View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

Stochastic Game Theory:
Adjustment to Equilibrium Under Noisy Directional Learning

Simon P. Anderson, Jacob K. Goeree, and Charles A. Holt*

Department of Economics
114 Rouss Hall
University of Virginia
Charlottesville, VA 22903-3328

July 1999

ABSTRACT

This paper presents a dynamic model in which agents adjust their decisions in the direction of higher payoffs, subject to random error. This process produces a probability distribution of players' decisions whose evolution over time is determined by the Fokker-Planck equation. The dynamic process is stable for all potential games, a class of payoff structures that includes several widely studied games. In equilibrium, the distributions that determine expected payoffs correspond to the distributions that arise from the logit function applied to those expected payoffs. This "logit equilibrium" forms a stochastic generalization of the Nash equilibrium and provides a possible explanation of anomalous laboratory data.

JEL Classification: C62, C73
Keywords: bounded rationality, noisy directional learning, Fokker-Planck equation, potential games, logit equilibrium.

* This research was funded in part by the National Science Foundation (SBR-9617784 and SBR-9818683). We wish to thank (without implicating) Kenneth Arrow, Harmen Bussemaker, Peter Hammond, Michihiro Kandori, Regis Renault, Robert Rosenthal, Thomas Sargent, and Arthur Schram for helpful comments and suggestions on earlier drafts.

1. INTRODUCTION

Small errors and shocks may have offsetting effects in some economic contexts, in which case there is not much to be gained from an explicit analysis of stochastic elements. In other contexts, a small amount of randomness can have a large effect on equilibrium behavior.[1] Regardless of whether random elements or "trembles" are due to preference shocks, experimentation, or actual mistakes in judgement, the effect can be particularly important when players' payoffs are quite sensitive to others' decisions, e.g. when payoffs are discontinuous as in auctions, or highly interrelated as in coordination games. Nor do errors cancel out when the Nash equilibrium is near a boundary of the set of feasible actions and noise pushes actions towards the interior, as in a public goods contribution game where the Nash equilibrium is at zero contributions (full free riding). Errors are more likely when payoff differences across alternatives are small, so the consequences of mistakes are minor. For example, when managers or agents are weakly motivated by profits to owners, they may not exert much effort to find the optimal action.

Stochastic elements have been incorporated successfully into a wide array of economic theories. These stochastic elements have been typically assumed to be driven by exogenous shocks.[2] Despite Simon's (1957) early work on modeling bounded rationality, the incorporation of noise in the analysis of economic games is relatively recent. Rosenthal (1989) and McKelvey and Palfrey (1995) propose noisy generalizations of the standard Nash equilibrium.[3] McKelvey and Palfrey's "quantal response equilibrium" allows a wide class of probabilistic choice rules to be substituted for perfect maximizing behavior in an equilibrium context. Other economists have introduced noise into models of learning and evolutionary adjustment; see for instance Foster and Young (1990), Fudenberg and Harris (1992), Kandori, Mailath, and Rob (1993), Binmore, Samuelson, and Vaughan (1995), and Chen, Friedman, and Thisse (1997). In particular, Foster and Young (1990) and Fudenberg and Harris (1992) use a Brownian motion process, similar to

---

[1] For example, in evolutionary models of coordination a small mutation rate may prevent the system from getting stuck in an equilibrium that is risk dominated (see e.g. Kandori, Mailath, and Rob, 1993, and Young, 1993). Similarly, a small amount of noise or "trembles" can be used to rule out certain Nash equilibria (Selten, 1975).

[2] For instance, real business cycle models and much econometric work make this assumption.

[3] See Smith and Walker (1993) and Smith (1997) for an alternative approach.

the one specified in section 2.

Our goal in this paper is to provide a unified approach to equilibrium and evolutionary dynamics for a class of models with continuous decisions. The dynamic model is based on an assumption that decisions are changed locally in the direction of increasing payoff, subject to some randomness. Specifically, we propose a model of noisy adjustment to current conditions that, in equilibrium, yields a steady-state probability distribution of decisions for each player. Our modeling approach is inspired by two strands of thought, directional adaptive behavior and randomness, both of which are grounded in early writings on bounded rationality.

Selten and Buchta's (1994) "learning direction theory" postulates that players are more likely to shift decisions in the direction of a best response to recent conditions. They show that such behavior was observed in an experimental trial of a first-price auction. However, Selten and Buchta (1994) expressly do not model the rate of adaption. One contribution of this paper is to operationalize learning direction theory by specifying an adjustment process. Our model is also linked to the literature on evolutionary game theory in which strategies with higher payoffs become more widely used. Such evolution can be driven by increased survival and fitness arguments with direct biological parallels (e.g. Foster and Young, 1990), or by more cognitive models in which agents learn to use strategies that have worked better for themselves (e.g., Roth and Erev, 1995, and Erev and Roth, 1998), or in which they imitate successful strategies used by others (Vega-Redondo, 1997, and Rhode and Stegeman, 1997). An alternative to imitation and adaptation has been to assume that agents move in the direction of best responses to others' decisions. This is the approach we take.[4]

In addition to "survival of the fittest," biological evolution is driven by mutation of existing types, which is the second element that motivates our work. In the economics literature, evolutionary mutation is often specified as a fixed "epsilon" probability of switching to a new decision that is chosen randomly from the entire feasible set (see the discussion in Kandori, 1997,

---

[4] Models of imitation and reinforcement-learning are probably more likely to yield good predictions in noisy, complex situations where players do not have a clear understanding of how payoffs are determined, but rather can see clearly their own and others' payoffs and decisions. Best-response and more forward-looking behavior is probably more likely in situations where the nature of the payoff functions is clearly understood. For example, in a Bertrand game in which the low-priced firm makes all sales, it is implausible that firms would be content merely to copy the most successful (low) price.

and the references therein). Instead of mutation via new types entering a population, we allow existing individuals to make mistakes with the probability of a mistake being inversely related to its severity. The assumption of error-prone behavior can be justified by the apparent noisiness of decisions made in laboratory experiments with financially motivated subjects. To combine the two strands of thought, we analyze a model of noisy adjustment in the direction of higher payoffs. The payoff component is more important when the payoff gradient is steep, while the noise component is more important when the payoff gradient is relatively flat.

The next step in the analysis is to translate this noisy directional adjustment into an operational description of the dynamics of strategic choice. For this step, we use a classic result from theoretical physics, namely the Fokker-Planck equation that describes the evolution of a macroscopic system that is subject to microscopic fluctuations (e.g., the dispersion of heat in some medium). The state of the system in our model is a vector of the individual players' probability distributions over possible decisions. The Fokker-Planck equation shows how the details of the noisy directional adjustment rule determine the evolution of this vector of probability distributions. These equations thus describe behavioral adjustment in a stochastic game, in which the relative importance of stochastic elements is endogenously determined by payoff derivatives.

The prime interest in the dynamical system concerns its stability and steady state (a vector of players' decision distributions that does not change over time). The adjustment rule is particularly interesting in that it yields a steady state in which the distributions that determine expected payoffs are those that are generated by applying a logit probabilistic choice rule to these expected payoffs. Our approach derives this *logit equilibrium* (McKelvey and Palfrey, 1995) from a completely different perspective than its usual roots. We then prove stability of the adjustment rule for an important class of games, i.e. "potential games" for which the Nash equilibrium can be found by maximizing some function of all players' decisions. In particular, the Liapunov function that is maximized in the steady state of our model is the expected value of the potential function plus the standard measure of entropy in the system, which is weighted by an error parameter. Finally, we show how the stability analysis bolsters the intuition behind comparative statics and dynamics properties.

The dynamic model and its steady state are presented in section 2. Section 3 contains an analysis of global stability for an interesting class of games, i.e. potential games, which include public goods, coordination, oligopoly, and two-person matrix game formulations. An application to the minimum-effort coordination game is presented in section 4, and section 5 concludes.

2. Evolution and Equilibrium with Stochastic Errors

In this section we specify a stochastic model in continuous time to describe the interaction of a finite number of players. In our model, players tend to move towards decisions with higher expected payoffs, but such movements are subject to random shocks. At any point in time, the state of the system is characterized by probability distributions of players' decisions. The steady-state equilibrium is a fixed point at which the distributions that determine expected payoffs have converged to distributions of decisions that are based on those expected payoffs. The importance of stochastic inputs is parameterized in a manner that yields the standard Nash equilibrium as a limiting case with no noise. The specific evolutionary process we consider shows an intuitive relationship between the nature of the adjustment and the probabilistic choice structure used in the equilibrium. In particular, with adjustments that are proportional to marginal payoffs plus normal noise, the steady state has a logit structure.

There are $n \geq 2$ players that make decisions in continuous time. At time t, player i selects an action $x_i(t) \in [\underline{x}, \bar{x}]$, where $i = 1,...,n$. Since actions will be subject to random shocks, behavior will be characterized by probability distributions. Let $F_i(x,t)$ be the probability that player i chooses an action less than or equal to x at time t. Similarly, let the vector of the n-1 other players' decisions and probability distributions be denoted by $x_{-i}(t)$ and $F_{-i}(x_{-i},t)$ respectively. The instantaneous expected payoff for player i at time t depends on the action taken and on the distributions of others' decisions:

$$\text{(1)} \qquad \pi^e{}_i(x_i(t),t) \;=\; \int \pi_i(x_i(t),x_{-i})\, dF_{-i}(x_{-i},t), \qquad i = 1,\ldots,n.$$

We assume that payoffs, and hence expected payoffs, are bounded from above. In addition, we assume that expected payoffs are differentiable in $x_i(t)$ when the distribution functions are. The latter condition is ensured when payoffs the $\pi_i(x_i,x_{-i})$ are continuous.[5]

To capture the idea of local adjustment to better outcomes, we assume that players move in the *direction* of increasing expected payoff, with the rate at which players change increasing in the marginal benefit of making that change.[6] This marginal benefit is denoted by $\pi^e_i{}'(x_i(t),t)$, where the prime denotes the partial derivative with respect to $x_i(t)$. However, individuals may make mistakes in the calculation of expected payoff, or they may be influenced by non-payoff factors. Therefore, we assume that the directional adjustments are subject to error, which we model as an additive disturbance, $w_i(t)$, weighted by a variance parameter σ_i:[7]

$$dx_i(t) \;=\; \pi^e_i{}'(x_i(t),t)\,dt \;+\; \sigma_i\,dw_i(t), \qquad i = 1,\ldots,n. \tag{2}$$

Here $w_i(t)$ is a standard Wiener (or white noise) process that is assumed to be independent across players and time. Essentially, dx_i/dt equals the slope of the individual's expected payoff function plus a normal error with zero mean and unit variance.

The deterministic part of the local adjustment rule (2) indicates a "weak" form of feedback in the sense that players react to the distributions of others' actions (through the expected payoff function), rather than to the actions themselves. This formulation is motivated by laboratory experiments that use a random matching protocol. Random matching causes players' observations of others' actions to keep changing even when behavior has stabilized. When players gain experience they will take this random matching effect into account and react to the "average observed decision" or the distribution of decisions rather than to the decision of

---

5 Continuity of the payoffs is sufficient but not necessary. For instance, in a first-price auction with prize value V, payoffs are discontinuous, but expected payoffs, $(V - x_i) \prod_{j\neq i} F_j(x_i)$, are twice differentiable when the F_j are twice differentiable. More generally, the expected payoff function will be twice differentiable even when the payoffs $\pi_i(x_i,x_{-i})$ are only piece-wise continuous.

6 Friedman and Yellin (1997) show that when adjustment costs are quadratic in the speed of adjustment, it is optimal for players to alter their actions partially and in proportion to the gradient of expected payoff.

7 This adjustment process is supplemented with so-called "reflecting boundary conditions" (see Gihman and Skorohod, 1972) to ensure that actions stay within the feasible region $[\underline{x}, \bar{x}]$.

their latest opponent.[8]

The stochastic part of the local adjustment rule in (2) captures the idea that such adaptation is imperfect and that decisions are subject to error. It is motivated by observed noise in laboratory data where adjustments are often unpredictable, and subjects sometimes experiment with alternative decisions. In particular, "errors" or "trembles" may occur because current conditions are not known precisely, expected payoffs are only estimated, or decisions are affected by factors beyond the scope of current expected payoffs, e.g. emotions like curiosity, boredom, inertia, or desire to change. The random shocks in (2) capture the idea that players may use heuristics or "rules of thumb" to respond to current payoff conditions. We assume that these responses are, on average, proportional to the correct expected payoff gradients, but that calculation errors, extraneous factors, and imperfect information require that a stochastic term be appended to the deterministic part of (2). Taken together, the two terms in (2) simply imply that a change in the direction of increasing expected payoff is more likely, and that the magnitude of the change is positively correlated with the expected payoff gradient.

The adjustment rule (2) translates into a differential equation for the distribution function of decisions, $F_i(x,t)$. This equation will depend on the density $f_i(x,t)$ corresponding to $F_i(x,t)$, and on the slope, $\pi^e_i{}'(x,t)$, of the expected payoff function. It is a well-known result from theoretical physics that the stochastic adjustment rule (2) yields the Fokker-Planck equation for the distribution function.[9,10]

---

[8] An alternative formulation results when the expected payoff in (2) is replaced by the instantaneous payoff, $\pi(x_1(t),...,x_n(t))$, at time t. One important difference is that the latter formulation gives rise to a *single* Fokker-Planck equation that describes the evolution of the joint density of $x_1(t),...,x_n(t)$. In contrast, (2) leads to a *separate* equation for the marginal density of each x_i (linked only through the expected payoff function), see Proposition 1.

[9] This result has been derived independently by a number of physicists, including Einstein (1905), and the mathematician Kolmogorov (1931). The first term on the right side of (3) is known as a drift term, and the second term is a diffusion term. The standard example of pure diffusion without drift is a small particle in a suspension of water; in the absence of external forces the particle's motion is completely determined by random collisions with water molecules (Brownian motion). A drift term is introduced, for instance, when the particle is charged and influenced by an electric field.

[10] Binmore, Samuelson, and Vaughan (1995) use the Fokker-Planck equation to model the evolution of choice probabilities in 2 × 2 matrix games. Instead of using the expected-payoff derivative as we do in (2), they use a non-linear genetic-drift function. Friedman and Yellin (1997) consider a one-population model in which all players get the same payoff from a given vector of actions, which they call "games of common interest." (This is a subset of the class of potential games discussed below.) They start out with the assumption that the distribution evolves according to (3), but without the error term (i.e. $\mu_i = 0$). This deterministic version of Fokker-Planck is used to show that behavior converges

Proposition 1. The noisy directional adjustment process (2) yields the Fokker-Planck equation for the distributions of decisions:

$$(3) \qquad \frac{\partial F_i(x,t)}{\partial t} = -\pi^{e\,\prime}_i(x,t)\; f_i(x,t) + \mu_i\; f_i^{\prime}(x,t), \quad i = 1,\ldots,n,$$

where $\mu_i = \sigma_i^2/2$.

A derivation of the Fokker-Planck equation is presented in Appendix A. Existence of a (twice differentiable) solution to the Fokker-Planck equation is demonstrated in most textbooks on stochastic processes (e.g., Smoller, 1994; Gihman and Skorohod, 1972). Notice that there is a separate equation for each player $i = 1,\ldots,n$, and that the individual Fokker-Planck equations are interdependent only through the expected payoff functions.[11]

The Fokker-Planck equation (3) has a very intuitive economic interpretation. First, players' decisions tend to move in the direction of greater payoff, and a larger payoff derivative induces faster movement. In particular, when payoff is increasing at some point x, lower decisions become less likely, decreasing $F_i(x,t)$. The rate at which probability mass crosses over at x depends on the density at x, which explains the $-\pi^{e\prime}_i(x,t)\, f_i(x,t)$ term on the right side of (3). The second term, $\mu_i f_i^{\prime}$, reflects aggregate noise in the system (due to intrinsic errors in decision making), which causes the density to "flatten out." Locally, if the density has a positive slope at x, then flattening moves mass toward lower values of x, increasing $F_i(x,t)$, and vice versa, as indicated by the second term on the right side of equation (3).

Since $\mu_i = \sigma_i^2/2$, the coefficient μ_i in (3) determines the importance of errors relative to payoff-seeking behavior for individual i. First consider the limiting case $\mu_i = 0$. If behavior in (3) converges, it must be the case that $\pi^{e\prime}_i(x)\, f_i(x) = 0$, which is the necessary condition for an interior Nash equilibrium: either the necessary condition for payoff maximization is satisfied at x, or else the density of decisions is zero at x. As μ_i goes to infinity in (3), the noise effect dominates and the Fokker-Planck equation tends to $\partial F_i/\partial t = \mu_i\, \partial^2 F_i/\partial x^2$, which is the "heat

---

to a (local) Nash equilibrium in such games.

[11] Replacing the expected payoff in (2) by the instantaneous payoff, $\pi(x_1(t),\ldots,x_n(t))$, results in a *single* Fokker-Planck equation that describes the evolution of the joint density of $x_1(t),\ldots,x_n(t)$.

equation" that describes how heat spreads out uniformly in some medium. In this limit, the steady state of (3) is a uniform density with $f_i' = 0$.

In a steady state of the process in (3), the right side is identically zero, which yields the equilibrium conditions:

$$f_i'(x) \;=\; \pi^{e\,\prime}_i(x)\; f_i(x)\,/\,\mu_i, \qquad i = 1,\ldots,n, \tag{4}$$

where the t arguments have been dropped since these equations pertain to a steady state. These equations can be simplified by dividing both sides by $f_i(x)$ and integrating, to obtain:

$$f_i(x) \;=\; \frac{\exp(\pi^e_i(x)/\mu_i)}{\int_{\underline{x}}^{\bar{x}} \exp(\pi^e_i(s)/\mu_i)\, ds}, \qquad i = 1,\ldots,n, \tag{5}$$

where the integral in the denominator is a constant, independent of x, which ensures that the density integrates to one.

The formula in (5) is a continuous analogue to the logit probabilistic choice rule. Since the expected payoffs on the right side depend on the distributions of the other players' actions (see (1)), the equations in (5) are not explicit solutions. Instead, these equations constitute *equilibrium conditions* for the steady state distribution: the probability distributions that determine expected payoffs must match the choice distributions determined by the logit formula in (5). In the steady-state equilibrium these conditions are simultaneously satisfied. The steady-state equilibrium is a continuous version of the quantal response equilibrium proposed by McKelvey and Palfrey (1995).[12] Thus we generate a logit equilibrium as a steady-state from a more primitive formulation of noisy directional learning, instead of imposing the logit form as a model of decision error. To summarize:

---

[12] Rosenthal (1989) proposed a similar equilibrium with endogenously determined distributions of decisions, although he used a linear probabilistic choice rule instead of the logit rule. McKelvey and Palfrey (1995) consider a more general class of probabilistic choice rules, which includes the logit formula as a special case. Our model with continuous decisions is similar to the approach taken in Lopez (1995).

Proposition 2. When players adjust their actions in the direction of higher payoff, but are subject to normal error as in (2), then any steady state of the Fokker-Planck equation (3) constitutes a logit equilibrium as defined by (5).

This derivation of the logit model is very different from the usual derivations. Luce (1959) uses an axiomatic approach to tie down the form of choice probabilities.[13] In econometrics, the logit model is typically derived from a "random-utility" approach.[14] Both of these derivations are static in nature. Here the logit model results from the behavioral assumption of directional adjustment with normal error.

Some properties of the equilibrium distributions can be determined from the structure of (4) or (5), independent of the specific game being considered. Equation (5) specifies the choice density to be proportional to an exponential function of expected payoff, so that actions with higher payoffs are more likely to be chosen, and the local maxima and minima of the equilibrium density will correspond to local maxima and minima of the expected payoff function. The error parameter determines how sensitive the density is to variations in expected payoffs. As the error parameter goes to infinity, the slope of the density in (4) goes to zero, and so the density in (5) becomes uniform, i.e. totally random and unaffected by payoff considerations. Conversely, as the error parameter becomes small, the density in (5) will place more and more mass on decisions with high expected payoffs. In the literature on stochastic evolution, it is common to proceed

---

[13] Luce (1959) postulated that decisions satisfy a "choice axiom," which implies that the ratio of the choice probabilities for two decisions is independent of the overall choice set containing those two choices (the Independence of Irrelevant Alternatives property). In that case, he shows that there exist "scale values" u_i such that the probability of choosing decision i is $u_i/\Sigma_j u_j$. The logit model follows when $u_i = \exp(\pi_i/\mu)$.

[14] This footnote presents the random-utility derivation of the logit choice rule for a finite number of decisions. Suppose there are m decisions, with expected payoffs $u_1,...,u_m$. A probabilistic discrete choice model stipulates that a person chooses decision k if: $u_k + \varepsilon_k > u_i + \varepsilon_i$, for all $i \neq k$, where the ε_i are random variables. The errors allow the possibility that the decision with the highest payoff will not be selected, and the probability of such a mistake depends on both the magnitude of the difference in the expected payoffs and on the "spread" in the error distribution. The logit model results from the assumption that the errors are i.i.d. and double-exponentially distributed. The probability of choosing decision k is then $\exp(u_k/\mu) / \Sigma_i \exp(u_i/\mu)$, where μ is proportional to the standard deviation of the error distribution. There are two alternative interpretations of the ε_i errors: they can either represent mistakes in the calculation or perception of expected payoffs, or they can represent unobservable preference shocks. These two interpretations are formally equivalent, although one implies bounded rationality and the other implies that seemingly incorrect decisions are really rational with respect to the unobserved preferences. See Anderson, de Palma, and Thisse (1992, chapter 2) for further discussion and other derivations of the logit model.

directly to the limiting case as the amount of noise goes to zero.[15] This limit is not our primary interest, for two reasons. First, econometric analysis of data from laboratory experiments (e.g., Capra et al., 1999) yields error parameter estimates that are significantly different from zero, which is the null hypothesis corresponding to a Nash equilibrium. Second, the limiting case of perfect rationality is generally a Nash equilibrium, and our theoretical analysis was originally motivated as an attempt to explain data patterns that are consistent with economic intuition but which are not predicted by a Nash equilibrium. As we have shown elsewhere, the (static) logit model (5) yields comparative static results that conform with both economic intuition and data patterns from laboratory experiments, but are not predicted by the standard Nash equilibrium (Anderson, Goeree, and Holt, 1998a, 1998b, 1999; Capra *et al*., 1999). The dynamic adjustment model presented here gives a theoretical justification for using the logit equilibrium to describe decisions when behavior has stabilized, e.g., in the final periods of laboratory experiments.

To summarize the main result of this section, the steady-state distributions of decisions that follow from the adjustment rule (2) satisfy the conditions that define a logit equilibrium. Therefore, when the dynamical system described by (3) is stable, the logit equilibrium results in the long run when players adjust their actions in the direction of higher payoff (directional learning), but are subject to error. In the next section, we use Liapunov function methods to prove stability and existence for the class of potential games.

3. Stability Analysis

So far, we have shown that any steady state of the Fokker-Planck equation (3) is a logit equilibrium. We now consider the dynamics of the system (3) and characterize sufficient conditions for a steady state to be attained in the long run. Specifically, we use Liapunov methods to prove stability for a class of games that includes some widely studied special cases. A Liapunov function is non-decreasing over time and has a zero time derivative only when the system has reached an equilibrium steady state. The system is (locally) stable when such a function exists. Although our primary concern is the effect of endogenous noise, it is instructive

---

[15] One exception is Binmore and Samuelson (1997), who consider an evolutionary model in which the mistakes made by agents (referred to as "muddlers") are not negligible. At the aggregate level, however, the effect of noise is washed out when considering the limit of an infinite population.

to begin with the special case in which there is no decision error and all players use pure strategies. Then it is natural to search for a function of all players' decisions that will be maximized (at least locally) in a Nash equilibrium. In particular, consider a function, $V(x_1,...,x_n)$, with the property $\partial V/\partial x_i = \partial \pi_i/\partial x_i$ for $i = 1,...,n$. When such a function exists, Nash equilibria can be found by maximizing V. The $V(\cdot)$ function is called the potential function, and games for which such a function exists are known as potential games (Monderer and Shapley, 1996).[16]

The usefulness of the potential function is not just that it is (locally) maximized at a Nash equilibrium. It also provides a direct tool to prove equilibrium stability under the directional adjustment hypothesis in (2). Indeed, in the absence of noise, the potential function itself is a Liapunov function (see also Slade, 1994):

$$(7) \qquad \frac{dV}{dt} = \sum_{i=1}^{n} \frac{\partial V}{\partial x_i}\frac{dx_i}{dt} = \sum_{i=1}^{n} \frac{\partial \pi_i}{\partial x_i}\frac{dx_i}{dt} = \sum_{i=1}^{n} (\partial \pi_i / \partial x_i)^2 \geq 0,$$

where the final equality follows from the directional adjustment rule (2) with no noise, i.e. $\sigma_i = 0$.[17] Thus the value of the potential function is strictly increasing over time unless all payoff derivatives are zero, which is a necessary condition for an interior Nash equilibrium. The condition that $dV/dt = 0$ need not generate a Nash equilibrium: the process might come to rest at a local maximum of the potential function that corresponds to a *local* Nash equilibrium from which large unilateral deviations may still be profitable.

Our primary interest concerns noisy decisions, so we will work with the expected value of the potential function. It follows from (1) that the partial derivatives of the expected value of the potential function correspond to the partial derivatives of the expected payoff functions:

$$(8) \qquad \pi^{e\,\prime}_i(x_i,t) = \frac{\partial}{\partial x_i}\int V(x_i,x_{-i})\, dF_{-i}(x_{-i},t), \qquad i = 1,\ldots,n.$$

Again, the intuitive idea is to use something that is maximized at a logit equilibrium to construct a Liapunov function, i.e. a function whose time derivative is non-negative and only equal to zero

---

[16] Rosenthal (1973) first used a potential function to prove the existence of a pure-strategy Nash equilibrium in congestion games.

[17] This type of deterministic gradient-based adjustment has a long history, see Arrow and Hurwicz (1960).

at a steady state. When $\mu_i > 0$ for at least one player *i*, then the steady state is not generally a Nash equilibrium, and the potential function must be augmented to generate an appropriate Liapunov function. Look again at the Fokker-Planck equation (3); the first term on the right side is zero at an interior maximum of expected payoff, and the $f_i'(x,t)$ term is zero for a uniform distribution. Therefore, we want to augment the Liapunov function with a term that is maximized by a uniform distribution. Consider the standard measure of noise in a stochastic system, entropy, which is defined as $-\Sigma^n_{i=1}\int f_i \log(f_i)$. It can be shown that this measure is maximized by a uniform distribution, and that entropy is reduced as the distribution becomes more concentrated. The Liapunov function we seek is constructed by adding entropy to the expected value of the potential function:

$$(9) \qquad L = \int_{\underline{x}}^{\overline{x}} .. \int_{\underline{x}}^{\overline{x}} V(x_1,..,x_n)\, f_1(x_1,t)\, ..f_n(x_n,t)\, dx_1 .. dx_n - \sum_{i=1}^{n} \mu_i \int_{\underline{x}}^{\overline{x}} f_i(x_i,t) \log(f_i(x_i,t))\, dx_i .$$

The μ_i parameters determine the relative importance of the entropy terms in (9), which is not surprising given that μ_i is proportional to the variance of the Wiener process in player *i*'s directional adjustment rule (2). Since entropy is maximized by a uniform distribution (i.e. purely random decision making), it follows that decision distributions that concentrate probability mass on higher-payoff actions will have lower entropy. Therefore, one interpretation of the role of the entropy term in (9) is that, if the μ_i parameters are large, then entropy places a high "cost" of concentrating probability on high-payoff decisions.[18]

We prove that the dynamical system described by (3) converges to a logit equilibrium, by showing that the Liapunov function (9) is non-decreasing over time.[19]

---

[18] The connection between entropy and the logit choice probabilities is well established in physics and economics. For example, Anderson, de Palma, and Thisse (1992) showed that logit demands are generated from a representative consumer with a utility function that has an entropic form.

[19] The notion of convergence used here is "weak convergence" or "convergence in distribution:" the random variable $x(t)$ weakly converges to the random variable X if $\lim_{t\to\infty} \text{Prob}[x(t) \le x] = \text{Prob}[X \le x]$ for all x. Proposition 3 thus implies that the random variable $x_i(t)$ defined in (2) weakly converges to a random variable that is distributed according to a logit equilibrium distribution, for any starting point $x_i(0)$.

Proposition 3. For the class of potential games, behavior converges to a logit equilibrium when players adjust their actions in the direction of higher payoff, but are subject to normal error as in (2).

Proof. In Appendix B we show that the Liapunov function is non-decreasing over time; by taking the time derivative of the Liapunov function, partially integrating, and using the Fokker-Planck equation, we can express this time derivative in a form that is analogous to (7)

$$\frac{dL}{dt} = \sum_{i=1}^{n} \int_{0}^{\bar{x}} \frac{(\partial F_i(x_i,t)/\partial t)^2}{f_i(x_i,t)} dx_i \geq 0 . \tag{10}$$

The entropy term in (9) is maximized by the uniform densities $f_i(x,t) = 1/(\bar{x} - \underline{x})$, $i = 1,...,n$. It follows from this observation that the maximum entropy is given by $\log(\bar{x} - \underline{x}) \sum_i \mu_i$, which is finite. The expected value of the potential function is bounded from above since, by assumption, expected payoffs are. Therefore, the Liapunov function, which is the sum of expected potential and entropy, is bounded from above. Since L is non-decreasing over time for any potential game, we must have $dL/dt \to 0$ as $t \to \infty$, so $dF_i/dt \to 0$ in this limit. By (3) this yields the logit equilibrium conditions in (4). The solutions to these equilibrium conditions are the logit equilibria defined by (5). Q.E.D.

When there are multiple logit equilibria, the equilibrium attained under the dynamical process (3) is determined by the initial distributions $F_i(x,0)$. We now show that (local) maxima of the Liapunov function correspond to (locally) stable logit equilibria.

Proposition 4. A logit equilibrium is locally (asymptotically) stable under the process (3) if and only if it corresponds to a strict local maximum of the Liapunov function in (9). When the logit equilibrium is unique, it is globally stable.

Proof. We first show that strict local maxima of the Liapunov function are locally (asymptotically) stable logit equilibria. Let $\underline{F}^*(x)$ denote a vector of distributions that constitutes a logit equilibrium which corresponds to a strict local maximum of the Liapunov function. Suppose that at $\underline{F}^*$ the Liapunov function attains the value L^*. Furthermore, let U be the set of distributions in the neighborhood of $\underline{F}^*$ for which $L \geq L^* - \varepsilon$, where $\varepsilon > 0$ is small. Since ε can be made arbitrarily small, we may assume that U contains no other stationary points of L. Note from (10) that L is non-decreasing over time, so no trajectory starting in U will ever leave it. Moreover, since $\underline{F}^*$ is the only stationary point of L in U, Proposition 3 implies that all trajectories starting in U necessarily converge to $\underline{F}^*$ in the limit $t \rightarrow \infty$, i.e., $\underline{F}^*$ is locally (asymptotically) stable. Hence, strict local maxima of L are locally stable logit equilibria.

Next, we prove that any locally (asymptotically) stable logit equilibrium, $\underline{F}^*$, is a strict local maximum of L. Since $\underline{F}^*$ is locally (asymptotically) stable, there exists a local neighborhood U of $\underline{F}^*$ that is invariant under the process (3), and whose elements converge to $\underline{F}^*$. The Liapunov function is strictly increasing along a trajectory starting from any distribution in U (other than $\underline{F}^*$ itself), so L necessarily attains a strict local maximum at $\underline{F}^*$. Finally, when the logit equilibrium is unique, it corresponds to the unique stationary point of L. Proposition 3 holds for any initial distribution, so the logit equilibrium is globally stable. Q.E.D.

It follows from (10) that $dF_i/dt = 0$ when the Liapunov function is (locally) maximized, which, by (3) and (4), implies that a logit equilibrium is necessarily reached. Recall that, in the absence of noise, a local maximum of the Liapunov function does not necessarily correspond to a Nash equilibrium; the system may come to rest at a local Nash equilibrium, for which "large" unilateral deviations are still profitable (see Friedman and Yellin, 1997). In contrast, with noise, local maxima of the Liapunov function always produce a logit equilibrium in which decisions with higher expected payoffs are more likely to be made. In fact, even (local) *minima* of the Liapunov function correspond to such equilibria, although they are unstable steady states of the dynamical system (3).

Propositions 3 and 4 do not preclude the existence of multiple locally stable equilibria. In such cases, the initial conditions determine which equilibrium will be selected. As shown in the proof of Proposition 4, if the initial distributions are "close" to those of a particular logit equilibrium, then that equilibrium will be attained under the dynamic process (3). The 2 × 2 coordination game example presented at the end of this section illustrates the possibility of multiple stable equilibria.

Since the existence of potential functions is crucial to the results of Proposition 3, we next discuss conditions under which such functions can be found. A necessary condition for the existence of a potential function is that $\partial^2\pi_i/\partial x_j\partial x_i = \partial^2\pi_j/\partial x_i\partial x_j$ for all i, j, since both sides are equal to $\partial^2 V/\partial x_i\partial x_j$. Hence, the existence of a potential function requires $\partial^2[\pi_i - \pi_j]/\partial x_j\partial x_i = 0$ for all i, j. Moreover, these "integrability" conditions are also sufficient to guarantee existence of a potential function. It is straightforward to show that payoffs satisfy the integrability conditions if and only if: $\pi_i(x_1,...,x_n) = \pi_c(x_1,...,x_n) + \theta_i(x_i) + \phi_i(x_{-i})$ for $i = 1,...,n$, where π_c is the same for players, hence it has no i subscript. To see that this class of payoffs solves the integrability condition, note that the common part, π_c, cancels when taking the difference of π_i and π_j, and the player specific parts, θ_i and ϕ_i, vanish upon differentiation. If we define $V(x_1,...,x_n) = \pi_c(x_1,...,x_n) + \Sigma^n_{i=1}\theta_i(x_i)$, we can write the above payoffs π_i as the sum of two components: a common component and a component that only depends on others' decisions

$$\pi_i(x_1,...,x_n) \quad = \quad V(x_1,...,x_n) + \alpha_i(x_{-i}), \qquad i = 1,...,n, \tag{11}$$

where we have defined $\alpha_i(x_{-i}) = \phi_i(x_{-i}) - \Sigma_{j\neq i}\theta_j(x_j)$. The common part, V, has no i subscript, and is the same function for all players, although it is not necessarily symmetric in the x_i. The individual part, $\alpha_i(x_{-i})$, may differ across players. The common part includes benefits or costs that are determined by one's own decision, e.g. effort costs. The $\alpha_i(x_{-i})$ term in (11) does not affect the Nash equilibrium since it is independent of one's own decision, e.g. others' effort costs or gifts received from others. It follows from this observation that the partial derivative of $V(x_1,...,x_n)$ with respect to x_i is the same as the partial derivative of $\pi_i(x_1,...,x_n)$ with respect to x_i for $i = 1,...,n$, so $V(\cdot)$ is a potential function for this class of payoffs. Proposition 3 then implies that behavior converges to a logit equilibrium for this class of games.

The payoff structure in (11) covers a number of important games. For instance, consider a linear public goods game in which individuals are given an endowment, ω. If an amount x_i is contributed to a public good, the player earns $\omega - x_i$ for the part of the endowment that is kept. In addition, every player receives a constant (positive) fraction m of the total amount contributed to the public good. Therefore, the payoff to player *i* is: $\pi_i = \omega - x_i + m X$, where X is the sum of all contributions including those of player *i*. The potential for this game is: $V(x) = \omega + m X - \Sigma_i x_i$, and $\alpha_i(x_{-i}) = \Sigma_{j\neq i} x_j$. Another example is the minimum-effort coordination game (see e.g. Bryant, 1983) for which: $\pi_i = \min_{j=1...N}\{x_j\} - cx_i$, where effort costs $c \in [0, 1]$. Here, $V(x) = \min_{j=1...N}\{x_j\} - \Sigma_i cx_i$ (see also section 4). In both of these applications the common part represents a symmetric production function, included once, minus the sum of all players' effort costs. In previous work on public goods and coordination games, we showed that the logit equilibrium is unique (Anderson, Goeree, and Holt, 1998b, 1999). Therefore, the directional adjustment process studied here is globally stable for these games.

It is also straightforward to construct potential functions for many oligopoly models. Consider a Cournot oligopoly with *n* firms and linear demand, so that $\pi_i = (a - b X)\, x_i - c_i(x_i)$, where X is the sum of all outputs and $c_i(x_i)$ is firm *i*'s cost function. Since the derivative of firm *i*'s profit with respect to its own output is given by $\partial\pi_i/\partial x_i = a - b X - b x_i - c_i'$, the potential function is easily derived as: $V = a X - b/2\, X^2 - b/2\, \Sigma_i x_i^2 - \Sigma_i c_i(x_i)$. Some non-linear demand specifications can also be incorporated.

As a final example, consider the class of symmetric two-player matrix games with two decisions, and payoffs shown in Table 1 below.

		Player 2	
		D_I	D_{II}
Player 1	D_I	*a*, *a*	*b*, *c*
	D_{II}	*c*, *b*	*d*, *d*

Table I. The Payoff Matrix (payoffs for player 1, payoffs for player 2)

Player *i* is characterized by a probability x_i of choosing decision D_I.[20] Thus the payoff to player

20 This formulation corresponds to the setting in some laboratory experiments when subjects are required to select probabilities rather than actions, with the experimenter performing the randomization according to the selected probabilities. This method is used when the focus is on the extent to which behavior conforms to a mixed-strategy Nash

i is linear in the probability x_i:

(12) $$\pi_i(x_i, x_{-i}) \;=\; d \;+\; (a - b - c + d)\, x_i\, x_{-i} \;+\; (b - d)\, x_i \;+\; (c - d)\, x_{-i}, \quad i = 1, 2.$$

It is straightforward to show that for this payoff structure the potential function is given by $V = (a - b - c + d)\, x_1 x_2 + (b - d)(x_1 + x_2)$.[21]

This example is useful to illustrate the possibility of multiple logit equilibria and their stability. Since the payoff in (12) is linear in a player's own probability, the expected payoff will also be linear. Therefore, it follows from (5) that the equilibrium densities are exponential:

(13) $$f_i(x) \;=\; \frac{\gamma_i \exp(\gamma_i x)}{\exp(\gamma_i) - 1}, \quad i = 1, 2,$$

where the parameter γ_i is defined as:

(14) $$\gamma_i \;=\; [(a - b - c + d)\, E_{-i} \;+\; b \;-\; d\,]/\mu_i, \quad i = 1, 2,$$

where E_{-i} denotes the expected value of x_{-i}. The expression in (13) is not an explicit solution since γ_i on the right side depends on the expected value of the other's choice, as indicated by (14). In order to obtain an equilibrium consistency condition, we use the densities in (13) to calculate the expected value E_i:

(15) $$E_i \;=\; \frac{\exp(\gamma_i)}{\exp(\gamma_i) - 1} \;-\; \frac{1}{\gamma_i}, \quad i = 1, 2.$$

The solutions, E_1^* and E_2^* (or, equivalently, γ_1^* and γ_2^*) to (14) and (15) determine the logit equilibrium densities in (13). Consider a symmetric coordination game with $a = 2$, $d = 1$, and

---

equilibrium. Ochs (1995) used this approach in a series of matching-pennies games. He reports that choice probabilities are sensitive to a player's own payoffs, contrary to the prediction of a mixed-strategy equilibrium. He finds some empirical support for the quantal response equilibrium.

[21] In an asymmetric game, the letters representing payoffs in (12) would have i subscripts, $i = 1, 2$. Asymmetries in the constant or final two terms pose no problems for the construction of a potential function, so the only difficulty is to make the $(a_i - b_i - c_i + d_i)$ coefficient of the interaction terms match for the two players. This can be accomplished by a simple rescaling of all four payoffs for one of the players. Rescaling by a positive factor will not affect the stability proof of Proposition 3.

$b = c = 0$, and assume furthermore that $\mu_1 = \mu_2 = \mu$. For these parameter values, there are two-pure strategy Nash equilibria plus one in mixed strategies. For low values of μ, there are also three logit equilibria. These are illustrated in Figure 1, where E_1^* is plotted as a function of μ. (Since $b = c$ for this example, we have $E_2^* = E_1^*$ for all μ.) The upper graph shows two of the equilibria, which exist only for $\mu \leq 0.085$. The lower graph shows the third logit equilibrium, which exists for all $\mu \geq 0$. The light line corresponds to the unstable equilibrium, and the dark lines to the stable ones.[22]

The existence of multiple *stable* equilibria in this example is perhaps striking because standard evolutionary models would always select the risk-dominant equilibrium (D_I,D_I), at which the potential is globally maximized (see e.g. Foster and Young, 1990, and Kandori, Mailath, and Rob, 1993). In our context, the equilibrium that is selected depends on the initial conditions. In this sense, "history" matters in the directional learning model but not in the standard evolutionary model.[23]

4. Patterns of Dynamic Adjustment

In this section, we describe in more detail how the system adjusts towards its steady state. In particular, we show how the Fokker-Planck equation can be used to analyze the evolution of players' decisions in a two-person minimum-effort coordination game. The payoff for each player is the minimum of the two efforts, minus the cost of the player's own effort: $\pi_i = \min\{x_1, x_2\} - cx_i$, where x_i is player i's effort level and $c < 1$ is a cost parameter. Notice that a unilateral increase from any common effort level is costly but does not affect the minimum. Similarly, a unilateral decrease from any common effort level will reduce the minimum by more than the cost

---

[22] This can be proved in the following manner: first consider the limit equilibria as $\mu \to 0$. In this limit, the equilibrium corresponding to the dashed line in Figure 1 is given (from (14) and (15)) by $f_i(x) = \gamma \exp(\gamma x)/(\exp(\gamma)-1)$, where $\gamma \approx -2.15$, and $i = 1, 2$. If this equilibrium were stable, arbitrary perturbations of f_1 and f_2 would reduce the value of the Liapunov function. However, consider the perturbations: $f_i(x) \to f_i(x) + \varepsilon(x - 1/2)$ for $i = 1, 2$, with ε small. It can then be shown that for this perturbation the Liapunov function *increases* by $\varepsilon^2/48$, so that the corresponding equilibrium is not locally stable by Proposition 3. Furthermore, at the other two limit equilibria, arbitrary perturbations reduce the Liapunov function because the negative change in the entropy term dominates, so these equilibria are locally stable. These stability properties extend to $\mu > 0$ until a bifurcation takes place, which is at $\mu = 0.085$.

[23] See, however, Binmore and Samuelson (1999) who show that adding small perturbations (or "drift" terms) to an evolutionary selection process can have a large effect on which equilibrium is selected when the payoff landscape has "flat valleys."

saving, since $c < 1$. Hence, *any* common effort level is a Nash equilibrium. In contrast, there is a *unique* logit equilibrium for the minimum-effort coordination game that is symmetric across players (see Anderson, Goeree, and Holt, 1999). The Fokker-Planck equation will thus be globally stable and produce a unique steady-state distribution and, hence, a unique prediction for the steady-state average effort levels. These predictions will be compared with data from laboratory experiments based on the minimum-effort coordination game (Goeree and Holt, 1999).

Note that while a low effort cost makes it relatively safe to choose a high effort, a high cost makes this action risky as it may not be matched by the other player, which suggests that actual behavior might be sensitive to changes in the cost parameter. Goeree and Holt (1999) report an experiment with randomly matched pairs of subjects who made effort choices from a *continuous* interval [110, 170]. They conducted three sessions with a low effort cost of $c = .25$ and three sessions with a high cost of $c = .75$. The initial period-one decisions were uniformly distributed over the range of feasible effort choices for both treatments. However, efforts tended to rise over time in the low-cost sessions while efforts declined over time when the effort cost was high. The time-sequences of average effort choices for three groups of 10 subjects in each treatment are given by the thin light lines in Figure 2, with an upward pattern for the low-effort cost treatment and an essentially symmetric downward adjustment for the high-effort cost treatment. The thick light lines show average efforts for each treatment. The strong treatment effect, which is consistent with simple intuition about the effects of effort costs, is not predicted by the Nash equilibrium.

The separation of effort levels for the two treatments conforms nicely with the notion of maximum potential, discussed above. First, consider the ordinary (deterministic) potential function for the two-player minimum-effort coordination game: $V = \min\{x_1, x_2\} - cx_1 - cx_2$. Maximization obviously requires equal effort levels, $x_1 = x_2 = x$. The potential then becomes $V = (1 - 2c)\,x$, which is maximized at $x = \bar{x} = 170$ when $c = .25$ and at $x = \underline{x} = 110$ when $c = .75$. The introduction of some noise pushes these predictions away from the boundaries towards the center of the range of feasible decisions.

To compare the patterns of adjustments in Figure 2 with those predicted by the noisy evolutionary model of this paper, we have to solve for the distribution of effort decisions using the Fokker-Planck equation in (3). To determine marginal payoffs, note that an increase in effort by player i will raise the minimum with probability 1 - F_j and increase the cost at rate c, so $\pi^{e\prime}_i$ = 1 - F_j - c. The Fokker-Planck equation becomes:

$$\frac{\partial F_i(x,t)}{\partial t} = -(1 - F_j(x,t) - c)\, f_i(x,t) + \mu_i\, f_i^{\prime}(x,t), \quad i = 1,2,\ i \neq j. \tag{16}$$

With an error parameter μ = 7.4 and uniform initial distributions (as reported in Goeree and Holt, 1999), equation (16) can be solved numerically. The dark lines in Figure 2 show the time paths of the average efforts thus found. Note that the evolutionary model reproduces the qualitative features of the experimental data and is capable of predicting the final period averages when behavior has settled down.[24]

5. Conclusion

Models of bounded rationality are appealing because the calculations required for optimal decision making are often quite complex, especially when optimal decisions depend on what others are expected to do. This paper begins with an assumption that decisions are adjusted locally toward increasing payoffs. These adjustments are sensitive to stochastic disturbances. When the process settles down, systematic adjustments no longer occur, although behavior remains noisy. The result is an equilibrium probability distribution of decisions, with errors in the sense that optimal decisions are not always selected, although more profitable decisions are more likely to be chosen. The first contribution of this paper is to use a simple model of noisy directional adjustments to derive an equilibrium model of behavior with endogenous decision errors that corresponds to the stochastic generalization of Nash equilibrium proposed by Rosenthal (1989) and McKelvey and Palfrey (1995). The central technical step in the analysis is to show that directional adjustments subject to normal noise yield a Fokker-Planck equation, with a steady state that corresponds to a "logit equilibrium." This equilibrium is described by a logit probabilistic choice function coupled to a Nash-like consistency condition.

---

[24] The (pooled) average efforts in the final three periods were 159(11) for the low-cost session and 126(4) for the high-cost session, where the number in parentheses denotes the standard deviation of the average. The limiting values predicted by the evolutionary model are 153 and 127 respectively.

The second contribution of this paper is to prove stability of the logit equilibrium for all potential games. We use Liapunov methods to show that the dynamic system is stable for a class of interesting payoff functions, i.e. those for potential games. This class includes minimum-effort coordination games, linear/quadratic public goods and oligopoly games, and two-person 2×2 matrix games in which players select mixed strategies. The process model of directional changes adds plausibility to the equilibrium analysis, and an understanding of stability is useful in deciding which equilibria are more likely to be observed.

Models with stochastic elements are of interest because they can explain behavior of human decision makers in complex, changing situations. The stochastic logit equilibrium provides an explanation of data patterns in laboratory experiments that are consistent with economic intuition but which are not explained by a Nash equilibrium analysis (McKelvey and Palfrey, 1995, and Anderson, Goeree, and Holt, 1998a,b, 1999). The effects of noise is important when the Nash equilibrium is near the boundary of the set of feasible decisions, so that errors are biased toward the interior. In addition, errors have non-symmetric effects when payoff functions are sensitive to noise in others' behavior, and the behavior is pushed toward "safer" configurations of behavior, like low-risk, low-effort outcomes in coordination games. In the presence of noise, equilibrium behavior is not necessarily centered around the Nash prediction; errors that push one player's decision away from a Nash decision may make it safer for others to deviate. In some parameterizations of a "traveler's dilemma" game, for example, the Nash equilibrium is at the lower end of the feasible set, whereas behavior in laboratory experiments conforms more closely to a logit equilibrium with a unimodal density located at the upper end (Capra, Goeree, Gomez, and Holt, 1999).

The stochastic elements in our model are intended to capture a variety of factors, such as errors, trembles, experimentation, and non-payoff factors such as emotions. In some contexts, behavior may be largely driven by a specific bias, like the "winner's curse" in common-value auctions. In these cases, it is probably better to model the specific bias explicitly. When there is not single identifiable bias, we prefer to follow the common practice of putting left-out factors into an error term. Adding an error term to a gradient adjustment rule yields a tractable model with a steady-state equilibrium that has appealing theoretical and empirical properties.

APPENDIX A: DERIVATION OF THE FOKKER-PLANCK EQUATION (3)

Recall that the directional adjustments are stochastic: $dx(t) = \pi^{e\prime}(x(t),t)dt + \sigma\, dw(t)$ (see (2)), where we have dropped the player-specific subscripts for brevity. Note that the payoff derivative $\pi^{e\prime}$ depends on time through the decision x and through other players' distribution functions. After a small time change, Δt, the change in a player's decision can be expressed as:

(A1) $$\Delta x(t) \equiv x(t+\Delta t) - x(t) = \pi^{e\prime}(x,t)\,\Delta t + \sigma\,\Delta w(t) + o(\Delta t)\,,$$

where $\sigma\Delta w(t)$ is a normal random variable with mean zero and variance $\sigma^2\Delta t$, and $o(\Delta t)$ indicate terms that go to zero faster than Δt (i.e., K is of $o(\Delta t)$ when $K/\Delta t \rightarrow 0$ as $\Delta t \rightarrow 0$). A player's decision, therefore, is a random variable $x(t)$ that has a time-dependent density $f(x,t)$. Let $h(x)$ be an arbitrary twice differentiable function that vanishes at the boundaries, as does its derivative. At time $t + \Delta t$, the expected value of $h(x)$ can be expressed directly as:

(A2) $$E\{h(x(t+\Delta t))\} = \int_{\underline{x}}^{\overline{x}} h(x)\, f(x, t+\Delta t)\, dx.$$

The directional adjustment rule in (A1) can be used to obtain an alternative expression for the expected value of $h(x)$ at time $t + \Delta t$:

(A3) $$E\{h(x(t+\Delta t))\} = E\{h(x(t) + \Delta x(t))\} \approx E\{h(x(t) + \pi^{e\prime}(x,t)\,\Delta t + \sigma\Delta w(t))\},$$

where we neglected terms of $o(\Delta t)$. The rest of the proof is based on a comparison of the expected values in (A2) and (A3). A Taylor expansion of (A3) will involve $h'(x)$ and $h''(x)$ terms, that can be partially integrated to convert them to expressions in $h(x)$. Since $h(\cdot)$ is arbitrary, one can equate equivalent parts of the expected values in (A2) and (A3), which yields the Fokker-Planck equation in the limit as Δt goes to zero.

Let $g(y)$ be the density of $\sigma\Delta w(t)$, i.e. a normal density with mean zero and variance $\sigma^2\Delta t$. The expectation in (A3) can be written as an integral over the relevant densities:

(A4) $$E\{h(x(t+\Delta t))\} = \int_{-\infty}^{\infty}\int_{\underline{x}}^{\overline{x}} h(x + \pi^{e\prime}(x,t)\Delta t + \sigma y)\, f(x,t)\, g(y)\, dx\, dy.$$

A Taylor expansion of the right side of (A4) yields:

$$\int_{-\infty}^{\infty}\int_{\underline{x}}^{\overline{x}} \{ h(x) + h'(x)\,[\pi^{e\prime}(x,t)\,\Delta t + \sigma y] + \frac{1}{2} h''(x)\,[\pi^{e\prime}(x,t)\,\Delta t + \sigma y]^2 + \cdots \}\, f(x,t) g(y)\, dx\, dy ,$$

where the dots indicate terms of o(Δt). Integration over y eliminates the terms that are linear in y, since it has mean zero. In addition, the expected value of y^2 is $\sigma^2 \Delta t$, so the result of expanding and integrating the above expression is:

$$\int_{\underline{x}}^{\overline{x}} h(x)\, f(x,t)\, dx \;+\; \Delta t \int_{\underline{x}}^{\overline{x}} h'(x)\, \pi^{e\prime}(x,t)\, f(x,t)\, dx \;+\; \Delta t\, \frac{\sigma^2}{2} \int_{\underline{x}}^{\overline{x}} h''(x)\, f(x,t)\, dx \;+\; o\,(\Delta t) .$$

The integrals containing the h' and h'' term can be integrated by parts to obtain integrals in $h(x)$:

(A5) $$\int_{\underline{x}}^{\overline{x}} h(x) f(x,t)\, dx - \Delta t \int_{\underline{x}}^{\overline{x}} h(x) (\pi^{e\prime}(x,t)\, f(x,t))'\, dx + \Delta t\, \frac{\sigma^2}{2} \int_{\underline{x}}^{\overline{x}} h(x) f''(x,t)\, dx ,$$

where a prime indicates a partial derivative with respect to x, and we used the fact that h and its derivative vanish at the boundaries. Since (A5) is an approximation for (A2) when Δt is small, take their difference to obtain:

(A6) $$\int_{\underline{x}}^{\overline{x}} h(x) \big[f(x, t + \Delta t) - f(x,t) \big]\, dx = \Delta t \int_{\underline{x}}^{\overline{x}} h(x) \big[-(\pi^{e\prime}(x,t)\, f(x,t))' + (\sigma^2/2) f''(x,t) \big]\, dx .$$

The terms in square brackets on each side must be equal at all values of x, since the choice of the $h(x)$ function is arbitrary. Dividing both sides by Δt, taking the limit $\Delta t \to 0$ to obtain the time derivative of $f(x,t)$, and equating the terms in square brackets yields:

(A7) $$\frac{\partial f(x,t)}{\partial t} \;=\; -(\pi^{e\prime}(x,t)\, f(x,t))' + \frac{\sigma^2}{2} f''(x,t) .$$

Since the primes indicate partial derivatives with respect to x, we can integrate both sides of (A7) with respect to x to obtain the Fokker-Planck equation in (3).

APPENDIX B: DERIVATION OF EQUATION (10)

The Liapunov function in (9) depends on time only through the density functions, since the x's are variables of integration. Hence the time derivative is:

$$\text{(B1)} \quad \frac{dL}{dt} = \sum_{i=1}^{n} \int_{\underline{x}}^{\bar{x}} .. \int_{\underline{x}}^{\bar{x}} V(x_1,..,x_n) \prod_{j\neq i} f_j(x_j,t) \frac{\partial f_i(x_i,t)}{\partial t} dx_1..dx_n - \sum_{i=1}^{n} \mu_i \int_{\underline{x}}^{\bar{x}} (1 + \log(f_i(x_i,t))) \frac{\partial f_i(x_i,t)}{\partial t} dx_i .$$

The next step is to integrate each of the expressions in the sums in (B1) by parts. First note that $\partial f_i/\partial t = \partial^2 F_i/\partial t \partial x_i$ and that the anti-derivative of this expression is $\partial F_i/\partial t$. Moreover, the boundary terms that result from partial integration vanish because $F_i(0,t) = 0$ and $F_i(\bar{x},t) = 1$ for all t, i.e. $\partial F_i/\partial t = 0$ at both boundaries. It follows that partial integration of (B1) yields:

$$\text{(B2)} \quad \frac{dL}{dt} = - \sum_{i=1}^{n} \int_{\underline{x}}^{\bar{x}} .. \int_{\underline{x}}^{\bar{x}} \frac{\partial V(x_1,..,x_n)}{\partial x_i} \prod_{j\neq i} f_j(x_j,t) \frac{\partial F_i(x_i,t)}{\partial t} dx_1..dx_n + \sum_{i=1}^{n} \mu_i \int_{\underline{x}}^{\bar{x}} \frac{f_i^{\prime}(x_i,t)}{f_i(x_i,t)} \frac{\partial F_i(x_i,t)}{\partial t} dx_i .$$

Equation (8) can be used to replace $\int \partial V/\partial x_i \, dF_{-i}$ with $\pi^{e\,\prime}_i$, and then the integrals in (B2) can be combined as:

$$\text{(B3)} \quad \frac{dL}{dt} = \sum_{i=1}^{n} \int_{\underline{x}}^{\bar{x}} \{- \pi^{e\,\prime}_i(x_i,t) + \mu_i \frac{f_i^{\prime}(x_i,t)}{f_i(x_i,t)} \} \frac{\partial F_i(x_i,t)}{\partial t} dx_i = \sum_{i=1}^{n} \int_{\underline{x}}^{\bar{x}} \frac{(\partial F_i(x_i,t)/\partial t)^2}{f_i(x_i,t)} dx_i ,$$

where the final equation follows from (3). Note that the right side of (B3) is strictly positive unless $\partial F_i/\partial t = 0$ for $i = 1,..,n$, i.e., when the logit conditions in (4) are satisfied.

References

Anderson, Simon P., André de Palma, and Jacques-François Thisse (1992) *Discrete Choice Theory of Product Differentiation*, Cambridge: MIT Press.

Anderson, Simon P., Jacob K. Goeree, and Charles A. Holt (1998a) "Rent Seeking with Bounded Rationality: An Analysis of the All-Pay Auction," *Journal of Political Economy*, August, *106*(4), 828-853.

Anderson, Simon P., Jacob K. Goeree, and Charles A. Holt (1998b) "A Theoretical Analysis of Altruism and Decision Error in Public Goods Games," *Journal of Public Economics*, November, *70*, 297-323.

Anderson, Simon P., Jacob K. Goeree, and Charles A. Holt (1999) "Minimum-Effort Coordination Games: An Equilibrium Analysis of Bounded Rationality," University of Virginia, Thomas Jefferson Center for Political Economy, Working Paper 273.

Arrow, Kenneth J., and Leonid Hurwicz (1960) "Stability of the Gradient Process in *n*-Person Games," *Journal of the Society of Industrial and Applied Mathematics*, *8*, 280-294.

Binmore, Ken, Larry Samuelson, and Richard Vaughan (1995) "Musical Chairs: Modeling Noisy Evolution," *Games and Economic Behavior*, *11*, 1-35.

Binmore, Ken, and Larry Samuelson (1997) "Muddling Through: Noisy Equilibrium Selection," *Journal of Economic Theory*, *74*, 235-265.

Binmore, Ken, and Larry Samuelson (1999) "Evolutionary Drift and Equilibrium Selection," *Review of Economic Studies*, *66,* 363-393.

Bryant, John (1983) "A Simple Rational Expectations Keynes-Type Model," *Quarterly Journal of Economics*, *98*, 525-528.

Capra, C. Monica, Jacob K. Goeree, Rosario Gomez, and Charles A. Holt (1999) "Anomalous Behavior in a Traveler's Dilemma?" *American Economic Review*, *89*(3), 678-690.

Chen, Hsiao-Chi, James W. Friedman, and Jacques-François Thisse (1997) "Boundedly Rational Nash Equilibrium: A Probabilistic Choice Approach," *Games and Economic Behavior*, *18*, 32-54.

Cyert, Richard M., and James G. March (1963) *A Behavioral Theory of the Firm*, Englewood Cliffs, NJ: Prentice-Hall Inc.

Einstein, Albert (1905) *Ann. Physik* (4) *17*, 549.

Erev, Ido, and Alvin E. Roth (1998) "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *American Economic Review*, *88*(4), 848-881.

Foster, Dean, and Peyton Young (1990) "Stochastic Evolutionary Game Dynamics," *Theoretical Population Biology*, *38*, 219-232.

Friedman, Daniel, and Joel Yellin (1997) "Evolving Landscapes for Population Games," University of California, Santa Cruz, draft.

Fudenberg, D., and C. Harris (1992) "Evolutionary Dynamics with Aggregate Shocks," *Journal of Economic Theory*, *57*(2), 420-441.

Gihman, I. I., and A. V. Skorohod (1972) *Stochastic Differential Equations*, Berlin: Springer-Verlag.

Goeree, Jacob K. and Charles A. Holt (1999) "An Experimental Study of Costly Coordination," working paper, University of Virginia.

Kandori, Michihiro (1997) "Evolutionary Game Theory in Economics," in D. Kreps and K. Wallis, eds., *Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress, Vol. 1,* Cambridge, U.K.: Cambridge University Press, 243-277.

Kandori, Michihiro, George Mailath, and Rafael Rob (1993) "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, *61*(1), 29-56.

Kolmogorov, A. N. (1931) *Math. Annalen 104*, 415.

Lopez, Gladys (1995) *Quantal Response Equilibria for Models of Price Competition*, unpublished Ph.D. dissertation, University of Virginia.

Luce, R. Duncan (1959) *Individual Choice Behavior*, New York: John Wiley & Sons.

McKelvey, Richard D., and Thomas R. Palfrey (1995) "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, *10*, 6-38.

Monderer, Dov, and Lloyd S. Shapley (1996) "Potential Games," *Games and Economic Behavior*, *14*, 124-143.

Ochs, Jack (1995) "Games with Unique Mixed Strategy Equilibria: An Experimental Study," *Games and Economic Behavior*, *10*, 202-217.

Rhode, Paul, and Mark Stegeman (1997) "Non-Nash Equilibria of Darwinian Dynamics (with Applications to Duopoly)," unpublished draft, Virginia Tech.

Rosenthal, Robert W. (1973) "A Class of Games Possessing Pure-Strategy Nash Equilibria," *International Journal of Game Theory*, *2*, 65-67.

Rosenthal, Robert W. (1989) "A Bounded Rationality Approach to the Study of Noncooperative Games," *International Journal of Game Theory*, *18*, 273-292.

Roth, Alvin, and Ido Erev (1995) "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games and Economic Behavior*, *8*, 164-212.

Selten, Reinhard (1975) "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Form Games," *International Journal of Game Theory*, *4*, 25-55.

Selten, Reinhard, and Joachim Buchta (1994) "Experimental Sealed Bid First Price Auctions with Directly Observed Bid Functions," University of Bonn Discussion Paper B-270.

Simon, Herbert A. (1957) *Models of Man*, New York: John Wiley & Sons.

Slade, Margaret E. (1994) "What Does an Oligopoly Maximize?," *Journal of Industrial Economics*, *58*, 45-61.

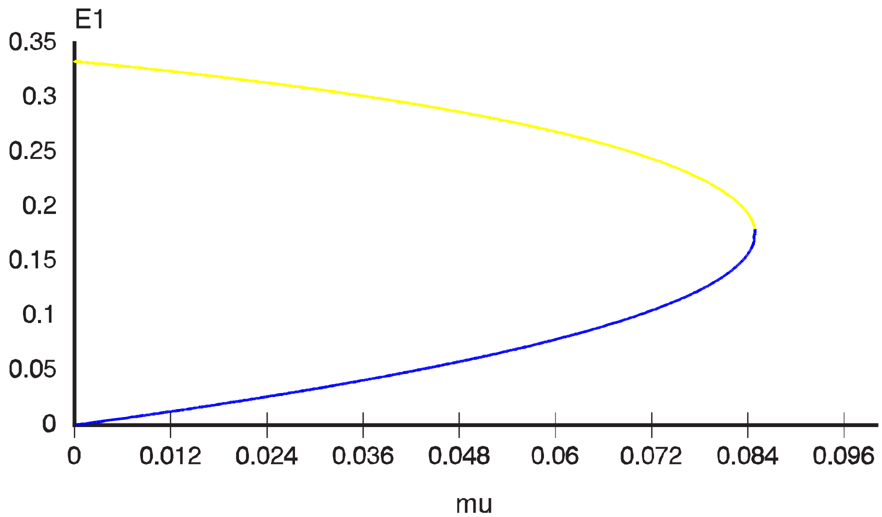
Smith, Vernon L. (1997) "Monetary Rewards and Decision Cost in Experimental Economics: An Extension," working paper University of Arizona.

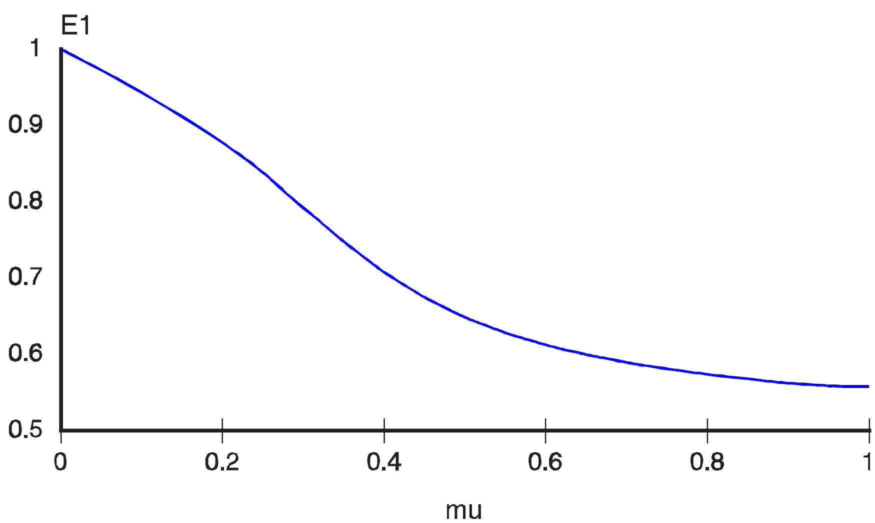
Smith, Vernon L., and James M. Walker (1993) "Monetary Rewards and Decision Cost in Experimental Economics," *Economic Inquiry*, *31*, 245-261.

Smoller, Joel (1994) *Shock Waves and Reaction-Diffusion Equations*, Berlin: Springer-Verlag.

Vega-Redondo, Fernando (1997) "The Evolution of Walrasian Behavior," *Econometrica*, *65*, 375-384.

Young, Peyton (1993) "The Evolution of Conventions," *Econometrica*, *61*, 57-84.





Figures 1a and 1b. The three different solutions, E_1^*, as functions of μ.
The dark lines correspond to stable logit equilibria and the light line to the unstable logit equilibrium.

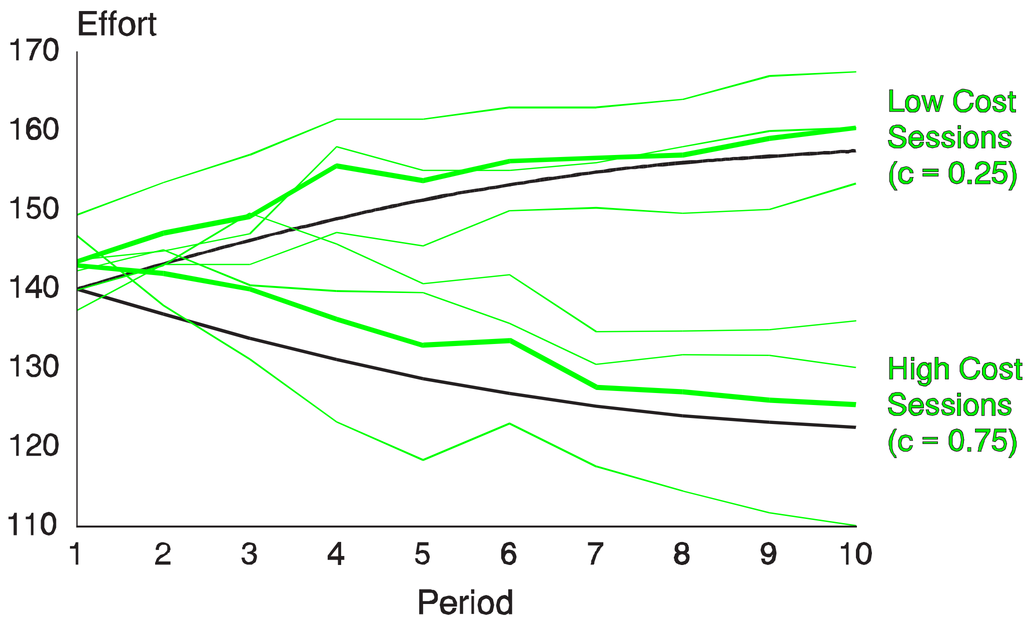


Figure 2. Coordination Game: Average Effort Decisions by Period
Key: Thin light lines are session averages, thick light lines are treatment averages, dark lines are predictions of evolutionary model.