Journal of Econometrics 153 (2009) 183–195

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Learning in a multilateral bargaining experiment

Guillaume R. Fréchette

New York University, Department of Economics and CESS, 19 West Fourth Street, 6th floor, New York, NY 10012, United States

ARTICLE INFO

Article history:
Available online 27 June 2009

JEL classification:
C92
D83
C12

Keywords:
Learning
Bargaining
Estimation
Hypothesis tests

ABSTRACT

This paper analyzes data from an investigation of a majoritarian bargaining experiment. A learning model is proposed to account for the evolution of play in this experiment. It is also suggested that an adjustment must be made to account for the panel structure of the data. Such adjustments have been used in other fields and are known to be important as unadjusted standard errors may be severely biased downward. These results indicate that this adjustment also has an important effect in this application. Furthermore, an efficient estimator that takes into account heterogeneity across players is proposed. A unique learning model to account for the paths of play under two different amendment rules cannot be rejected with the standard estimator with adjusted standard errors, however it can be rejected using the efficient estimator. The data and the estimated learning model suggest that after proposing "fair" divisions, subjects adapt and their proposals change rapidly in the treatment where uneven proposals are almost always accepted. Their beliefs in the estimated learning model are influenced by more than just the most recent outcomes.

1. Introduction

Bargaining is clearly a central issue in economics, but it is also very important in both political science and psychology. As a result, it is a topic where research has been shared across disciplines and has, for instance, been at the root of some key papers in behavioral economics, where experimental methods have served as an important tool in the development of concepts such as other-regarding preferences (Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999; Charness and Rabin, 2002) and learning (Roth and Erev, 1995; Cooper et al., 1997; Camerer and Ho, 1999). The basic bilateral bargaining model of Rubinstein (1982) was extended to a multilateral setting by Baron and Ferejohn (1989) with the goal of modeling legislative bargaining. Since then the model has been used to study many aspects of politics and is cited in countless papers. According to Persson and Tabellini (2000): "The model has since become one of the workhorse models in the rational choice approach to US congressional politics."(Persson and Tabellini (2000), p. 110.)

Of course, a multilateral bargaining model is of interest to economists. Some of the recent applications and developments of the model in economics include theoretical work on the division of a stochastic amount of money (Eraslan and Merlo, 2002), and applied work on the effect of representation on the economic growth of states (Levitt and Poterba, 1999).

In both political science and economics there have been empirical tests of the theory with mixed results (for instance Warwick and Druckman (2001), Knight (2002) and Ansolabehere et al. (2005)).[1] Some find support for the existence of proposer power, while others find that distributions of payoffs, contrary to what theory suggests, seem to be proportional to the number of votes controlled. Clearly, there are many factors that make field data difficult to analyze. Experimental methods have a distinct advantage by allowing the researcher to vary only the amendment rules. They can also eliminate problems such as the varying importance or salience of different portfolios of ministries that arise in studying coalition governments (Warwick and Druckman, 2001), repeated play effects, the effects of the selection rule (Fréchette et al., 2005b), and varying bargaining power (Fréchette et al., 2005a).

This paper reports results from an experiment testing the comparative static predictions between two amendment rules: closed and open. When studying the Baron and Ferejohn model of legislative bargaining, Fréchette et al. (2003) found that although qualitative predictions of the model were supported, the stationary subgame perfect equilibrium point predictions tended to be off the mark. Subjects' behavior evolved very rapidly in a way that suggests adaptive behavior. This paper uses that same data to explain the evolution of proposals by an adaptive learning model. The focus of this paper is the process by which proposals for the division of the money evolve, thus it will be of very little relevance to political economy, and focuses on learning and the estimation problem. It explains the evolution of behavior by using a belief

[1] Many other papers precede Baron and Ferejohn (1989) but give us information about the model's performance in field data, for instance Browne and Franklin (1973).

E-mail address: frechette@nyu.edu.

0304-4076/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2009.06.004

based learning model. Furthermore, it is shown that one cannot reject the hypothesis of a unified model explaining the data from the two different games using a typical estimator while adjusting the variance of the estimates to account for the panel nature of the data. On the other hand, when a more efficient estimator that directly accounts for heterogeneity in the data is used, this hypothesis can be rejected. These two observations point to the importance of taking into account the panel structure of the data, both for the purpose of hypothesis testing and for the estimation of parameters.

The Baron and Ferejohn model is a "divide the dollar"game with more than two players and a majority voting rule. The experiment explores a closed and an open rule setting. The data from the experiment show that subjects under both treatments originally behave the same, offering an almost equal split of the pie to all subjects. However, under the closed rule, they move toward a situation where $\frac{n-1}{2}$ subjects are excluded from the benefits, and the proposer keeps a greater share of the pie for herself. On the other hand, at the end of the open rule sessions, about half of the proposals distribute money to all subjects, while the other half excludes fewer than $\frac{n-1}{2}$ members from benefits. For example, although nobody gave zero to another subject in election 1 of the closed rule, by election 15 nearly 50% of all offers gave nothing to 2 out of 5 members. In contrast, under the open rule, there were no offers in election 15 giving nothing to 2 members.

The end-of-session behavior is vastly different from actions at the beginning. This result is quite striking considering the fact that the experiments only lasted fifteen elections. The average share the proposer takes for herself under the closed rule almost doubles between election 1 and election 15.

Using the standard approach to estimating learning where one treats repeated observations as independent, one can reject the hypothesis that a single learning model explains the evolution in both closed and open rule games. Such a result is consistent with most evidence from previous studies which shows that data should not be pooled across games. However, since estimation of learning models employs panel data, it is likely that observations across periods for a given individual are correlated in a way that is not accounted for by the learning model. In ignoring the fact that one has repeated measures for an individual when performing estimation, it is often the case that standard errors are biased downward. Once the variance–covariance matrix is adjusted to account for correlation across observations for a given individual, the hypothesis that both games can be explained by the same learning model cannot be rejected. Adjustments of this sort have, to my knowledge, never been performed when estimating learning models and may account for the common finding that parameter estimates are not stable across similar games (Stahl, 1996; Camerer and Ho, 1999; Erev and Haruvy, 2001). Previous estimation of learning models have varied the specific dynamics estimated but also details of the estimation, especially how the priors are obtained. The effects of the adjustment to the variance–covariance matrix proposed here is shown to be robust to a few different specifications of the priors. On the other hand, the results of the key hypothesis tests are affected if heterogeneity in priors is incorporated in the estimator. A more efficient estimator proposed in this paper, which takes the heterogeneity across subjects directly into account, leads to the opposite conclusion, namely that the hypothesis of a common learning model can be rejected. Hence, the results highlight the fact that this type of adjustment of the variance–covariance matrix can have an important impact on the power of tests. However, the relatively high rates of type I error that will be observed in simulations using the more efficient estimator also point to the difficulties inherent in using experimental data, in which samples tend to be relatively small.

The paper proceeds as follows. First, the model on which this experiment is based is presented, followed by the experimental design, and the experimental results. Then the learning model, the estimation method, and the estimation of the variance–covariance matrix is described. This is followed by an explanation of the simplifications of the model that are made for estimation purposes. Results are presented and more specifically, it is shown that one cannot reject the hypothesis that both closed and open rules result from the same learning process after the variance–covariance is appropriately adjusted. Finally, an estimator that allows for heterogeneous priors is proposed, and estimates are presented, followed by simulation results about the performance of the different estimators.

2. Theoretical model and predictions

The Baron–Ferejohn (BF, 1989) model is in essence a "divide the dollar" game where proposals are approved by majority rule. It is intended to reflect a legislative setting. Four elements define a legislature: n members, a recognition rule, an amendment rule, and a voting rule. Members of the legislature have risk-neutral preferences and derive utility solely from benefits allocated to their district. Finally, there is perfect information – all actions are observable – and also there is common knowledge of preferences and legislative rules. Both the recognition rule and the amendment rule are assumed to select a person at random.

At the beginning of an election,[2] each member i has a probability $p_i = 1/n$ of being recognized, and if recognized, makes a proposal specifying how benefits will be distributed. A proposal x^i is a distribution $x^i = \left(x_1^i, \ldots, x_n^i\right)$ such that $\sum_{j=1}^n x_j^i \leq 1$ where x_j^i is the share that i allocates to voter j. This proposal is then the motion on the floor. The status quo corresponds to no allocation of benefits, $x = (0, \ldots, 0)$.

Under a closed amendment rule the motion is voted on immediately (against the status quo). If the proposal is approved, the legislature adjourns. If it is not approved, the amount of benefits to be divided is discounted, and the legislature moves to the next round and the process repeats itself.

Under an open rule, a member j is selected with probability $p_j = \frac{1}{n-1}$, after a member (i) has been recognized and made a proposal. This member can accept the motion on the floor, in which case it is voted on. If the motion on the floor is approved, the legislature adjourns. Otherwise the benefits are discounted, the game moves to the next round, and the process repeats itself. If the motion is amended (by j), the newly selected member (j) proposes an alternative distribution of benefits, and the legislature votes between the two proposals. The winning proposal becomes the motion on the floor; the benefits are discounted; and the legislature moves to the next round. The process repeats itself until an agreement is reached.

Preferences of member j are represented by the utility function $u^j\left(x^k, t\right) = \rho^{t-1} x_j^k$ where t is the round in which the legislature adopts the distribution x^k, and $\rho \leq 1$ is the discount factor intended to reflect time preferences or capture the probability of re-election. The specific parameter values used in the experiment were: $\rho = 0.8$ and $n = 5$.

Because there are so many equilibria (including multiple subgame perfect equilibria), BF focus on stationary subgame perfect equilibria (SSPE). The SSPE predictions for the parameter values used in the experiment are presented in Table 1.[3] The SSPE for the closed rule only includes a minimal winning coalition. To see this, suppose $\underline{x}$ is the smallest share one would accept to be in a

[2] My use of the term election corresponds to Baron and Ferejohn's use of legislative session, and rounds corresponds to their use of elections.

[3] For a detailed exposition of the equilibrium and its proof, see (Baron and Ferejohn, 1989).

Table 1
Theoretical predictions for stationary subgame equilibrium outcome with 5 subjects and a discount factor of 0.8.

Predictions	Closed rule	Open rule
Number of voters receiving a positive payoff besides the proposer	2	2
Number of voters receiving zero payoff	2	2
Share to the proposer ($amount – round 1)	0.68 ($17)	0.52 ($13)
Share to coalition members ($amount – round 1)	0.16 ($4)	0.24 ($6)
Probability of proposal being approved in the 1st round	1	0.5

coalition. Then the proposal $\left(1-2\underline{x}, \underline{x}, \underline{x}, 0, 0\right)$, where the first entry is the share to the proposer, maximizes the proposer's utility since giving to a fourth member does not increase the probability of the proposal being accepted and only reduces the proposer's share. Proposals are always accepted in round 1. The share that the proposer takes for herself depends on the number of members and the discount factor.

Under the open rule, not only does the number of members and the discount factor affect the distribution of shares, but it also affects the number of members included in the coalition. Given the parameters implemented, the SSPE for the open rule is also one with a minimal coalition. With a minimal winning coalition, since $\frac{n-1}{2}$ members are excluded from the coalition, these excluded members will amend it. Thus there is a 50% chance that a proposal will be accepted at every round. The positive probability of being included in a future coalition increases the continuation value of the game for the non-proposers, so that they need to be offered more in order to accept a proposal. This explains why the proposer keeps a smaller share of the pie in equilibrium.

3. Experimental design

Five subjects were recruited for each experimental session which consisted of fifteen elections. The amendment rule remained the same within a given experimental session but differed between sessions (four closed rule sessions and four open rule sessions).[4] To minimize the possibility of any repeated play game effects, at the start of each election each of the five "legislators" were randomly assigned a new subject number. This subject number was known only to that individual legislator and changed across elections but not between rounds of a given election.

At the start of each election each subject filled out a proposal form for allocating $25.00 among the five voters by their subject numbers. Once the proposal forms were completed and collected, a roll of a five-sided die was used to determine which proposal would take the floor. This proposal, along with the subject number of the proposer, was posted on the blackboard for everyone to see.

In closed rule sessions, each subject would next complete a voting form indicating whether they accepted or rejected the proposed division. The voting forms were then collected and tabulated. If a simple majority accepted the proposal then the payoff was implemented and the election ended. If the proposal was rejected, each voter had the opportunity to propose a new division after applying the discount rate of 0.8 to the total benefits. (Discounting was done by the experimenters, with the total amount of money to be allocated in the next round posted on the blackboard.) Results were posted on the blackboard underneath each proposal. The blackboard contained information from the last several elections.

In an open rule session, after a proposal had been selected, each legislator completed a form either seconding or amending the standing proposal. When a voter chose to amend a proposal, she was required to propose an alternative distribution of benefits. Although these "seconding" forms were collected from everyone (to preserve the anonymity of the proposer), a roll of a four-sided die determined which legislator, other than the proposer, would be recognized to second or amend the standing proposal. If this legislator seconded the proposal, an election was held following the same procedures as in the closed rule sessions. If the legislator amended the proposal, the amended proposal was posted on the blackboard, along with the original proposal, and a run-off election was held. The winner of the run-off election was the standing proposal in the next round of the election. Benefits following a run-off election were subject to discounting, so that the shares in the standing proposal in the next round of the election were multiplied by 0.8. These shares where posted (in dollar amounts) along with the total amount of money to be allocated. Each open rule election continued in this way until a proposal was both seconded and approved by a simple majority.

Subjects were recruited through announcements in undergraduate classes and advertisements in student newspapers at the University of Pittsburgh and Carnegie Mellon University. This resulted in recruiting a broad cross section of graduate and undergraduate students from both campuses. At the end of each experimental session, four elections were randomly selected (by four rolls of a fifteen-sided die), with subjects paid the sum of their earnings in the four elections selected. Each subject also received a participation fee of $5.00.

In each session an additional subject was recruited to roll the dice. Having a subject roll the dice helps assure subjects that the outcomes are indeed randomly determined. This subject received a fixed fee of fifteen dollars.

In each session, practice elections were held first to familiarize subjects with the procedures and accounting rules. All experimental sessions were conducted using pencil and paper.[5]

4. Experimental results

The main results are organized around three key observations.[6]

Observation 1. *The average share that proposers take for themselves increases dramatically under the closed rule but decreases slightly under the open rule.*

Although behavior under both rules is very similar at the beginning, it diverges significantly by the end. Specifically, proposers in the closed rule take more for themselves by the end – an increase of almost 43% – whereas proposers take slightly less at election 15 than at election 1 under the open rule.

Observation 2. *Although proposals at the beginning of the experiment involve supermajorities under both rules, the coalitions at the end are, more often than not, minimal winning coalitions under the closed rule.*

The data reveal that under both treatments, when they first begin, the proposals almost always include all subjects in the coalition; this strategy will be referred to as the "almost even"(AE) strategy. As the sessions evolve, different patterns emerge under both treatments. In the closed rule, it steadily closes in on the equilibrium outcome of a minimal winning coalition in which two voters receive zero, or near zero allocations. The strategy which almost totally excludes two members will be referred to as the "almost double zero"(ADZ) strategy. For the last five elections, about 67% of the subjects play ADZ, and 29% play AE in round 1. For the open rule, we see a similar, though not nearly as pronounced, decline in the popularity of the AE strategy, but it is not replaced

4 The first session in each treatment had ten elections. Looking at the data from these sessions it was clear that behavior was still evolving, so we extended all subsequent sessions to fifteen elections.

5 Instructions are available at http://homepages.nyu.edu/~gf35/print/Frechette_2003a_inst.pdf.

6 The reader interested in more details is directed to Fréchette et al. (2003).

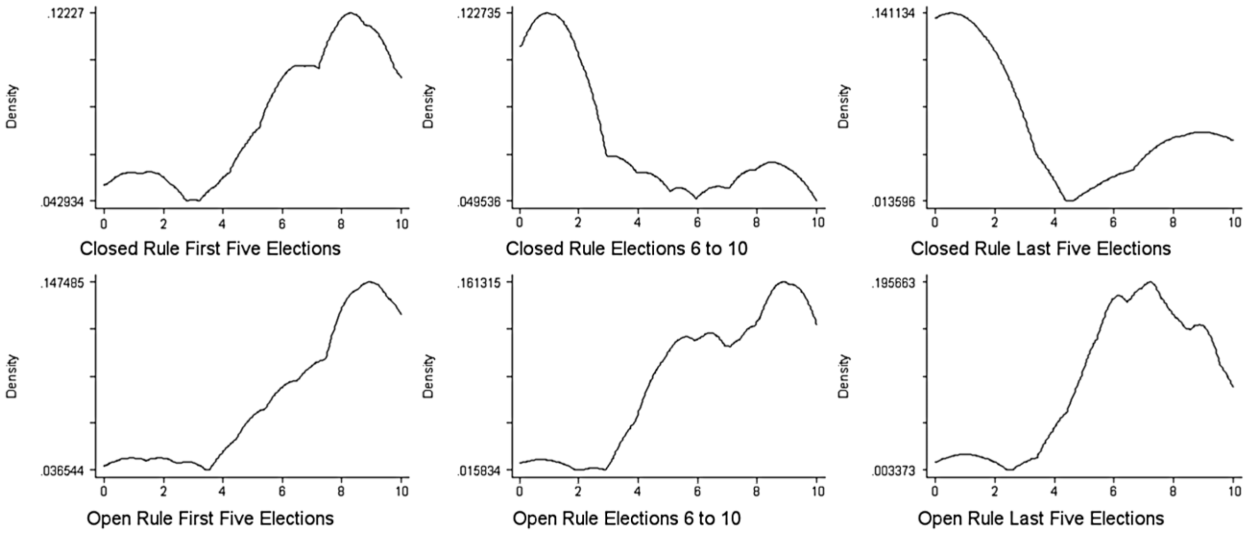

Fig. 1. Sum of the two lowest amounts.

by the ADZ strategy. Rather, subjects make offers that almost exclusively include four members in the coalition; this strategy will be referred to as the "almost single zero"(ASZ) strategy. For the last five rounds 55% of subjects play ASZ, and 40% play even in round 1.[7]

The evolution described above can be seen more readily by looking at the sum of the two smallest shares. The sum of the two smallest shares is an interesting statistic because in both treatments, the equilibrium prediction is that the sum of the two lowest amounts should be zero. Fig. 1 presents kernel density estimates (Silverman, 1986) of the sum of the two smallest shares (presented in dollar amounts) in round 1 for sets of 5 consecutive elections. Note that since subjects were dividing $25 per election, a sum of $10 for the two lowest amounts represents an even split between all subjects. For the first five elections, the graphs for the closed and open rules are virtually identical, with most proposals distributing the money nearly evenly under both the closed and open rules. However, by the last 5 elections the majority of offers totally exclude two members under the closed rule and even more proposals are evenly distributed under the open rule.

In summary, the subjects' behavior in this experiment changes mainly along two dimensions: the average share the proposers take for themselves and the number of members included in the winning coalitions. Further, patterns of play along both these dimensions were essentially identical under both treatments at the beginning.

Observation 3. *Voting behavior does not change over time and is mostly determined by the share one is offered.*

One's share seems to be the only robust factor of major importance. For instance, simply using the following rule, accept if my share is greater than or equal to 0.16, can explain 93% of the closed rule data and 96% of the open rule data. Clearly, there is very little room for anything else to matter besides share. This indicates that little is to be gained empirically from modeling learning on the voting side since voting behavior does not evolve over the course of the experiment. Furthermore, not doing so substantially simplifies the analysis.

5. Learning & estimation procedure

The learning model used in this paper is a two parameter, belief based learning model in the style of Cheung and Friedman (1997) which allows learning ranging from Cournot to fictitious play.[8] No attempt is made to determine whether this is the best model within the growing number of learning models available. It is, however, a fairly standard choice, and as will be seen, it performs well enough not to prompt looking for an alternative.[9] The model is described in detail below, using the following notation

subject $: i \in \{1, \dots, 5\}$

proposal types $: j \in J = \{1, \dots, 9\}$

vote $: v \in \{a, r\}$

proposer $: P \in \{1, \dots, 5\}$

where a stands for accept and r stands for reject. The space of possible proposals is divided into nine representative proposal types, as explained later.

Throughout, utility is taken to equal monetary payoffs and consequently expected utility to equal expected monetary payoffs.[10]

7 These numbers assume that giving somebody $3.00 or less is excluding him from the coalition. Note that most of these are much closer to zero than three and that no one votes in favor of such shares. Thus any proposal where the two lowest shares are $3.00 or less are counted as ADZ, and if only the lowest is $3.00 or less, it counts as ASZ. The even strategy includes all proposals where the lowest amount is greater than $3.00. Half of the ADZ offers actually gave exactly zero to two members and 49% of the ASZ offers gave zero to one member. Increasing the cutoff would have only a minor impact on the categorization. For instance, changing the cutoff share from 0.12 to 0.16 would not change any of the categorization for the closed rule data while it would increase the fraction of ADZ by one percentage point and decrease both the ASZ and AE by half of one percentage point.

8 For more on belief based learning models, the interested reader is referred to Fudenberg and Levine (1999).

9 Other popular models include Experience-Weighted Attraction or EWA learning (Camerer and Ho, 1999) and reinforcement learning (Roth and Erev, 1995). This choice is not central to the argument that the estimator of the variance–covariance matrix should be adjusted. The adjustment that will be proposed would have similar effect on estimates from alternative learning models as long as the correlation across error terms is similar.

10 Cooper and Stockman (2002) show how fairness considerations might be relevant for learning models while Armantier (2006) shows how the evolution of play in ultimatum games with unequal endowments moves behavior away from the "fair"divisions observed at the beginning of his experiments. In the case of the current experiment, fairness considerations of the type modeled by Bolton and Ockenfels (2000) or Fehr and Schmidt (1999) are not likely to be the driving force for proposal behavior since there is a movement away from "fair"distributions of payoffs. Furthermore, this movement is not the result of efficiency considerations since the amount divided is unaffected by the distribution of shares. To see how efficiency might interact with fairness, see Charness and Rabin (2002). Note also that it is just as "easy"to make a "fair"offer in either treatment; if there exists an asymmetry from a fairness perspective, it is on the responder side.

Subjects have beliefs about the probability of each proposal being accepted or rejected. These beliefs are represented by weights assigned to a proposal being accepted or rejected. Defining $1\{\cdot\}$ as an indicator function which takes on the value 1 if the statement inside the bracket is true and 0 otherwise, the weight assigned to the acceptance (rejection) of a proposal of type j at time t by subject i is given by:

$$w_{ij}^{t}(c) = (1-\delta)\, w_{ij}^{t-1}(c) + 1\left\{v = c \text{ and } Y_i^t = j\right\}; \quad c \in \{a, r\} \tag{1}$$

where c is a dummy variable for accepted (a) or rejected (r), δ is a forgetting parameter that will be described in more detail later, and Y_i^t is the type of the proposal on the floor at time t in subject i's session. Thus the weight assigned to the acceptance (rejection) of a proposal of type j is equal to the weight assigned to its acceptance (rejection) in the previous round,[11] multiplied by $(1-\delta)$, to which you add 1 if the proposal on the floor this round was of type j and was accepted (rejected) and 0 otherwise. The forgetting parameter $\delta = 0$ if all past actions have equal importance, and if $\delta = 1$, only the last period of play matters. The former is sometimes referred to as fictitious play while the latter is known as Cournot learning or Cournot adjustment. Note that in the open rule, the type of proposal which is selected in the run-off election (if there is one) has the weight tied to its acceptance updated by one, and the defeated strategy's weight of rejection is updated by one. The probability assigned to the acceptance of a proposal of type j is the ratio of the weight put on it being accepted to the sum of the weight put on its acceptance and the weight put on its rejection; also known as the strength. Formally, subject i assigns probability μ that a proposal of type j will be accepted at time t:

$$\mu_{ij}^{t}(a) = \frac{w_{ij}^{t}(a)}{w_{ij}^{t}(a) + w_{ij}^{t}(r)}. \tag{2}$$

Subjects choose to propose a strategy of type j if that strategy maximizes their subjective expected utility which is equal to the utility they derive from a proposal of type j, that is u_j,[12] multiplied by the probability they assign to its acceptance at that time, $\mu_{ij}^{t}(a)$, plus the continuation value if that proposal is rejected, C_{ij}^{t}, and plus an error term $\lambda\varepsilon_{ij}^{t}$ where λ is a scaling factor that will be described shortly. Thus, subject i assigns to a proposal of type j at time t, given his beliefs about the probability of acceptance of such a proposal, the subjective expected value

$$EU_i^t\left(j \mid \mu_{ij}^{t}\right) = u_j\mu_{ij}^{t}(a) + C_{ij}^{t} + \lambda\varepsilon_{ij}^{t}. \tag{3}$$

In what follows, it will be assumed that the continuation value C is the same, at a given time and for a given individual, for all strategies, that is $C_i^t(j) = C_i^t(s)\ \forall j, s \in J$.[13] Thus, without loss of generality, subjective expected utility can be treated simply as

$$EU_i^t\left(j \mid \mu_{ij}^{t}\right) = u_j\mu_{ij}^{t}(a) + \lambda\varepsilon_{ij}^{t}. \tag{4}$$

For simplicity, $u_j\mu_{ij}^{t}(a)$ will be represented by $\overline{EU}_{ij}^{t}$. It is assumed that subjects choose j such that $EU_i^t\left(j \mid \mu_{ij}^{t}\right)$ is the greatest. The λ parameter controls for the level of use of available information. If $\lambda \to \infty$ then all strategies are equally likely, and as $\lambda \to 0$, the model approaches its deterministic version. Thus, if $\overline{EU}_{ij}^{t} - \overline{EU}_{ik}^{t}$ is re-written as $\overline{EU}_{i,jk}^{t}$ the probability that alternative j is selected at time t by subject i, denoted p_{ij}^{t}, is given by[14]

$$\begin{aligned} p_{ij}^{t} &= \mathrm{pr}\left(Y_i^t = j\right) \\ &= \mathrm{pr}\begin{bmatrix} \varepsilon_{i1}^{t} < \frac{1}{\lambda}\overline{EU}_{i,j1}^{t} + \varepsilon_{ij}^{t}, \ldots, \varepsilon_{ij-1}^{t} < \frac{1}{\lambda}\overline{EU}_{i,jj-1}^{t} + \varepsilon_{ij}^{t}, \\ \varepsilon_{ij+1}^{t} < \frac{1}{\lambda}\overline{EU}_{i,jj+1}^{t} + \varepsilon_{ij}^{t}, \ldots, \varepsilon_{iJ}^{t} < \frac{1}{\lambda}\overline{EU}_{i,jJ}^{t} + \varepsilon_{ij}^{t} \end{bmatrix} \\ &= \mathrm{pr}\left[\varepsilon_{ik}^{t} < \frac{1}{\lambda}\overline{EU}_{i,jk}^{t} + \varepsilon_{ij}^{t} \text{ for all } k \neq j\right]. \end{aligned} \tag{5}$$

Let $f\left(\varepsilon_{i1}^{t}, \ldots, \varepsilon_{iJ}^{t}\right) = f_i^t(\varepsilon)$ be the joint density function of ε_{ij}^{t}, and $F\left(\varepsilon_{i1}^{t}, \ldots, \varepsilon_{iJ}^{t}\right)$ the associated cumulative distribution function. Then

$$p_{ij}^{t} = \int_{-\infty}^{\infty}\int_{-\infty}^{\frac{1}{\lambda}\overline{EU}_{i,j1}^{t}+\varepsilon_{ij}^{t}} \cdots \int_{-\infty}^{\frac{1}{\lambda}\overline{EU}_{i,jj-1}^{t}+\varepsilon_{ij}^{t}} \tag{6}$$

$$\times \int_{-\infty}^{\frac{1}{\lambda}\overline{EU}_{i,jj+1}^{t}+\varepsilon_{ij}^{t}} \cdots \int_{-\infty}^{\frac{1}{\lambda}\overline{EU}_{i,jJ}^{t}+\varepsilon_{ij}^{t}} f_i^t(\varepsilon)\,\mathrm{d}\varepsilon_{iJ}^{t}\ldots\mathrm{d}\varepsilon_{i1}^{t} \tag{7}$$

$$= \int_{-\infty}^{\infty} F_j\begin{pmatrix} \varepsilon_{ij}^{t}, \frac{1}{\lambda}\overline{EU}_{i,j1}^{t} + \varepsilon_{ij}^{t} \ldots \frac{1}{\lambda}\overline{EU}_{i,jj-1}^{t} + \varepsilon_{ij}^{t}, \\ \frac{1}{\lambda}\overline{EU}_{i,jj+1}^{t} + \varepsilon_{ij}^{t}, \ldots \frac{1}{\lambda}\overline{EU}_{i,jJ}^{t} + \varepsilon_{ij}^{t} \end{pmatrix}\mathrm{d}\varepsilon_{ij}^{t}. \tag{8}$$

It is often assumed that ε_{ij}^{t} is distributed with a Type I extreme value function, which gives rise to the usual logistic form for the probabilities.[15] This paper will diverge from the usual assumption about the distribution of the error term by assuming that ε_{ij}^{t} is independently and identically distributed with standard normal distribution.[16] The PDF and CDF associated with the standard normal will be denoted ϕ and Φ respectively. Given this distribution

$$p_{ij}^{t} = \int_{-\infty}^{\infty}\phi\left(\varepsilon_{ij}^{t}\right)\begin{pmatrix} \Phi\left(\frac{1}{\lambda}\overline{EU}_{i,j1}^{t} + \varepsilon_{ij}^{t}\right)\ldots\Phi\left(\frac{1}{\lambda}\overline{EU}_{i,jj-1}^{t} + \varepsilon_{ij}^{t}\right) \\ \times\Phi\left(\frac{1}{\lambda}\overline{EU}_{i,jj+1}^{t} + \varepsilon_{ij}^{t}\right)\ldots\Phi\left(\frac{1}{\lambda}\overline{EU}_{i,jJ}^{t} + \varepsilon_{ij}^{t}\right) \end{pmatrix}\mathrm{d}\varepsilon_{ij}^{t} \tag{9}$$

$$= \int_{-\infty}^{\infty}\phi\left(\varepsilon_{ij}^{t}\right)\prod_{k\neq j}\Phi\left(\frac{1}{\lambda}\overline{EU}_{i,jk}^{t} + \varepsilon_{ij}^{t}\right)\mathrm{d}\varepsilon_{ij}^{t}. \tag{10}$$

Those familiar with Gauss–Hermite integration or with random effects probit/logit estimation (Butler and Moffitt, 1982) will notice that this is amenable to such an approximation technique. Using a change of variable, the above can be re-written as:

$$p_{ij}^{t} = \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{\frac{-\left(\varepsilon_{ij}^{t}\right)^2}{2}}\prod_{k\neq j}\Phi\left(\frac{1}{\lambda}\overline{EU}_{i,jk}^{t} + \varepsilon_{ij}^{t}\right)\mathrm{d}\varepsilon_{ij}^{t}$$

11 Clearly this requires weights at time 0, or priors, to get started. These priors will be estimated from the data, and they are constrained to be strictly positive.

12 Since utility will be treated simply as payoff, one can think of this as the amount subject i gives herself in a proposal of type j. For estimation, these proposal types will be categories which cover an interval of amounts to the proposer, and thus they are approximated by using the average amount proposers were offering themselves in proposals of that type.

13 This means that the continuation value of the game is perceived to be the same regardless of the strategy which is used in the first place. Although this is probably incorrect, it seems to be a good first approximation. For instance, the fact that subjects do not change the way they vote across the experiment suggests that they do not use the information they get to change their estimate of the continuation value of the game.

14 The derivation below draws heavily on Hausman and Wise (1978).

15 The logistic form would give rise to the following probabilities $p_{ij}^{t} = \frac{\exp\left(\frac{1}{\lambda}\overline{EU}_{ij}^{t}\right)}{\sum_J \exp\left(\frac{1}{\lambda}\overline{EU}_{ij}^{t}\right)}$.

16 This is sometimes referred to as the independent normal as opposed to the more general multivariate normal distribution. Although the normal distribution is used in many applications, it has not been used in the estimation of learning models. There is no obvious reason why the logistic distribution is a more appropriate choice for this application, and thus showing how one would use the normal instead might prove useful to researchers who want to compare the two.

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\left(r_{ij}^t\right)^2} \prod_{k \neq j} \Phi\left(\frac{1}{\lambda}\overline{EU}_{i,jk}^t + \sqrt{2} r_{ij}^t\right) dr_{ij}^t$$

$$\text{where} : r_{ij}^t = \frac{\varepsilon_{ij}^t}{\sqrt{2}}$$

$$\simeq \frac{1}{\sqrt{\pi}} \sum_{m=1}^{M} w_m \prod_{k \neq j} \Phi\left(\frac{1}{\lambda}\overline{EU}_{i,jk}^t + \sqrt{2} a_m\right)$$

$$\text{where} : w_m \text{ and } a_m \text{ are the quadrature weights and quadrature abscissas.} \quad (11)$$

The quadrature weights and quadrature abscissas for Gauss–Hermite integration are available in Abramowitz and Stegun (1972).

The pseudo-likelihood can be written as

$$d_{ij}^t = 1\{Y_i = j \mid t\}$$

$$\ln L = \sum_t \sum_i \sum_j d_{ij}^t \ln p_{ij}^t. \quad (12)$$

5.1. Estimation of the variance–covariance matrix

The data used in learning models involves tracking the same individuals over time. It is a well-known fact that neglecting to adjust the variance when using clustered samples tends to lead to downward bias in the estimated variance. Discussion of this problem and of ways to correct for it can be found in many fields and applications: for example Altonji (1986) and Ham (1986) in labor economics, Moulton (1986, 1990)) on the use of aggregate variables on micro data, Beck and Katz (1995) in political science, Deaton (1997) on the use of survey data, and Ham et al. (2005) in experimental economics. These examples mainly address problems and cases which do not use maximum likelihood, but methods proposed in these papers can be extended to maximum likelihood estimation,[17] although the interpretation of the results may be slightly different.[18] The way in which the interpretation differs will be discussed below, after giving more specific details about this correction.

The variance–covariance calculation employed allows observations for a given individual to be correlated with each other.[19] The variable we are solving for is $\hat{\theta} \equiv \arg\max_\theta \ln L(\theta; Y)$. Let $\widehat{V}(\theta) = \left(\frac{-\partial^2 \ln L}{\partial \theta^2}\right)^{-1}$, i.e. $\widehat{V}(\theta)$ is the conventional (or standard) estimate of the variance–covariance matrix of the parameter estimate θ. Suppose there are I individuals, then the adjusted variance–covariance is estimated by

$$\widehat{AV}(\theta) = \left(\frac{I}{I-1}\right) \widehat{V}(\theta) \left(\sum_{i=1}^{I} u_i u_i'\right) \widehat{V}(\theta)$$

$$\text{where} : u_i = \frac{\partial \ln L_i}{\partial \theta}.$$

In other words, u_i is the contribution of individual i to the vector of scores, and the first term simply is a finite sample adjustment. Thus, the variance–covariance of the error term – stacked by individuals – is a block diagonal matrix.[20] $\widehat{AV}(\theta)$ is asymptotically equivalent to $\widehat{V}(\theta)$ if u_i is i.i.d. and is a consistent estimator of the variance if u_{it} is serially correlated with $Eu_iu_i' = \Omega$.

An alternative to using the above adjustment is to assume a certain form for the unobserved heterogeneity, construct the appropriate likelihood function, and maximize it. I adopt this procedure below, and show that it has a significant effect on the inferences one draws for the parameter estimates.[21]

6. Simplifications

When discretizing the continuous proposal space it is inevitable that arbitrary decisions have to be made. However, the robustness of the results to these choices can, and will, be tested by comparing the results to sensible alternatives.[22] The goal when choosing how to discretize is to choose a parsimonious representation that is meaningful. Indeed, each additional proposal type adds one parameter to be estimated, and could easily lead to overfitting. At the same time, there must be enough proposal types to capture the features of the data that are observed. One crucial aspect that will be exploited is the fact that patterns of play are virtually identical at the beginning of the experiment under both treatments. This suggests that subjects have similar priors when they begin playing. This observation, combined with the fact that both games have exactly the same proposal space, allows the priors of the learning model to be constrained so that they are the same in a meaningful way. The robustness of the results to these simplifications will be assessed by re-estimating the model under alternative simplifications (Section 8).

The simplifications are the following. First, the initial weight of each proposal type is the same for both the closed and open rules. In a section on robustness (Section 8) that follows the estimation, it will be shown that the main results are not affected if this restriction is relaxed.

Second, the original choices are $\left(x_1^i, x_2^i, x_3^i, x_4^i, x_5^i\right)$ where $x_j^i \in [0, 1]\, \forall i, j$, $\sum_{j=1}^{5} x_j^i \le 1 \forall i$, x_j^i is the share assigned by proposer i to player j.[23] In what follows, superscripts are omitted for convenience. These proposed share allocations will be divided into nine basic proposal types in the following way. Since this paper is interested in the evolution of proposals that exclude a different number of subjects under each rule, there will be AE, ASZ, and ADZ proposal types. These are defined by the values of x_4 and x_5. Furthermore, conditional on the proposal being of the AE, ASZ, or ADZ types, there will be three cases: one where the proposed average share to coalition members is high, one where it is average, and one where it is low. These are defined by the values of x_2 and x_3 for ADZ proposals. It also includes x_4 for ASZ proposals, and it includes all but x_1 in the case of AE proposals. These are the two elements that the simplification of the proposal space tries to preserve: the number of subjects included and the generosity of the offer to those included. More specifically, a proposal is considered AE if x_4 and x_5 are greater than 0.12, ASZ if $x_5 \le 0.12$ and $x_4 > 0.12$, and ADZ if x_4 and x_5 are both smaller than or equal to 0.12. To determine if a proposal is high, medium, or low, the average coalition share is computed, i.e., $\frac{x_2+x_3+x_4+x_5}{4}$ if a proposal

17 For instance, Sakata (2002) discusses such estimator in a quasi-maximum likelihood environment and points out that these are relevant for very common data sources such as the Current Population Survey, the Panel Study of Income Dynamics, and the Health and Retirement Study. The complex sample designs used in the construction of these data sets imply a violation of the common i.i.d. assumption.

18 For references on such variance estimators applied to pseudo-likelihoods, see Binder (1983), Greene (2000), pp. 480–491, 823–824, Judge et al. (2000), Skinner (1989a,b), and StataCorp (2001).

19 Its origin may be traced back to Huber (1967) and a more recent discussion can be found in White (1994).

20 This is, in fact, the sandwich estimator where the data from a given individual is treated as a super-observation.

21 Recent papers studying individual heterogeneity in the context of a different learning model (EWA) are Cabrales and Garcia-Fontes (2000) and Wilcox (2006).

22 Discretization is used in virtually all studies of learning with continuous choice space. For work on estimation without discretization see Armantier (2004).

23 Proposals will always be presented with the highest share to the left and the lowest to the right.

Table 2
Proposal types.

	Average share offered to coalition members		
	First tier	Second tier	Third tier
AE	*Proposal Type 1* ≤ 0.19 (16)	*Proposal Type 2*(0.19, 0.2] (20)	*Proposal Type 3* > 0.2 (2)
ASZ	*Proposal Type 4* ≤ 0.227 (9)	*Proposal Type 5*(0.227, 0.24] (9)	*Proposal Type 6* > 0.24 (8)
ADZ	*Proposal Type 7* ≤ 0.28 (17)	*Proposal Type 8*(0.28, 0.3] (9)	*Proposal Type 9* > 0.3 (10)

Numbers in parentheses are the percentage of play falling in each category.

is AE, $\frac{x_2+x_3+x_4}{3}$ if a proposal is ASZ, and $\frac{x_2+x_3}{2}$ if a proposal is ADZ. The value of the observations for the $33\frac{1}{3}$ and $66\frac{2}{3}$ percentiles are obtained, and these determine the cut-offs for the three tiers. This is presented in Table 2. Let us consider a specific example to make this clearer. For ASZ, the average share offered to coalition members corresponding to the $33\frac{1}{3}$ percentile is 0.227. Thus one proposal type will be all proposals where one person is excluded from the division of benefits, and the other three people included in the coalition receive an average share of 0.227 or less. Looking only at the average share offered is a clear simplification and could raise some concerns, but this does not seem to be much of a problem. Looking at the difference between the highest share and the lowest share offered to coalition members indicates that both the mode and median difference is zero. More specifically, 59% of offers allocate exactly the same amount to all members included in the coalition (the difference between highest and lowest share is zero). The average is slightly above zero, more exactly 0.03, which represents about 75¢ in round one. To increase confidence in the results, an alternative specification with a finer division of the proposal space is also estimated. Those results will be presented in the robustness section.

There is a third restriction also meant to reduce the number of parameters. Theoretically, belief based learning does not impose a restriction on the strengths, the sum of the weights on acceptance and rejection, of the beliefs across proposal types (we will denote the strength assigned to proposal type j by subject i at time t by s_{ij}^t). More specifically they do not have to be the same.[24]However, they have been constrained to be equal for the first election to make the optimization problem more manageable.

As previously mentioned, the priors for both the open and closed rules are assumed to be the same. This simplification is motivated by the observation that patterns of play are nearly identical at the start of the experiment. Thus, throughout this paper, the main parameter of interest is δ and to a lesser extent λ. δ defines the learning process: Cournot ($\delta = 1$), fictitious play ($\delta = 0$), or a mixture, while λ gives us a measure of how well subjects use the information available to them. In the present case λ for the closed and open rules should be comparable as the decision space and payoffs are the same under both rules.

7. Estimation results

Before examining the results, let us formulate the hypotheses of interest. Remember that the original hypothesis that learning might explain the data comes from the fact that for very similar proposals (of the ADZ type), subjects receive very different feedback under the two treatments. Under the closed rule, these are passed all the time, whereas under the open rule, they are amended half the time. Thus I will formulate and test two hypotheses, ranked in order of importance for parameter stability. First, I test if the same forgetting parameter can account for behavior in both treatments, that is:

Table 3
Summary of maximum likelihood estimates with unadjusted (in parentheses) and adjusted [in brackets] standard errors.

Belief based learning	
δ closed	0.258
	(0.043)
	[0.087]
δ open	0.151
	(0.027)
	[0.044]
λ closed	0.489
	(0.060)
	[0.095]
λ open	0.510
	(0.088)
	[0.193]
Log Lik.	−1214.640

$$H_0 : \delta \text{ closed} = \delta \text{ open}. \quad \text{(H1)}$$

Remember that δ defines the learning model (Cournot vs. fictitious play), and thus this is the main test that determines if the hypothesis of a unique learning model can be rejected. Second, it would also be interesting to know if subjects under both closed and open rules use information similarly (in addition to using the same forgetting parameter). Thus, a single learning model with the same level of use of information can account for the paths of play. This gives rise to hypothesis 2:

$$H_0 : \delta \text{ closed} = \delta \text{ open}, \qquad \lambda \text{ closed} = \lambda\text{open}. \quad \text{(H2)}$$

Note that the λ's cannot always be compared across games, but since the proposal space is constrained to be the same, they should be in this case. Let us now turn to the actual estimates.

Table 3 presents the estimation results for the important parameters.[25] First note that estimates of λ are very close. The δ's are further apart but still reasonably close with a difference of about ten percentage points. Second, the estimates of δ are relatively low, implying a learning process close to fictitious play. Third, the adjustment to the variance–covariance has a significant impact: on average it increases the standard errors by an average factor between 1.5 and 2.0.

H1 and H2 are tested using Wald tests. In both cases, the null hypothesis cannot be rejected (p-values are 0.254 and 0.195 for H1 and H2 respectively). It is noteworthy however, that both H1 and H2 would have been rejected if it was not for the adjustment of the variance–covariance matrix. The p-values for these tests are less than 1% without the correction.

An additional hypothesis that is suggested by the low values of δ's is that the learning process observed is fictitious play. Let us

24 That is, in general, $s_{ij}^t \equiv w_{ij}^t(a) + w_{ij}^t(r) \neq w_{ik}^t(a) + w_{ik}^t(r)$ for $k \neq j$ and for any t.

25 The complete results can be found in the Appendix. The estimates were obtained using the Matlab constrained optimization routine. Other programs that have been used to compute results reported in this paper are written by Mario J. Miranda and Paul L. Fackler as part of the COMPECON toolbox (http://www4.ncsu.edu/~pfackler/compecon/) and m-files written by James P. LeSage as part of the Econometrics toolbox (http://www.spatial-econometrics.com/). Other programs to perform hypothesis testing can be found at http://homepages.nyu.edu/~gf35/html/code.htm.

refer to this hypothesis as H3, that is:

$$H_0 : \delta \text{ closed} = \delta \text{ open} = 0. \quad \text{(H3)}$$

That hypothesis, however, is rejected.

Analysis of the results so far reveals some interesting conclusions. First, the use of a simple adaptive learning model can generate different paths of play under closed and open amendment rules that are consistent with what is observed in the experiment. Second, adjusting the variance–covariance matrix to account for the panel structure of the data affects the conclusions reached. Finally, in this experiment, once the panel structure of the data is taken into account, the hypothesis that a single model can account for the observed proposals cannot be rejected. Furthermore, that single model that best fits the data is very close to fictitious play.

8. Robustness of the results

One could worry that the results presented so far are driven by the specifics of the estimation approach chosen here. Prior studies of learning have sometimes used other estimation methods. For instance, it is somewhat common to assume uniform priors over the strategies (proposal types in this case).[26] Others "burn-in"the priors using the frequency of play from the early part of the experiment.[27] One could be concerned that the effects of $\widehat{AV}(\theta)$ instead of $\widehat{V}(\theta)$ are fragile to such estimation choices, or that estimates can change enough that results from hypothesis tests be affected. Table 4 reports estimates when priors are either burned-in or assumed uniform and the p-values for the hypothesis tests H1, H2 and H3. When priors are burned-in, the frequency of play in the first two periods is used. Clearly the effect of using $\widehat{AV}(\theta)$ instead of $\widehat{V}(\theta)$ is unchanged, namely the standard errors are increased.[28] Once again the adjustment affects the results. Specifically, H2 would be rejected if $\widehat{V}(\theta)$ was used. Note however, that for the case of uniform priors, H2 is rejected even when using $\widehat{AV}(\theta)$. Also, the fact that the estimates are very similar to the ones reported in Table 3 suggests that the results are robust. For instance, the δ parameters are still small and close in value but big enough to reject the hypothesis of fictitious play. Similarly, the λ's are small and close in value. Consequently, we cannot reject the hypothesis that a unique learning model fits the evolution of play in both of these treatments, at least in terms of δ.

Other concerns about the robustness of the results might pertain to the validity of the simplifications. For instance, the specific proposal types might affect the results. That is, the division could be too coarse and miss some important features. To test this, the model is re-estimated with a division of the share offered in quartiles instead of tiers. This generates eleven proposal types and is described in Table 5.[29] Another simplification which can cause concerns is that of equal priors for the same proposal types in the open and closed amendment rules. To test this, the model is re-estimated with different strength and priors for each treatment. This also allows us to directly test the hypothesis of the equality of priors. Those results are reported in Table 6. For the case with different priors for each amendment rule, H1, H2, and H3 are tested jointly with the equality of priors. Table 6 also reports the results of the test on the equality of priors by itself. Once again, using $\widehat{AV}(\theta)$ instead of $\widehat{V}(\theta)$ has a similar effect: standard errors are greater,

Table 4
Summary of maximum likelihood estimates with unadjusted (in parentheses) and adjusted [in brackets] standard errors.

	Uniform priors	Burned-In priors
δ closed	0.206	0.221
	(0.055)	(0.049)
	[0.108]	[0.098]
δ open	0.194	0.185
	(0.032)	(0.032)
	[0.046]	[0.049]
λ closed	0.394	0.338
	(0.051)	(0.038)
	[0.083]	[0.063]
λ open	0.282	0.268
	(0.040)	(0.040)
	[0.074]	[0.088]
Log Lik.	−1275.546	−1278.986
H1	(0.842)	(0.496)
	[0.909]	[0.694]
H2	(0.001)	(0.005)
	[0.034]	[0.137]
H3	(0.000)	(0.000)
	[0.000]	[0.001]

For H1, H2, and H3 unadjusted p-values are in Parentheses and adjusted p-values in square brackets.

and this affects the interpretation of the data. Clearly, the equality of priors cannot be rejected. It is also clear that changing the way the proposal types are categorized does not affect the qualitative results. As with the other estimates, a unique model to explain the data cannot be rejected at any conventional level. The only noticeable difference is that H3 is not rejected at the 10% level when different priors are allowed, but this is mainly driven by the fact that here H3 is tested jointly with the assumption of equal priors.

Hence, the results are robust to many alternative specifications. This is not to say that the estimates do not vary; the estimates of λ, for instance, are much higher in the specification with 11 proposal types than in the others. What is meant rather is that in most cases, the conclusions of hypothesis tests are unaffected by the details of the specification. In all cases using $\widehat{AV}(\theta)$ instead of $\widehat{V}(\theta)$ increases the standard errors and affects some of the results of hypothesis tests. Overall, using the available data and the typical estimator for such learning models with adjusted variance–covariance matrix, it seems safe to say that a unique learning model is not rejected by the data, and that this model has low δ's, although not low enough to be exactly fictitious play. The low δ's imply that past outcomes, not just the most recent ones, affect beliefs.

9. An efficient estimator

The adjustment to the variance–covariance matrix discussed in this paper, although it provides for a consistent estimator of the variance–covariance matrix of the usual estimator, it does not generate efficient estimates. Hence, the observation that we can no longer reject the null hypothesis that the parameters are the same for the closed and open rules could be the result of not having an efficient estimator. In this section, a way to specify the problem which would account for within subject correlation and be amenable to estimation is proposed. This will clearly come at the cost of more restrictive assumptions on the data generating process.

Two of the simplifications we have made are to assume every subject has the same priors and that the strength (the sum of the weight on acceptance and rejection) is originally the same for every possible choice. Thus, the strength is the same for every choice and every subject before the game begins (remember that $s_{ij}^t \equiv w_{ij}^t(a) + w_{ij}^t(r)$ and this meant $s_{ij}^0 = s_j^0 \quad \forall i$). Instead, we could allow the strength to vary across subjects at

26 See for instance Erev and Roth (1998).

27 See for instance Ho et al. (2002).

28 This is not very surprising as only negative correlation within cluster would lead to smaller standard errors.

29 The value of the 50th percentile observation is the same as the 75th percentile for the AE strategy. Hence, only one type is defined for these two quartiles.

Table 5
Proposal types.

Coalition type	Average share offered to coalition members			
	1st quartile	2nd quartile	3rd quartile	4th quartile
AE	$\leq$ 0.180 (10)	(0.180–0.200] (26)	> 0.200 (2)	
ASZ	$\leq$ 0.220 (7)	(0.220–0.240] (11)	(0.240–0.243] (2)	> 0.243 (6)
ADZ	$\leq$ 0.260 (11)	(0.260–0.300] (16)	(0.300–0.320] (7)	> 0.320 (3)

Numbers in parentheses are the percentage of play falling in each category.

Table 6
Summary of maximum likelihood estimates with unadjusted (in parentheses) and adjusted [in brackets] standard errors.

	11 Proposal types	Different priors across treatments
δ closed	0.131	0.259
	(0.029)	(0.061)
	[0.046]	[0.128]
δ open	0.130	0.111
	(0.014)	(0.037)
	[0.025]	[0.058]
λ closed	0.969	0.422
	(0.132)	(0.048)
	[0.230]	[0.076]
λ open	0.790	0.431
	(0.081)	(0.137)
	[0.142]	[0.277]
Log lik.	−1251.946	−1193.145
H1 (and	(0.979)	(0.092)
Equal priors)	[0.990]	[0.849]
H2 (and	(0.070)	(0.017)
Equal priors)	[0.370]	[0.573]
H3 (and	(0.000)	(0.000)
Equal priors)	[0.000]	[0.088]
Equal priors	*n/a*	(0.561)
	n/a	[0.991]

For hypothesis tests unadjusted *p*-values are in parentheses and adjusted *p*-values in square brackets.

the beginning, but put some structures on these differences. We will assume that $s_i^0 = s^0 + \varsigma_i^0$, and ς_i^0 has a lognormal distribution, i.e., $\ln\left(\varsigma_i^0\right) \sim N(\mu_s, \sigma_s)$. Now denoting the joint likelihood for the data $\Pr\left(Y \mid \delta, \lambda, w_1^0(a), \ldots, w_J^0(a), s^0, \mu_s, \sigma_s\right) = f\left(Y \mid \delta, \lambda, w_1^0(a), \ldots, w_J^0(a), s^0, \mu_s, \sigma_s\right)$ and the pdf of ς_i^0 by h; using the assumed distribution of ς_i^0 we obtain:

$$\begin{aligned}
&\Pr\left(Y \mid \delta, \lambda, w_1^0(a), \ldots, w_J^0(a), s^0, \mu_s, \sigma_s\right)\\
&= \int_{s_i^0 \in \mathbb{R}^+} f\left(Y_i \mid \delta, \lambda, w_1^0(a), \ldots, w_J^0(a), s^0, \varsigma_i^0\right)\\
&\quad \times h\left(\varsigma_i^0 \mid \mu_s, \sigma_s\right) \partial \varsigma_i^0\\
&= \prod_{i=1}^{N} \int_{s_i^0 \in \mathbb{R}^+} l\left(Y_i \mid \delta, \lambda, w_1^0(a), \ldots, w_J^0(a), s^0, \varsigma_i^0\right)\\
&\quad \times \frac{\exp\left(-\frac{\left(\ln \varsigma_i^0 - \mu_s\right)^2}{2\sigma_s^2}\right)}{\varsigma_i^0 \sigma_s \sqrt{2\pi}} \partial \varsigma_i^0\\
&= \prod_{i=1}^{N} \int_{s_i^0 \in \mathbb{R}^+} \prod_{t=1}^{T} l\left(Y_i^t \mid \delta, \lambda, w_1^0(a), \ldots, w_J^0(a), s^0, \varsigma_i^0\right)\\
&\quad \times \frac{\exp\left(-\frac{\left(\ln \varsigma_i^0 - \mu_s\right)^2}{2\sigma_s^2}\right)}{\varsigma_i^0 \sigma_s \sqrt{2\pi}} \partial \varsigma_i^0
\end{aligned}$$

Table 7
Summary of maximum likelihood estimates with standard errors (in parentheses).

Belief based learning	
δ closed	0.269
	(0.053)
δ open	0.160
	(0.040)
λ closed	0.498
	(0.100)
λ open	0.411
	(0.127)
μ_s	0.980
	(1.886)
σ_s	2.834
	(1.349)
Log Lik.	−1207.219

$$\begin{aligned}
&= \prod_{i=1}^{N} \int_{s_i^0 \in \mathbb{R}^+} \prod_{t=1}^{T} l\left(Y_{ij}^t \mid \delta, \lambda, w_1^0(a), \ldots, w_J^0(a), s^0, \varsigma_i^0\right)^{d_{ij}^t}\\
&\quad \times \frac{\exp\left(-\frac{\left(\ln \varsigma_i^0 - \mu_s\right)^2}{2\sigma_s^2}\right)}{\varsigma_i^0 \sigma_s \sqrt{2\pi}} \partial \varsigma_i^0,
\end{aligned} \quad (13)$$

where $l\left(Y_{ij}^t \mid \delta, \lambda, w_1^0(a), \ldots, w_J^0(a), s^0, \varsigma_i^0\right)$ is given by (11). Substituting $\varsigma_i^0 = \exp\left(\sqrt{2}\sigma_s S_i + \mu_s\right)$ in (13) results in an expression also amenable to Gauss–Hermite integration as is shown below (the $\prod_{i=1}^{N}$ is ignored).

$$\begin{aligned}
&\int_{s_i^0 \in \mathbb{R}^+} \prod_{t=1}^{T} l\left(Y_{ij}^t \mid \delta, \lambda, w_1^0(a), \ldots, w_J^0(a), s^0, \varsigma_i^0\right)^{d_{ij}^t}\\
&\quad \times \frac{\exp\left(-\frac{\left(\ln \varsigma_i^0 - \mu_s\right)^2}{2\sigma_s^2}\right)}{\varsigma_i^0 \sigma_s \sqrt{2\pi}} \partial \varsigma_i^0\\
&= \frac{1}{\sqrt{\pi}} \int_{s_i^0 \in \mathbb{R}^+} \mathrm{e}^{-(S_i)^2} \prod_{t=1}^{T} l\left(Y_{ij}^t \mid \delta, \lambda, w_1^0(a), \ldots, w_J^0(a), s^0, S_i\right)^{d_{ij}^t} \partial S_i\\
&\simeq \frac{1}{\sqrt{\pi}} \sum_{m=1}^{M} w_m \prod_{t=1}^{T} l\left(Y_{ij}^t \mid \delta, \lambda, w_1^0(a), \ldots, w_J^0(a), s^0, S_i\right)^{d_{ij}^t},
\end{aligned}$$

where as before w_m and a_m are the quadrature weights and quadrature abscissas. Taking the natural logarithm of (13) allows us to sum over the N subjects.

Table 7 presents the estimation results for the important parameters using the same proposal types as for the results presented in Table 3.[30] The loglikelihood is substantially improved by allowing for random effects in the prior, and as a consequence, using a likelihood ratio test, we can firmly reject the hypothesis that there is no heterogeneity in priors.

Comparing these estimates of δ and λ to those of Table 3 indicates that both δ's change slightly with a small increase in the

30 The complete results can be found in Table 12 of the Appendix.

Table 8
Rate of type I and type II errors for simulated data.

		No random effects		With random effects
		Unadjusted	Adjusted	
Type I	H1	0.299	0.177	0.273
	H2	0.424	0.371	0.402
Type II	H1	0.606	0.765	0.648
	H2	0.508	0.566	0.546

difference between the two. λ changes more, and the difference between them increases significantly. For the four parameters of interest, the standard errors are greater than the unadjusted standard errors of Table 3 while three out of four parameters have smaller standard errors than the adjusted ones from Table 3. As for our main hypotheses H1 and H2: they can all be rejected. Similarly, H3 can be rejected.

Although improved efficiency is desirable, the impact of using this more efficient estimator is not clear in small samples. To illustrate the impact of small samples, a Monte Carlo experiment is performed. The simulated data is generated under the assumptions that the learning model is correctly specified (including the simplifications described in Section 6) and that the distributional assumptions on the error terms in the estimation are correct. Two simulated data sets are generated: one with the parameter values from the estimation with random effects (reported in Table 7), and one with the δ_{closed} and δ_{open} as well as the λ_{closed} and λ_{open} set equal to the mid-point between the values of those same estimates. The simulation results reported will rely on 1000 simulated experiments of the size of those conducted in the laboratory. Using those, the rate of Type I and Type II errors when testing H1 and H2 at the 5% level are computed for each estimator. The results of those simulations are reported in Table 8.[31]

The first observation is that adjusting the variance–covariance matrix has the expected effect and reduces the rate of Type I error. The second observation is that, as expected, the estimator with random effects improves the rate of Type II error. On the other hand, this comes at the cost of important increases in Type I error. Another observation is that the rates of Type I error are all higher than 5%. This can be substantially improved if the criterion for rejection is extended to require rejection also with alternative specifications of the priors, that is uniform and burned-in priors (as explored in Section 8). In that case, that is if the criterion for rejection is that the hypothesis must be rejected for all three specifications, the rates of Type I error fall to 0.04 and 0.16 for H1 and H2 respectively using the adjusted variance–covariance matrix (0.04 and 0.20 if one uses the unadjusted). However this comes at the expense of extremely high rates of Type II error: 0.94 and 0.77 for H1 and H2 respectively. One possibility would be to use a much more restrictive "nominal"significance level that yields an actual rate of type I error of 5%. However, for this particular simulation, this does not seem promising as one would need to use a "nominal"significance level much below 1% to achieve the desired rate of type I errors.

Clearly one should not attribute too much to the specific numbers from this one Monte Carlo experiment as the parameters would vary depending on the assumptions on the underlying parameter estimates. For instance, if δ's and λ's were assumed to be further apart, the rate of Type I and Type II errors would improve. Nonetheless, they do highlight the trade-offs at stake and the limitations imposed by the sample size.

10. Discussion of the results

The overall evolution of proposals can be divided in three components. First, there is the fact that in election 1 proposals are similar under both amendment rules, and these involve mostly offers for supermajorities (AE offers). Second, under the closed rule, proposals of the ADZ type become more and more common, and by the end they represent the majority of proposals. Third, under the open rule, there is a slight increase in the number of ASZ proposals, and there is definitely no gain in popularity of the ADZ kind of proposals. The fact that behavior in election 1 is almost identical under both amendment rules is reminiscent of results in bargaining with no or few repetitions. One such result is the near equal splits in Ultimatum game experiments. In the Ultimatum game, a subject proposes (roles are usually randomly assigned) to another a division of a certain amount of money. That subject can either accept or reject that offer, if he accepts, the division is implemented, otherwise, they both receive nothing. The subgame perfect Nash equilibrium of this game is for the proposer to offer (almost) nothing to the other subject, and for that subject to accept the offer. However, typical offers are in the 40% to 50% range, and offers of low positive shares are often rejected (see Roth (1995) for a review of the literature). Binmore et al. (1985) offer this explanation: "subjects, faced with a new problem, simply choose "equal division"as an "obvious"and "acceptable"compromise."[32] They go on to say "We suspect [...] that such considerations are easily displaced by calculations of strategic advantage, once players fully appreciate the structure of the game."[33] This is, to a certain extent, what is observed in this experiment. Namely, after starting with obvious offers in election 1, subjects learn the strategic interactions at play under each rule, and this leads them, at least under the closed rule, to make very different offers by the end. Others have argued that individuals "have difficulty incorporating the implications of the committee's decision rule into their allocation decisions."[34] Messick et al. (1997) study a multiparty Ultimatum game (six subjects per group) where the proposal[35] needs either to be accepted by one of the five responders or by all of the five responders. Subjects played the game only once. They find that offers are unaffected by the different rules.[36] This is once again similar to the proposal behavior observed in election 1. This view suggests that as they repeat the experiment, subjects are learning, or understanding, the rules of the game through the feedback they obtain. Given the nature of the games, they obtain very different feedback, and this is why they make different proposals under both treatments at the end of the fifteen elections.

One question that arises then is whether a learning model can account for such evolution, and simply because of the variations in feedback, would the two rules lead to different paths of play. The question of parameter stability has received much attention recently in the learning literature. As pointed out by Erev and Haruvy (2001, 2005), statistical evidence at least dating back

31 The results of the simulations seem fairly stable, even after only 200 repetitions, the results are very similar to what is reported in Table 8. For instance, results after 200 simulated experiments for the estimator without random effects are the following: Type I error rates for H1 are 0.29/0.21 (unadjusted/adjusted), and 0.41/0.34 for H2; Type II error rates for H1 are 0.60/0.72 and 0.52/0.52 for H2. The differences between those numbers and the ones reported in Table 8 do not have an important effect on how the results would be interpreted.

32 Binmore et al. (1985), p. 1180.

33 Binmore et al. (1985), p. 1180.

34 Messick et al. (1997), p. 87.

35 They had two treatments, one where proposals need to allocate the same amount to everyone but to the proposer himself and one where there was no such restriction.

36 In the treatments where proposals to others do not have to be the same, they do find a statistically significant effect of the rules, but it is very small in size. Thus they claim "Neither study managed to produce differences in allocation decisions."p.97.

to Stahl (1996) suggest that parameters cannot be pooled across games. In the case of this experiment, we show that correcting the variance–covariance to account for the panel structure of the data affects the conclusions one would reach about parameter stability. However, when we account directly for this by allowing heterogeneity in priors, both hypotheses can be rejected. This suggests that to account for the evolution of play in this experiment, a different model for each treatment is needed, or in other words, that the different feedbacks are not, in themselves, enough to generate the different evolution of play.

11. Conclusion

Experiments on multilateral bargaining with open and closed amendment rules show that play under the two treatments began with similar choices but evolved into strikingly different types of choices over time. This clearly suggests adaptive behavior on the part of subjects. This view is confirmed by the fitting of a simple belief based learning model to the data. Furthermore, it is shown that if the variance–covariance matrix is appropriately adjusted, the hypothesis that the results from the two different games can be explained by a unified learning model, a model with the same key parameters, cannot be rejected. Such learning, although it lies somewhere between Cournot and Fictitious play, is very close to the latter. However, if the problem is tackled more directly by allowing different subjects to have different priors, then the hypothesis of a unique learning model can be rejected.

More generally, this paper has confirmed what has now been shown repeatedly, namely that learning models can help explain experimental data and data on bargaining experiments in particular. However, when performing hypothesis tests, one needs to account for the fact that such data involves repeated measures for the same individuals.[37] A way to perform such adjustments is suggested. One drawback of such an adjustment, which is demonstrated using simulations, is a decrease in power when performing hypothesis tests. This result extends to other applications where these types of adjustments, sometimes referred to as clustering, are performed.

Another solution which accounts for the source of this problem more directly, i.e., different priors, is also proposed. Although it is more efficient if correct, it is less robust than the first approach described in the paper, namely adjusting to the variance–covariance matrix. For this particular application, it is also shown that the efficient estimator results in high rates of Type I error due to the relatively small sample size typical of experimental data sets.

From a broader point of view, these results suggest that the previously observed variability in parameters might be overstated. Nonetheless, in the end, we also reject the hypothesis of a single learning model. One of the key disappointments of learning models thus far has been that every game seems to require different parameter values, even for similar games.[38] If learning models capture something about how individuals adapt to their environment and the feedback they get from their actions, then one should not need to rely on different parameter values for very similar games. This suggests at least two explanations. Either the two games are more different then they appear on the surface, or the model is not appropriately specified in terms of the "true"underlying structural parameters. For instance, it may be that δ and λ should be functions that depend on the variance in the feedback subjects receive. Attempts in that direction have been made, for instance, by Ho et al. (2002).

[37] This is true even though the current experiment has a relatively low (by experimental standards) number of observations per subject. As the number of observations per subject is increased we can expect that the (downward) bias in the standard errors would become even more important (see Pepper (2002)).

[38] I am not suggesting that parameters should be the same for all possible cases. However one would expect that for similar tasks and environments, parameters should not vary.

Table 9
Maximum likelihood estimates.

	Estimated priors	Uniform priors	Burned-In priors
δ closed	0.258	0.206	0.221
	[0.087]	[0.108]	[0.098]
δ open	0.151	0.194	0.185
	[0.044]	[0.046]	[0.049]
λ closed	0.489	0.394	0.338
	[0.095]	[0.083]	[0.063]
λ open	0.510	0.282	0.268
	[0.193]	[0.074]	[0.088]
s^0	22.520	12.723	11.204
	[17.582]	[7.990]	[8.354]
$w_1^0(a)$	8.860		
	[9.307]		
$w_2^0(a)$	17.563		
	[14.186]		
$w_3^0(a)$	8.972		
	[11.681]		
$w_4^0(a)$	8.399		
	[8.190]		
$w_5^0(a)$	9.703		
	[9.085]		
$w_6^0(a)$	10.066		
	[11.468]		
$w_7^0(a)$	7.821		
	[7.010]		
$w_8^0(a)$	4.243		
	[5.793]		
$w_9^0(a)$	6.265		
	[6.844]		

Adjusted standard errors in square brackets.

Table 10
Maximum likelihood estimates of the priors.

	11 proposal types
s^0	33.976
	[1.289]
$w_1^0(a)$	10.809
	[1.556]
$w_2^0(a)$	23.863
	[2.167]
$w_3^0(a)$	21.633
	[3.725]
$w_4^0(a)$	12.092
	[0.696]
$w_5^0(a)$	15.498
	[1.973]
$w_6^0(a)$	9.557
	[2.410]
$w_7^0(a)$	17.845
	[3.071]
$w_8^0(a)$	9.993
	[1.397]
$w_9^0(a)$	10.743
	[1.073]
$w_{10}^0(a)$	10.126
	[1.458]
$w_{11}^0(a)$	8.210
	[1.949]

Adjusted standard errors in square brackets.
This table completes the results reported in Table 6.

Acknowledgements

This paper would have never been possible without the help, support, and suggestions of John Kagel. It also benefited from the comments of the editor, an associate editor and three referees as

Table 11
Maximum likelihood estimates of the priors.

Different priors across treatments	Closed	Open
Strength	23.056	10.563
	[23.312]	[12.566]
PT 1	5.639	3.684
	[6.543]	[6.067]
PT 2	15.103	7.326
	[17.178]	[8.090]
PT 3	0.807	1.164
	[6.668]	[11.937]
PT 4	5.724	3.405
	[6.949]	[5.186]
PT 5	0.000	3.929
	[0.516]	[5.561]
PT 6	0.000	3.892
	[1.358]	[6.738]
PT 7	7.579	1.000
	[10.316]	[4.508]
PT 8	3.113	0.000
	[4.375]	[4.181]
PT 9	5.114	0.800
	[6.104]	[4.450]

PT stands for proposal type.
Adjusted standard errors in square brackets.
This Table completes the results reported in Table 6.

Table 12
Maximum likelihood estimates.

Belief based learning with random effects		
δ closed	0.269	(0.053)
δ open	0.160	(0.040)
λ closed	0.498	(0.100)
λ open	0.411	(0.127)
s^0	11.512	(5.870)
μ_s	0.980	(1.886)
σ_s	2.834	(1.349)
$w_1^0\,(a)$	3.909	(2.723)
$w_2^0\,(a)$	10.268	(3.943)
$w_3^0\,(a)$	0.000	(4.954)
$w_4^0\,(a)$	3.195	(2.567)
$w_5^0\,(a)$	3.653	(3.252)
$w_6^0\,(a)$	2.636	(3.906)
$w_7^0\,(a)$	2.919	(1.820)
$w_8^0\,(a)$	0.312	(2.748)
$w_9^0\,(a)$	1.679	(2.035)

well as discussions with David Cooper, John Ham, and Lung-Fei Lee to whom I am very grateful. Finally I wish to thank Stephen Cosslett, Matthew Embrey, Jim Engle-Warnick, Ido Erev, David Hineline, Steven F. Lehrer, Dan Levin, Jim Peck, Shinichi Sakata, Chloe Tergiman, and participants at the Ohio State University Applied Microeconomics workshop, at the Harvard University Experimental/Behavioral seminar, at the Learning session of the Public Choice, Economic Science Association 2001 meetings and at the Markets and Bargaining session of the 2001 North American Meetings of the Econometric Society for helpful comments. I gratefully acknowledge the support of NSF via grants SES-0519045 and SES-0721111 as well as by the Center for Experimental Social Science and the C.V. Starr Center. Any opinions, findings, and conclusions or recommendations in this material are those of the author and do not necessarily reflect the views of the National Science Foundation or the other funding agencies.

Appendix. Complete estimation results

see Tables 9–12

References

Abramowitz, M., Stegun, I.A. (Eds.), 1972. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. US Government Printing Office, Washington, D.C.

Altonji, J.G., 1986. Intertemporal substitution in labor supply: Evidence from micro data. The Journal of Political Economy 94, S176–S215.

Ansolabehere, S., Snyder, J.M.J., Strauss, A.B., Ting, M.M., 2005. Voting weights and voter advantages in the formation of coalition governments. American Journal of Political Science 49 (3), 550–563.

Armantier, O., 2004. Does observation influence learning?. Games and Economic Behavior 46 (2), 221–239.

Armantier, O., 2006. Do wealth differences affect fairness considerations. International Economic Review 47 (2), 391–429.

Baron, D.P., Ferejohn, J.A., 1989. Bargaining in legislatures. American Political Science Review 83 (4), 1181–1206.

Beck, N., Katz, J.N., 1995. What to do (and not to do) with time-series cross-section data. American Political Science Review 89, 634–647.

Binder, D.A., 1983. On the variances of asymptotically normal estimators from complex surveys. International Statistical Review 51, 279–292.

Binmore, K., Shaked, A., Sutton, J., 1985. Testing noncooperative bargaining theory: A preliminary study. The American Economic Review 75, 1178–1180.

Bolton, G.E., Ockenfels, A., 2000. Erc: A theory of equity, reciprocity, and competition. The American Economic Review 90 (1), 166–193.

Browne, E.C., Franklin, M., 1973. Aspects of coalition payoffs in European parliamentary democraties. The American Political Science Review 67 (2), 453–469.

Butler, J.S., Moffitt, R., 1982. A computationally efficient quadrature procedure for the one-factor multinomial probit model. Econometrica 50, 761–764.

Cabrales, A., Garcia-Fontes, W., 2000. Estimating learning models from experimental data. mimeo.

Camerer, C.F., Ho, T.-H., 1999. Experience-weighted attraction learning in normal form games. Econometrica 67, 827–874.

Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. The Quarterly Journal of Economics 17 (3), 817–869.

Cheung, Y.-W., Friedman, D., 1997. Individual learning in normal form games: Some laboratory results. Games and Economic Behavior 19 (1), 46–76.

Cooper, D.J., Garvin, S., Kagel, J.H., 1997. Signalling and adaptive learning in an entry limit pricing game. RAND Journal of Economics 28, 662–683.

Cooper, D.J., Stockman, C.K., 2002. Fairness and learning: An experimental examination. Games and Economic Behavior 41 (1), 26–45.

Deaton, A. (Ed.), 1997. The Analysis of Household Surveys. World Bank, Washington, DC.

Eraslan, H., Merlo, A., 2002. Majority rule in a stochastic model of bargaining. Journal of Economic Theory 103, 31–48.

Erev, I., Haruvy, E., 2001. On the application and interpretation of learning models. In: Advances in Experimental Business Research. Kluwer Academic Publishers, Boston.

Erev, I., Haruvy, E., 2005. On the potential uses and current limitations of data driven learning models. Journal of Mathematical Psychology 49 (5), 357–371.

Erev, I., Roth, A.E., 1998. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. American Economic Review 88 (4), 848–881.

Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition and cooperation. Quarterly Journal of Economics 114 (3), 817–868.

Fréchette, G.R., Kagel, J.H., Lehrer, S.F., 2003. Bargaining in legislatures: An experimental investigation of open versus closed amendment rules. American Political Science Review 97 (2), 221–232.

Fréchette, G.R., Kagel, J.H., Morelli, M., 2005a. Gamson's law versus non-cooperative bargaining theory. Games and Economic Behavior 51 (2), 365–390.

Fréchette, G.R., Kagel, J.H., Morelli, M., 2005b. Nominal bargaining power, selection protocol and discounting in legislative bargaining. Journal of Public Economics 89 (8), 1497–1517.

Fudenberg, D., Levine, D.K., 1999. Learning and Evolution in Games. The MIT Press, Cambridge, Massachusetts.

Greene, W.H., 2000. Econometric Analysis. Prentice Hall, New Jersey.

Ham, J.C., 1986. Testing whether unemployment represents intertemporal labour supply behaviour. Review of Economic Studies 53, 559–578.

Ham, J.C., Kagel, J.H., Lehrer, S.F., 2005. Randomization, endogeneity and laboratory experiments: The role of cash balances in private value auctions. Journal of Econometrics 125 (1–2), 175–205.

Hausman, J.A., Wise, D.A., 1978. A conditional probit model for qualitative choice: Discrete decisions recognizing interdependence and heterogeneous preferences. Econometrica 46 (2), 403–426.

Ho, T.-H., Camerer, C.F., Chong, J.-K., 2002. Functional ewa: A one parameter theory of learning in games. mimeo.

Huber, P.J., 1967. The behavior of maximum likelihood estimates under nonstandard conditions. In: Cam, L.M.L., Neyman, J. (Eds.), Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1. University of California Press, Berkeley, CA, pp. 221–233.

Judge, G.G., Miller, D.J., Mittlehammer, R.C., 2000. Quasi-maximum likelihood estimation and inference. In: Econometric Foundations. Cambridge University Press, Cambridge, United Kingdom, pp. 245–280 (Chapter 11).

Knight, B., 2002. Endogenous federal grants and crowd-out of state government spending: Theory and evidence from the federal highway aid program. The American Economic Review 92 (1), 71–92.

Levitt, S.D., Poterba, J.M., 1999. Congressional distributive politics and state economic performance. Public Choice 99, 185–216.
Messick, D.M., Moore, D.A., Bazerman, M.H., 1997. Ultimatum bargaining with a group: Underestimating the importance of the decision rule. Organizational Behavior and Human Decision Processes 69 (2), 87–101.
Moulton, B.R., 1986. Random group effects and the precision of regression estimates. Journal of Econometrics 32, 385–397.
Moulton, B.R., 1990. An illustration of a pitfall in estimating the effects of aggregate variables on micro units. Review of Economics and Statistics 72, 334–338.
Pepper, J.V., 2002. Robust inferences from random clustered samples: An application using data from the panel study of income dynamics. Economic Letters 75, 341–345.
Persson, T., Tabellini, G., 2000. Political Economics: Explaining Economic Policy. The MIT Press, Cambridge, Massachusetts.
Roth, A.E., 1995. Introduction. In: Handbook of Experimental Economics. Princeton University Press, Princeton, pp. 3–109.
Roth, A.E., Erev, I., 1995. Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. Games and Economic Behavior 8 (1), 164–212.
Rubinstein, A., 1982. Perfect equilibrium in a bargaining model. Econometrica 50, 97–110.
Sakata, S., 2002. Quasi-maximum likelihood estimation with complex survey data. mimeo.
Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Chapman and Hall, London.
Skinner, C.J., 1989a. Introduction to Part A. In: Analysis of Complex Surveys. Wiley, England, pp. 23–58.
Skinner, C.J., 1989b. Domain means, regression and multivariate analysis. In: Analysis of Complex Surveys. Wiley, England, pp. 59–84.
Stahl, D.O., 1996. Boundedly rational rule learning in a guessing game. Games and Economic Behavior 16, 303–330.
StataCorp, 2001 Stata Statistical Software: Release 7.0, Stata Corporation, College Station, TX, chapter 23.11 Obtaining Robust Variance Estimates, pp. 254–258.
Warwick, P.V., Druckman, J.N., 2001. Portfolio salience and the proportionality of payoffs in coalition government. British Journal of Political Science 31 (4), 627–649.
White, H., 1994. Estimation, Inference and Specification Analysis. Cambridge University Press, New York.
Wilcox, N.T., 2006. Theories of learning in games and heterogeneity bias. Econometrica 74 (5), 1271–1292.