

Learning and Belief Based Trade¹

First Version: October 31, 1994

This Version: September 13, 2005

Drew Fudenberg

David K. Levine²

Abstract: We use the theory of learning in games to show that no-trade results do not require that gains from trade are common knowledge nor that play is a Nash equilibrium.

¹ We thank Felipe Zurita for comments and encouragement. We are grateful to NSF grants SES-01-12018, SES-03-14713, and SES-04-26199 for financial support.

² Department of Economics Harvard University, and UCLA/Federal Reserve Bank of Minneapolis.

1. Introduction

The idea of speculation as trading based on information differences is a widespread one both inside and outside of economics. Such phenomenon as betting on horse races, not to speak of speculation in the stock market, are difficult to imagine in a world in which everyone has identical beliefs. Indeed, authors such as Hirshleifer [1975] have argued that the very idea of speculation is meaningless unless there are differences in beliefs. Yet the idea of speculation as information based trading runs quickly afoul of various no-trade theorems. The simplest such result is that if agents are risk averse and have a common prior, and the initial allocation is Pareto-optimal, then in a Nash equilibrium there must be no trade. This follows from the fact that if there were an equilibrium with trade, each agent would at least weakly improve his utility, contradicting the assumption that the initial allocation was optimal. Kreps [1977] and Tirole [1982] prove extensions of this result to rational expectations equilibria with risk-neutral traders. Milgrom and Stokey [1982] show that the assumption of Nash equilibrium can be replaced by the assumption that it is common knowledge that all players have prefer the proposed allocation to the initial one. Thus, either Nash equilibrium or common knowledge of agreement to trade, along with a common prior and risk averse agents, implies that there cannot be trade solely on the basis of differences in beliefs.

From the viewpoint of non-equilibrium learning theory, though, both the assumption of a common prior on Nature's moves and the assumption of a Nash equilibrium (that is, a common belief on players' strategies) may be too strong. In the theory of learning in games, the assumption of exogenous knowledge about the distribution of moves is replaced with the idea that players acquire knowledge through learning. Thus common beliefs about either Nature's moves or the play of other players may or may not arise, depending on the environment. Consequently, the steady states of standard learning processes correspond not to the Nash equilibria but to the larger class of self-confirming equilibria that we introduced in Fudenberg and Levine [1993].³ In simultaneous-move complete-information games, if players observe the profiles of

³ See also Battigalli (1987), Fudenberg and Kreps [1988,1995], and Rubinstein and Wolinsky (1994).

actions played in each round, the self-confirming equilibria coincide with the set of Nash equilibria of the game.⁴ By contrast, as argued in Dekel et al [2004], in games of incomplete information, if players begin with inconsistent priors there are broad classes of games in which the self-confirming equilibria (and hence the steady states of standard learning processes) do not coincide with the Nash equilibria.

Nevertheless, there are important classes of incomplete-information games where the steady states of learning models do coincide with the Nash equilibria. For example, Dekel at al showed that this is the case when players observe one another's actions and there are independent private values. In the trading games that we consider here, it is not plausible that all agents observe one another's actions. Never-the-less the equivalence of Nash and self-confirming equilibria still holds, because the games have the property that each agent knows his own utility function and hence knows the payoff he will get from not trading. As we show, it is this “known security level” property that underlies the no-trade results.

In addition, we show that not even self-confirming equilibrium is needed for the no-trade conclusion. Specifically, while the steady states of standard learning processes must be self-confirming equilibrium, there is no guarantee that even well-behaved learning procedures necessarily converge to a steady state. For this reason, we also examine the notion of “marginal best response distributions” introduced by Fudenberg and Levine [1995]. If all players follow learning procedures that are moderately rational, then the joint distribution of play must at least converge to the set of these distributions. In both cases, we show that the no-trade theorem applies. The intuition is simple: if agents are risk averse, the only possibility of trade is based on information differences, and trade takes place, then there must be an agent who would do better not to trade. A player need not be a terribly clever learner to discover that he is doing poorly, all that is required is that he know the utility he would get by not trading. So in the long run, all trade must stop.

⁴ We will not formally model the dynamics of learning, but we have in mind “belief-based” processes in which players base their actions on their beliefs about opponents' play. Fudenberg and Kreps [1995] and Fudenberg and Levine [1993b] showed that the long-run outcomes of such processes correspond to the self-confirming equilibria; they considered general extensive form games and supposed that the signals corresponded to the terminal nodes of the game.

We should emphasize that we are not claiming that in practice there is no trade based on information differences. Rather we are claiming that there must be some other underlying reason for trade, such as portfolio balance, joy of betting on the horses, noise traders who are not rational, before it becomes possible to trade based on information differences. See for example, Zurita (2004) for a model in which underlying gains to balancing portfolios allows trading based on information differences in a model with common knowledge.

2. The Model

There are n traders $i = 1, \dots, n$. Each trader has finitely many possible types, with trader i 's type denoted θ_i . The profile of types θ is called the state. There are m goods, so the consumption bundle consumed by trader i is $x_i \in \mathbb{R}^m$. Trader i 's endowment is $\bar{x}_i(\theta_i)$ depending his type; note that endowments do not depend on the types of other players.⁵ Utility $u_i(x_i, \theta)$ is Von Neumann-Morgenstern and comes from the consumption of goods and may also depend on the state. We assume strict risk aversion:

Assumption 1: $u_i(\cdot, \theta)$ is strictly concave.

The final allocation is determined from endowments by a finite simultaneous-move game.⁶ Each trader i observes his own type θ_i then chooses an action $a_i \in A_i$ from a finite set. Mixed actions are denoted by α_i . The final allocation is given by $x_i = f_i(\theta, a)$, and is assumed to be socially feasible

Assumption 2: $\sum_i f_i(\theta, a) = \sum_i \bar{x}_i(\theta)$.

Each trader has the option of not trading, denoted by a_i^0 .

Assumption 3: $f_i(\theta, a_i^0, a_{-i}) = \bar{x}_i(\theta_i)$.

If learning by traders is to be possibly, the economy must meet repeatedly. We assume that each time the economy meets the state is determined by an independent draw

⁵ Since a player's type is supposed to encapsulate all private information available to him, and since we presume players know their own endowments before beginning trading, a player's own type should determine his endowment.

⁶ Or the game may be an elaborate dynamic game, in which case our simultaneous move game represents the strategic form.

from a fixed (objective) probability distribution ρ that is unknown to the traders. Traders do not necessarily observe the realized value of θ , so if they start out with incorrect beliefs about ρ , it is not obvious that they will learn the true distribution.

Since we are interested only in trade due to differences in beliefs, we must rule out other reasons for trade. Consequently we assume that the endowment is *ex ante* Pareto efficient; that is

Assumption 4: *There exist weights $\lambda_i > 0$ such that if $\sum_i x_i(\theta) \leq \sum_i \bar{x}_i(\theta_i)$*

$$\sum_i \lambda_i \sum_{\theta} u_i(\bar{x}_i(\theta_i), \theta) \rho(\theta) \geq \sum_i \lambda_i \sum_{\theta} u_i(x_i(\theta), \theta) \rho(\theta).$$

We consider two equilibrium concepts that relax Nash equilibrium. The key components of self-confirming (and Nash) equilibrium are each player i 's *beliefs* about Nature's move, her *strategy*, and her *conjecture* about the strategies used by her opponents. Player i 's beliefs, denoted by $\hat{\mu}^i$, are a point in the space $\Delta(\Theta)$ of distributions over Nature's move, and her strategy is a map $\sigma_i : \Theta_i \rightarrow \Delta(A_i)$. The space of all such strategies is denoted Σ_i , and the player's conjectures about opponents' play are assumed to be a $\hat{\sigma}_{-i} \in \times_{-i} \Sigma_j$, that is, a strategy profile of i 's opponents. The notation $\hat{\mu}^i(\cdot | \theta_i)$ refers to the conditional distribution derived from $\hat{\mu}^i$, conditional on the private type θ_i , while $\sigma_i(a_i | \theta_i)$ denotes the probability that $\sigma_i(\theta_i)$ assigns to a_i .

Of course, what players might learn from repeated play depends on what they observe at the end of each round of play. To model this, the equilibrium concepts suppose that after each play of the game, players receive private signals $y_i = y_i(a, \theta)$. As the notation indicates, these signals are a deterministic function of a and θ . We assume that each player observes her own private signal y_i , along with her own action and own type, so these are their only sources of information about Nature's and their opponents' moves.⁷

Our equilibrium concept is a variation on the type of self -confirming equilibrium defined in Fudenberg and Levine [1993] and Dekel el al [2004]

⁷ We consider the case in which knowledge of opponents' play comes *only* from learning by observation and updating, and not from deduction based on opponents' rationality, so we do not require that players know their opponents' utility functions or beliefs. Rubinstein and Wolinsky (1994), Battigalli and Guaitoli (1997) and Dekel, Fudenberg and Levine (1999) present solution concepts based on steady states in which players do make deductions based on rationality of the other players.

Definition 1: A strategy profile σ is an ε -self-confirming equilibrium with conjectures $\hat{\sigma}_{-i}$ and beliefs $\hat{\mu}_i$ iff for each player i ,

$$(i) \quad \text{for all } \theta_i, \hat{\mu}_i(\cdot | \theta_i) = \rho(\theta_i)$$

and for any pair θ_i, \hat{a}_i such that $\hat{\mu}^i(\theta_i) \cdot \sigma_i(\hat{a}_i | \theta_i) > 0$ both the following conditions are satisfied

$$(ii) \quad \sum_{a_{-i}, \theta_{-i}} u_i(\hat{a}_i, a_{-i}, \theta_i, \theta_{-i}) \hat{\mu}^i(\theta_{-i} | \theta_i) \hat{\sigma}_{-i}(a_{-i} | \theta_{-i}) + \varepsilon \geq \max_{a_i} \sum_{a_{-i}, \theta_{-i}} u_i(a_i, a_{-i}, \theta_i, \theta_{-i}) \hat{\mu}^i(\theta_{-i} | \theta_i) \hat{\sigma}_{-i}(a_{-i} | \theta_{-i}),$$

and for any \bar{y}_i in the range of y_i

$$(iii) \quad \sum_{\{a_{-i}, \theta_{-i}: y_i(\hat{a}_i, a_{-i}, \theta_i, \theta_{-i}) = \bar{y}_i\}} \hat{\mu}^i(\theta_{-i} | \theta_i) \hat{\sigma}_{-i}(a_{-i} | \theta_{-i}) \\ = \sum_{\{a_{-i}, \theta_{-i}: y_i(\hat{a}_i, a_{-i}, \theta_i, \theta_{-i}) = \bar{y}_i\}} \rho(\theta_{-i} | \theta_i) \sigma_{-i}(a_{-i} | \theta_{-i}).$$

We say that σ is a *self-confirming equilibrium* if there is some collection $(\hat{\mu}^i, \hat{\sigma}_{-i})_{i \in I}$ such that (i), (ii) and (iii) are satisfied.⁸

Our key assumption is that each trader observes enough information to determine her utility from the no-trade action. For example, if the endowment represents some complicated stock portfolio, and the trader engages in a complicated series of trades, if the trader does not observe the prices of stocks that were held in positive quantities in her endowment, but were traded away, then she may not be able to determine the utility of not having traded at all.

Assumption 5: (Known Security Levels):

$$u_i(\bar{x}_i(\theta_i), \theta) \text{ depends only on } \theta_i, y_i. \quad ^9$$

⁸ It is appropriate to have a single $\hat{\sigma}_{-i}$ for each player i in the definition because we assume that there is a single agent in each player role. This is called the “unitary” version of self-confirming equilibria; when we consider large populations and matching in Section 4 we allow for heterogeneous beliefs.

Note that i ’s beliefs about opponents’ play take the form of a strategy profile as opposed to a probability distribution over strategy profiles. The complications that arise due to correlations in conjectures are discussed in Fudenberg and Kreps (1988) and Fudenberg and Levine (1993a); we simplify by ignoring them here. Given this restriction, there is no further loss of generality in taking beliefs to be point conjectures. Battigalli (1987) defined a similar concept to the one above, as did Kalai and Lehrer (1993).

⁹ That is, if for some a , $\theta_i = \theta'_i$ and $y(a, \theta) = y(a, \theta')$ then $u_i(\bar{x}_i(\theta_i), \theta) = u_i(\bar{x}_i(\theta_i), \theta')$.

This immediately implies the following sufficient condition for an ε -self-confirming equilibrium, which underlies our first result:

Lemma 1: If a profile of mixed actions α is a ε -self-confirming equilibrium, then $\alpha_i(a_i) > 0$ implies

$$\sum_{\theta, a_{-i}} u_i(f_i(\theta, a_i, a_{-i}), \theta) \alpha_{-i}(a_{-i}) \rho(\theta) + \varepsilon \geq \sum_{\theta, a_{-i}} u_i(\bar{x}_i(\theta_i), \theta) \rho(\theta).$$

This says that the expected utility from the action actually taken gives within ε of the utility from the endowment.

The idea of self-confirming equilibrium is that we do not require that players beliefs about what they did not see opponents do be correct. However, there is no general theorem guaranteeing the global convergence of a sensible class of learning procedures to a self-confirming equilibrium. This leads us our second “equilibrium” notion, a variation the idea of a marginal best response distribution introduced in Fudenberg and Levine [1995].

Definition 2: A joint distribution μ over pure action profiles is an ε -marginal best response distribution if

$$\sum_{\theta, a} u_i(f_i(\theta, a), \theta) \mu(a) \rho(\theta) + \varepsilon \geq \max_{a_i} \sum_{\theta, a_{-i}} u_i(f_i(\theta, a), \theta) \mu_{-i}(a_{-i}) \rho(\theta)$$

where μ_{-i} is the marginal over all actions by players other than player i .

This says that the utility that player i actually gets is at least within ε of the most he could get against the marginal distribution of opponents actions; that is, correlations are ignored. The significance of this notion is that there exist a broad class of approximately universally consistent learning strategies and if players use such strategies, asymptotic play will be close to an approximate marginal best response distribution even if it never converges. From the definition, it appears that it is necessary that players observe their opponents actions. However, Fudenberg and Levine [1998] and Hart and Mas-Colell [2001] show that there are learning procedures that give this result when players observe only their own action and own utility. In particular, Assumption 5 need not be satisfied for these learning procedures to work. In other words, marginal best

response distributions capture long-run non-equilibrium play under the very weak assumption that players know their past actions and payoffs.

3. The Result

Our conclusion is that in the limit as $\varepsilon \rightarrow 0$ for either self-confirming or marginal best-response there is convergence to no-trade. The idea is that under our assumption of strict concavity of the utility functions, any probability distribution over socially feasible allocations that Pareto dominates the endowment must involve no-trade. As $\varepsilon \rightarrow 0$ both ε -self confirming equilibria and ε -marginal best response distributions give each trader at least the utility that they could get from their endowment, and so the limiting allocation must weakly Pareto dominate the endowment.

First we show that socially feasible allocations that weakly Pareto dominate the endowment involve no trade.

Lemma 2: *If μ is a joint probability distribution over actions such that*

$$\sum_{\theta} u_i(f_i(\theta, a), \theta) \mu(a) \rho(\theta) \geq \sum_{\theta} u_i(\bar{x}_i(\theta_i), \theta) \rho(\theta)$$

then

$$\mu(\{a \mid f_i(\theta, a) \neq \bar{x}_i(\theta_i)\}) = 0$$

Proof: Since \bar{x} is *ex ante* Pareto efficient, and the u_i 's are strictly concave the only socially feasible allocation that weakly Pareto dominates \bar{x} is \bar{x} itself. Consequently, any probability distribution over socially feasible allocations that weakly Pareto dominates \bar{x} must choose \bar{x} with probability one. The result now follows from the fact that $f(\cdot, a)$ is socially feasible. ■

Our main results now say that in the limit both self-confirming equilibria and marginal best response equilibria involve no trade. In the case of self-confirming, the fact that each trader gets at least the endowment utility in the limit follows from upper hemi-continuity of the ε -equilibrium correspondence and the fact that traders know the endowment utility.

Theorem 1: *If α^ε is a sequence of ε -self-confirming equilibrium then*

$$\lim_{\varepsilon \rightarrow 0} \alpha^\varepsilon(\{a \mid f_i(\theta, a) \neq \bar{x}_i(\theta_i)\}) = 0,$$

where $\alpha^\varepsilon(\cdot)$ is the joint probability distribution over actions induced by α^ε .

Proof: If not there is an action a , a state θ and a subsequence such that $\alpha^\varepsilon(\cdot) \rightarrow \alpha(\cdot)$, $\alpha(a) > 0$ and $f_i(\theta, a) \neq \bar{x}_i(\theta)$. Since $\alpha^\varepsilon(\cdot) \rightarrow \alpha(\cdot)$ and $\varepsilon \rightarrow 0$, it follows from continuity and the Lemma 1 that

$$\sum_{\theta, a_{-i}} u_i(f_i(\theta, a_i, a_{-i}), \theta) \alpha_{-i}(a_{-i}) \rho(\theta) \geq \sum_{\theta, a_{-i}} u_i(\bar{x}_i(\theta), \theta) \rho(\theta).$$

This contradicts Lemma 2. ■

In the case of ε -marginal best response distribution the fact that each trader gets at least the endowment utility in the limit follows from the fact that a marginal best response distribution gives each player at least the minmax.

Theorem 2: *If μ^ε is a sequence of ε -marginal best response distributions then*

$$\lim_{\varepsilon \rightarrow 0} \mu^\varepsilon(\{a \mid f_i(\theta, a) \neq \bar{x}_i(\theta)\}) = 0.$$

Proof: As in the proof of Theorem 1, we may use the definition of an ε -marginal best response distribution to conclude that $\mu^\varepsilon \rightarrow \mu$ with

$$\sum_{\theta, a} u_i(f_i(\theta, a), \theta) \mu(a) \rho(\theta) \geq \max_{a_i} \sum_{\theta, a_{-i}} u_i(f_i(\theta, a), \theta) \mu_{-i}(a_{-i}) \rho(\theta).$$

Since

$$\begin{aligned} & \max_{a_i} \sum_{\theta, a_{-i}} u_i(f_i(\theta, a), \theta) \mu_{-i}(a_{-i}) \rho(\theta) \\ & \geq \sum_{\theta, a_{-i}} u_i(f_i(\theta, a_i^0, a_{-i}), \theta) \mu_{-i}(a_{-i}) \rho(\theta) \\ & = \sum_{\theta} u_i(\bar{x}_i, \theta) \rho(\theta) \end{aligned}$$

this again contradicts Lemma 2. ■

References

- Battigalli, Pierpaolo (1987) “Comportamento Razionale Ed Equilibrio Nei Giochi E Nelle Situazioni Sociali,” unpublished undergraduate dissertation, Bocconi University, Milano.
- Battigalli, Pierpaolo and D. Guaitoli [1997] “Conjectural Equilibria and Rationalizability in a Game with Incomplete Information,” in Decision, Games and Markets, P. Battigalli, A. Montesano and F. Panunzi, Eds., Dordrecht: Kluwer Academic Publishers.
- Dekel, Eddie, Drew Fudenberg and David K. Levine (1999) “Payoff Information and Self-Confirming Equilibrium,” *Journal of Economic Theory*; 89(2), 165-85.
- Dekel, Eddie, Drew Fudenberg, and David K. Levine (2004), “Learning to Play Bayesian Games,” *Games and Economic Behavior*, **46**: 282-303.
- Fudenberg, Drew and David Kreps (1988) “A Theory of Learning, Experimentation, and Equilibrium in Games,” unpublished mimeo.
- Fudenberg, Drew and David Kreps (1995) “Learning in Extensive-Form Games. I. Self-Confirming Equilibria,” *Games and Economic Behavior*; 8(1), 20-55.
- Fudenberg, Drew and David K. Levine (1993), “Self-Confirming Equilibrium,” *Econometrica*, **61**: 523-546.
- Fudenberg, Drew and David K. Levine (1995), “Consistency and Cautious Fictitious Play,” *Journal of Economic Dynamics and Control*, **19**: 1065-1090.
- Fudenberg, Drew and David K. Levine (1998), *The Theory of Learning in Games*, MIT Press, Cambridge, MA.
- Hart, Sergiu and Andreu Mas-Colell (2001), “A General Class of Adaptive Strategies,” *Journal of Economic Theory*, **98**: 26-54.
- Hirshleifer, Jack (1975), “Speculation and Equilibrium: Information, Risk, and Markets,: *The Quarterly Journal of Economics*, **89**: 519-542.
- Kalai, Ehud and Ehud Lehrer (1993) “Rational Learning Leads to Nash Equilibrium,” *Econometrica*; 61(5), 1019-45.
- Milgrom, Paul and Nancy Stokey (1982), “Information, Trade and Common Knowledge,” *Journal of Economic Theory*, **26**: 17-27.
- Tirole, Jean (1982) “On the Possibility of Trade under Rational Expectations,” *Econometrica*, **50**: 1163-1182.

Zurita, Felipe (2004), “On the limits to Speculation in Centralized versus Decentralized Market Regimes,” *Journal of Financial Intermediation*, **13**: 378-408.