The Interplay between Analytics and Computation in the Study of Congestion Externalities: The Case of the El Farol Problem

EDUARDO ZAMBRANO

University of Notre Dame and Banco Central de Venezuela

Abstract

In this paper I study the El Farol problem, a deterministic, boundedly rational, multi-agent model of a resource subject to congestion externalities that was initially studied computationally by Arthur (1994). I represent the interaction as a game, compute the set of Nash equilibria in mixed strategies of this game, and show analytically how the method of inductive inference employed by the agents in Arthur's computer simulation leads the empirical distribution of aggregate attendance to be like those in the set of Nash equilibria of the game. This set contains only completely mixed strategy profiles, which explains why aggregate attendance appears random in the computer simulation even though its set-up is completely deterministic.

1. Introduction

In this paper I study a famous problem from the field of computational economics known as the El Farol Problem, which is a model of bounded

rationality posed by Brian Arthur (1994). In this problem agents employ a set of predictors to forecast the actions of others over time and decide whether to attend the "El Farol" bar based on their most accurate predictor of aggregate attendance at each period. When each boundedly rational agent forecasts that the bar is going to be too crowded, the agent decides not to go to the bar, and vice versa.

The problem has the feature that no forecasting rule of deterministic actions can be at the same time correct and available to all agents, and that the sequence of outcomes that one obtains through (deterministic) computer simulation of the problem looks more like a random process with a stationary mean than a deterministic function.

There are several reasons why the El Farol problem is worth studying. The first one is that the empirical distribution of aggregate behavior resulting from the system of *deterministic* forecasting and decision is realization-equivalent to the empirical distribution of aggregate behavior resulting from the mixed-strategy Nash equilibria of the underlying game, thereby providing a clear link between equilibria in randomized strategies and the long-run behavior of a collection of boundedly rational agents.

The second reason is that the bar in this problem is essentially a resource subject to congestion, therefore making the problem a stylized version of the central problems in public economics, such as those of traffic congestion and congestion on computer networks like the Internet.[1] As congestion of this kind is routinely experienced in modern economies, a deeper understanding of what transpires in computer simulations of the El Farol problem may help us develop methods for understanding how resources subject to congestion can be better managed in modern economies.

The third reason why the problem is worth studying is that it is regarded as paradigmatic of those in the complex adaptive systems literature, for it involves a fairly large number of intelligent and adaptive agents with local information whose "forecasts (...) act to create the world they are trying to forecast."[2] These systems include multi-agent systems (like the Internet), ecological systems, common pool resources, and financial markets. For this reason, the El Farol problem has generated a lot of attention from physicists and computer scientists. Indeed, the mathematician John Casti has asserted that "a decent mathematical formalism within which to describe and analyze

---

[1]While there has been some research in the area, as evidenced by the seminal work by Vickrey (1963, 1969), Rosenthal (1973), Arnott, de Palma, and Lindsey (1993), Varian and Mackie-Mason (1994), Shenker (1995), and Wolpert et al. (1999) much remains to be done. In the words of Ted Bergstrom: "In my opinion neither of these areas has received the attention from economists that is merited by its importance and interest. Both seems to me areas in which economic theory and econometrics are likely to be powerful tools." See http://www.econ.ucsb.edu/~tedb/econ230b.html.

[2]Casti (1996, p. 8).

the (. . .) El Farol problem would go a long way toward the creation of a workable scientific *theory* of (complex, adaptive) systems."[3]

Physicists and computer scientists work mainly in the context of a game inspired by the El Farol problem deemed the minority game—a game in which an odd number of players must join one of two groups. Each player wins when he or she joins the group with a minority of the players.

Interestingly, many of the computer scientists and physicists involved in this kind of research have had some resistance in employing the standard tools of mathematical economics and game theory to study them. The reason is that it has been argued that computational agents lack the kind of sophistication that players are supposed to have according to game theory and therefore game-theoretic predictions about what would arise in interactions among computer agents would not be of use.[4]

This paper shows that this need not be the case, as the kind of algorithms that define the manner in which computer agents interact over time can be given a game-theoretic interpretation in a way that makes the behavior of a collection of boundedly rational agents consistent with the equilibria of the underlying game.

This is what the research reported in this paper has done for the case of the El Farol problem. This suggests that the standard tools of game theory should be of tremendous utility in understanding the rapidly growing world of network-based computation and interaction.

To summarize, in this paper I study the El Farol problem, a deterministic, boundedly rational, multi-agent model of a resource subject to congestion that was initially studied computationally by Arthur (1994). I represent the interaction as a game, compute the set of Nash equilibria in mixed strategies of this game, and then show analytically how the method of inductive inference employed by the agents in Arthur's computer simulation leads the empirical distribution of aggregate attendance to be like those distributions in the set of Nash equilibria of the game. This set contains only completely mixed strategy profiles, which helps to explain why aggregate attendance appears random even though the set-up of Arthur's computer simulation is completely deterministic.

The ability to do this illustrates the nice interplay that exists between computational modeling and mathematical analysis. Both the El Farol problem and the Minority Game have been the subject of many computational models. In this paper, I show how we can use mathematics to explain what happens in these computational models. This should not be interpreted as

---

[3]Casti (1996, p. 9).

[4]Challet and Zhang (1997), for example, when introducing the minority game to the physics community, stated: "the approach traditionally used in economics is not convenient to generalise to include irrationality," and methods of statistical physics are used by these and other authors instead.

a demonstration that mathematics always trumps computation. On the contrary, the two are complementary. The computational models of Arthur and others open a new set of questions and generate phenomena that we would like to understand fully. The mathematics I present provides an explanation of some dimensions of those phenomena. They erase some of the mystery. In doing so, they enhance, not diminish, the discovery of the phenomena themselves.

The structure of the rest of this paper is the following. In Section 2, I present the details of the El Farol problem, in Section 3 I rewrite the El Farol problem in game-theoretic terms, and in Section 4 I provide analytic results that explain some characteristics of the empirical distribution of outcomes over time. I provide a discussion of related work in Section 5 and offer conclusions in Section 6.

2. The El Farol Problem

2.1. The El Farol Model

One hundred inhabitants of Santa Fe decide independently each week whether to go to the bar El Farol, where it is known that a good evening awaits if not too many people show up at the same time. Concretely, each inhabitant goes to the bar if she expects fewer than 60 to show up, and stays home otherwise. The only information available to every inhabitant at each week is the number of people that went to the bar in the previous weeks.

To decide whether or not to go to the bar each inhabitant needs a method to forecast attendance. However, the problem has the feature that no forecasting rule of deterministic attendance can be at the same time correct and available to all agents. To see why assume that the forecasting rule asserts that 60 or more individuals will show up at El Farol. Then nobody will go, thus invalidating the forecast. Similarly, if the forecasting rule asserts that less than 60 individuals will show up at the bar, then all will go, which again invalidates the forecast. To deal with this difficulty Arthur (1994) allows each agent to "form several predictors or hypotheses that map the past d weeks attendance figures into next week's."[5] Examples of possible predictors include the following:

- Predict next week's attendance to be the same as last week's;
- Predict an average of the last 4 weeks;
- Predict the same as 2 weeks ago.

At each period each individual uses the predictor that is "currently the most accurate (...) in her set"[6] (the so-called *active* predictor), and makes a decision as to whether to attend the bar or not based on it. Arthur (1994) has

---

[5]Arthur (1994, p. 409).
[6]Arthur (1994).

the agents compute the accuracy of the predictors as follows: If s^i is one of the predictors for agent i, its accuracy at the end of period t is computed as follows:

$$U_t(s^i) = \lambda U_{t-1}(s^i) + (1 - \lambda)|s^i(d(h_{t-1})) - y_t|$$

where h_{t-1} is the history of attendance up to period $t - 1$, $d(h_{t-1}) \in D$, is the attendance for the last d weeks (where d is a fixed number), D is the set of all possible attendance profiles for the last d weeks, $s^i(d(h_{t-1}))$ is the prediction of predictor s^i for period t given history h_{t-1}, y_t is the actual attendance for period t, and λ is a number strictly between zero and one. In other words, Arthur (1994) computes the accuracy of each predictor at a given period as a weighted average of the accuracy of the predictor in the previous period and the absolute difference between the predictor's last prediction and the actual attendance.[7]

In this set-up, the collection of decisions of the agents determines each week's attendance figure, which is reported to all agents in the next period. Then the figure is used by each agent to update the accuracies of her monitored predictors, and a new decision is made.

To analyze this model Arthur designs a computer experiment where he first creates a finite but fairly large family of possible predictors, of which a number of them are independently and identically distributed to each of the 100 agents. The forecasting rules given to the agents are diverse enough that at every period each agent has at least a rule that forecasts an attendance level below 60, and a rule that forecasts an attendance level of 60 or above. Then, given starting conditions and the fixed set of predictors available to each agent the overall dynamics is completely determined.

2.2. The Puzzle

When performing computer experiments of the "El Farol" problem Arthur (1994) obtains the following:

Fact 1. Mean attendance always converges to 60, and
Fact 2. On average, 40% of the active predictors are forecasting above 60 and 60% below 60.

Casti in turn reports the third interesting feature of the problem, namely, that the sequence of outcomes looks more like a random process than a deterministic function, even though the set-up of the problem is entirely deterministic.[8] This, according to Casti, leads to the following conjectures:

---

[7]I am grateful to Bruce Edmonds for kindly providing me with a copy of Arthur's original code.
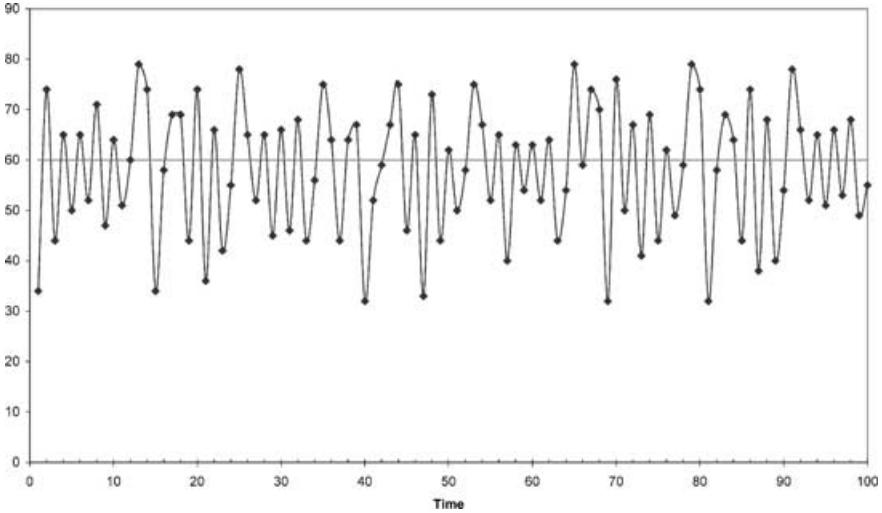[8]Casti (1996, p. 8).

Figure 1: One hundred weeks of aggregate attendance according to the El Farol simulation

> *Conjecture 1.* The average number of people who actually go to the bar converges to the threshold value as the number of periods becomes large.
>
> *Conjecture 2.* The time series on attendance levels is a deterministic random process, i.e., it is "chaotic."[9]

There is a major obstacle to a resolution to these conjectures: Casti asserts "there currently exists no mathematical formalism within which to even meaningfully phrase these questions."[10]

2.3. The Statistics

One of the messages of the present paper is that game theory is a useful formalism for thinking about problems like the one studied by Arthur (1994). Before presenting the precise game-theoretic representation of the interaction underlying the El Farol problem I briefly analyze statistically the outcome of a typical run from Arthur's code.

Figure 1 displays 100 weeks of aggregate attendance according to a run of Arthur's model. The mean attendance in the entire run (which consists of 10,000 "weeks") is 59.07 and the standard deviation is 12.8. The mean is remarkably stable early in the run, as Figure 2 shows. It is noteworthy that, while the mean is close to 60 (more precisely: within one to 60), it is not 60,

---

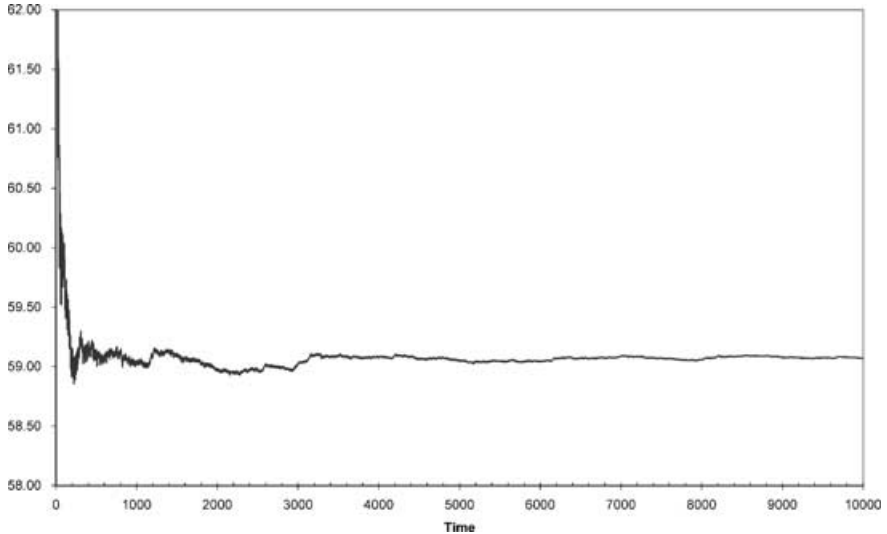[9]Casti (1996).

[10]Casti (1996, p. 9).

Figure 2: Average of aggregate attendance over time in the El Farol simulation

as Fact 1 and Conjecture 1 would suggest. The long-run *median*, however, is exactly 60. One could then rewrite Conjecture 1 as follows:

> *Conjecture 1′*. The median number of people who actually go to the bar converges to the threshold value as the number of periods becomes large. The average number of people converges to a number that is within one to the threshold.

Conjecture 2 suggests that a run obtained from Arthur's code is "as if" it was a realization of a stochastic process, even though Arthur's setup is entirely deterministic.

To verify this claim I ran a non-parametric test of whether in a typical run from Arthur's code the observations for aggregate attendance occur in random order. If aggregate attendance is positively serially correlated it will tend to remain above or below its median for several observations in a row. On the other hand, if aggregate attendance is negatively serially correlated observations above the median are likely to be followed by observations below the median. Therefore, the number of times the time series of aggregate attendance "crosses" the median cannot be either too high or too low under the random order hypothesis (Swed and Eisenhart 1943).

The number of times aggregate attendance crosses the median in our sample run of Arthur's code of 10,000 weeks is 7,901 while the expected value of crossings under the the null hypothesis of random order is about 5,000, numbers that are statistically different to any reasonable significance level (the test's z-statistic, which is approximately normal, was equal to 58.01).
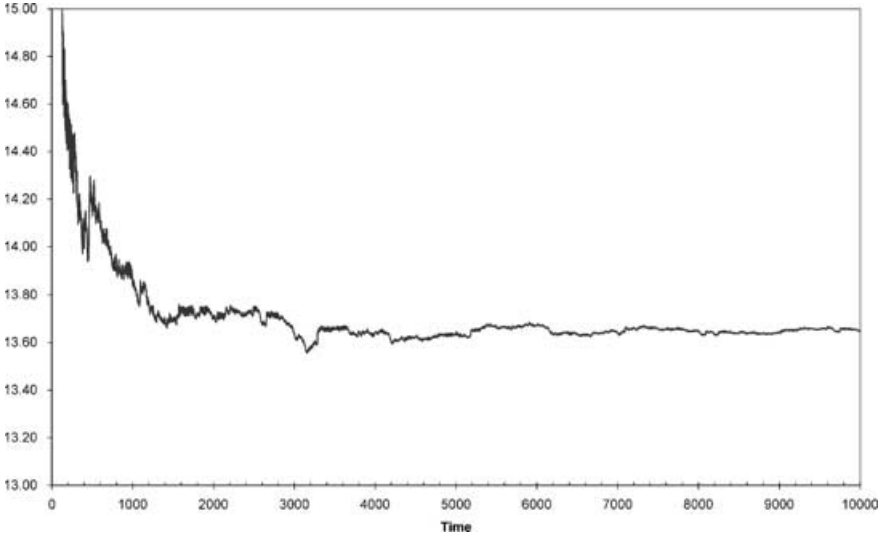
Figure 3: Standard deviation of aggregate attendance over time in the El Farol simulation

Despite this, persistent fluctuations in the data remain, as evidenced by the fact that the standard deviation of aggregate attendance over time is bounded away from zero (Figure 3). Hence, while Conjecture 2 does not seem to hold, the question remains as to why the persistent fluctuations in the data do not vanish over time. One could write this as follows:

Conjecture 2′. The standard deviation of aggregate attendance is stable over time and bounded away from zero.

3. Game Theory

3.1. The Prediction Game

A game $G = \langle N, (A_i), (u_i) \rangle$ consists of a set N of players and, for each player i, a finite *action set* A_i and a *payoff function* $u_i : A_i \times A_{-i} \to \Re$, where $A_{-i} := \Pi_{j \neq i} A_j$, is the Cartesian product of the action sets of every player but i.

One can view the "El Farol" model as a game where the decision that each player makes is one of either going or not going to the bar, subject to the player's forecast, and where the payoff of attending (respectively, not attending) when the bar is uncrowded (respectively, crowded) is greater than the payoff of attending (respectively, not attending) when the bar is crowded (respectively, uncrowded).[11] Call this the *attendance game*. A more useful

---

[11]In fact, this is how many in the literature represent the interaction. See the discussion of the related literature in Section 5.

representation, for the purpose of obtaining analytical results, is the following representation:

- The strategy set for each player is the set of integers from 0 to 100.
- The payoffs for each player are measured as the negative of the absolute difference between the player's choice of strategy, and the number of (other) players that chose a number smaller than 60.[12] The number of (other) players that chose a number smaller than 60 will be called the "aggregate attendance."

This description of the game underlying the "El Farol" model is called the *prediction game*. The interpretation is, of course, that each player is trying to predict how many other players chose, as their own prediction, a number below 60.

While this representation seems substantially different than the attendance game, it turns out that they can be made to be equivalent when (i) the game is repeated, (ii) the players are using Arthur's method of belief revision in the attendance game, and (iii) the players are using exponentially weighted beliefs in the prediction game. This will be explored formally in Theorem 3. Before investigating any of this, I consider some properties of the equilibria of the prediction game.

LEMMA 1: *The prediction game has no pure strategy equilibria.*

Proof: See the Appendix.

That the prediction game has no pure-strategy Nash equilibria is exactly what the statement that "no forecasting rule of deterministic actions can be at the same time correct and available to all agents" refers to. This is no proof that "there is no deductively rational solution—no correct expectational model [for this problem],"[13] however, because there are many forecasting rules of stochastic actions that can be at the same time correct and available to all agents, namely, the mixed-strategy Nash equilibria of the game.

LEMMA 2: *The median number of players who choose a number less than 60 in any mixed Nash equilibrium of the prediction game is exactly 60. The distance between the median and the mean in any mixed Nash equilibrium is at most equal to 1.*

Proof: See the Appendix.

3.2. The Repeated Prediction Game

Now that we know a number of features of the prediction game let us consider the repeated El Farol. The stage game is played every week, and to keep full

[12]Each player, as in a model with a continuum of players, ignores his or her effect on this number. This is exactly as in Arthur's set-up.
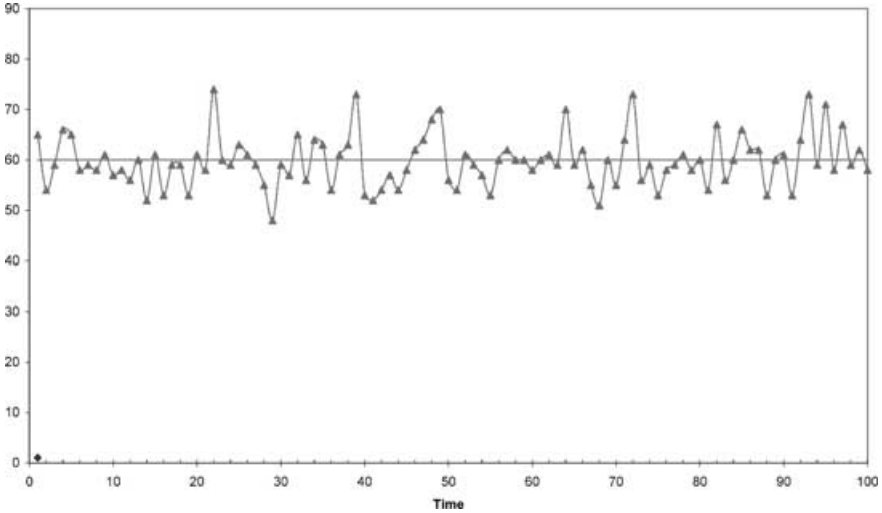
[13]Arthur (1994, p. 409).

Figure 4: One hundred weeks of aggregate attendance according to the symmetric Nash equilibrium of the underlying game

consistency with Arthur's framework, the players discount the future completely. At the beginning of the week the only information available to every player is the attendance in the previous weeks.

It is not hard to see that if any Nash equilibrium of the stage game is played at every period, the path of play that one would observe would show an average attendance of about 60, and obviously exhibits a variance of aggregate attendance over time that is bounded away from zero. To illustrate this point I have simulated the aggregate attendance that arises in a Nash equilibrium of the prediction game played repeatedly (Figure 4). That is, this sequence exhibits similar properties that the empirical distribution of aggregate attendance of the El Farol problem. Moreover, under the representation of the El Farol problem used in this paper, the statements "Mean attendance always converges to 60," and "On average, 40% of the active predictors are forecasting above 60 and 60% below 60," are, in fact, equivalent, which implies that an explanation of why Fact 1 arises is at the same time an explanation of why Fact 2 arises.

The statements in this last paragraph alone do not solve the problem posed by Arthur's simulation, however, because players in the computer experiments of the El Farol model done by Arthur (1994) do not use equilibrium strategies of the repeated prediction game. This discussion, however, clearly situates the problem as a game-theoretic one and, in particular, one for which a satisfactory solution exists. The key to understanding how the game-theoretic machinery can be used to shed light on a problem of this nature is explained below.

3.2.1. Playing a Repeated Game

That agents in a repeated game are endowed with Bayesian rationality, but not mutual knowledge of the strategies chosen by others, means that they are endowed with a prior over the relevant uncertainty space of their decision problems.

It turns out that the method of forecasting and decision used by the agents in the El Farol model allows us to recover *exactly* the preferences and beliefs underlying each agent's choice.

THEOREM 1: *The agents in the El Farol model behave exactly as if they are playing the prediction game period after period, with beliefs about aggregate attendance given by a variant of what is known in the literature of learning in games as "exponentially weighted fictitious play."*

Proof: Notice that if each player i uses at each period the most accurate predictor in her set, this means choosing s^i to maximize $U_t(s^i)$ at every period t. But notice that $U_t(s^i) = \lambda U_{t-1}(s^i) + (1 - \lambda)|s^i(d(h_{t-1})) - y_t|$ can be written as

$$U_t(s^i) = \sum_{\tau=1}^{t} \lambda^{t-\tau} u(s^i(d(h_{\tau-1})), y_\tau)$$

where $u(s^i(d(h_{\tau-1})), y_\tau) = |s^i(d(h_{\tau-1})) - y_\tau|$. Maximizing $U_t(s^i)$ by choice of s^i is of course equivalent to maximizing $\frac{\lambda}{1-\lambda^t} U_t(s^i)$. Notice, however, that if we define f_t to be a probability distribution on the set $D \times \{0, 1, 2, \ldots, 100\}$ such that the forecast probability of attendance y at date $t + 1$ and an attendance profile for the last d weeks of δ (up to week t), is

$$\frac{\lambda}{1 - \lambda^t} \sum_{\tau=0}^{t-1} \lambda^{t-\tau} 1_{\{y_\tau = y, d(h_{\tau-1}) = \delta\}}$$

then $\frac{\lambda}{1-\lambda^t} U_t(s^i)$ is the expectation of $u(s^i(\tilde{d}), \tilde{y})$, with respect to f_t. Call this expectation $V(s^i, f_t)$. Distribution f_t is the exponentially weighted fictitious play belief function, so that players at the end of every period t choose strategy $\hat{s}_{t+1}^i$ from the set $\arg \max_{s_i} V(s^i, f_t)$; that is, they basically behave according to the model known in the literature on learning in games as "exponentially weighted fictitious play." ■

Nachbar (1999), Kalai, Lehrer, and Smorodinsky (1996), Nyarko (1997), and Jordan (1997) have shown that any model of best-responding subject to adaptive forecasts can be given a full Savage-Bayesian decision-theoretic representation. *The contribution of Theorem 1 goes beyond this in that it provides an exact decision-theoretic representation of the behavioral rules that the players in the El Farol model employ.* Therefore, individual behavior in the El Farol problem is consistent with Savage-Bayesian rationality. But there is more: the empirical

distribution of aggregate attendance over time in the model is consistent with Nash equilibrium. This I will show in the following section.

4. When Learning Leads to Equilibrium

To study the long-run properties of the sequence of attendance it is useful to define the regret that a player experiences from having played the actions that he or she has actually chosen over time. Using an idea *adapted* from Hart and Mas-Colell (2001), define the conditional *expected* regret from strategy s^j to s^k at time t

$$DC_t^i(s^j, s^k) = \frac{1}{t} \sum_{\tau \leq t:\ \hat{s}_\tau^i = s^j,} \left[V(s^k, f_{\tau-1}) - V(s^j, f_{\tau-1}) \right].$$

Notice that each bracketed term in the summation is non-positive since players are best responding at every period τ to $f_{\tau-1}$. Since expected payoffs are linear in the probabilities,

$$DC_t^i(s^j, s^k) = V\left(s^k, \frac{1}{t} \sum_{\tau \leq t:\ \hat{s}_\tau^i = s^j} f_{\tau-1} \right) - V\left(s^j, \frac{1}{t} \sum_{\tau \leq t:\ \hat{s}_\tau^i = s^j} f_{\tau-1} \right) \leq 0.$$

and by a proposition in Hart and Mas-Colell (2000, p. 1133), the *average belief*, $\frac{1}{t} \sum f(t)$, is a correlated equilibrium distribution (projected onto the aggregate attendance levels) of the game being played.

From here it quickly follows that:

THEOREM 2: *The empirical distribution of aggregate attendance converges to the set of correlated equilibria of the prediction game.*

Proof: Exponentially weighted fictitious play beliefs evolve over time according to the formula

$$f(t+1) = f(t) + \frac{(1 - \lambda)}{1 - \lambda^{t+2}} [y_{t+1} - f(t)].$$

For large t, the denominator of the second summand can be ignored. Therefore,

$$\frac{1}{T} \sum_{\tau=t}^{T} f(\tau + 1) - \frac{1}{T} \sum_{\tau=t}^{T} f(\tau) \approx (1 - \lambda) \left[\frac{1}{T} \sum_{\tau=t}^{T} y_{\tau+1} - \frac{1}{T} \sum_{\tau=t}^{T} f(\tau) \right].$$

Since the average belief converges to a correlated equilibrium of the game, the left-hand side of the previous equation converges to zero, and hence the empirical distribution $\frac{1}{T} \sum_{\tau=t}^{T} y_{\tau+1}$ converges together with the average belief to the set of correlated equilibria, as well. ∎

The last step in understanding the behavior of the El Farol model is the following:

THEOREM 3: *The set of Nash equilibria and the set of correlated equilibria of the prediction game coincide.*

Proof: See the Appendix.

The contribution of the last three theorems is that together they show that the empirical distribution of attendance in the El Farol model converges to the set of Nash equilibria of the prediction game. One should therefore expect a median attendance equal to 60, an average attendance within 1 to 60, which explains Fact 1, Fact 2, and Conjecture 1′ as outlined in section 2 of the paper. Moreover, all of these equilibria are in mixed strategies, which explains why the standard deviation of aggregate attendance is bounded away from zero, as posed in Conjecture 2′. This is so even though all players use pure strategies period by period.

4.1. Remarks

It is important to comment on what the result does *not* say. First, it does not say that the players have learned to predict the strategies chosen by the others because players cannot recover the strategies chosen by others solely by observing the time series of aggregate attendance.

Second, it does not even say that the players have learned to predict the continuation path of play. This is in general a much stronger result that holds for beliefs with very restricted support, as explained by Nachbar (1997, 1999).

Third, it does not say that the beliefs of the players converge to the set of Nash equilibria of the game. Beliefs that are an exponentially smoothed variant of fictitious play beliefs converge to a Nash equilibrium only if the equilibrium is a pure strategy equilibrium. But this is impossible for the prediction game, since there is no pure strategy Nash equilibrium.

Fourth, it does not say that the players end up using the mixed strategies that form the Nash equilibrium of the repeated game, or that the process of aggregate attendance is random, for that matter. It does say, however, that the empirical distribution of aggregate attendance has features that are just as the features that arise in the empirical distribution induced by the set of Nash equilibria of the game. This set is entirely composed of mixed strategies, even though the forecasting rules are entirely deterministic. Therefore, the analytic results from the present section explain the puzzling facts about the El Farol model simulations, as are outlined in section 2.

While the techniques are somewhat different, the method used for obtaining the result in Theorem 3 is related to results on the convergence of empirical distributions to Nash and correlated equilibrium by Jordan (1997) and Nyarko (1994), respectively. Those results do not apply to the present paper directly, as some key hypotheses of those theorems regarding mutual absolute continuity of the beliefs of the players are not readily verifiable for the El Farol model. Those results, however, are the direct inspiration behind the analysis in the present paper.

5. Related Literature

To enable the reader to contrast my results to those already available in the literature on this problem, let me recall them here briefly. I have written the El Farol model as a repeated game of prediction where beliefs about other player's predictions evolve according to exponentially smooth fictitious play beliefs. This is not a variant on the El Farol problem: my representation yields the exact behavioral rules used for decision making and prediction by the players in Arthur (1994)'s simulation of the El Farol model. My representation also has the advantage that it allows me to analytically demonstrate why aggregate attendance oscillates around 60 in the simulations, and why the empirical distribution of outcomes appears to be generated by a random process even though the set-up is completely deterministic. This is, in short, the contribution of this paper.

The papers more directly related to the present work are Arthur (1994), which I have already discussed, Greenwald et al. (2000), Edmonds (1999), Bell et al. (2003), Farago et al. (2002), Johnson (1998), and Leady (2002). All these papers contain either simulations, or analytical results to variants on the El Farol problem, or both. None of them contain analytical results related to the original simulation done by Arthur (1994).

Greenwald et al. (2000) study the El Farol problem with the purpose of showing that rationality and predictability are incompatible in this game, in the spirit of Nachbar (1997)'s seminal work. While they succeed in what they attempt to do, this does not contradict the findings of the present paper. This is so because players' beliefs in the present paper, consistent with their findings, do not converge to the truth. *The average* of the beliefs over time, however, does converge to the set of correlated equilibria, and this is, *in this game*, sufficient for the empirical distribution of attendance to converge to the set of Nash equilibria of the game.

Edmonds (1999) endows the agents with an evolving set of predictors, allows communication between the players, and studies the outcome of simulations very much like the one by Arthur. His findings are consistent with Arthur's original findings regarding attendance oscillating around 60 in a seemingly random fashion. While Edmonds suggests that the results in the simulations are inconsistent with Bayesian rationality and equilibrium in the limit, he focuses, like Greenwald et al. (2000) on actual play and actual beliefs of the game, as opposed to average beliefs and the empirical distribution of play, as in the present paper.

Bell et al. (2003) and Farago et al. (2002) treat the El Farol problem as an engineering problem. Instead of trying to explain the outcome of the simulations done by Arthur in game-theoretic terms, they are more interested in studying what type of learning algorithms players could use for the system to find their way to efficient and fair equilibria, and they succeed in finding both deterministic and stochastic strategies for the players that do just this. They also focus on the *attendance game representation* of the El Farol model.

Johnson et al. (1998) consider how the variance of attendance in the simulations of the El Farol problem change in response to the number of predictors available in the entire system and the number of predictors that each agent selects.

Interestingly, Leady (2002) conducted experiments with human subjects playing the *attendance game representation* of the El Farol model and got results qualitatively similar to those from the simulations and the analytical results presented in this paper: while the sequence of actions did not approach any of the equilibria of the game, *average* attendance to the bar over time was about 60%.

5.1. The Minority Game

An important paper in the literature is that of Challet and Zhang (1997), which studies the variant on the El Farol problem discussed in the introduction called the "minority game" using tools from statistical physics. The papers devoted to the minority game are too numerous to discuss here in any detail (a website dedicated to the minority game, http://www.unifr.ch/econophysics/minority/papers.html, contains more than 110 references to papers devoted to its study). Let me recall briefly the contribution of three papers from that strand of the literature that are relevant for understanding what transpires in the El Farol problem.

Marsili, Challet, and Zecchina (1999) study the minority game under alternative assumptions about whether the agents take into account their own effect into aggregate outcomes. When the agents do take into account their own effect into aggregate outcomes they obtain convergence of outcomes to the Nash equilibria of the minority game while this does not happen when agents ignore their own effect into aggregate outcomes. In this latter case convergence is to a notion they call *naive agents' equilibrium*, a point that is associated with the maximization of a potential function, as in Hart and Mas Colell (2001).

Shalizi and Albers (2003) use symbolic dynamics to study discrete adaptive games like the minority game and show that there cannot be deterministic chaos in the canonical model of a such game.

It is nevertheless of interest that the results that are reported on many of the minority game papers regarding how an efficient use of the resources in the model (i.e., the groups the agents intend to join) depend on the richness of the space of predictors that the agents can use to form their forecasts. The most efficient use of the resources seems to obtain when the pool of predictors is neither too large nor too small, a condition described by physicists as a "phase shift." A canonical example of the type of papers, which report findings of this sort, is that of Savit, Manuca, and Riolo (1997).

A concise summary of what is known about this game can be found in Farmer (1999). I conjecture that many of the analytical methods used in the

present paper could be used to shed light on the puzzling facts of the minority game simulations.

5.2. The El Farol Problem and Congestion Externalities

The purpose of this sub-section is to remark that this paper contributes to the study of the kinds of congestion problems that arise through the interaction of computer agents in a network like the Internet.

The reader may easily see how the El Farol problem falls within the class of congestion externality problems: There is a resource (the bar) that provides a better service when not so many people use it. Roads and network services, among others, are resources subject to exactly the same kind of congestion. The case of network services and the Internet are especially important since computer scientists have had some resistance in employing the standard tools of mathematical economics and game theory to study them. Interestingly, the reason is that it has been argued that computational agents lack the kind of sophistication that players are supposed to have according to game theory and that therefore game-theoretic predictions about what would arise in interactions among computer agents would not be of use.

This paper shows that this need not be the case, as the kind of algorithms that define the manner in which computer agents interact over time can be given a Savage-Bayesian interpretation. This opens the door for the tools of the literature on learning in games to be used to study those interactions in the hope that they provide an accurate prediction of the unfoldings of the interaction. This is, in fact, what the research reported in this paper has done for the case of the El Farol problem, which suggests that the standard tools of game theory should be of tremendous utility in understanding the rapidly growing world of network-based computation and interaction.

6. Conclusions

Proponents of bounded rationality models often argue that what makes rational choice theory unbelievable is not that agents are supposed to have well-defined objectives and that they act to achieve them, but that the agents' choice occurs in a context in which they have detailed information about the environment, perhaps under the form of a prior function defined over a large and complicated state space. It is then reasonable for many to build models with agents that select a standard *best response* subject to their forecasts about the environment, and then having the forecasts be determined by some model of adaptive learning, such as least squares learning, genetic algorithms, pattern recognition, and so on.

The main point of this paper is to remark that this class of models of bounded rationality can easily be made consistent with the standard Bayesian decision-theoretic framework. This is so because the forecasting systems used in these bounded rationality models induce a (possibly misspecified) prior

over the set of possible future outcomes that, conditional on any partial history, assigns probability 1 to the set of possible futures that are consistent with the outcome generated by the forecasting system given that history, and probability zero elsewhere. As a consequence, these models of bounded rationality are just as Bayesian rational as those models where the prior is explicitly spelled out.

While this point has been made repeatedly in the theoretical literature on Bayesian learning in games (see, e.g., Nachbar 1999, Kalai, Lehrer, and Smorodinsky 1996, Nyarko 1997, and Jordan 1997), this insight has not fully permeated into applied research conducted in the bounded rationality camp. *For this reason the contribution of this paper is to make this point by providing an example of a famous model of bounded rationality from the computational economics literature that fits into the standard Bayesian decision-theoretic framework.*

Providing a solution to the El Farol problem turns out to be important in its own right. Arthur's model represents very well the central problem from the public economics literature of understanding how resources that are subject to congestion are used in society.

As the incredible interest that this problem has generated in the computer science and in the physics literature demonstrates, the results of this paper open up many new avenues for research because the issues raised by problems like the El Farol, and a deeper understanding of the relationship between endogenous uncertainty, bounded rationality, congestion externalities, and strategic interaction are far from being completely understood for the general case. The reader is strongly encouraged to check Blume and Easley (1998) and Nachbar (1997) to see why.

To conclude, let me reiterate my belief in that the research reported in this paper illustrates the interplay that exists between computational modeling and mathematical analysis. Both the El Farol problem and the minority game have been the subject of many computational models. In this paper, I am showing how we can use mathematics to explain what happens in these computational models. This should not be interpreted as a demonstration that mathematics always trumps computation. On the contrary, the two are complementary. The computational models of open a new set of questions and generate phenomena that we would like to understand fully. The mathematics I present provides an explanation for some dimensions of these phenomena. They erase some of the mystery. In doing so, they enhance, not diminish, the discovery of the phenomena themselves.

Appendix

Proof of Lemma 1:   Assume that all players choose pure strategy a^i, a number between 0 and 100. Let $F = \#\{i : a^i < 60\}$. Notice that for each player i to be best responding to the actions of all, i.e., to be maximizing the payoff $-|a^i - F|$, he or she would choose $a^i = F$. Therefore, if there was an equilibrium in pure strategies, it would have to have all players choosing

the same number. Therefore, F is either 0 or 100. If $F = 0$ then all players choose $a^i = 0$, so $F = 100$, which cannot be. If $F = 100$ then all players choose $a^i = 100$, so $F = 0$, which cannot be either. Hence, there is no Nash equilibrium in pure strategies. ∎

Proof of Lemma 2: Let p be a probability distribution on aggregate attendance, i.e., on the set $\{0, 1, \ldots, 100\}$ that arises from the probabilistic choices of all the players, each player i choosing action k with probability q_k^i. I first show that there cannot be any mixed equilibria with $\sum_{j=0}^{60} p(j) > \frac{1}{2}$.

The expected payoff of a player choosing action k from the set $\{0, 1, \ldots, 100\}$ given distribution p is $u(k, p) = -\sum_{j=0}^{100} p(j)|k - j|$. Now compare $u(60, p)$ with $u(k, p)$ for $k > 60$:

$$u(k, p) - u(60, p) = -\left[\sum_{j=0}^{100} p(j)|k - j| - \sum_{j=0}^{100} p(j)|60 - j| \right]$$

$$= -\left\{ \begin{array}{l} (k - 60)\left[\sum_{j=0}^{60} p(j) - \sum_{j=k+1}^{n} p(j) \right] \\ + (60 + k) \sum_{j=61}^{k} p(j) - 2 \sum_{j=61}^{k} p(j)j \end{array} \right\}$$

$$\leq -\left\{ (k - 60)\left[\sum_{j=0}^{60} p(j) - \sum_{j=61}^{n} p(j) \right] \right\} < 0.$$

This means that when $\sum_{j=0}^{60} p(j) > \frac{1}{2}$ no player would want to choose a number k above 60. Intuitively, this is so because when $\sum_{j=0}^{60} p(j) > \frac{1}{2}$ the median of the distribution p must be at most 60, and when players attempt to minimize expected absolute loss function, as in this case, they will want to choose a number as close to the median as possible.[14] Here, any number k above 60 must be farther to the median than 60 is. This means that, given p, players will put probability zero on any strategy $k > 60$. This, of course, means that $p(100) = 1$ and, consequently, $\sum_{j=0}^{60} p(j) = 0$, which contradicts our hypothesis.

A similar argument shows that there cannot be any mixed equilibria with $\sum_{j=0}^{60} p(j) < \frac{1}{2}$. In this case no player would want to choose a number k below 60, which would make $p(0) = 1$, contradicting our hypothesis. Therefore, in any mixed equilibrium of this game, $\sum_{j=0}^{60} p(j) = \frac{1}{2}$ and 60 is the *median* of the distribution over $\{0, 1, \ldots, 100\}$ of any mixed equilibrium.

To show that the mean is just about 60, we first need to understand a few more properties of the mixed equilibria of this game. We have shown above that if $k \neq 60$ then $u(k, p) - u(60, p) \leq 0$. It turns out that when $\sum_{j=0}^{60} p(j) = \frac{1}{2}$ this can be refined, using arguments identical

[14]See, for example, exercise 2.19 in Casella and Berger (1990).

to those given above, to: if $k \neq 60, 61$ then $u(k, p) - u(60, p) < 0$, $u(61, p) = u(60, p)$. I ommit the details here.

This means that in any mixed equilibrium all players put positive probability only on the numbers 60 and 61, and the distribution p is such that $\sum_{j=0}^{60} p(j) = \frac{1}{2}$. Therefore, the equilibrium distribution p arises from the random choice of 100 players between actions 60 and 61, each player randomizing independently with probabilities $(q_{60}^i, 1 - q_{60}^i)$. In other words, the 100 random choices can be thought as of 100 independent heterogeneous (e.g., not identically distributed) Bernoulli trials with p being the distribution of the sum of the 100 trials. A remarkable fact about heterogeneous Bernoulli trials, which is due to Jogdeo and Samuels $(1968)^{15}$ is that if the mean of the distribution p is an integer, then the mean is the median. If the mean is not an integer, then the median is one of the two adjacent integers. Therefore, the mean is within a distance of one to the median. ■

Proof of Theorem 3: Let P be a correlated equilibrium distribution of the prediction game, and p the induced distribution over the number of players who choose a number below 60. An argument similar to the one used in the proof of Lemma (2) shows that if $P(s^i = k) > 0$ then $\sum_{j=0}^{60} p(j \mid s^i = k) = \frac{1}{2}$ for all i. From this it follows that only 60 and 61 can have positive probability as strategies recommended to each player i under P. The proof is complete when one shows that the prediction about attendance according to p, and the prediction about attendance according to p, conditional on s^i, are the same. This follows from the Law of Total Probability as follows:

$$\sum_{j=0}^{60} p(j) = \sum_{j=0}^{60} p(j \mid s^i = 60) P(s^i = 60)$$

$$+ \sum_{j=0}^{60} p(j \mid s^i = 61) P(s^i = 61)$$

$$= \frac{1}{2} (P(s^i = 60) + P(s^i = 61)) = \frac{1}{2}.$$

Therefore, the Nash and the correlated equilibrium distributions coincide. ■

References

ARTHUR, B. (1994) Inductive reasoning and bounded rationality, *American Economic Review* **84**(2), 406–411.

---

[15]See, for example, Siegel (1992).

ARNOTT, R., A. DE PALMA, and R. LINDSEY (1993) A structural model of peak-period congestion: A traffic bottleneck with elastic demand, *American Economic Review* **83**, 161–179.

BELL, A., W. SETHARES, and J. BUCKLEW (2003) Coordination failure as a source of congestion in information networks, *IEEE Transactions on Signal Processing*, **51**.

BLUME, L., and DAVID EASLEY (1998) Rational expectations and rational learning, in *Organizations with Incomplete Informations. Essays in Economic Analysis: A Tribute to Roy Radner*, M. Majumdar, ed. Cambridge: Cambridge University Press.

CASTI, J. (1996) Seeing the Light at El Farol, *Complexity*, **1**, 7–10.

CASELLA, G., and R. BERGER (1990) *Statistical Inference*. Belmont CA: Duxbury Press.

CHALLET, D., and Y.-C. ZHANG (1997) Emergence of cooperation and organization in an evolutionary game, *Physica A* **246**, 407.

EDMONDS, B. (1999) Modelling bounded rationality in agent-based simulations using the evolution of mental models, in *Computational Techniques for Modelling Learning in Economics*, T. Brenner, ed. Dordrecht: Kluwer, 305–332.

FARAGO, J., A. GREENWALD, and K. HALL (2002) Fair and efficient solutions to the santa fe bar problem, manuscript, Brown University.

FARMER, J. (1999) Physicists attempt to scale the ivory towers of finance, *Computing in Science & Engineering*, **1**, 26–39.

GREENWALD, A., B. MISHRA, and R. PARIKH (2000) On the impossibility of rational learning in the santa fe bar problem, manuscript, Brown University.

HART, S., and A. MAS-COLELL (2000) A simple adaptive procedure leading to correlated equilibrium, *Econometrica* **68**, 1127–1150.

HART, S., and A. MAS-COLELL (2001) A general class of adaptive strategies, *Journal of Economic Theory* **98**, 26–54.

JOGDEO, K., and S. SAMUELS (1968) Monotone convergenge of binomial probabilities and a generalization of Ramanujan's equation, *Annals of Mathematical Statistics* **39**, 1191–1195.

JOHNSON, N., S. JARVIS, R. JONSON, P. CHEUNG, Y. KWONG, and P. HUI (1998) Volatility and agent adaptability in a self-organized market, *Physica A*, **256**, 230–236.

JORDAN, J. (1997) Bayesian learning in games: A non-bayesian perspective, in *The Dynamics of Norms*, C. Bicchieri, R. Jeffrey, and B. Skyrms, eds. Cambridge, UK: Cambridge University Press.

KALAI, E., E. LEHRER, and R. SMORODINSKY (1996) Calibrated forecasting and merging, manuscript, Kellog Graduate School of Management.

LEADY, J. (2002) If nobody is going there anymore because it's too crowded, then who is going? Experimental evidence of learning and imitation in the El Farol coordination game, manuscript, University of Michigan.

MARSILI, M., D. CHALLET, and R. ZECCHINA (1999) Exact solution of a modified El Farol's bar problem, http://xxx.lanl.gov/pdf/cond-mat/9904392.

NACHBAR, J. (1997) Prediction, optimization and learning in repeated games, *Econometrica* **65**, 275–309.

NACHBAR, J. (1999) Rational Bayesian learning in repeated games, manuscript, Washington University, St. Louis.

NYARKO, Y. (1994) Bayesian learning in repeated games leads to correlated equilibria in normal form games, *Economic Theory* **4**, 821–841.

NYARKO, Y. (1997) Savage-Bayesian agents play a repeated game, in *The Dynamics of Norms*, C. Bicchieri, R. Jeffrey, and B. Skyrms, eds. Cambridge, UK: Cambridge University Press.

ROSENTHAL, R. (1973) A class of games possessing pure-strategy Nash equilibria, *International Journal of Game Theory* **2**, 65–67.

SAVIT, R., R. MANUCA, and R. RIOLO (1997) Adaptive competition, market efficiency, phase transitions and spin-glasses, manuscript, University of Michigan.

SHALIZI, C., and D. ALBERS (2003) Symbolic dynamics for discrete adaptive games, manuscript, University of Michigan.

SHENKER, S. (1995) Making greed work in networks: A game theoretic analysis of switch service disciplines, *IEEE/ACM Transactions on Networking* **3**, 819–831.

SIEGEL, A. (1992) Toward a usable theory of chernoff bounds for heterogeneuos and partially dependent random variables, manuscript, New York University.

SWED, F., and C. EISENHART (1943) Tables for testing randomness of grouping in a sequence of alternatives, *Annals of Mathematical Statistics* **14**, 66–87.

VARIAN, H., and J. MACKIE-MASON (1994) *Economics of the Internet*, http://www-personal.umich.edu/~jmm/papers/Economics_of_Internet.pdf.

VICKREY, W. (1963) Pricing in urban and suburban transport, *American Economic Review* **53**, 452–465.

VICKREY, W. (1969) Congestion theory and transport investment, *American Economic Review* **59**, 251–260.

WOLPERT, D., K. WHEELER, and K. TUMER (1999) Collective intelligence for control of distributed dynamical systems, *Tech Report: NASA Ames Research Center*, NASA-ARC-IC-99-44.

ZAMBRANO, E. (1997) The revelation principle of bounded rationality, *SFI Working Paper 97-06-060E*, The Santa Fe Institute.