimum costs of transitions driven by a single population.

So we see that in Example 1, the least cost transition from E_1 to E_2 requires errors to be made by agents in both populations. This is due to the behavior of the process close to the boundary of the basin of attraction of E_1. After β-agents make errors, the cost of errors by α-agents is reduced. A single error by a β-agent has a lower cost than the consequent reduction in the cost of two errors by α-agents. Two errors by β-agents reduce the cost further still. However, after two errors have been made by β-agents, subsequent errors by β-agents no longer have a linear effect on the cost of an error by an α-agent due to the zero lower bound on $V(\cdot,\cdot)$. Following two errors by β-agents, the cost of a third error by a β-agent is higher than the cost of two errors by α-agents. A moment's consideration leads one to see that, for any given population size, examples can be constructed for which least cost transitions involve errors by both populations.

Fortunately, for exponential revision protocols, a class that includes the logit choice rule, we shall show that when the population size is large, starting from a convention E_i, the least cost transition path out of the basin of attraction of E_i has a cost approximately equal to the least cost such transition from the restricted class of paths which only involve a single population making errors. The class of exponential revision protocols is defined (Sandholm, 2010b) as the processes satisfying

$$\log\left(\frac{p^\eta_\gamma(i|j,x)}{p^\eta_\gamma(j|i,x)}\right) = \eta^{-1}\left(\pi(j,x_{\gamma-}) - \pi(i,x_{\gamma-})\right).$$

where $\gamma^- := \alpha$ for $\gamma = \beta$ and $\gamma^- := \beta$ for $\gamma = \alpha$. This is a flexible class of rules which includes the Baker and Metropolis better reply dynamics, and the logit choice rule (see Appendix A).

Definition 5. *Let $\overrightarrow{c}(.,.)$ be a cost function restricted to minimize resistance over paths satisfying (i) errors are only made by one of the populations, and (ii) only a single alternative strategy is ever played in error. That is,*

$$\overrightarrow{c}(x,x') := \min_{\{x^1,...,x^T\}\in\overrightarrow{\mathcal{P}}(x,x')}\sum_{\tau=1}^{T-1}V(x^\tau,x^{\tau+1}).$$

14

where $\overrightarrow{\mathcal{P}}(x, x')$ is the set of paths from x to x' such that for any $\{x^1, \ldots, x^T\} \in \overrightarrow{\mathcal{P}}(x, x')$, there exists some $\gamma \in \{\alpha, \beta\}$, k, j, such that for any x^t, x^{t+1} such that $V(x^t, x^{t+1}) > 0$, we have that $x^{t+1} = (x^t)^{\gamma, k, j}$.

For the remainder of the paper, to aid conciseness we use the following notation.

Definition 6. *The relations $\approx$ and $\lesssim$ are defined such that, for a and b dependent on N, $a \approx b$ if and only if for all $\epsilon > 0$, there exists $\bar{N}$ such that for all $N > \bar{N}$, $a_N \in (b_N - \epsilon, b_N + \epsilon)$. Likewise, $a \lesssim b$ if and only if for all $\epsilon > 0$, there exists $\bar{N}$ such that for all $N > \bar{N}$, $a_N \leq b_N + \epsilon$.*

The main theorem of this section can now be stated. Under exponential revision protocols, including the logit choice rule, lowest cost transitions between conventions of contract games can be approximated by the lowest cost transitions which involve errors being made by agents in only one of the populations, and those agents making only one type of error.

Theorem 1. *Let Λ, indexed by L, be the set of strict Nash equilibria of a contract game. Let $|\Lambda| \geq 2$. Let $i \in L$ be fixed. Let the strategy revision rule be an exponential revision protocol. Then*

$$\frac{1}{N} \min_{j \neq i} c(E_i, E_j) \approx \frac{1}{N} \min_{j \neq i} \overrightarrow{c}(E_i, E_j) \tag{4}$$

The proof (see Appendix B) relies on explicitly bounding transition costs from below. This is achieved by showing that, from E_i, there is always a least cost transition path to outside of $D(E_i)$ that involves the first error on the path being repeated consecutively until the state is close to the border of $D(E_i)$. The cost of this path segment bounds the total cost of the path from below. Moreover, as N gets large, the cost of the path segment approaches the cost of the least cost path involving errors in only a single population.

Using Theorem 1, we obtain a simple estimation of the least cost transition path away from any convention.

Corollary 1. *Consider a contract game. Let the strategy revision rule be an exponential revision protocol. Then*

$$\frac{1}{N} \min_{j \neq i} c(E_i, E_j) \approx \frac{1}{N} \min_{j \neq i} \pi_\alpha(i) \left\lceil N \frac{\pi_\beta(i)}{\pi_\beta(i) + \pi_\beta(j)} \right\rceil \wedge \frac{1}{N} \min_{j \neq i} \pi_\beta(i) \left\lceil N \frac{\pi_\alpha(i)}{\pi_\alpha(i) + \pi_\alpha(j)} \right\rceil. \tag{5}$$

An almost immediate implication of Corollary 1 is that for two strategy, two player coordination games with two strict Nash equilibria, under the logit dynamic any stochastically stable convention must correspond to a risk dominant Nash equilibrium. This result has previously been arrived at by Staudigl (2012) using a different methodology based on optimal control problems.

Results of this section in hand, we now return to our application.

5. Application to contract games

In this section we characterize the stochastically stable conventions. The stochastically stable convention for each process is associated with a bargaining solution. Transition costs between conventions are estimated, following which, the stochastically stable conventions can be characterized. Finally, the stochastically stable conventions are shown to approximate bargaining solutions.

The following lemma gives the overall costs of transitions between conventions for undirected processes. These transition costs do not depend on the destination convention, only on the origin. This is because least cost transitions are caused by extreme actions by agents in one of the populations and after enough such actions have occurred, the process can transit to *any* convention for zero additional cost.

Lemma 3. *For $i, j \in \{1, \ldots, n-1\}$, $i \neq j$,*

$$\frac{1}{N}c^U(E_i, E_j) = \frac{1}{N}\left\lceil N\frac{f(\delta i)}{f(\delta i) + \bar{s}_\beta}\right\rceil \wedge \frac{1}{N}\left\lceil N\frac{\delta i}{\delta i + \bar{s}_\alpha}\right\rceil, \tag{6}$$

$$\frac{1}{N}c^L(E_i, E_j) \approx \frac{1}{N}\delta i\left\lceil N\frac{f(\delta i)}{f(\delta i) + \bar{s}_\beta}\right\rceil \wedge \frac{1}{N}f(\delta i)\left\lceil N\frac{\delta i}{\delta i + \bar{s}_\alpha}\right\rceil \tag{7}$$

For $c^U(E_i, \cdot)$, the expression to the left of the $\wedge$ and inside the $\lceil . \rceil$ is the number of errors that α-agents must make to induce a β-agent to best respond with an action other than i_β. The equivalent expression to the right of the $\wedge$ is the number of errors that β-agents must make to induce an α-agent to best respond with an action other than i_α. For $c^L(E_i, \cdot)$, these costs are adjusted by the cost of the individual errors. For example, if every β-agent is playing i_β, the cost of an α-agent making a mistake and choosing an action other than i_β is equal to δi, as the α-agent in question would obtain a payoff of 0, instead of δi which is the payoff from best responding.

The next lemma gives overall costs of least cost transitions between states in Λ for directed processes. For these processes, the destination convention does matter. Least cost transitions are caused by agents from one of the populations demanding an incremental increase in their payoffs. That is, the lowest cost transition from a convention E_i is to one of the adjacent conventions E_{i-1}, E_{i+1}.

Lemma 4. *For $i \in \{1, \ldots, n-1\}$, $Z \in \{U, L\}$,*

$$\frac{1}{N} \min_{j \neq i} \hat{c}^Z(E_i, E_j) \approx \frac{1}{N} \hat{c}^Z(E_i, E_{i-1}) \wedge \frac{1}{N} \hat{c}^Z(E_i, E_{i+1}),$$

$$\frac{1}{N} \min_{j \neq i} \hat{c}^U(E_i, E_j) = \frac{1}{N} \left\lceil N \frac{i}{2i-1} \right\rceil \wedge \frac{1}{N} \left\lceil N \frac{f(\delta i)}{f(\delta(i+1)) + f(\delta i)} \right\rceil, \tag{8}$$

$$\frac{1}{N} \min_{j \neq i} \hat{c}^L(E_i, E_j) \approx \frac{1}{N} f(\delta i) \left\lceil N \frac{i}{2i-1} \right\rceil \wedge \frac{1}{N} \delta i \left\lceil N \frac{f(\delta i)}{f(\delta(i+1)) + f(\delta i)} \right\rceil. \tag{9}$$

Once again, the expressions inside the $\lceil . \rceil$ on either side of the $\wedge$ in (8) are the number of errors required to induce a best response which differs from that of the initial convention. Again, the logit expressions in (9) are adjusted for the cost of the individual errors. Note that the expressions to the left hand side of the $\wedge$ in (6), (8) and (9) are decreasing in i, and the expressions to the right hand side of the $\wedge$ are increasing in i. Furthermore, note that for $i = 1$, the left hand side in (6), (8), (9) is larger than the right hand side, and the converse is true for $i = n - 1$. Note that the expressions on neither side of the $\wedge$ in (7) are monotonic, but both sides are concave in i.

An i-graph is a directed graph on L such that every vertex except for i has exactly one exiting edge and the graph has no cycles. Let $\mathcal{G}(i)$ denote the set of i-graphs. For a graph g, let $(j \to k) \in g$ denote an edge from j to k in g. Define:

$$\mathcal{V}(i) := \min_{g \in \mathcal{G}(i)} \sum_{(j \to k) \in g} c(E_j, E_k).$$

We know from Freidlin and Wentzell (1984, chap.6), Young (1993a) that:

$$\mu(E_i) > 0 \Leftrightarrow i \in \arg \min_{j \in L} \mathcal{V}(j).$$

This result will be used in our characterization theorem of section 5.2. The observant reader will, however, have noticed that four stochastic processes and only three bargaining solutions have been defined in the paper so far. As a final step before the characterization, we define another bargaining solution.

5.1. The Q4 bargaining solution

We now proceed to define a fourth bargaining solution, the $Q4$ bargaining solution. The reason we believe this solution to be interesting and relevant will become clear with Theorem 2 of the next section.

Definition 7. *The Q4 bargaining solution.*

$$t^{Q4} := \arg \max_{0 \leq t \leq \bar{s}_\alpha} t \, f(t) \, \phi(t), \qquad where \quad \phi(t) := \frac{1}{t + \bar{s}_\alpha} \wedge \frac{1}{f(t) + \bar{s}_\beta}$$

17

That is, the $Q4$ bargaining solution maximizes an adjusted Nash product. Our assumptions on $f(.)$ guarantee that $t\,f(t)\,\phi(t)$ is strictly concave, so t^{Q4} is unique. We will see that the solution is characterized by a three part piecewise function, with one part differing considerably from the other two parts. The properties of the $Q4$ bargaining solution are analyzed further and compared to existing bargaining solutions in section 6. First, the characterization theorem of the paper is given.

5.2. Characterization theorem

The theorem presented in this section characterizes the selection results of the evolutionary processes discussed above. To reiterate, the processes are all perturbed best response processes and differ only in their perturbation structure. The perturbations analyzed vary along two dimensions: they can be undirected or directed, they can be uniform or logit. First, a lemma is given which characterizes the stochastically stable states of the model for given δ and given large population size. The stochastically stable states are conventions which maximize the expressions (6), (7), (8), (9).

Lemma 5. *For large N, $c \in \{c^U, c^L, \hat{c}^U, \hat{c}^L\}$, $\mu(E_i) > 0$ implies that $j = i$ maximizes the approximation of $\min_{k \in L \setminus \{j\}} c(E_j, E_k)$ given in (6), (7), (8), (9). When this maximizer is unique, $\mu(E_i) > 0$ if and only if*

$$i \in \arg\max_{j \in L} \min_{k \in L \setminus \{j\}} c(E_j, E_k). \tag{10}$$

This characterizes the stochastically stable states for the problem. The principle theorem of the paper now approximates these states for large N and small δ, linking them to bargaining solutions. The theorem states that for a fine discretization (small δ) and large populations (large N), the stochastically stable states of our four processes correspond to our four bargaining solutions. The content of the theorem is summarized by table 1 given in the introduction to the paper. The results of the first row of table 1 (uniform errors) are known from Young (1998a) and Naidu et al. (2010). The results of the second row (logit errors) are, to the best of the authors' knowledge, new.

Theorem 2. *For any $\varepsilon > 0$, there exists $\bar{\delta}$ such that for all $\delta < \bar{\delta}$, there exists $N_\delta \in \mathbb{N}$ such that for all $N \geq N_\delta$, $\mu(E_i) > 0 \implies |\delta i - t^*| < \varepsilon$, where*

$$t^* = \begin{cases} t^{KS} & \textit{if } P^\eta \textit{ is uniform-undirected.} \\ t^{Q4} & \textit{if } P^\eta \textit{ is logit-undirected.} \\ t^{NB} & \textit{if } P^\eta \textit{ is uniform-directed.} \\ t^E & \textit{if } P^\eta \textit{ is logit-directed.} \end{cases}$$

Uniform and logit errors are by far the most common errors used in the stochastic stability literature. Under uniform errors, the (order of the) probability of a given error is state

independent, so selection is determined by how easy it is to cause a change in the best response of one population when mutations occur in the other population. Directed errors truncate the error distribution so that some errors are completely disallowed. This is one way of varying the probabilities with which errors are made, and allowing the ease with which errors are made to influence selection. A more gradual way is to use logit perturbations, under which errors which are more costly (in terms of payoffs) to the erring agent occur with lower probability. It is gratifying that the three classic bargaining solutions are selected under three of the four combinations covered by the theorem. However, the solution in our fourth quadrant, the $Q4$ solution, is quite a different object: the two drivers of selection, (i) ease of making errors, and (ii) ease of responding to errors, combine to create non-monotonicities that give the solution unusual properties. These shall be discussed in section 6.

6. Properties of the $Q4$ solution

In this section we examine the properties of the $Q4$ bargaining solution. First, we rewrite t^{Q4} as

$$t^{Q4} = \arg \max_{0 \leq t \leq \bar{s}_\alpha} h_1(t) \wedge h_3(t), \tag{11}$$

where

$$h_1(t) := \frac{tf(t)}{t + \bar{s}_\alpha}, \qquad h_3(t) := \frac{tf(t)}{f(t) + \bar{s}_\beta}$$

Not coincidentally, h_1, h_3 are the functions either side of the $\wedge$ in expression 7. We denote the maximizers of these functions by t_1, t_3 respectively.

$$t_l := \arg \max_{0 \leq t \leq \bar{s}_\alpha} h_l(t), \quad l = 1, 3.$$

When $h_1(t)$ and $h_3(t)$ intersect for $0 \leq t \leq \bar{s}_\alpha$, that is for $\frac{1}{2} \leq \frac{\bar{s}_\alpha}{\bar{s}_\beta} \leq 2$, we let t_2 be the value of t for which this intersection occurs. That is, t_2 solves

$$t_2 + \bar{s}_\alpha = f(t_2) + \bar{s}_\beta.$$

Remark 1. *The $Q4$ bargaining solution solves*

$$t^{Q4} := \begin{cases} t_1 & \text{if } h_1(t_1) < h_3(t_1), & \text{(Case 1)} \\ t_3 & \text{if } h_3(t_3) < h_1(t_3), & \text{(Case 3)} \\ t_2 & \text{otherwise.} & \text{(Case 2)} \end{cases}$$

The cases of the solution are numbered by the order in which they occur as the ratio $\bar{s}_\alpha/\bar{s}_\beta$ moves from high to low values. For high values of $\bar{s}_\alpha/\bar{s}_\beta$, the maximum of $h_1(\cdot)$ lies underneath the curve of $h_3(\cdot)$. This is when Case 1 holds and is illustrated for a linear bargaining frontier in figure 4. For low values of $\bar{s}_\alpha/\bar{s}_\beta$, the maximum of $h_3(\cdot)$ lies underneath the curve of $h_1(\cdot)$ and we are in Case 3. For values of $\bar{s}_\alpha/\bar{s}_\beta$ close to 1, the maximizer of (11) is determined by the intersection of $h_1(\cdot)$ and $h_3(\cdot)$. The decomposition of $Q4$ in Remark 1 facilitates analysis, as Case 2 exhibits very different properties to Cases 1 and 3.

In table 2 we list the axioms satisfied by the four bargaining solutions in the paper. Efficiency is achieved by every solution, and implies that if a solution gives t^* to Player α, it gives $f(t^*)$ to Player β. Other axioms are as follows.

Definition 8. *Let g, f be two bargaining frontiers, t_g^*, t_f^* their associated solutions.*

IIA	$g \geq f,\ g(t_g^*) \leq f(t_g^*) \implies t_g^* = t_f^*.$
Invariance	$g(x) = f(ax),\ a \in \mathbb{R} \implies t_g^* = \frac{1}{a}t_f^*.$
Monotonicity	$g \geq f \implies t_g^* \geq t_f^*.$
Individual Monotonicity	$g \geq f,\ g(0) = f(0) \implies t_g^* \geq t_f^*.$
Stretch-monotonicity	$g(x) = f(ax),\ a \in \mathbb{R},\ a < 1 \implies t_g^* \geq t_f^*.$

We include a non-standard axiom: stretch-monotonicity. This holds when a stretch of the bargaining frontier parallel to the axis measuring player γ's payoffs will (weakly) increase the payoff of player γ. Stretch-monotonicity is weaker than individual monotonicity, which is in turn weaker than monotonicity. Furthermore, stretch-monotonicity is implied by invariance, and is therefore satisfied by *all* of the existing major bargaining solutions. It can be seen in table 2 that the $Q4$ bargaining solution is highly irregular in that it does not comply with many of the axioms.

Axiom	Nash	K-S	Egalitarian	$Q4$
Symmetry	Yes	Yes	Yes	Yes
Efficiency	Yes	Yes	Yes	Yes
IIA	Yes	No	Yes	No
Invariance	Yes	Yes	No	No
Monotonicity	No	No	Yes	No
Individual Monotonicity	No	Yes	Yes	No
Stretch-monotonicity	Yes	Yes	Yes	No

Table 2: Axioms satisfied by bargaining solutions.

The presence of $\bar{s}_\alpha$ and $\bar{s}_\beta$ in the $Q4$ solution means that IIA is violated. In Case 1 and Case 3, the $Q4$ solution is similar to the Nash solution, but adjusted to take into account

the best possible outcome for one of the players. If $\bar{s}_\alpha > \bar{s}_\beta$ and the conditions for Case 1 are satisfied, then Player α does worse than he does under the Nash solution. Moreover, an increase in $\bar{s}_\alpha$ results in player α achieving a higher payoff: his payoff is increasing in his best possible outcome. These facts can be seen by comparing the first order condition for the Nash bargaining solution:

$$t^{NB} f'(t^{NB}) + f(t^{NB}) = 0$$

to the first order condition for the $Q4$ solution in Case 1:

$$t^{Q4} f'(t^{Q4}) + \frac{\bar{s}_\alpha}{t^{Q4} + \bar{s}_\alpha} f(t^{Q4}) = 0.$$

The increase of Player α's payoff in his best possible outcome differs from the similar effect in the Kalai-Smorodinsky solution. The effect in the latter depends on the ratio of $\bar{s}_\alpha$ and $\bar{s}_\beta$, whereas in Case 1 of the $Q4$ solution, changes in $\bar{s}_\beta$ have no direct effect. Symmetrically, in Case 3 the solution depends on $f(.)$ and $\bar{s}_\beta$, but not directly on $\bar{s}_\alpha$.

When $\bar{s}_\alpha$ and $\bar{s}_\beta$ are relatively close to one another and the solution is in Case 2, player γ's payoff does not necessarily increase with $\bar{s}_\gamma$, and can even decrease. In fact, the solution is an Egalitarian solution with a notional disagreement point of $(\bar{s}_\beta, \bar{s}_\alpha)$. The disagreement point is wholly notional as it lies outside of the bargaining set. The players equalize their losses from this notional disagreement point. Somewhat bizarrely, this notional disagreement point for a player is equal to the maximum attainable payoff of the other player (see figure 2). This creates nonmonotonicities: it can be seen immediately from the expression for the solution in Case 2, and the illustration in figure 2, that holding the bargaining frontier fixed close to the current solution, an improvement in the best possible outcome for a player will result in his achieving a lower payoff.

Proposition 1. *Considering t^{Q4} as a function of $\bar{s}_\alpha$, $\bar{s}_\beta$, and $f(.)$ in the neighborhood of the solution, we have that in Case 1, $\frac{\partial t^{Q4}}{\partial \bar{s}_\alpha} > 0$, in Case 2, $\frac{\partial t^{Q4}}{\partial \bar{s}_\alpha} < 0$, in Case 3, $\frac{\partial t^{Q4}}{\partial \bar{s}_\alpha} = 0$.*

Moreover, we can make a stronger statement about non-monotonicity. We shall shortly see by means of an example that stretch-monotonicity is violated by the $Q4$ bargaining solution. Consequently, the $Q4$ bargaining solution satisfies neither individual monotonicity nor invariance.[6]

---

[6]It may be argued that invariance should be understood as a simple rescaling of payoffs, and that therefore error probabilities should also be rescaled. The authors agree that should everything be rescaled, then invariance will result. However, invariance as an axiom is more than just a statement about rescaling. It is also a normative statement about how wealth affects bargaining power. It is this interpretation of invariance that justifies an analysis of rescaled payoffs without a corresponding rescaling of error probabilities.
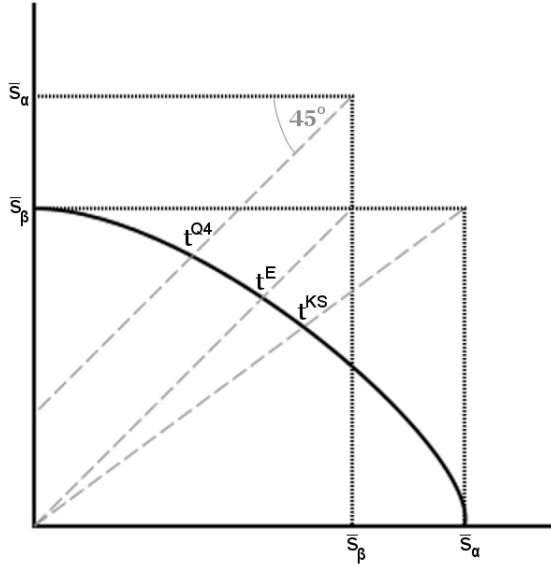
Figure 2: Case 2 of the $Q4$ solution, also illustrating Egalitarian and Kalai-Smorodinsky solutions for comparison.

6.1. Example: linear bargaining frontier

The case of a linear bargaining frontier is now analyzed. The frontier is given by the equation $f(t) = \bar{s}_\beta - t\frac{\bar{s}_\beta}{\bar{s}_\alpha}$. Conditions under which each case of the solution pertains and explicit solutions for each case are given in table 3. An increase in $\bar{s}_\alpha$ is equivalent to a stretch of the bargaining frontier parallel to the horizontal axis. It can be seen that when Case 2 pertains, an increase in $\bar{s}_\alpha$ results in a reduction in t^*, even though the section of the bargaining frontier where the solution lies does not remain constant under the stretch. Figure 3 shows how, fixing $\bar{s}_\beta$, the payoff of Player α varies with $\bar{s}_\alpha$. Plots of the least

Case	Condition	Solution
1	$\bar{s}_\alpha > \left(\frac{3\sqrt{2}}{2} - 1\right)\bar{s}_\beta$	$t^{Q4} = (\sqrt{2} - 1)\bar{s}_\alpha$
2	$\left(\frac{3\sqrt{2}}{2} - 1\right)^{-1}\bar{s}_\beta \leq \bar{s}_\alpha \leq \left(\frac{3\sqrt{2}}{2} - 1\right)\bar{s}_\beta$	$t^{Q4} = \frac{(2\bar{s}_\beta - \bar{s}_\alpha)\bar{s}_\alpha}{\bar{s}_\alpha + \bar{s}_\beta}$
3	$\bar{s}_\alpha < \left(\frac{3\sqrt{2}}{2} - 1\right)^{-1}\bar{s}_\beta$	$t^{Q4} = \left(2 - \sqrt{2}\right)\bar{s}_\alpha$

Table 3: Explicit expressions for the $Q4$ bargaining solution when the frontier is linear.

resistance exit from a convention $(t, f(t))$ driven by the mistakes of each agent type under all four of our error specifications (that is, plots of the expressions in equations (6), (7), (8), (9)) are given in figure 4 and 5 for t ranging from zero to $\bar{s}_\alpha$. It can be observed that for three of the specifications, the least resistances for transitions induced by the errors of any given agent type are monotone functions of t. Thus, the solution lies at the intersection of
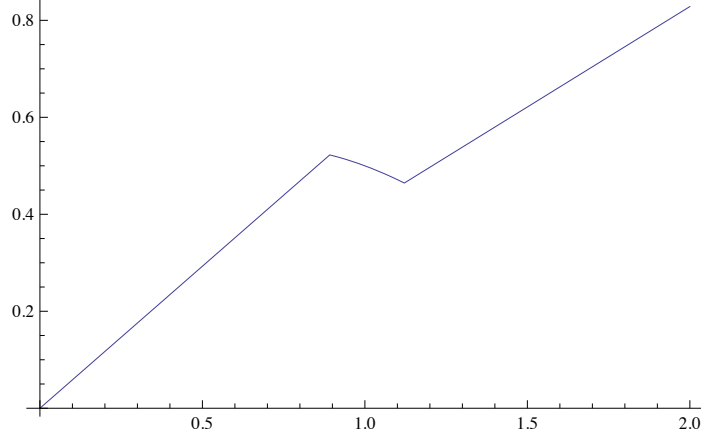
22

Figure 3: t^{Q4} by $\bar{s}_\alpha$, keeping $\bar{s}_\beta = 1$.

the two lines. For the undirected-logit specification the least resistances are non-monotonic. Two effects compete: the ease with which errors are made, and the ease of responding to errors. When either of these things are too easy, least resistances are low. The solution will now not necessarily lie at the intersection of the two lines. When the ratio $\bar{s}_\alpha/\bar{s}_\beta$ differs significantly from 1, the maximum of one of the curves lies below the other curve. This is Case 1 and Case 3 of our analysis, and Case 1 is illustrated in figure 4. When both maxima of the curves lie above the other curve, then the solution is the intersection of the curves. This is Case 2 of our analysis, and is illustrated in figure 5.

7. Discussion

This paper studies interactions by which standard axiomatic bargaining solutions emerge from the non-cooperative play of minimally forward looking individuals. It develops a method of determining minimum transition costs under processes for which mutation probabilities are log-linearly state dependent. Since Theorem 1 holds trivially for uniform mistake models, this result can be regarded as a useful generalization of the existing method of computing minimum transition costs. This method is used to study perturbed adaptive play of contract games – coordination games over points in a discretized bargaining set. The analysis highlights the interaction of two forces that drive evolutionary selection, namely (i) the ease of making errors, and (ii) the ease of responding to errors. Three major bargaining solutions are justified by plausible behavioral rules. The logit choice rule and other exponential strategy revision protocols give rise to a new bargaining solution with interesting features: the $Q4$ solution.

An important feature of the $Q4$ solution is that it arises from any exponential revision protocol. The logit choice rule requires that an agent compare his expected payoffs from all

Figure 4: Least resistances by t for a linear frontier, $\bar{s}_\alpha = 1.5$, $\bar{s}_\beta = 1$.

Undirected − uniform

Directed − uniform

Undirected − logit

Directed − logit

Undirected − uniform

Directed − uniform

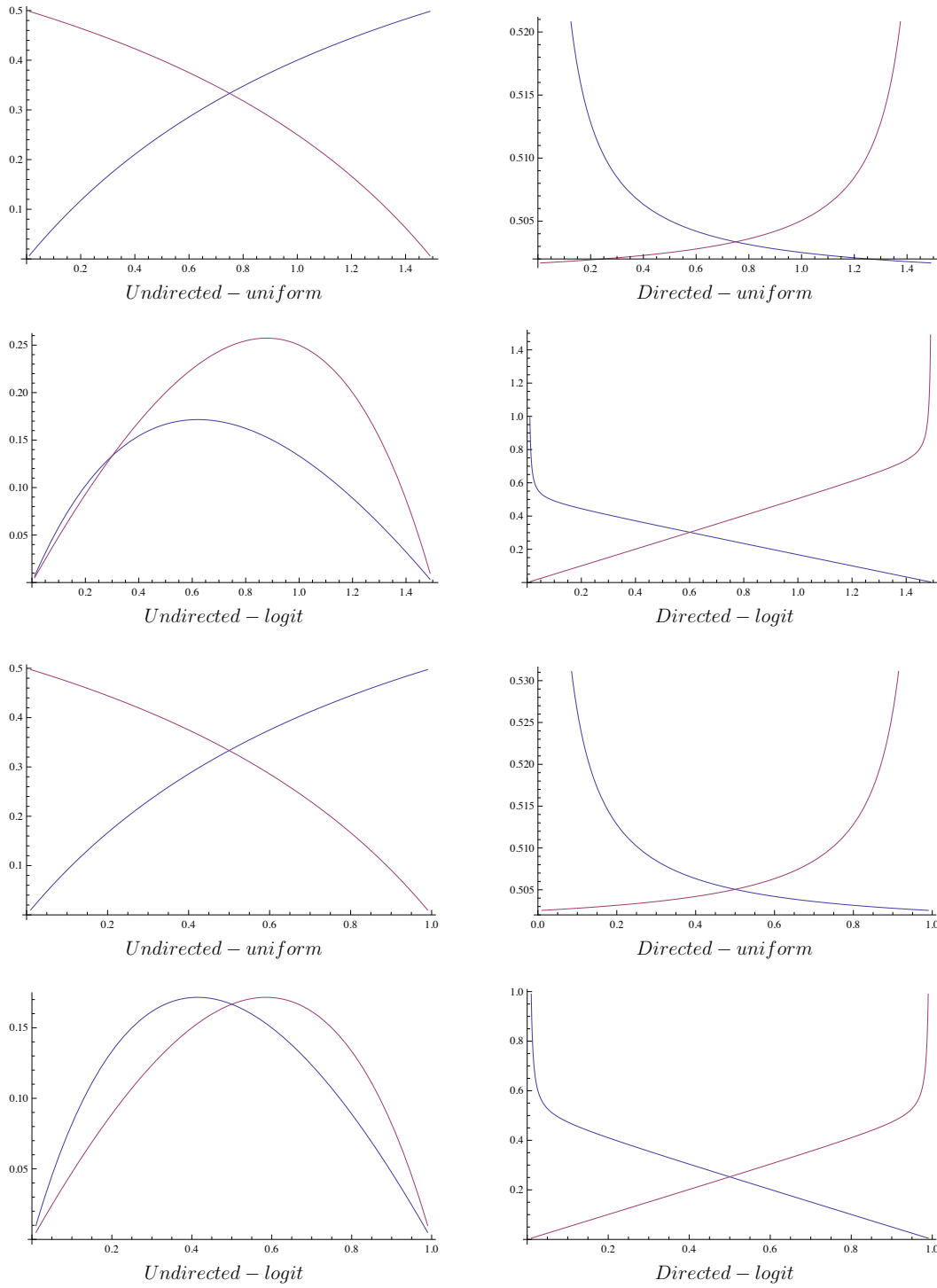*Undirected − logit*

Directed − logit

Figure 5: Least resistances by t for a linear frontier, $\bar{s}_\alpha = \bar{s}_\beta = 1$.

possible strategies. Thus, the informational requirements of the logit choice rule are great when there are numerous alternatives from which to choose. However, the class of exponential revision protocols includes some rules with very limited informational requirements, such as better reply rules, by which, in any given period, an agent only compares the expected payoff from his current strategy to that of a single alternative strategy (see Sandholm, 2010a). That the $Q4$ solution emerges from every member of this popular class of dynamics is a result which contrasts with popular misconceptions of the "anything can happen" result of Bergin and Lipman (1996). Log-linear error probabilities with full support suffice to select $Q4$, irrespective of the finer details of the dynamic.

It can be asked whether intuitively appealing axioms can be found which uniquely characterize the $Q4$ solution. Such an attempt would be complicated by the piecewise nature of the $Q4$ solution. When players' best possible outcomes differ considerably, the $Q4$ solution is an adjusted Nash bargaining solution. When best possible outcomes are similar, the $Q4$ solution is similar to the Egalitarian solution, equalizing the losses of each player from a notional payoff equal to the best possible outcome of the opposing player. Therefore, any axiomatic characterization is likely to be similarly hybrid.

Bibliography

Agastya, M., 1999. Perturbed adaptive dynamics in coalition form games. Journal of Economic Theory 89, 207–233.

Beggs, A., 2005. Waiting times and equilibrium selection. Economic Theory 25, 599–628.

Belloc, M., Bowles, S., 2013. The persistence of inferior cultural-institutional conventions. American Economic Review 103, 93–98.

Bergin, J., Lipman, B.L., 1996. Evolution with state-dependent mutations. Econometrica 64, 943–56.

Binmore, K., Samuelson, L., Young, P., 2003. Equilibrium selection in bargaining models. Games and Economic Behavior 45, 296 – 328.

Blume, L.E., 1993. The statistical mechanics of strategic interaction. Games and Economic Behavior 5, 387 – 424.

Blume, L.E., 1996. Population Games. Working Papers 96-04-022. Santa Fe Institute.

van Damme, E., Weibull, J.W., 2002. Evolution in games with endogenous mistake probabilities. Journal of Economic Theory 106, 296–315.

Dembo, A., Zeitouni, O., 1998. Large Deviations: Techniques and Applications. Springer, New York.

Freidlin, M.I., Wentzell, A.D., 1984. Random perturbations of dynamical systems, isbn 9780387983622 430 pp., 2nd ed.(1998). Springer .

Kalai, E., 1977. Proportional solutions to bargaining situations: Interpersonal utility comparisons. Econometrica 45, pp. 1623–1630.

Kalai, E., Smorodinsky, M., 1975. Other solutions to nash's bargaining problem. Econometrica 43, 513–18.

Kandori, M., Mailath, G.J., Rob, R., 1993. Learning, mutation, and long run equilibria in games. Econometrica 61, 29–56.

Naidu, S., Hwang, S.H., Bowles, S., 2010. Evolutionary bargaining with intentional idiosyncratic play. Economics Letters 109, 31 – 33.

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

Nax, H.H., Pradelski, B.S.R., 2013. Decentralized dynamics and equitable core selection in assignment games. Working Paper .

Newton, J., 2012a. Coalitional stochastic stability. Games and Economic Behavior 75, 842–54.

Newton, J., 2012b. Recontracting and stochastic stability in cooperative games. Journal of Economic Theory 147, 364–81.

Sandholm, W., 2010a. Pairwise comparison dynamics and evolutionary foundation for nash equilibrium. Games 1, 3–17.

Sandholm, W.H., 2010b. Population games and evolutionary dynamics. Economic learning and social evolution, Cambridge, Mass. MIT Press.

Shapley, L., Shubik, M., 1971. The assignment game i: The core. International Journal of Game Theory 1, 111–130.

Shiryaev, A.N., 1995. Probability (2nd ed.). Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Staudigl, M., 2012. Stochastic stability in asymmetric binary choice coordination games. Games and Economic Behavior 75, 372 – 401.

Weibull, J., 1995. Evolutionary game theory. MIT Press, Cambridge, Mass. [u.a.].

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

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

Young, H.P., 1998a. Conventional contracts. Review of Economic Studies 65, 773–92.

Young, H.P., 1998b. Individual strategy and social structure. Princeton University Press.

Appendix A. Examples of revision rules

Here we present some revision rules under which Theorem 1 holds (see Appendix B).

Better Reply (Baker) Dynamic	$p_\gamma^\eta(l'\|l,x) \propto \frac{\exp(\eta^{-1}\pi_\gamma(l',x))}{\exp(\eta^{-1}\pi_\gamma(l',x))+\exp(\eta^{-1}\pi_\gamma(l,x))}$
Better Reply (Metropolis) Dynamic	$p_\gamma^\eta(l'\|l,x) \propto \frac{\exp(\eta^{-1}\pi_\gamma(l',x))}{\exp(\eta^{-1}\pi_\gamma(l',x))\vee\exp(\eta^{-1}\pi_\gamma(l,x))}$
Incomplete Logit	$p_\gamma^\eta(l'\|l,x) = \frac{\exp(\eta^{-1}\pi_\gamma(l',x))}{\sum_{\tilde{l}\in C}\exp(\eta^{-1}\pi_\gamma(\tilde{l},x))}$ where $\{l,l'\} \subseteq C \subset S$
Fixed comparison	$p_\gamma^\eta(l'\|l,x) \propto \frac{\exp(\eta^{-1}\pi_\gamma(l',x))}{\exp(\eta^{-1}\pi_\gamma(l',x))+\exp(\eta^{-1}M)}$ for some $M > 0$

Then using lemma 1, we can find the following V functions for the Baker and Metropolis, logit, and fixed comparison dynamics respectively. Baker, Metropolis and logit with full support are examples of exponential revision protocols (Sandholm, 2010b).

$$V(x, x^{\gamma,i,j}) = \pi_\gamma(i,x) \vee \pi_\gamma(j,x) - \pi_\gamma(j,x) = [\pi_\gamma(i,x) - \pi_\gamma(j,x)]_+$$
$$V(x, x^{\gamma,i,j}) = \max_{\tilde{l}\in C} \pi_\gamma(\tilde{l},x) - \pi_\gamma(j,x)$$
$$V(x, x^{\gamma,i,j}) = [M - \pi_\gamma(j,x)]_+ .$$

Appendix B. Proof of Theorem 1 and Corollary 1

This section gives sufficient conditions for Theorem 1 and Corollary 1 to hold. These conditions are satisfied by exponential revision protocols. To express transitions by agents from one strategy to another more succinctly, we write

$$e_i^\alpha := ((0,\cdots,N,\cdots,0),(0,\cdots,0,\cdots,0)), \quad N \text{ in } i\text{th position}$$
$$e_j^\beta := ((0,\cdots,0,\cdots,0),(0,\cdots,N,\cdots,0)), \quad N \text{ in } j\text{th position}.$$

Throughout this section we consider transitions from E_i to some other convention for fixed i. We define

$$\Xi_\gamma := \left\{ x_\gamma \in \mathbb{R}^n : \sum_i x_\gamma(i) = N \right\}$$

and it is easy to see that $\Xi = \Xi_\alpha \times \Xi_\beta$.

We consider V functions which satisfy the following properties:

C1: *Irrelevance of own population*
$\quad V(x, x^{\gamma,k,l}) = V(y, y^{\gamma,k,l})$ for $x_\gamma = y_\gamma$, $\gamma = \alpha, \beta$.

C2: *Affine Linearity*

For all i, j, γ, there exists affine linear functions $v^{\gamma, i, j} : \Xi_{\gamma^-} \to \mathbb{R}$ and constants $\lambda_\gamma^{i,j} > 0$ and sets $C_\gamma^i \subseteq \{0, \dots, n\}$ such that

$$
V(x, x^{\gamma, i, j}) = \begin{cases} 0 & \text{if} \quad x \in \{y \in \Xi : y_{\gamma^-}(i) \leq N - \lambda_\gamma^{i,j}, \ y_{\gamma^-}(i) + y_{\gamma^-}(j) = N_{\gamma^-}\} \\ v^{\gamma, i, j}(x_{\gamma^-}) & \text{if} \quad x \in \{y \in \Xi : y_{\gamma^-}(i) > N - \lambda_\gamma^{i,j}\} \text{ and } j \in C_\gamma^i \\ \infty & \text{if} \quad x \in \{y \in \Xi : y_{\gamma^-}(i) > N - \lambda_\gamma^{i,j}\} \text{ and } j \notin C_\gamma^i \end{cases} .
$$

C3: *Coordination*

For all i, j, k, γ,

$$
v^{\gamma, i, j}(e_i^{\gamma^-}) = v^{\gamma, i, k}(e_i^{\gamma^-}),
$$

where $v^{\gamma, i, j}$ is increasing w.r.t. $x_{\gamma^-}(i)$ and decreasing w.r.t. $x_{\gamma^-}(j)$, and $v^{\gamma, i, j}(e_j^{\gamma^-}) < 0 < v^{\gamma, i, j}(e_i^{\gamma^-})$.

Condition 1 merely requires that play occurs between two populations, for which payoffs of each population depend on the strategy profile of the opponent population. Condition 2 requires that costs of transition vary linearly within the basin of attraction of each Nash equilibrium (described by $\lambda_\gamma^{i,j}$) and thus coincide with some affine function $v^{\gamma, i, j}$. Condition 3 is satisfied when the underlying game is a contract game. Property **C2** implies that we can define

$$
\bar{\lambda}_\alpha^{i,j} = \bar{\theta}_{ij} := \max \left\{ \theta \in \mathbb{N} : v^{\beta, i, j}\left(\left(1 - \frac{\theta}{N}\right) e_i^\alpha + \frac{\theta}{N} e_j^\alpha \right) > 0 \right\} \tag{B.1}
$$

$$
\bar{\lambda}_\beta^{i,j} = \bar{\zeta}_{ij} := \max \left\{ \zeta \in \mathbb{N} : v^{\alpha, i, j}\left(\left(1 - \frac{\eta}{N}\right) e_i^\beta + \frac{\eta}{N} e_j^\beta \right) > 0 \right\}.
$$

Rules that satisfy **C1-C2**, and satisfy **C3** when payoffs are determined by an underlying coordnation game include the logit choice rule, the Baker dynamic, the Metropolis dynamic, generalized logit, and logit with limited domain of choice (See Appendix A).

Using the affine linearity of v, we have

$$
v^{\beta, i, j}\left(\left(1 - \frac{\theta}{N}\right) e_i^\alpha + \frac{\theta}{N} e_j^\alpha \right) > 0 \implies \left(1 - \frac{\theta}{N}\right) v^{\beta, i, j}(e_i^\alpha) + \frac{\theta}{N} v^{\beta, i, j}(e_j^\alpha) > 0.
$$

Observe that **C2** implies that $v^{\beta, i, j}(e_i^\alpha) > 0$ and $v^{\beta, i, j}(e_j^\alpha) < 0$. It follows that

$$
\bar{\theta}_{ij} := \left\lceil N \frac{v^{\beta, i, j}(e_i^\alpha)}{v^{\beta, i, j}(e_i^\alpha) - v^{\beta, i, j}(e_j^\alpha)} \right\rceil - 1, \text{ and } \bar{\zeta}_{ij} := \left\lceil N \frac{v^{\alpha, i, j}(e_i^\beta)}{v^{\alpha, i, j}(e_i^\beta) - v^{\alpha, i, j}(e_j^\beta)} \right\rceil - 1.
$$

When V is given by the logit (or Baker) dynamic, we have

$$v^{\beta,i,j}(e_i^\alpha) = \pi_\beta(i), \ \ v^{\beta,i,j}(e_j^\alpha) = -\pi_\beta(j).$$

Theorem 3. *Suppose that V satisfies* **C1-C3**. *Let i be given. Then*

$$\frac{1}{N}\min_j c^N(E_i, E_j) \approx \frac{1}{N}\min_{j \in C_\alpha^i} v^{\alpha,i,j}(e_i^\beta)\bar{\theta}_{ij} \wedge \frac{1}{N}\min_{j \in C_\beta^i} v^{\beta,i,j}(e_i^\alpha)\bar{\zeta}_{ij}.$$

Using Theorem 3, Theorem 1 and Corollary 1 can be proven.

Proof of Theorem 1. It follows from (B.1) and **C2** that $^1/_N\,\bar{\theta}_{ij} \approx {}^1/_N\,(\bar{\theta}_{ij}+1)$ and that for $j \in C_\alpha^i$, $v^{\alpha,i,j}(e_i^\beta)(\bar{\theta}_{ij}+1)$ is the cost of a path of transitions from E_i to $\Xi \setminus D(E_i)$ in which the only errors involve α players switching to j. As this is true for any $j \in C_\alpha^i$, and a similar statement applies for $v^{\beta,i,j}(e_i^\alpha)(\bar{\zeta}_{ij}+1)$, the proof is complete. ∎

Proof of Corollary 1. Observe that

$$v^{\alpha,i,j}(x_\beta) = \frac{x_\beta(i)}{N}\pi_\alpha(i) - \frac{x_\beta(j)}{N}\pi_\alpha(j) \text{ and } v^{\beta,i,j}(x_\alpha) = \frac{x_\alpha(i)}{N}\pi_\beta(i) - \frac{x_\alpha(j)}{N}\pi_\beta(j)$$

$$\lambda_\alpha^{i,j} := \bar{\theta}_{ij} = \left\lceil N\frac{\pi_\beta(i)}{\pi_\beta(i)+\pi_\beta(j)}\right\rceil - 1 \text{ and } \lambda_\beta^{i,j} := \bar{\zeta}_{ij} = \left\lceil N\frac{\pi_\alpha(i)}{\pi_\alpha(i)+\pi_\alpha(j)}\right\rceil - 1$$

satisfies C2. ∎

The idea of the proof of Theorem 3 is as follows. To estimate the minimum bound for the lowest cost transitions, we study the minimization problem of the cost function over all possible paths escaping E_i. Estimation of such minima is complicated when the cost function of a given path loses linearity at the boundary of the basin of attraction, as is illustrated by Example 1. To overcome this problem, we explicitly estimate the size of the basin of attraction (Lemma 6) and construct a "truncated" path which has the same or lower cost than the original path, retaining linearity (Lemma 7).

For a path $\Gamma = (x_1, x_2, \cdots, x_L)$, we write $V(\Gamma) := \sum_{l=1}^{L-1} V(x_l, x_{l+1})$. We let

$$\bar{D}(E_j) := \{x \in \Xi : \text{there exists a path } \Gamma \text{ from } x \text{ to } E_j \text{ such that } V(\Gamma) = 0\}$$

$$\underline{\theta}_i := \min\left\{\bar{\theta}_{ij} : j \in S\right\}, \qquad \underline{\zeta}_i := \min\left\{\bar{\zeta}_{ij} : j \in S\right\}.$$

We shall use the notation $(\alpha, k, l : \theta)$ to denote a number θ of α-agents switching, in succession, from action k to action l. Similarly, let $(\beta, k', l' : \zeta)$ denote a number ζ of β-agents switching from action k to action l. Suppose that a path escaping E_i, $\Gamma = (x_1, x_2, \cdots, x_L)$,

consists of the following transitions

$$(\alpha, k_1, l_1 : \theta_1) \rightarrow (\alpha, k_2, l_2 : \theta_2) \rightarrow (\beta, k_1', l_1' : \zeta_1) \rightarrow (\alpha, k_3, l_3 : \theta_3) \rightarrow$$
$$\cdots \rightarrow (\beta, k_{L'}', l_{L'}' : \zeta_{L'}) \rightarrow (\alpha, k_L, l_L : \zeta_L),$$

such that $x_1 = E_i$, $x_1, \cdots, x_{L-1} \in \bar{D}(E_j)^c$ and $x_L \in \bar{D}(E_j)$.

Lemma 6. *Suppose that V satisfies $\mathbf{C1}$-$\mathbf{C3}$. Let $\{v^{\gamma, i, k}\}_k$ be given by $\mathbf{C2}$. Then the following statements hold.*
(1) Let y be a state in Γ immediate after the transition $(\alpha, k_{\tilde{m}}, l_{\tilde{m}}; \theta_{\tilde{m}})$. If $v^{\beta, i, k}(y_\alpha) \leq 0$ for some k, then $\sum_{\{m: k_m = i \text{ and } m \leq \tilde{m}\}} \theta_m \geq \underline{\theta}_i$
(2) Let y be a state in Γ immediate after the transition $(\beta, k_{\tilde{q}}', l_{\tilde{q}}'; \zeta_{\tilde{q}})$. If $v^{\alpha, i, k}(y_\beta) \leq 0$ for some k, then $\sum_{\{q: k_q' = i \text{ and } q \leq \tilde{q}\}} \zeta_q \geq \underline{\zeta}_i$
(3) $\sum_{\{m: k_m = i\}} \theta_m \geq \underline{\theta}_i$ or $\sum_{\{q: k_q' = i\}} \zeta_q \geq \underline{\zeta}_i$

Proof. We first show that (1) holds. First we establish that if $y_\alpha(i) > N - \underline{\theta}_i$, then $v^{\beta, i, k}(y_\alpha) > 0$ for all k. Let $y \in \Xi$ such that $y_\alpha(i) > N - \underline{\theta}_i$. If $y = (e_i^\alpha, y_\beta)$, then from $\mathbf{C2}$, $v^{\beta, i, k}(y_\alpha) > 0$ for all k and we are done. Thus suppose that $y_\alpha(i) \neq N$. We define

$$c_j := \frac{y_\alpha(j)}{N - y_\alpha(i)}.$$

for $j = 1, 2, \cdots, i - 1, i + 1, \cdots, n$. Then

$$\sum_{j \neq i} c_j = 1 \text{ and}$$

$$y_\alpha = c_1(N - y_\alpha(i), 0, \cdots, 0, y_\alpha(i), 0, \cdots, 0) + c_2(0, N - y_\alpha(i), \cdots, 0, y_\alpha(i), 0, \cdots, 0)$$
$$+ c_n(0, \cdots, 0, y_\alpha(i), 0, \cdots, N - y_\alpha(i))$$
$$= \sum_{j \neq i} c_j(\frac{y_\alpha(i)}{N} e_i^\alpha + \frac{N - y_\alpha(i)}{N} e_j^\alpha).$$

By $\mathbf{C2}$ and $\mathbf{C3}$,

$$v^{\beta, i, k}(y_\alpha) = v^{\beta, i, k}(\sum_{j \neq i} c_j(\frac{y_\alpha(i)}{N} e_i^\alpha + \frac{N - y_\alpha(i)}{N} e_j^\alpha)) = \sum_{j \neq i} c_j v^{\beta, i, k}(\frac{y_\alpha(i)}{N} e_i^\alpha + \frac{N - y_\alpha(i)}{N} e_j^\alpha)$$
$$> v^{\beta, i, k}(\frac{y_\alpha(i)}{N} e_i^\alpha + \frac{N - y_\alpha(i)}{N} e_k^\alpha) > 0$$

where the first inequality follows from the fact that $v^{\beta, i, k}(x)$ is decreasing in $x_\alpha(k)$. This shows that if $y_\alpha(i) > N - \underline{\theta}_i$, $v^{\beta, i, k}(y_\alpha) > 0$ for all k. Thus if $v^{\beta, i, k}(y_\alpha) \leq 0$ for some k, $y_\alpha(i) \leq N - \underline{\theta}_i$. Let y be the state in Γ immediately after $(\alpha, k_{\tilde{m}}, l_{\tilde{m}}; \theta_{\tilde{m}})$. If θ_m's are the

31

number of transitions by α-agents prior to y, then

$$y_\alpha(i) = N - \sum_{\substack{\{m:k_m=i, \\ m \leq \tilde{m}\}}} \theta_m + \sum_{\substack{\{m:l_m=i, \\ m \leq \tilde{m}\}}} \theta_m$$

So

$$\sum_{\substack{\{m:k_m=i, \\ m \leq \tilde{m}\}}} \theta_m = N - y_\alpha(i) + \sum_{\substack{\{m:l_m=i, \\ m \leq \tilde{m}\}}} \theta_m \geq N - y_\alpha(i) \geq N - (N - \underline{\theta}_i) = \underline{\theta}_i.$$

Thus if $v^{\beta,i,k}(y_\alpha) \leq 0$ for some k, $\sum_{\substack{\{m:k_m=i, \\ m \leq \tilde{m}\}}} \theta_m \geq \underline{\theta}_i$. Then (2) follows similarly. (3) follows from the fact that if $y = x_L \in \bar{D}(E_j)$, then $V(y, y^{\alpha,i,k}) = 0$ for some k or $V(y, y^{\beta,i,k'}) = 0$ for some k'. Thus from (1) and (2), (3) follows. ∎

Consider again a path Γ from E_i to $\bar{D}(E_j)$. We seek a lower bound for $V(\Gamma)$. To do this we ignore terms $V(x, x^{\alpha,k,l})\theta$ for $k \neq i$ (or $V(x, x^{\beta,k',l'})\zeta$ for $k' \neq i$) which represent transitions from $k \neq i$ to l. Then we bound the remaining terms from below. If x is the state in Γ immediately after $(\alpha, k_{\tilde{m}}, l_{\tilde{m}}; \theta_{\tilde{m}})$, then by **C3**

$$V(x, x^{\beta,i,l}) = v^{\beta,i,l}(x_\alpha) \geq v^{\beta,i,l}(y^{\tilde{m}}), \text{ where } y_\alpha^{\tilde{m}}(i) = N - \sum_{\{m:k_m=i, \ m \leq \tilde{m}\}} \theta_m, \ y_\alpha^{\tilde{m}}(l) = \sum_{\{m:k_m=i, \ m \leq \tilde{m}\}} \theta_m.$$

Let

$$r_m(\eta_1, \cdots, \eta_{q_m}) := v^{\alpha,i,l_m}(y_\beta^{q_m}), \quad u_q(\delta_1, \cdots, \delta_{m_q}) := v^{\beta,i,l'_q}(y_\alpha^{m_q}).$$

Then, omitting any terms related to transitions other than those from i, $V(\Gamma)$ is bounded below by

$$\varphi(\theta, \zeta) := r_1\theta_1 + u_1(\theta_1)\zeta_1 + \cdots + r_{m_{\tilde{q}}}(\zeta_1, \cdots, \zeta_{\tilde{q}-1})\theta_{m_{\tilde{q}}} + u_{\tilde{q}}(\theta_1, \cdots, \theta_{m_{\tilde{q}}})\zeta_{\tilde{q}} + \cdots + r_L(\zeta_1, \cdots, \zeta_{L'})\theta_L. \tag{B.2}$$

We will consider the following minimization problem:

$$\min\{\varphi(\theta, \zeta) : 0 \leq \theta_m \leq \bar{\theta}_{k_m,l_m}, \ 0 \leq \zeta_q \leq \bar{\zeta}_{k'_q,l'_q} \text{ for all } m = 1, \cdots L, q = 1, \cdots, L', \tag{B.3}$$

$$r_m \geq 0 \text{ for all } m \text{ and } u_q \geq 0 \text{ for all } q, \ \sum_{m=1}^L \theta_m \geq \underline{\theta}_i\}.$$

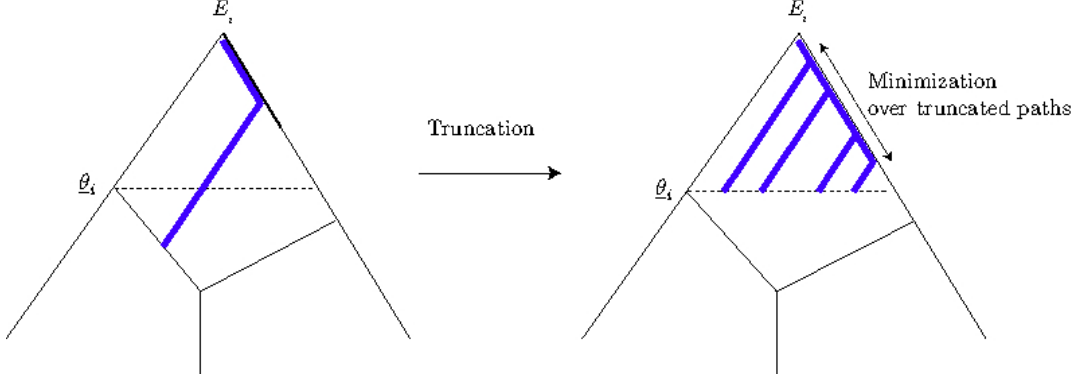Similar problems can be defined for ϕ functions whose last term has a $u_{L'}(.)$ rather than a

32

Figure B.6: **Illustration of the truncation lemma.**

$r_L(.)$.

Lemma 7 (Truncation). *Consider the minimization problem given by (B.2) and (B.3). Let (θ^*, ζ^*) be the optimal choices. If $u_{\tilde{q}}(\theta^*) = 0$ in (B.2), for*

$$\phi(\theta, \zeta) := r_1 \theta_1 + u_1(\theta_1)\zeta_1 + \cdots + r_{m_{\tilde{q}}}(\zeta_1, \cdots, \zeta_{\tilde{q}-1})\theta_{m_{\tilde{q}}},$$

we have

$$\varphi(\theta^*, \zeta^*) \geq \min\{\phi(\theta, \zeta) : 0 \leq \theta_m \leq \bar{\theta}_{k_m, l_m}, \ 0 \leq \zeta_q \leq \bar{\zeta}_{k'_q, l'_q}, \ r_m \geq 0, \ u_q \geq 0$$
$$\text{for all } 1 \leq \ m \leq m_{\tilde{q}}, 1 \leq q \leq \tilde{q} - 1, \quad \sum_{\{m:m \leq m_{\tilde{q}}\}} \theta_m \geq \underline{\theta}_i\}.$$

A similar result holds if $r_{\tilde{m}}(\zeta^*) = 0$ in (B.2).

Proof. From Lemma 6, if $u_{\tilde{q}}(\theta^*) = 0$, $\sum_{\{m:m \leq m_{\tilde{q}}\}} \theta_i^* \geq \underline{\theta}_i$. Thus for $r = 1, 2, 3, 4$ we have

$$\varphi_r(\theta^*, \zeta^*) \geq r_1 \theta_1^* + u_1(\theta_1^*)\zeta_1^* + \cdots + r_{m_{\tilde{q}}}(\zeta_1^*, \cdots, \zeta_{\tilde{q}-1}^*)\theta_{m_{\tilde{q}}}^*$$
$$\geq \min\{\phi(\theta, \zeta) : 0 \leq \theta_m \leq \bar{\theta}_{k_m, l_m}, \ 0 \leq \zeta_q \leq \bar{\zeta}_{k'_q, l'_q}, \ r_m \geq 0, \ u_q \geq 0$$
$$\text{for all } 1 \leq \ m \leq m_{\tilde{q}}, 1 \leq q \leq \tilde{q} - 1, \quad \sum_{\{m:m \leq m_{\tilde{q}}\}} \theta_m \geq \underline{\theta}_i\}.$$

∎

Proof of Theorem 3. We let

$$\bar{c}_N := \min_{j \in C_\alpha^i} v^{\alpha, i, j}(e_i^\beta)(\bar{\theta}_{ij} + 1) \wedge \min_{j \in C_\beta^i} v^{\beta, i, j}(e_i^\alpha)(\bar{\zeta}_{ij} + 1)$$

$$\underline{c}_N := \min_{j \in C_\alpha^i} v^{\alpha, i, j}(e_i^\beta)\bar{\theta}_{ij} \wedge \min_{j \in C_\beta^i} v^{\beta, i, j}(e_i^\alpha)\bar{\zeta}_{ij}$$

We will show that $\bar{c}_N \geq \min_j c^N(E_i, E_j) \geq \underline{c}_N$. We suppose that Γ is a path from E_i to

$D(E_j)$ and that the final step in the path is a transition by an α-agent (the other case follows similarly). Consider a φ function based on Γ and the minimization problem given by B.3 before Lemma 7. Let (θ^*, ζ^*) be the optimal choices. If $r_m(\zeta^*) = 0$ for some m, or $u_q(\theta^*) = 0$ for some q, we can apply Lemma 7 and start over. Hence we suppose that at (θ^*, ζ^*), $r_m(\theta^*, \zeta^*) > 0$, $u_q(\theta^*, \zeta^*) > 0$ for all m, q. First, suppose that $\sum_m \theta_m^* > \underline{\theta}_i$. Recall that θ_m is associated with switches from i to l_m, and ζ_m is associated with switches from i to l'_m. Because of affine linearity, we must have that (1) $\theta_m^* = 0$ or $\bar{\theta}_{il_m}$, (2) $\zeta_q^* = 0$ or $\bar{\zeta}_{il'_q}$ for all m, q, and (3) at least one $\theta_{\tilde{m}}^* = \bar{\theta}_{il_{\tilde{m}}}$.

$$
\begin{aligned}
&\text{if } \theta_1^* = \bar{\theta}_{il_1}, && \tilde{\varphi}(\theta^*, \zeta^*) \geq v^{\alpha,i,l_1}(e_i^\beta)\bar{\theta}_{il_1} > v^{\alpha,i,l_1}(e_i^\beta)\underline{\theta}_i \\
&\text{if } \theta_1^* = 0, \zeta_1^* = \bar{\zeta}_{il'_1}, && \tilde{\varphi}(\theta^*, \zeta^*) \geq v^{\beta,i,l'_1}(e_i^\alpha)\bar{\zeta}_{il'_1} > v^{\alpha,i,l'_1}(e_i^\alpha)\underline{\zeta}_i \\
&\text{if } \theta_1^* = 0, \zeta_1^* = 0, \zeta_2^* = \bar{\zeta}_{il'_2}, && \tilde{\varphi}(\theta^*, \zeta^*) \geq v^{\beta,i,l'_2}(e_i^\alpha)\bar{\zeta}_{il'_2} > v^{\alpha,i,l'_2}(e_i^\alpha)\underline{\zeta}_i \\
&\quad\vdots && \quad\vdots \\
&\text{if } \theta_1^* = 0, \theta_2^* = 0, \cdots, \theta_{\tilde{m}}^* = \bar{\theta}_{il_{\tilde{m}}} \\
&\quad \zeta_1^* = 0, \zeta_2^* = 0, \cdots, \zeta_{q_{\tilde{m}}}^* = 0 && \tilde{\varphi}(\theta^*, \zeta^*) \geq v^{\alpha,i,l_{\tilde{m}}}(e_i^\beta)\bar{\theta}_{il_{\tilde{m}}} > v^{\alpha,i,l_{\tilde{m}}}(e_i^\beta)\underline{\theta}_i
\end{aligned}
\tag{B.4}
$$

Thus we obtain the desired lower bound $\underline{c}_N$. Next, suppose that $\sum_m \theta_m^* = \underline{\theta}_i$. Further divide into two cases: (i) $\theta_m^* < \underline{\theta}_i$ for all m and (ii) $\theta_{\tilde{m}}^* = \underline{\theta}_i$ for some $\tilde{m}$. Consider case (i). In this case there are $\theta_{\tilde{m}}^*$ and $\theta_{\tilde{m}'}^*$ such that $0 < \theta_{\tilde{m}}^*, \theta_{\tilde{m}'}^* < \underline{\theta}_i$. Since $\sum_m \theta_m = \underline{\theta}_i$ is linear with respect to θ_m's, the affine linearity of φ w.r.t. θ_m implies that either $(\theta_{\tilde{m}}^* + o, \theta_{\tilde{m}'}^* - o)$ or $(\theta_{\tilde{m}}^* - o, \theta_{\tilde{m}'}^* + o)$ for small o gives lower or equal φ than θ^*. A lower φ value contradicts optimality of θ_m^*. If the new value is equal, repeat the argument until some $\theta_m^* = \underline{\theta}_i$ for some m. Now consider case (ii). If $\theta_{\tilde{m}}^* = \underline{\theta}_i$, then evaluating as in (B.4), we obtain the desired lower bound.

Concerning the upper bound, let j_α^* and j_β^* be the states to which the direct escaping costs are minimal, that is j_α^* solves $\min_j v^{\alpha,i,j}(e_i^\beta)\bar{\theta}_{ij}$, and j_β^* solves $\min_j v^{\beta,i,j}(e_i^\alpha)\bar{\zeta}_{ij}$. The upper bound follows by either choosing a path consisting solely of $\bar{\theta}_{ij_\alpha^*} + 1$ transitions by α-agents from i to j_α^*, or a path comprising $\bar{\zeta}_{ij_\beta^*} + 1$ transitions by β-agents from i to j_β^*. $\blacksquare$

Appendix C. Proofs of other Lemmas and Theorems

Proof of Lemma 3. If, for $i, j \in \{1, \ldots, n-1\}$, a sequence $\{x^0, \ldots, x^T\}$, $x^0 = E_i$, $x^T = E_j$, is such that $\sum_{\tau=0}^{T-1} V(x^\tau, x^{\tau+1}) = c(E_i, E_j)$. For $\gamma = \alpha, \beta$, define

$$\tau_\gamma = \min\{\tau : V(x^\tau, x^{\tau+1}) = 0, \ x^{\tau+1} = (x^\tau)^{\gamma, j, k}, \ k \neq i\} \wedge T.$$

It must be that

$$(N - x_\alpha^{\tau_\beta}(i))\bar{s}_\beta \geq x_\alpha^{\tau_\beta}(i) f(\delta i),$$

where the left hand side is an upper bound on $\pi_\beta(j_\beta, x_\alpha^{\tau_\beta})$, $j \neq i$, and the right hand side equals $\pi_\beta(i_\beta, x_\alpha^{\tau_\beta})$. Rearranging, we obtain

$$N - x_\alpha^{\tau_\beta}(i) \geq \left\lceil N \frac{f(\delta i)}{f(\delta i) + \bar{s}_\beta} \right\rceil =: \xi_\alpha.$$

Similarly,

$$N - x_\beta^{\tau_\alpha}(i) \geq \left\lceil N \frac{\delta i}{\delta i + \bar{s}_\alpha} \right\rceil =: \xi_\beta.$$

If $\tau_\beta < \tau_\alpha$, then at least ξ_α α-agents need to have made errors before time τ_β. Similarly, if $\tau_\alpha < \tau_\beta$, then at least ξ_β β-agents need to have made errors before time τ_α.

<u>Uniform-undirected</u>

For $\tau = 0, \ldots, \xi_\alpha$, let $x_\alpha^{\tau+1}(0) = x_\alpha^\tau(0) + 1$, noting that $V(x^\tau, x^{\tau+1}) = 1$. For $\tau = \xi_\alpha + 1, \ldots, \xi_\alpha + N$, let $x_\beta^{\tau+1}(0) = x_\beta^\tau(0) + 1$, noting that $V(x^\tau, x^{\tau+1}) = 0$. For $\tau = \xi_\alpha + N + 1, \ldots, N + N$, let $x_\alpha^{\tau+1}(0) = x_\alpha^\tau(0) + 1$, noting that $V(x^\tau, x^{\tau+1}) = 0$. Note that $x^{2N} = E_0$. We have shown that $c(E_i, E_0) \leq \xi_\alpha$. For $\tau = N + N + 1, \ldots, 3N$, let $x_\alpha^{\tau+1}(j) = x_\alpha^\tau(0) + 1$, noting that $V(x^\tau, x^{\tau+1}) = 0$. For $\tau = 3N + 1, \ldots, 2N + 2N$, let $x_\alpha^{\tau+1}(j) = x_\alpha^\tau(j) + 1$, noting that $V(x^\tau, x^{\tau+1}) = 0$. We have shown that $c(E_0, E_j) = 0$. Therefore $c(E_i, E_j) \leq c(E_i, E_0) + c(E_0, E_j) \leq \xi_\alpha$. A similar construction shows that $c(E_i, E_j) \leq \xi_\beta$. As at least ξ_α or ξ_β errors are required before $\tau = \min\{\tau_\alpha, \tau_\beta\}$, it must be that $c(E_i, E_j) \geq \min\{\xi_\alpha, \xi_\beta\}$. The above construction shows that this is also an upper bound on $c(E_i, E_j)$.

<u>Logit-undirected</u>

Applying Corollary 1 and noting that minima are obtained at $j_\alpha = 0$ and $j_\beta = n$ we have

$$\left| \frac{1}{N} \min_j c^N(E_i, E_j) - \frac{1}{N}\delta i \left\lceil N \frac{f(\delta i)}{f(\delta i) + \bar{s}_\beta} \right\rceil \wedge \frac{1}{N} f(\delta i) \left\lceil N \frac{\delta i}{\delta i + \bar{s}_\alpha} \right\rceil \right| < \epsilon.$$

35

Now note that this approximation can be exactly attained for transitions to either E_0 or E_n. For example, if α-agents make ξ_α errors to play 0_α, then each of these transitions has a cost of δi, and the transitions suffice to reach E_0 in the same manner as in the proof of the uniform-undirected case. From E_0, any E_j can be reached at zero cost, so we can state the stronger result that for all $j \neq i$

$$\left| \frac{1}{N} c^N(E_i, E_j) - \frac{1}{N} \delta i \left\lceil N \frac{f(\delta i)}{f(\delta i) + \bar{s}_\beta} \right\rceil \wedge \frac{1}{N} f(\delta i) \left\lceil N \frac{\delta i}{\delta i + \bar{s}_\alpha} \right\rceil \right| < \epsilon.$$

∎

Proof of Lemma 4. If, for $i, j \in \{1, \ldots, n-1\}$, a sequence $\{x^0, \ldots, x^T\}$, $x^0 = E_i$, $x^T = E_j$, is such that $\sum_{\tau=0}^{T-1} V(x^\tau, x^{\tau+1}) = c(E_i, E_j)$. and letting τ_α, τ_β, be defined as in the proof of Lemma 3, it must be that, when $\tau_\beta < \tau_\alpha$,

$$(N - x_\alpha^{\tau_\beta}(i)) f(\delta(i+1)) \geq x_\alpha^{\tau_\beta}(i) f(\delta i),$$

where the left hand side is an upper bound on $\pi_\beta(j_\beta, x_\alpha^{\tau_\beta})$, $j \neq i$, as $\Delta_\alpha(x^\tau) = \{i+1, \ldots, n\}$ for $\tau < \tau_\alpha$. The right hand side equals $\pi_\beta(i_\beta, x_\alpha^{\tau_\beta})$. Rearranging, we obtain

$$N - x_\alpha^{\tau_\beta}(i) \geq \left\lceil N \frac{f(\delta i)}{f(\delta(i+1)) + f(\delta i)} \right\rceil =: \hat{\xi}_\alpha.$$

Similarly, when $\tau_\alpha < \tau_\beta$,

$$N - x_\beta^{\tau_\alpha}(i) \geq \left\lceil N \frac{i}{2i-1} \right\rceil =: \hat{\xi}_\beta.$$

Uniform-directed

For $\tau = 0, \ldots, \xi_\alpha$, let $x_\alpha^{\tau+1}(i+1) = x_\alpha^\tau(i+1) + 1$, noting that $V(x^\tau, x^{\tau+1}) = 1$. For $\tau = \xi_\alpha + 1, \ldots, \xi_\alpha + N$, let $x_\beta^{\tau+1}(i+1) = x_\beta^\tau(i+1) + 1$, noting that $V(x^\tau, x^{\tau+1}) = 0$. For $\tau = \xi_\alpha + N + 1, \ldots, 2N$, let $x_\alpha^{\tau+1}(i+1) = x_\alpha^\tau(i+1) + 1$, noting that $V(x^\tau, x^{\tau+1}) = 0$. Note that $x^{N+N} = E_{i+1}$. We have shown that $c(E_i, E_{i+1}) \leq \xi_\alpha$. A similar construction shows that $c(E_i, E_{i-1}) \leq \xi_\beta$. As at least ξ_α or ξ_β errors are required before $\tau = \min\{\tau_\alpha, \tau_\beta\}$, it must be that $c(E_i, E_j) \geq \min\{\xi_\alpha, \xi_\beta\}$. The above construction shows that this bound is attained for some $j \in \{i-1, i+1\}$. ∎

Logit-directed

Applying Corollary 1 and noting that minima are obtained at $j_\alpha = i+1$ and $j_\beta = i-1$ we

have

$$\left| \frac{1}{N} \min_j c^N(E_i, E_j) - \frac{1}{N} f(\delta i) \left\lceil N \frac{i}{2i-1} \right\rceil \wedge \frac{1}{N} \delta i \left\lceil N \frac{f(\delta i)}{f(\delta(i+1)) + f(\delta i)} \right\rceil \right| < \epsilon.$$

Now note that this approximation can be exactly attained for transitions to either E_{i-1} or E_{i+1}. For example, if α-agents make ξ_α errors to play $(i+1)_\alpha$, then each of these transitions has a cost of δi, and the transitions suffice to reach E_{i+1} in the same manner as in the proof of the uniform-undirected case.

Proof of lemma 5. Let i maximize the relevant one of (6), (7), (8), (9). For $c = c^U, c^L$, let $g = \{j \to i : j \in L, j \neq i\}$. For $c = \hat{c}^U, \hat{c}^L$, let $g = \{j \to j+1 : j \in L, j < i\} \cup \{j \to j-1 : j \in L, j > i\}$. Note that $k \to l \in g$ implies that $\min_{j \in L} c(E_k, E_j) \approx c(E_k, E_l)$. Also note that by our choice of i, for all $k \neq i$, $\min_{j \in L} c(E_k, E_j) \lesssim \min_{j \in L} c(E_i, E_j)$. This implies that $\mathcal{V}(i) \lesssim \mathcal{V}(j)$ for all $j \in L$ if and only if i is a maximizer of the relevant expression. So $\mathcal{V}(i) \leq \mathcal{V}(j)$ for all $j \in L$ implies that i is a maximizer. ∎

Proof of Theorem 2. To prove the Theorem we use the following two lemmas.

Lemma 8. *Suppose that φ is a real valued function and S is a finite set. Then there exists $\hat{N}$ such that for all $N > \hat{N}$*

$$\arg\max_{t \in S} \frac{1}{N} \lceil N\varphi(t) \rceil = \arg\max_{t \in S} \varphi(t).$$

The proof of the above lemma readily follows from the pointwise convergence of $1/N \lceil N\varphi(t) \rceil$ to $\varphi(t)$.

Lemma 9. *Suppose φ is a continuous function which admits a unique maximum. Suppose φ_δ such that φ_δ converges uniformly to φ as $\delta \to 0$ and $t^* \in \arg\max \varphi(t)$ and $i^* \in \arg\max_i \varphi_\delta(i\delta)$. Then for all $\epsilon > 0$, there exists $\bar{\delta} > 0$ such that for all $\delta < \bar{\delta}$, we have $|i^*\delta - t^*| < \epsilon$.*

Proof. By the definitions of $t^* \in \arg\max_t \varphi(t)$ and $i^* \in \arg\max_i \varphi_\delta(i\delta)$, we have $\varphi(t^*) \geq \varphi(i^*\delta)$ and $\varphi_\delta(i^*\delta) \geq \varphi_\delta(t)$. Let $\epsilon > 0$. By uniform convergence we can choose $\delta < \bar{\delta}$, such that $|\varphi_\delta(t^*) - \varphi(t^*)| < \epsilon$ and $|\varphi_\delta(i^*\delta) - \varphi(i^*\delta)| < \epsilon$. For $\delta < \bar{\delta}$, we have $\varphi(i^*\delta) \leq \varphi(t^*) \leq \varphi_\delta(t^*) + \epsilon \leq \varphi_\delta(i^*\delta) + \epsilon < \varphi(i^*\delta) + 2\epsilon$. Thus we have that

For all $\tilde{\epsilon} > 0$, there exists $\bar{\delta}$ such that for all $\delta < \bar{\delta}$, we have $|\varphi(t^*) - \varphi(i^*\delta)| < \tilde{\epsilon}$. (C.1)

Without loss of generality we suppose that $i^*\delta < t^*$ and let $\epsilon > 0$ be given. Then for $\epsilon > 0$ we can choose $\bar{\rho}$ such that for all $\rho < \bar{\rho}$

$$\varphi(t^*) - \rho < y < \varphi(t^*) \text{ implies } |\varphi^{-1}(y) - t^*| < \epsilon, \tag{C.2}$$

where φ^{-1} is the inverse function for φ defined in a neighborhood of t^* except t^*. Now let $\epsilon > 0$. Choose $\bar{\rho}$ satisfying (C.2) first. Then for $\tilde{\epsilon} = \rho < \bar{\rho}$, choose $\bar{\delta}$ satisfying (C.1). Then for ρ and for $\delta < \bar{\delta}$, we have $|\varphi(i^*\delta) - \varphi(t^*)| < \rho$. Also since $\rho < \bar{\rho}$, by (C.2) we have $|i^*\delta - t^*| < \epsilon$. Thus we show that for all $\epsilon > 0$, there exists $\bar{\delta} > 0$ such that for all $\delta < \bar{\delta}$, we have $|i^*\delta - t^*| < \epsilon$. ∎

Note that for large enough N the values taken by the expressions of the form $\lceil a \rceil \wedge \lceil b \rceil$ in (6), (7), (8), (9) equal $\lceil a \wedge b \rceil$. Lemma 8 then implies that for large N we can ignore the ceiling function in (6), (7), (8), (9) when determining the stochastically stable states. Replacing δi by t in expressions (6), (7) and taking the limit as $\delta \to 0$ gives

$$\frac{f(t)}{f(t) + \bar{s}_\beta} \wedge \frac{t}{t + \bar{s}_\alpha} \tag{C.3}$$

$$t\frac{f(t)}{f(t) + \bar{s}_\beta} \wedge f(t)\frac{t}{t + \bar{s}_\alpha} \tag{C.4}$$

respectively. Using Lemma 9, and noting that these functions are maximized at t^{KS}, t^{Q4}, respectively, we have the results for the cases of uniform-undirected and logit-undirected pertubations. For the case of logit-directed perturbations, the expression in (9) takes the form $a(t) \wedge b(t)$ after the ceiling function has been removed. Continuity of $f(\cdot)$ implies that there exist $\varepsilon > 0$, $\hat{\delta} > 0$ such that for all $\delta < \hat{\delta}$, $a(t) > b(t)$ for all $t < \varepsilon$, and $a(t) < b(t)$ for all $t > \bar{s}_\alpha - \varepsilon$. Therefore, the following function equals (9) at all $t = \delta i$, $i = 1, \ldots, n-1$.

$$\varphi_\delta(t) = \begin{cases} b(t) & \text{if } t < \varepsilon. \\ a(t) \wedge b(t) & \text{if } \varepsilon \leq t \leq \bar{s}_\alpha - \varepsilon. \\ a(t) & \text{if } t > \bar{s}_\alpha - \varepsilon. \end{cases}$$

Note that φ_δ converges uniformly to φ as $\delta \to 0$, where

$$\varphi(t) = \frac{f(t)}{2} \wedge \frac{t}{2}.$$

This expression is maximized at t^E so by Lemma 9 we have the result for logit-directed perturbations.

For the case of uniform-directed pertubations, we cannot apply Lemma 9, since the function in the cost expression does not converge to a function with a unique maximum. To address this case, we observe that Lemma 8 implies that for large N

$$i^* \in \arg\max_i \frac{\delta i}{2\delta i - \delta} \wedge \frac{f(\delta i)}{f(\delta(i+1)) + f(\delta i)}.$$

Writing $t = \delta i$ and noting that the LHS of the $\wedge$ is decreasing in t, and the RHS is increasing in t, it can be seen that δi^* can be approximated by $\tilde{t}$ given by

$$\frac{\tilde{t}}{2\tilde{t} - \delta} = \frac{f(\tilde{t})}{f(\tilde{t} + \delta) + f(\tilde{t})} \iff \tilde{t}\frac{f(\tilde{t} + \delta) - f(\tilde{t})}{\delta} + f(\tilde{t}) = 0$$

which approaches the first order condition for t^{NB} as $\delta \to 0$. Hence $\delta i^* \to t^{NB}$ (See detailed argument in Naidu et al., 2010). ∎