GAMES AND ECONOMIC BEHAVIOR **25**, 1–33 (1998)
ARTICLE NO. GA970633

Reinterpreting Arbitration's Narcotic Effect: An Experimental Study of Learning in Repeated Bargaining

Gary E. Bolton* and Elena Katok

Smeal College of Business, Pennsylvania State University, University Park, Pennsylvania

Received December 13, 1996

Field evidence suggests that arbitration increases negotiation dispute rates. We study repeated bargaining in a laboratory to understand the reasons why. Our results represent a reinterpretation of an explanation known as the narcotic effect. The standard interpretation assumes that the probability of dispute without arbitration is constant across negotiations, but field evidence suggests that experienced bargainers have fewer disputes. To properly assess arbitration's impact, we compare bargainer learning with and without arbitration, under otherwise comparable laboratory conditions, and develop a model to measure learning. We find strong evidence that learning occurs in both cases, but is slower with arbitration. *Journal of Economic Literature* Classification Numbers: C78, C92, D77, J52.

1. INTRODUCTION

We report here on a set of laboratory tests aimed at a long-standing puzzle: why having arbitration as a fallback influences the chances that bargainers will reach agreement on their own. The puzzle is closely linked to the question of why bargaining sometimes ends in dispute, and so the work speaks to this larger issue as well. We find that learning about one's bargaining partner plays a key role in explaining the incidence of dispute in repeated bargaining interactions, where one has the same partner over a series of negotiations. Introducing arbitration as a fallback for failed negotiations retards, but does not stop, the learning process. Thus our results cast learning as a pivotal piece of the arbitration puzzle. This provides a fresh perspective on past field research, and suggests new paths for theoretical work. Our findings are consistent with Schelling's view that

* To whom correspondence should be addressed, at Department of MSIS, 310 Beam, Penn State University, University Park, PA 16802. E-mail: geb3@psu.edu.

0899-8256/98 $25.00

bargaining is a struggle to establish commitment. We are able to characterize several important features of the learning process behind the struggle.

Arbitration saves losses that would otherwise occur when a negotiation ends in dispute (saving the cost of a strike, for example). Arbitration has long been used to resolve wage disputes in the U.S. public sector, and has grown in the private sector as well (see Allison, 1990). There are many variations on the theme, but arbitration procedures always involve an outside party, an arbitrator, with the authority to impose a binding settlement. Arbitration usually acts as a safety net, invoked after negotiations fail. We say that there is a ‘‘dispute’’ when bargaining ends without a voluntary settlement. When we refer to ‘‘bargaining with arbitration,’’ we mean a negotiation in which arbitration is the fallback after a dispute. We focus on arbitration in the context of repeated bargaining situations, the context most common to field research in the area.

While arbitration is a sure way of saving the negotiation pie, there is substantial field evidence that bargaining with arbitration lessens the likelihood that bargainers will reach a settlement on their own (e.g., Currie and McConnell, 1991; Lester, 1984). In this sense, arbitration tends to be ‘‘overused,’’ and this may have undesirable consequences. Arbitrator awards may not be as efficient as voluntary settlements (Crawford, 1982). Establishing an efficient multilevel wage schedule, for example, may require detailed information about both management and union preferences, information that, for strategic reasons, may be difficult for an arbitrator to elicit. Chronic overuse may serve to multiply the inefficiency.

Concern for the consequences of overuse has influenced arbitration in practice (Bloom, 1981). Final-offer arbitration, famous for its role in American baseball salary disputes, provides an instructive illustration. The rationale for final offer begins with the claim that, when allowed to decide freely, arbitrators tend to split the difference between final bargaining positions, thereby chilling the incentive to compromise during negotiations. The final-offer procedure proffers a straightforward remedy: forbid split-the-difference awards; constrain the arbitrator to choosing between bargainers’ final proposals. Unfortunately, there is substantial evidence that arbitrators do not behave as posited (Ashenfelter and Bloom, 1984; Farber and Bazerman, 1986), and whether final offer mitigates overuse is debatable (Brams et al., 1991).

But while final-offer may not be the solution to overuse, its story underscores the importance of the issue we will pursue here: any attempt to ease overuse must come to grips with why disputes are higher with arbitration. We believe that field studies provide important clues. Two explanations weave their way through the field literature we survey in the next section. We make use of both explanations, albeit in modified form.

One explanation has to do with the hypothesis that lower dispute costs raise the probability of a dispute. As we will explain, field evidence supports this hypothesis. But the same evidence suggests that there is a countervailing learning trend that, over time, tends to diminish disputes. So the level of dispute cost can be, at most, part of the story. Moreover, the dispute cost hypothesis is often motivated by bargaining theories that attribute disputes to private information. But our lab negotiations reproduce many aspects of the field data on disputes, even though there is complete information over those variables the theories typically assume are private information (such as payoffs). This raises some questions as about whether private information is a necessary ingredient to the disputes we see in the field.[1] Work we cite in the next section offers a different account of disputes, and the dispute cost hypothesis.

The other explanation for overuse that plays a role here is known as the *narcotic effect*. The classic statement of this hypothesis was put forward by Wirtz (1963): "[Bargainers] will turn to [arbitration] as an easy and habit forming release from the ... obligation of hard, responsible bargaining." The causation mechanism is vague—a real problem if the goal is to design a remedy for overuse—but the implication is clear: going to arbitration engenders dependence on the procedure. Of course, the meaning of "dependence" must be made precise. Field evidence suggests that the received definition has some shortcomings, but the evidence also offers clues that we exploit to formulate an alternative definition—one that involves learning.

Testing the hypotheses implied by the field data requires observations from negotiations both with and without arbitration, but gathered under otherwise comparable conditions. Finding comparable negotiations in the field is difficult.[2] We create comparable negotiations in the lab. Before describing the experiment, however, we need to discuss the relevant field literature.

[1] In a recent review of the experimental bargaining literature, Roth (1995) finds that dispute rates reported by complete and incomplete information experiments and those reported by field studies are all quite similar. He concludes that this similarity "raises some question about whether the incomplete information models are focusing on the underlying cause of disagreement" p. 294.

[2] A recent controversy concerning the Currie and McConnell (1991) study illustrates how difficult identifying comparable field negotiations can be. Gunderson et al. (1996) argue that Currie and McConnell misclassified over 30% of Canadian public sector negotiations with respect to the dispute resolution mechanism (they also find that Currie and McConnell's main results are robust to the reclassification). In their reply, Currie and McConnell (1996) concede some errors, but also note that "there is room for disagreement about the correct coding of the law variables in a number of cases," and go on to explain that, "Determining the correct coding of the legislation in effect for a given contract at a given time can be a daunting task. ... Even after we had determined the legislation in effect, it was not always easy to classify the legislation into one of our five legislative categories" p. 327.

2. FIELD LITERATURE: THE NARCOTIC EFFECT AND LEARNING

A considerable amount of effort has been directed at verifying the existence of a narcotic effect. Early empirical work was later criticized for methodological shortcomings (Anderson, 1981). An influential study by Butler and Ehrenberg (1981) offers corrective measures, as well as what we think to be a telling finding. Butler and Ehrenberg begin by clearly stating their interpretation of dependence: ‘‘We ask . . . whether the probability of going to impasse in one contract round is positively related to the extent to which the parties used impasse procedures in previous rounds’’ (p. 4). They then reanalyze data from Kochan and Baderschneider’s (1978) study of police and firefighter impasse proceedings. Butler and Ehrenberg find a narcotic effect for the initial period covered by the data. But for the later period they report what they refer to as a negative narcotic effect; that is, at some point use of the procedure tended to diminish the probability of dispute. They conjecture that dispute rates fall because bargainers become dissatisfied with arbitration, either out of disappointment with the awards or because the novelty of using arbitration wears off.

Butler and Ehrenberg’s interpretation of the narcotic effect has gained wide acceptance (e.g., Currie, 1989). There is, however, a strong implicit assumption. Note that the observed dependence of the probability of dispute on previous negotiating history is taken as evidence that arbitration is influencing the negotiation process. This assumes that the probability of dispute in negotiations without arbitration is independent of previous negotiating history; otherwise, there is no reason to attribute observed dependence to arbitration. (Note the influence of this assumption on Butler and Ehrenberg’s explanation of the negative narcotic effect.)

Contrary to the assumption, there is field evidence that experience influences dispute rates. Reder and Neumann (1980) begin their investigation by positing that repeated bargaining allows bargainers to ‘‘learn about one another’s behavior patterns during the bargaining process and develop conventions (protocols) to guide subsequent bargaining activity’’ (p. 868). Protocols are essentially coordination devices that cut the probability of a dispute. As such, protocol learning lowers dispute costs over the sequence of negotiations. Greater negotiating experience allows a bargainer to learn more about her partner’s behavior across a greater number of contingencies. More experienced bargaining units should therefore tend to have more elaborate protocols and lower dispute rates. Reder and Neumann examine strike data for U.S. manufacturing industries and find that experience does indeed have this effect. Successful protocol learning requires that each bargainer ‘‘limits his own behavior in any specified situation to what is compatible with the established behavior pattern already learned by the bargaining partner'' (p. 869). So protocol learning involves concessions, and we would expect concessions to be slower the lower the cost of resistance, or what is equivalent, the lower the expected cost of dispute. In fact, Reder and Neumann find that dispute rates are inversely related to dispute costs.

Reder and Neumann's study involved bargaining without arbitration; as such, their finding for learning casts doubt on the independence assumption underpinning Butler and Ehrenberg's interpretation of the narcotic effect. Specifically, if negotiations without arbitration exhibit history dependence, then there is no reason to attribute dependence found in negotiations with arbitration to arbitration.

That said, Reder and Neumann's conceptual framework also suggests an alternative interpretation of Wirtz's notion of narcotic-like dependence on the arbitrator. Specifically, because arbitration reduces but does not eliminate dispute costs (Bloom, 1981), there might still be protocol learning with arbitration, but the reduced cost might slow the rate. By this interpretation, arbitration involves dependence—a narcotic effect—in the sense that arbitrator intervention retards the process by which bargainers learn to resolve their differences by themselves. This interpretation offers a new perspective on what has been observed in the field: what Butler and Ehrenberg refer to as a negative narcotic effect may actually be a (positive) narcotic effect, depending on whether the observed learning is slower or faster or unchanged relative to what it would be absent arbitration. (Of course, Butler and Ehrenberg also find a period of rising dispute rates. As we explain in Section 7, we think that the process of learning is capable of explaining these episodes as well.)

Lester's (1989) study of police and firefighter arbitration in New Jersey, 1978 to 1987, provides further evidence that bargainers learn even with arbitration. Lester's main conclusion is that ''[t]he statistics for negotiations and awards . . . show that, over the decade, the parties increasingly achieved new collective agreements without the need for an award'' (p. 17). Lester goes on to examine reasons for the improvement, and concludes that the principal explanation is that, with increased experience, the parties better understood both the system and how others behaved in the system (p. 21); in other words, the key is learning.

3. THE EXPERIMENT: NA AND ELA TREATMENTS

We examine bargainer learning in the lab. Our experiment shares many features with an experiment reported by Ashenfelter et al. (1992). They compared various arbitration mechanisms, concluding that disputes are higher with arbitration than without. Section 3.5 provides a detailed comparison of the two studies.

We focus on conventional arbitration, a commonly employed procedure that gives the arbitrator the authority to award whatever settlement she deems fair. We begin by describing two treatments. (We discuss further tests in Section 6.) In the NA (no arbitrator) treatment, bargaining was always without arbitration; in the ELA (equal likelihood arbitrator) treatment, bargaining was always with arbitration.

3.1. *Experimental Design*

In both treatments, bargainers were anonymously paired with the same partner for 12 rounds of negotiation. We describe, in turn, the bargaining game and the arbitration mechanism.

Bargaining pairs played a simple bargaining game, that we call the *deadline game*.[3] It is a two-person, two-stage negotiation concerning the division of 100 chips:

Stage 1: Bargainers A and B simultaneously propose a division of the chips; call the proposals v^A and v^B. If offers are exactly compatible, then the game ends in agreement on the implied division; otherwise the game proceeds to stage 2.

Stage 2: After viewing both stage 1 offers, bargainers play the bimatrix game in Fig. 1, where a denotes accepting the partner's offer, and m denotes maintaining ones own offer. If both bargainers play a, then a coin flip decides which offer becomes the settlement.

Treatments are distinguished by what happens in the case of a dispute (both bargainers play m). In NA, both bargainers receive 0 chips. In ELA, an arbitrator awards a settlement.

3.2. *Arbitration*

As with Ashenfelter et al. (1992), arbitrator awards were modeled as a random draw. The technique is motivated by the findings of field studies (Ashenfelter and Bloom, 1984; Ashenfelter, 1987). Field bargainers usually

[3] This game was analyzed by Harsanyi (1977). Crawford (1982) explains bargaining disputes using a modified version. Roth and Schoumaker (1983) use the game in a lab study of bargainer expectations. Bolton (1997) identifies limit evolutionary stable equilibria.

		bargainer *B*	
		a	*m*
bargainer *A*	*a*	v^A with prob 1/2 v^B with prob 1/2	v^B
	m	v^A	NA: 0,0 ELA: arbitrator imposes a settlement

FIG. 1. Stage 2 of the Deadline Game.

have some say about who will arbitrate, so arbitrators who are predictably biased toward either party relative to other arbitrators tend to be vetoed by the disadvantaged party. Therefore, while different arbitrators may make different awards, we expect no predictable difference; acceptable arbitrators tend to be statistically interchangeable. The stochastic award distribution used in the experiment captures this feature.

Stochastic awards imply risk. So how bargainers evaluate the prospect of going to arbitration depends crucially on their risk profile. If both bargainers are risk averse, then each attaches a certainty equivalent to the prospect of arbitration that has lower value than the mean arbitrator award that he can expect; consequently, there are voluntary settlements that both bargainers would prefer to arbitration. Both, for example, would prefer voluntarily settling on the mean arbitrator award. But if bargainers are not risk averse, things can be quite different. Suppose, for example, that the negotiation involves 100 chips and the mean arbitrator award is 50−50. Suppose both bargainers attach a certainty equivalent of 51 chips to the prospect of arbitration. A voluntary settlement that gives both bargainers at least 51 chips is not feasible, so there is no voluntary settlement that is mutually preferred to arbitration.

From the discussion, we see that the important consequence of risk aversion is that it implies the existence of voluntary settlements that bargainers mutually prefer to arbitration. Since field bargainers are generally thought to be risk averse, risk aversion is the case of greatest interest. It is not clear, however, that lab subjects playing for relatively small stakes will behave in a risk-averse manner. In fact, Ashenfelter et al. conclude that their data ‘‘do[es] not indicate strong evidence of risk aversion’’ (p. 1430).[4] The ELA arbitrator does not induce risk aversion, but it does implement the important consequence; that is, the ELA arbitrator ensures a voluntary settlement that bargainers, regardless of risk profile, will mutually prefer to arbitration.[5]

ELA awards were determined as follows. An integer, x, was drawn from a normal distribution truncated at 0 and 100, centered at 50, with a standard deviation of 5.4. Each player was then awarded chips equal to $\min\{x, 100 - x\}$. This procedure implies an award distribution with a support of 0 to 50 chips, a mean of 45, and a standard deviation of 3.25. Since 50 chips is the maximum award, and there is a positive probability of receiving less, all bargainers regardless of risk posture should prefer a voluntary settlement of equal division over going to arbitration.

The added advantage of this particular award distribution is that it allows us to investigate the effect of increasing the cost associated with arbitration: we simply increase the standard deviation of the normal distribution from which x is drawn. Reflection (or the algebra) shows that the probability of obtaining any particular award is strictly lower for the new distribution, with the exception of the award of 0 chips, which has a strictly higher probability. Therefore any bargainer, independent of risk posture, should prefer the ELA award distribution over the new one. Section 6 reports treatments involving this manipulation.

3.3. *Presentation of the Arbitrator*

Information about the arbitrator was presented in a manner similar to the way in which it is available in the field. Bargainers in the ELA treatment were given a history of the arbitrator’s past awards, actually a

[4] The study also found that ‘‘dispute rates are inversely related to the uncertainty costs of disputes, indicating that some bargainers behave as if they were risk averse’’ (p. 1407). While the behavior is indeed consistent with risk aversion, it is also consistent with the proposition that subjects have a taste for small, not large, gambles. To see this, suppose as before that each bargainer attaches a certainty equivalent of 51 chips to arbitration, thereby precluding a voluntary settlement. Now suppose, as in the Ashenfelter et al. experiment, we increase the uncertainty associated with the award distribution. If bargainers prefer small gambles to large gambles, then certainty equivalents will decrease, say, to 48 chips. A 50−50 voluntary settlement would then be mutually preferred, even though bargainers are not consistently risk averse.

[5] Alternatively, we could attempt to induce risk-averse preferences among subjects. The efficacy of the method for doing so is much debated. Our more direct method captures the essential features while side-stepping the difficulties.

list of 100 random draws from the appropriate normal distribution. They were told to expect similar decisions in the future. The list included the average for the 100 draws (45 for both bargainers). No mention was made of the actual randomization process.

3.4. *Laboratory Protocol, Subject Pool, and Information Given to Subjects*

All subjects were Penn State University students, recruited through billboards posted around the University Park campus. Participation required appearing at a special place and time, and was restricted to one session. Cash was the only incentive offered. There was a total of 25 bargaining pairs (50 subjects) in each treatment.

Games were played through a computer interface. Upon arriving at the lab, participants read instructions (copy in Appendix) and then played practice games with the computer as partner. So as not to bias later play, the computer made decisions entirely at random (participants knew this). A brief quiz was given to check for understanding of computer operation, and the written instructions were read aloud. Participants were then randomly and anonymously paired. To avoid end-game effects, the total number of games was not revealed.

It was publicly announced that each chip had a value of $0.24. The computer automatically displayed the cash values of all offers. At the conclusion of the session, one game was selected by lottery. Participants were paid their earnings for the selected game plus a $5 show-up fee (payment in cash).[6]

3.5. *Comparison with Ashenfelter et al.* (1992)

The experimental design parallels that of Ashenfelter et al. (1992), with three notable exceptions. First, while bargaining in both experiments concerns a simple pie-splitting task, Ashenfelter et al. permitted bargainers to make and modify demands until they matched or time ran out. The deadline game is a simplification of this procedure, with the advantage that it has a concise extensive form that permits the sharp analysis of bargainer behavior presented in Section 5. Second, we give bargainers complete information about their partner's monetary chip value. Complete informa-

[6] Sessions were run from fall 1993 to spring 1994. Each involved 8–18 participants, and began between 3 and 5 PM, lasting about 90 min. Average earning (including the show-up fee) was $15.76, about $10.50 per hour, with a standard deviation of $4.53.

tion eliminates any confounding that would result from having to speculate about bargainer expectations over true payoffs.[7] Third, all bargainers in the Ashenfelter et al. experiment bargained without arbitration prior to bargaining with arbitration; Section 6 analyzes treatments that have the same half no arbitrator−half arbitrator structure.

As it turns out, despite the differences, data from the two experiments are quite comparable in terms of dispute rates and voluntary settlements. We draw explicit parallels in the analysis below. The one substantial difference is that Ashenfelter et al. do not report the round effects we find. However, their testing was restricted to looking for evidence that dispute rates in the final round differ from rates in other rounds (pp. 1415−1416). This test fails to pick up round effects when applied to many of our treatments. It is only when we consider the rate of change across all rounds that the round effects become clear. Charness (1998) reports round effects in an arbitration experiment with a design that is virtually identical to that of Ashenfelter et al., including information structure, payoff schedule, and bargaining game.

3.6. *Preliminary Analysis*

The data are summarized by Table 1 and Figs. 2 and 3. Each exhibit provides a somewhat different perspective on the pattern of disputes.

Table I displays early-round (first six) and late-round (last six) average dispute rates. Figure 2 displays dispute rates round by round. Comparing across treatments, we see that disputes are generally higher in ELA. In fact, a t-test strongly rejects equality across either early or late rounds (p-value < 0.0001 for both). That the difference is evident from the very first rounds suggests that arbitration induces a higher initial propensity to dispute. Ashenfelter et al. (1992) found that arbitration induces higher dispute rates of a quite similar magnitude.

Figure 2 illustrates that within each treatment, early-round dispute rates are higher than late-round rates. This is formally demonstrated by a matched-pair test on the difference in early- and late-round average

[7] For example, incomplete information about payoffs would have left the fall in disputes across rounds open to the interpretation that bargainers are playing tough early on to convince their partner that their monetary payoffs are too insubstantial to settle for less than a very large share of the chips. By this hypothesis, disagreements are strategic and have nothing to do with learning about the other bargainer. Complete information rules this hypothesis out. (We rule out a more general hypothesis about reputation building in Section 5.)

TABLE I
NA and ELA Treatments Summary of Dispute Rates (Standard Error)

Treatment	Number of observations	Rounds 1–6	Rounds 7–12	Difference
NA	25	0.180 (0.0332)	0.067 (0.0289)	−0.113 (0.0369)
ELA	25	0.367 (0.471)	0.220 (0.0516)	−0.147 (0.0539)
Difference		−0.187 (0.0115)	−0.153 (0.0118)	0.034 (0.0130)

disputes of Table I (two-tail p-value of 0.005 for NA, and 0.012 for ELA). So there is a pronounced learning trend evident both with and without arbitration.

Figure 3 exhibits the number of disputes by bargaining pair, and suggests a high degree of bargainer heterogeneity with respect to the propensity to dispute. Ashenfelter et al. (1992) also note the high degree of bargainer heterogeneity in dispute rates. Formally, a rank correlation test finds that bargaining pairs with a relatively higher dispute rate in early rounds are likely to have a relatively higher dispute rate in late rounds (two-tail p-value of 0.021 for NA and 0.003 for ELA). The test does not indicate the source of the heterogeneity. It may, for example, be due to

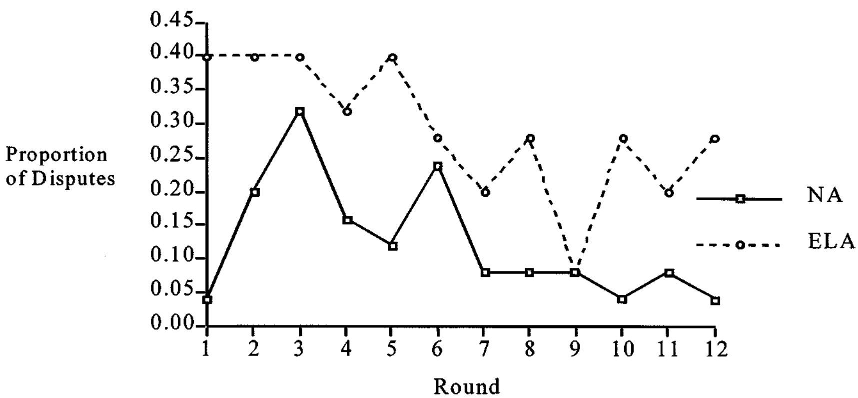

FIG. 2. Dispute data by rounds: NA and ELA.

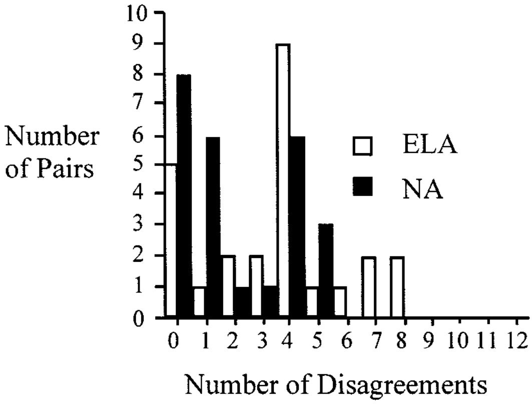

FIG. 3. Dispute rates by bargaining pair.

different initial propensities to dispute, or it may reflect history dependence.[8]

All three observations find a parallel in field studies. Observation of higher propensities to dispute with arbitration and the learning trend are discussed in Sections 1 and 2.[9] Kochan and Baderschneider (1978) find substantial heterogeneity in dispute rates across bargaining units.

We observe a downward trend in disputes for both the NA and ELA treatments. The addition of arbitration is apparently not in and of itself responsible for learning. There remains a question, however, as to whether adding arbitration substantially alters the learning pattern. From Table I we can calculate that from early to late rounds, the dispute rate falls by 63% for NA but by only 40% for ELA. These estimates, while crude, are nevertheless suggestive of a slower learning rate—a type of narcotic effect.

4. ANALYSIS OF THE NA AND ELA DATA

We would like to formally estimate and compare learning rates across treatments. Direct measurement is confounded by the higher initial

[8] Another way to see the heterogeneity: the data set taken as a whole, the present data along with those in Section 6, include 131 bargaining pairs who negotiated without arbitration for six (early) rounds. A contingency table test comparing dispute rates across individual treatments is unable to reject the hypothesis that pairs are drawn from the same population (p-value of 0.495). A goodness-of-fit test applied to the pooled distribution easily rejects a Bernoulli process model implying a constant probability of dispute across bargaining pairs (p-value of 0.004).

[9] Bazerman and Neale (1982) report an experiment in which bargainer training increased resolution frequency.

propensity to dispute in the presence of arbitration. A formal analysis requires a model of the interaction between the two effects, and must provide a control for bargainer heterogeneity. We describe two alternative specifications and compare their fit. The better model is then used to test for a narcotic effect.

4.1. *Round Effects versus Outcome Learning*

The first model treats learning as a simple round effect:

$$d_{i,j,t} = \text{probability that pair } i \text{ in treatment } j \text{ has a dispute at time } t$$
$$= \theta_{i,j} + \beta_j(\text{round} - 1) + \epsilon_{i,j,t} \tag{4.1}$$

where $\theta_{i,j}$ is a pair i fixed effect in treatment j; the β_j's are parameters to be estimated; and $\epsilon_{i,j,t}$ is the error term.

The round effect model is a natural baseline. It allows for the learning trend in a general way, without being precise about the source of the learning. So while the model implies that experience with bargaining influences the probability of having a dispute, it does not specify what aspect of this experience causes the change. (It may be something as straightforward as increased familiarity with the rules of the game.)

A natural, somewhat more precise hypothesis would be that bargainers learn from past negotiation outcomes, adapting their behavior in response to past success or failure.[10] Reder and Neumann argue that it is the cost of dispute that motivates bargainers to adopt settlement-facilitating protocols (Section 2). A dispute would then induce bargainers to modify their behavior in search of a protocol, thereby lowering the probability of a future dispute. The influence of a past settlement is less clear, although there might be some.[11] The outcome learning model captures the hypothesis that past negotiation outcomes motivate learning:

$$d_{i,j,t} = \theta_{i,j} + \delta_j D_{i,j,t} + \sigma_j S_{i,j,t} + \epsilon_{i,j,t} \tag{4.2}$$

[10] Gale et al. (1995) and Roth and Erev (1995) study adaptive learning in the context of the ultimatum bargaining game.

[11] To illustrate the ambiguity here, suppose that a bargainer plays a mixed strategy and observes a settlement. This could conceivably reinforce the mixed strategy, tending to stabilize the probability of dispute. Alternatively, it might reinforce playing the particular action chosen (as opposed to the entire mixed strategy), tending to reduce the probability of dispute. Or it may embolden the bargainer to experiment with riskier, more aggressive play in an attempt to increase his share of the settlement, thereby increasing the probability of dispute.

where $D_{i,j,t}$ is the cumulative number of disputes that pair i in treatment j had in rounds 1 through $t-1$. $S_{i,j,t}$ is the cumulative number of voluntary settlements through round $t-1$, θ's are pair fixed effects, ϵ's are error terms, and the δ_j's and σ_j's are parameters to be estimated.

The guiding consideration behind the (4.2) formulation is that the model must be capable of capturing the persistent learning trend on display in Fig. 2. The approach we take was originally suggested by Butler and Ehrenberg to capture this sort of persistence.[12] The approach in (4.2) captures the phenomenon with some economy, under the assumption that the marginal effect of an additional dispute (settlement) is roughly constant for the duration of the experiment.

A second advantage of the formulation in (4.2) is that it permits a direct test of the explanatory power of the outcome learning model versus the round effect model. Note that

$$D_{i,j,t} + S_{i,j,t} \equiv \text{round} - 1.$$

Therefore, (4.1) is equivalent to (4.2) if $\delta_j = \sigma_j$. Rejecting this equality would imply that outcome learning has greater explanatory power than round effects. To determine whether $\delta_j = \sigma_j$, define

$$\text{outcome}_{i,j,t} \equiv D_{i,j,t} - S_{i,j,t},$$

and, through a series of substitutions,[13] rewrite (4.2) as

$$d_{i,j,t} = \theta_{i,j} + \beta_j(\text{round} - 1) + \lambda_j \text{ outcome}_{i,j,t} + \epsilon_{i,j,t}. \qquad (4.3)$$

We then test the null hypothesis $\lambda_j = 0$; rejection is equivalent to rejecting the round effect model in favor of the outcome model.

Our focus is on estimating learning. To keep the exposition clear, we do not report the estimates for the pair fixed effects, but they are available upon request. (Remember that including fixed effects in the estimation is important to account for both bargainer heterogeneity and different initial propensities to dispute across treatments.)

Different underlying assumptions imply different estimation procedures. In Table II we present three different estimations of (4.3), based on three common sets of assumptions. The first column displays the least-squares

[12] Butler and Ehrenberg restrict their attention to the role of past disputes and do not consider past settlements, reflecting their definition of the narcotic effect. They say that the effect of past experience with arbitration "could be modeled by including lagged values . . . or a variable representing the number of times the unit went to impasse over a given number of previous rounds . . . . Which formulation one would assume to be the correct one to use would depend upon whether one believed the effects of previous impasse experiences depreciate over time" p. 9.

[13] The more straightforward substitution of $S = (\text{round} - 1) - D$ leads to an equation with both (round − 1) and D as regressors. Since these are positively correlated, there is a potential multicollinearity problem. The method in the text avoids this difficulty.

TABLE II
NA and ELA Treatments: Round vs. Outcome Learning (Standard Error)[a]

	LSDV	AR1	Probit
Round−NA	−0.074	−0.099	−0.302
	(0.015)	(0.015)	(0.068)
Round−ELA	−0.043	−0.046	−0.123
	(0.0079)	(0.0088)	(0.030)
Outcome−NA	−0.085	−0.107	−0.409
	(0.019)	(0.020)	(0.112)
Outcome−ELA	−0.056	−0.059	−0.203
	(0.013)	(0.013)	(0.062)
R^2	0.29	0.33	—
Log likelihood	—	—	−213.3
$\phi(\cdot)$ at the sample means	—	—	0.0636
Two-tail p-value for H_O: outcome−NA = outcome−ELA = 0	4.3E-09	2.4E-10	—
$n = 600$ (12 rounds for 50 pairs)			

[a] All three model estimates included pair fixed effects as regressors.

dummy variable estimates (LSDV). The second column corrects for first-order autocorrelation, using the generalized least-squares procedure (AR1). This provides a check on whether (4.2) adequately captures the lag effects of past games. The third column presents probit estimates. Another common estimation approach, treating $\theta_{i,j}$ as random effects, was discarded. In the linear case, it has been shown that random-effects models can produce biased and inconsistent estimates when, as is the case for (4.3), the conditional mean of the $\theta_{i,j}$ is dependent on other regressors (Mundlak, 1978). For probit, the data proved to be inconsistent with a random-effects formulation.[14]

For all three estimation procedures, the coefficients of the outcome variables, the λ_{NA} and λ_{ELA} parameters, are statistically different from 0

[14] Mixing probit and fixed effects when the data include units that always register the same dummy value can sometimes adversely effect estimate precision (Chamberlain, 1980). In terms of our model, the potential difficulty arises from bargaining units that either always had settlements or always had disputes. As it turns out, the probit estimates of the fixed-effect standard errors are indeed high, but the standard errors of the learning variables reported here in the tables appear unaffected. Specifically, we ran the probit, dropping the indicated bargaining units from the sample. The only difference in results was that the standard errors of the fixed effects were more in line with those for the LSDV and AR1 procedures; estimates of the learning coefficients along with the associated standard errors were virtually unchanged. The one exception is the LA45 treatment, reported in Section 6; in this case, the probit algorithm did not converge without removing the indicated units. The reported LA45 probit estimates are from dropping the indicated bargaining units.

($p < 0.01$ in all cases). We conclude that the outcome learning model describes the learning process better than the round-effect model.

4.2. *Testing for a Narcotic Effect in the Outcome Learning Model*

Table III displays estimates for the outcome learning model (4.2). The Settlements−NA row provides estimates of the influence of past settlements on NA dispute probabilities (σ_{NA} in Eq. (4.2)). All of the estimates are relatively small, and most are not significantly different from 0. The $\Delta\sigma_{arb}$ row measures the difference in past settlement influence across treatments. These estimates are all very small, and none are significant. We conclude that past settlements have little influence on learning either with or without arbitration, and do not account for differences in learning rates across treatments.

In contrast, the estimates of Disputes−NA (δ_{NA}) are all negative and all very significant, indicating that past disputes are the source of learning in NA ($p < 0.0001$ in all cases). The reader can obtain an estimate of δ_{ELA} by adding $\Delta\delta_{arb}$ to δ_{NA}. All of these are negative (and apparently significant), indicating that past disputes are also the source of learning in ELA.

Since past settlements are an insignificant factor, any difference in the rate of learning across treatments would have to come from differences in the response to past disputes. Consequently, the $\Delta\delta_{arb}$ variable estimates the (learning) narcotic effect. (Estimates in Section 6 control more carefully for past settlements, and confirm the present conclusions.) All of the

TABLE III
NA and ELA Treatments: Outcome Learning Model (Standard Error)[a]

	LSDV	AR1	Probit
Disputes−NA	−0.159	−0.206	−0.712
	(0.033)	(0.035)	(0.175)
Settlements−NA	0.010	0.0077	0.106
	(0.008)	(0.009)	(0.061)
$\Delta\delta_{arb}$ = arbitration effect	0.059	0.100	0.385
on learning from disputes	(0.038)	(0.040)	(0.194)
$\Delta\sigma_{arb}$ = arbitration	0.0023	0.0055	−0.026
effect on learning from settlements	(0.012)	(0.013)	(0.081)
R^2	0.29	0.33	—
Log likelihood	—	—	−213.3
$\phi(\cdot)$ at the sample means	—	—	0.06362
$n = 600$ (12 rounds for 50 pairs)			

[a] All three model estimates included pair fixed effects as regressors.

$\Delta\delta_{arb}$ estimates are positive. While the LSDV estimate is not quite significant ($p = 0.118$), it is significant for the two more sophisticated estimation procedures ($p = 0.013$ for AR1 and 0.047 for probit). So we have some evidence for a narcotic effect.[15]

In summary, we draw four preliminary conclusions from our analysis of the NA and ELA data. First, the outcome learning model is a better description of the learning process than is the round-effect model. Second, occurrence of a dispute drives down a bargaining pair's probability of having a dispute in the future; in short, bargainers learn from their mistakes. Third, introduction of the arbitrator tends to slow down, but does not stop, this learning; in this sense, arbitration induces a narcotic effect. Fourth, settlements play no significant role in learning.

5. SETTLEMENTS AND STAGE STRATEGIES

We can gain some further insight into the learning process by taking a closer look at the stage play of the deadline game. We begin by looking at settlements. Table IV (next page) displays the distribution of settlements for both treatments. As one might expect, equal division is an important protocol in these simple bargaining games. Ashenfelter et al. (1992) also observed settlements clustering around equal division. For both NA and ELA, the clustering is a bit tighter in the later rounds—enough to be weakly significant (one-tail z-test, $p = 0.073$ and 0.096, resp.).

We next examine changes in strategy choice for each stage of bargaining. Define the *offer spread* as the difference between what bargainer A offers B and what bargainer B offers A. The offer spread provides a concise summary of stage 1 play. (Nearly all offers allocated all 100 chips, so measuring offer spread by what was offered to B yields virtually identical results).

To determine whether the learning we identified is related to changes in stage 1 play, we regress offer spread on total past disputes and total past settlements, together with pair fixed effects. Table V shows very similar results, regardless of whether the model is fit using LSDV or AR1. The offer spread tends to shrink after a bargaining pair has a dispute (Disputes−NA, $p < 0.001$ for both estimates). The rate of shrinkage is substantially and significantly slower for ELA ($\Delta\delta_{arb}$, $p < 0.001$ for both esti-

[15] The reader may feel that a proper measure of the narcotic effect should account for differences in the aggregate level of disputes across treatments. Given that aggregated disputes are higher under arbitration, doing so would increase the measure of the effect, making our measure more conservative; i.e., if there is a narcotic effect by our measure, it will continue to be so when adjusted for the aggregate differences.

TABLE IV
Negotiated Settlements by Distance from Equal Division (Proportion)

	Distance from equal division (in chips)[a]							
	0	≤ 5	≤ 10	≤ 15	≤ 20	≤ 25	> 25	Mean (SD)
NA early rounds $n = 123$	0.59	0.21	0.13	0.02	0.01	0.02	0.02	2.93 (5.61)
NA late rounds $n = 140$	0.64	0.22	0.06	0.06	0	0	0.01	1.85 (4.95)
ELA early rounds $n = 95$	0.70	0.23	0	0.04	0	0.01	0.02	2.03 (4.19)
ELA late rounds $n = 117$	0.74	0.25	0	0	0.01	0	0.01	1.08 (3.32)

[a] Categories are mutually exclusive. The 0 category includes only equal divisions; ≤ 5 includes all divisions in which the larger share was 55 chips or less but not 50 chips, etc.

mates). Past settlements show no significant influence (Settlements−NA and $\Delta\sigma_{arb}$). The initial average offer spreads, as measured by the average pair fixed effects, differed little across treatments. The average for NA was 19.9 chips (LSDV estimate). The value for ELA was only 0.2 chips higher, very insignificant ($p = 0.932$). In sum, the one and only substantial difference in stage 1 behavior across treatments is that bargainers are slower to moderate their demands in ELA.

By the rules of the game, play continues to stage 2 only if stage 1 offers have a nonzero offer spread. In analyzing the stage 2 data, we discovered that the bargainer whose offer is closer to equal division behaves quite differently from the one whose offer is farther away. Refer to the bargainer whose offer is farthest from equal division as "far," and the other as "close" (for ties, arbitrarily assign one bargainer to each category).

Table VI presents estimates of the probabilities that far and close bargainers will play m in stage 2 as a function of the stage 1 offer spread, total past disputes, total past settlements, and pair fixed effects. Far and close bargainers respond to changes in offer spread in very similar ways: a smaller offer spread decreases the probability of playing m (Spread−NA, $p < 0.03$ for all estimates). (This is as we would expect, since a smaller offer spread decreases the benefits of playing m.) The substantial difference between far and close bargainers has to do with the initial propensity to play m, as measured by the pair fixed effects (not reported in Table VI). For NA, the average far bargainer fixed effect was 0.13 (LSDV estimate), meaning that the average far bargainer played m about 13% of the time when the offer spread was close to zero. In contrast, the average close bargainer fixed effect in NA was 0.67. For ELA, the average far fixed

TABLE V
NA and ELA: Stage 1 Offer Spread (Standard Error)[a]

	LSDV	AR1
Disputes−NA	−9.444	−9.824
	(1.418)	(1.546)
Settlements−NA	0.148	0.069
	(0.342)	(0.375)
$\Delta\delta_{arb}$ = arbitration effect on learning from disputes	7.584	8.226
	(1.624)	(1.780)
$\Delta\sigma_{arb}$ = arbitration effect on learning from settlements	0.598	0.567
	(0.527)	(0.572)
R^2	0.46	0.47
$n = 600$ (12 rounds for 50 pairs)		

[a] Both model estimates included pair fixed effects as regressors.

TABLE VI
NA and ELA: Stage 2 Propensity to Play m (Standard Error)[a]

	Far bargainer			Close bargainer		
	LSDV	AR1	Probit	LSDV	AR1	Probit
Offer spread−NA	0.006	0.005	0.025	0.006	0.005	0.026
	(0.002)	(0.002)	(0.008)	(0.002)	(0.002)	(0.009)
$\Delta spread$ = arbitration effect on offer spread	0.0003	0.001	−0.004	−0.002	−0.003	0.430
	(0.0026)	(0.003)	(0.011)	(0.002)	(0.002)	(0.143)
Disputes−NA	0.002	−0.003	−0.004	0.015	0.023	0.003
	(0.018)	(0.020)	(0.082)	(0.015)	(0.016)	(0.093)
Settlements−NA	−0.003	−0.008	0.006	−0.018	−0.011	−0.030
	(0.016)	(0.018)	(0.082)	(0.013)	(0.013)	(0.069)
$\Delta\delta_{arb}$ = arbitration effect on learning from disputes	−0.023	−0.021	0.043	−0.021	−0.032	0.460
	(0.024)	(0.026)	(0.100)	(0.020)	(0.020)	(0.246)
$\Delta\sigma_{arb}$ = arbitration effect on learning from settlements	0.014	0.013	0.018	0.025	0.016	−0.179
	(0.025)	(0.030)	(0.117)	(0.021)	(0.023)	(0.191)
R^2	0.35	0.35	—	0.33	0.35	—
Log likelihood	—	—	−180.8	—	—	−106.4
$\phi(\cdot)$ at the sample means	—	—	0.2376	—	—	5.1×10^{-8}
$n = 418$ (second stage data only)						

[a] All three model estimates included pair fixed effects as regressors.

effect was 0.31, and the average close fixed effect was 0.82. In sum, the initial propensity to play *m* is much higher for close bargainers than far bargainers, and the initial propensity for both far and close is higher in ELA than in NA.

But while the initial propensities to play *m* differ across treatments, there is little difference in offer spread influence (Δspread), and variables concerning past disputes and settlements are all insignificant. Table V estimates raise concern about whether the influence of offer spread and past disputes is meaningfully measured in the stage 2 model, because of the potential collinearity between offer spread and past disputes. To check this, we estimated the model without dispute and settlement variables. Estimates of the offer spread remained stable, independent of estimation technique. We also ran the model with dispute and settlement variables, but without the offer spread. None of the dispute or settlement variables were significant, independent of estimation technique. These results provide some assurance that the Table VI estimates are not biased by collinearity. In sum, we conclude that, fixing an offer spread, arbitration increases the initial propensity to play *m* in stage 2. Initial propensities are little influenced by experience, independent of whether there is an arbitrator.

Prior to analyzing the stage 2 data, we might have postulated that falling dispute rates are the result of bargainers attempting to establish a reputation for toughness in the early rounds, rather than the result of learning. Note, however, that the stage 2 rule that bargainers use to determine whether they accept or maintain does not change as the game progresses. Table VI measures "game progress" in terms of past disputes and settlements. Alternatively, we can measure progress in terms of a round variable. Making the implied substitution in the Table VI model, we found that the round effect is not significant at any standard level. By either method of measuring game progress, we reject the reputation hypothesis.

The (aggregated) strategy we have described for both treatments is essentially a noisy version of the following subgame perfect equilibrium: on the equilibrium path, both bargainers offer and accept equal division. If one or both bargainers deviate and demand more than an equal share, the one who is farthest from equal division accepts, while the other maintains.

We summarize the results of the stage analysis as follows: voluntary settlements cluster around equal division. Bargainers react to a dispute in the previous round by moderating their demands. This diminishes the offer spread, which diminishes the probability of dispute because it lowers the probability bargainers play *m* in stage 2. Given the offer spread, stage 2 behavior is stable with respect to prior negotiation experience. Introducing arbitration has two influences. First, apparently in response to lower dispute costs, bargainers are more likely to insist on their stage 1 offer (play m). Second, the rate at which bargainers adjust their offers in response to a dispute is substantially lowered. So the slower decline in ELA disputes can be traced to the slower rate at which bargainers moderate their demands after a dispute.

6. ANALYSIS OF FOUR ADDITIONAL TREATMENTS

We now repeat the analysis of Section 4 for four new treatments. The additional data provide two important robustness checks. First, in the new treatments, the first six rounds are played with no arbitrator; an arbitrator is then introduced for the last six rounds. This check is important because field bargaining units often have experience negotiating with one another prior to the introduction of arbitration. Second, each new treatment features a distinct arbitrator award distribution. This checks whether slower learning is an artifact of the ELA award distribution.

6.1. *The Arbitrator Award Distribution and the Data*

Arbitrator award distributions vary in the extent to which they treat the bargainers equally, as well as the size and type of arbitration cost. Table VII summarizes the awards distributions for the new treatments. The distribution for LA45 is the same as that for ELA, except that the former draws a separate award for each bargainer. Hence LA45 checks whether nonidentical awards, as well as prior experience bargaining without arbitration, alter the results. The award distribution for LA40 has a lower mean (40) and a higher standard deviation (7.5) than LA45. These changes increase the cost of arbitration, independent of risk posture (see Section 3).

TABLE VII
SA, AA, LA40, and LA45 Treatments: Summary

	Arbitrator award	
Treatment	Player A	Player B
SA	$x \sim N(50, 15)$ truncated at 0 and 100	$100 - x$
AA	$x \sim N(65, 15)$ truncated at 0 and 100	$100 - x$
LA40	$x \sim N(40, 7.5)$ truncated at 0 and 50	$x \sim N(40, 7.5)$ truncated at 0 and 50
LA45	$x \sim N(45, 3.25)$ truncated at 0 and 50	$x \sim N(45, 3.25)$ truncated at 0 and 50

In SA, the award for player A is drawn from a normal distribution, with a mean of 50 and a standard distribution of 15; B receives 100 minus A's award. The SA arbitrator imposes a cost on risk-averse bargainers only (Section 3). All of these award distributions are unbiased; neither bargainer is favored. The arbitrator for AA is the same as for SA, except that the normal distribution for drawing A's award has a mean of 65, making the AA arbitrator biased in favor of A.

Table VIII summarizes early- and late-round dispute data for the new treatments. Dispute rates consistently rise with the introduction of arbitration, although the increase varies with the award distribution (one-tail match pair test, $p < 0.025$ in all cases). The magnitudes of the rise in disputes are generally in line with those observed by Ashenfelter et al. (1992). Also note that dispute rates in the late rounds of LA45 are higher than for LA40 (rank correlation test, $p = 0.013$), consistent with Reder and Neumann's finding on the influence of dispute costs.

6.2. *Estimating the Model with the New Data*

Table VIII shows that the propensity to dispute jumps with the introduction of arbitration. To account for this, we modify (4.1) and (4.2) by including two fixed effects per bargaining pair, one for early rounds without arbitration and one for late rounds with arbitration.[16]

The outcome (λ) estimates in Table IX strongly favor outcome learning over round effects for all four treatments (SA, AA, LA40, LA45). The result is mostly independent of the estimation method ($p < 0.05$ in all cases, except the AR1 estimate for outcome−arb in LA40 treatment, where $p = 0.084$; the AR1 estimate for AA is not significant).

Table X displays estimates for the outcome learning model (4.2). The estimates of the influence of past dispute are all negative, and all but one is significant (Disputes−NA, $p < 0.05$ except the AR1 estimate for AA). In contrast, all of the estimates of the influence of past settlements, Settlements−NA, while positive, are small relative to Disputes−NA. Many of the Settlements−NA estimates are significant, although not uniformly so. Overall, the evidence indicates that past disputes are the major factors behind NA learning. Past settlements appear to have at most a minor influence.

As before, we interpret the $\Delta\delta_{\text{arb}}$ as estimates of the narcotic effect. Excluding the AA treatment, estimates for the narcotic effect are generally positive and significant ($p < 0.05$, except the LSDV estimates for SA and

[16] We tested a restriction in which the fixed effects under arbitration are equal to those without arbitration plus a constant (same constant for all bargaining pairs in a treatment). The restriction implies that the distribution of fixed effects with arbitration is equal to that without after a location shift. The data, however, reject the restriction.

TABLE VIII
SA, AA, LA40, LA45: Summary of Dispute Rates (Standard Error)

Treatment	Observations	Rounds 1–6	Rounds 7–12	Difference
SA	26	0.128	0.436	0.308
		(0.0373)	(0.0705)	(0.0618)
AA	26	0.141	0.673	0.532
		(0.0400)	(0.0676)	(0.0654)
LA40	27	0.117	0.222	0.105
		(0.0354)	(0.0590)	(0.0488)
LA45	27	0.204	0.395	0.191
		(0.0430)	(0.0611)	(0.0633)
All	**106**			

LA45, $p = 0.051$ and 0.069, resp., and the probit estimate for SA, $p = 0.12$). The AA data hint at a narcotic effect—the LSDV and AR1 estimates are positive—but none of the estimates are significant. Hence with the exception of the AA treatment, the new data add to our earlier evidence for a narcotic effect.

Of course, taking $\Delta\delta_{\text{arb}}$ as the measure of the narcotic effect assumes that differences in the response to past settlements, $\Delta\sigma_{\text{arb}}$, are negligible. In fact, the estimates in the $\Delta\sigma_{\text{arb}}$ row are all small and mostly insignificant, indicating that there is little difference across treatments in response to past settlements. As a further check, we reestimated (4.2), holding the coefficient for past settlements constant; that is, $\sigma_j = \sigma$ for all treatments j. This perhaps provides a fairer comparison of the narcotic effect variables. The results, displayed in Table XI, only strengthen our conclusions: the LSDV estimate of the narcotic effect in NA/ELA and the probit estimate for SA are now significant ($\Delta\delta_{\text{arb}}$, $p = 0.057$ and $p = 0.027$, respectively). The probit estimate of the narcotic effect in AA is now positive, but remains insignificant. We also estimated Table XI by dropping the settlement variables altogether; the conclusions do not change.

To summarize: the additional analysis strengthens the Section 4 conclusions. We have strong evidence that the outcome learning model has more explanatory power than round effects. In other words, bargainers learn from the outcomes of earlier negotiations, and primarily from past disputes. Introducing arbitration tends to slow, but not stop, learning. With the exception of the AA treatment, there is strong evidence for a narcotic effect. Finally, there is little evidence that settlements play anything more than a minor role in learning.

One further observation: the arbitrator award distribution was very similar for ELA and LA45, but ELA bargaining units were given prior

TABLE IX

SA, AA, LA40, LA45 Treatments: Round Effect vs. Outcome Learning (Standard Error)[a]

	SA	AA	LA40	LA45
LSDA				
Round−NA	−0.100	−0.042	−0.125	−0.083
	(0.300)	(0.028)	(0.025)	(0.029)
Round−arbitration	0.083	0.051	0.069	0.046
	(0.034)	(0.033)	(0.039)	(0.036)
Outcome−NA	−0.134	−0.075	−0.160	−0.191
	(0.035)	(0.033)	(0.028)	(0.047)
Outcome−arbitration	0.076	−0.061	−0.067	−0.098
	(0.019)	(0.022)	(0.020)	(0.033)
R^2	0.56	0.63	0.53	0.38
Two tail p-value for H_O: Outcome−NA = Outcome−arbitrator = 0	9.0×10^{-7}	2.1×10^{-3}	3.3×10^{-9}	8.1×10^{-6}
AR1				
Round−NA	−0.103	−0.025	−0.129	−0.119
	(0.034)	(0.031)	(0.028)	(0.036)
Round−arbitration	0.082	0.030	0.081	0.096
	(0.040)	(0.038)	(0.035)	(0.045)
Outcome−NA	−0.132	−0.051	−0.151	−0.231
	(0.037)	(0.035)	(0.031)	(0.053)
Outcome−arbitration	−0.069	−0.017	−0.036	−0.125
	(0.021)	(0.023)	(0.021)	(0.037)
R^2	0.64	0.70	0.66	0.46
Two-tail p-value for H_O: Outcome−NA = Outcome−arbitrator = 0	1.0×10^{-4}	3.846×10^{-1}	1.2×10^{-5}	6.9×10^{-6}
Probit				
Round−NA	−0.406	−0.127	−1.116	−0.660
	(0.163)	(0.124)	(0.330)	(0.170)
Round−arbitration	0.421	0.271	0.789	0.510
	(0.189)	(0.169)	(0.351)	(0.186)
Outcome−NA	−1.222	−0.636	−3.420	−1.500
	(0.343)	(0.223)	(0.963)	(0.336)
Outcome−arbitration	−0.854	−1.057	−1.099	−0.452
	(0.199)	(0.262)	(0.305)	(0.138)
Log likelihood	−69.221	−71.485	−41.125	−94.421
$\phi(\cdot)$ at the sample means	0.0125	0.2484	0.0243	0.3107
Rounds	12	12	12	12
Pairs	26	26	27	27
n	312	312	324	324

[a] All three model estimates included pair fixed effects as regressors.

TABLE X
SA, AA, LA40, LA45 Treatments: Outcome Learning Model (Standard Error)[a]

	SA	AA	LA40	LA45
LSDV				
Disputes−NA	−0.235	−0.117	−0.286	−0.274
	(0.063)	(0.058)	(0.052)	(0.067)
Settlements−NA	0.034	0.032	0.034	0.107
	(0.017)	(0.018)	(0.014)	(0.041)
$\Delta\delta_{arb}$ = arbitration effect	0.141	0.065	0.162	0.140
on learning from disputes	(0.069)	(0.062)	(0.062)	(0.077)
$\Delta\sigma_{arb}$ = arbitration effect	0.024	0.037	−0.022	−0.046
on learning from settlement	(0.029)	(0.038)	(0.021)	(0.058)
R^2	0.56	0.63	0.53	0.46
AR1				
Disputes−NA	−0.235	−0.077	−0.286	−0.350
	(0.069)	(0.063)	(0.057)	(0.078)
Settlements−NA	0.029	0.026	0.021	0.112
	(0.022)	(0.023)	(0.018)	(0.048)
$\Delta\delta_{arb}$ = arbitration effect	0.146	0.065	0.196	0.202
on learning from disputes	(0.075)	(0.068)	(0.068)	(0.091)
$\Delta\sigma_{arb}$ = arbitration effect	0.019	−0.0030	−0.032	−0.010
on learning from settlement	(0.034)	(0.043)	(0.026)	(0.067)
R^2	0.64	0.69	0.66	0.46
Probit				
Disputes−NA	−1.629	−0.763	−4.537	−2.160
	(0.470)	(0.308)	(1.254)	(0.481)
Settlements−NA	0.816	0.509	2.303	0.839
	(0.261)	(0.188)	(0.709)	(0.229)
$\Delta\delta_{arb}$ = arbitration effect	0.789	−0.149	3.110	1.558
on learning from disputes	(0.514)	(0.389)	(1.309)	(0.509)
$\Delta\sigma_{arb}$ = arbitration effect	0.053	0.691	−1.531	−0.536
on learning from settlement	(0.349)	(0.378)	(0.750)	(0.274)
Log likelihood	−69.221	−71.485	−41.125	−94.421
$\phi(\cdot)$ at the sample means	0.0125	0.2485	0.0243	0.3107
Rounds	12	12	12	12
Pairs	26	26	27	27
n	312	312	324	324

[a] All three model estimates included pair fixed effects as regressors.

TABLE XI
All Treatments: Outcome Learning Model with Fixed Settlements (Standard Error)[a]

	NA/ELA	SA	AA	LA40	LA45
LSDV					
Disputes−NA	−0.161	−0.244	−0.125	−0.276	−0.256
	(0.031)	(0.062)	(0.058)	(0.051)	(0.063)
$\Delta\delta_{arb}$ = arbitration effect	0.063	0.155	0.077	0.147	0.112
on learning from disputes	(0.022)	(0.067)	(0.061)	(0.060)	(0.069)
R^2	0.28	0.56	0.63	0.53	0.38
AR1					
Disputes−NA	−0.211	−0.239	−0.073	−0.276	−0.347
	(0.033)	(0.068)	(0.063)	(0.057)	(0.074)
$\Delta\delta_{arb}$ = arbitration effect	0.108	0.150	0.062	0.186	0.197
on learning from disputes	(0.035)	(0.075)	(0.067)	(0.068)	(0.084)
R^2	0.32	0.64	0.70	0.66	0.46
Probit					
Disputes−NA	−0.681	−1.669	−0.963	−2.649	−1.626
	(0.145)	(0.393)	(0.307)	(0.556)	(0.339)
$\Delta\delta_{arb}$ = arbitration effect	0.341	0.842	0.299	0.830	0.909
on learning from disputes	(0.142)	(0.381)	(0.312)	(0.492)	(0.339)
Log likelihood	−213.35	−69.223	−73.25	−44.009	−128.78
$\phi(\cdot)$ at the sample means	0.0643	0.0197	0.2057	0.0908	0.3416
Rounds	12	12	12	12	12
Pairs	50	26	26	27	27
n	600	312	312	324	324

[a] All three model estimates included pair fixed effects as regressors.

experience negotiating without arbitration. Nevertheless, initial dispute rates for negotiations with arbitration were very similar for both treatments. This suggests that the learning we observe is context dependent. There is substantial evidence in the cognitive psychology literature that context dependence is a general phenomenon (Perkins and Salomon, 1989, survey the literature.)

7. SUMMARY AND DISCUSSION

Our main findings can be summarized as follows:

1. Our experiment reproduces observations reported in a variety of field and lab studies. There is strong evidence for bargainer learning, both with and without arbitration. This is consistent with the field findings of

Reder and Neumann (1980), Butler and Ehrenberg (1981), and Lester (1989). Our data also confirm the Currie and McConnell (1991) field finding and the Ashenfelter et al. (1992) lab finding that dispute rates are higher with arbitration than without. Like Kochan and Baderschneider's (1978) field study and Ashenfelter et al., we find strong evidence for bargainer heterogeneity with respect to propensity to dispute. We find that dispute rates are sensitive to dispute costs; Reder and Neumann, and Currie and McConnell find the same.[17]

2. Virtually all of the learning we observe derives from past dispute incidents. We tested a version of an outcome learning model suggested by Butler and Ehrenberg (1981) against a baseline round-effects model. The outcome learning model was robustly the better performer. A dispute decreases the probability a dispute will happen in succeeding rounds. Past settlements have little effect.

3. Learning with arbitration tends to be slower than it is without; in this sense, arbitration has a narcotic effect. This result is robust to prior experience negotiating without arbitration, and to a variety of arbitrator award distributions.

4. Bargainers react to a prior round dispute by moderating their demands. The pattern of conceding to/maintaining demands is stable in the sense that it depends only on the demands made, not on prior negotiation experience. Arbitration has two influences: first, bargainers are more likely to insist on their demand. Second, the rate at which bargainers moderate their demands in response to a dispute is substantially lowered. The pattern of learning was unaffected by prior experience negotiating without arbitration, suggesting that the learning is context dependent.

The findings on the nature of bargainer learning (2–4) represent the unique contribution of this study. The fact that the simple lab negotiation captures so many features of field negotiations (finding 1) holds out the promise that field and lab negotiations are governed by a common underlying mechanics. That said, lab negotiations are simpler than most field negotiations. Some findings must therefore be extrapolated to the field by analogy. A literal reading of Table IV, for example, would imply that field bargainers avoid disputes by dividing 50–50. A more sophisticated reading involves an analogy with the general concept of protocol. Field protocols are no doubt dependent on a variety of considerations (e.g., precedent,

[17] Another parallel: In the AA treatment, where arbitrator awards were biased in favor of one bargainer, voluntary settlements moved significantly toward the advantaged bargainer ($p < 0.0001$). Olson and Rau (1997) find that voluntary teacher–school board settlements move in the direction of prior round arbitration awards.

culture); equal division is the particular protocol we observe in the lab. Casual observation suggests that protocols are indeed important to field negotiations. Schelling (1963) observes

> the remarkable frequency with which long negotiations over complicated quantitative formulas or *ad hoc* shares in some costs or benefits converge ultimately on something as crudely simple as equal shares, shares proportionate to some common magnitude (gross national product, population, foreign-exchange deficit, and so forth), or the shares agreed on in some previous but logically irrelevant negotiation. (p. 67; quote excerpted from 1980 reprint)

In the field, disputes do not always decrease with time; remember that Butler and Ehrenberg identified a period of rising as well as falling disputes. We think protocol learning is consistent with this episode. Recall that the rise in disputes occurred immediately after arbitration was introduced. With new institutions, bargainers may see new strategic opportunities. If old protocols unravel, disputes may rise (as happened in our experiment). In complicated negotiations, unraveling may take some time. But as bargainers learn about one another's behavior under the new institution, disputes should eventually drop, as Butler and Ehrenberg found.

The learning process can also explain why disputes sometimes go up outside the introductory phase. To give a specific example, Currie (1989) analyzes data that track the history of wage negotiations between British Columbia school boards and local branches of the Teachers' Federation (BCTF). The data spans 35 years of negotiation with arbitration. There is a good deal of variation in annual dispute rates (including two long periods in which rates tend to decrease). We focus on a dramatic rise from 1971, when disputes were at an all-time low of about 3%, to 1972, when the rate was over 60%. Reder and Neumann reason that the ideal protocol would cover all possible states of the world, but "In many of these states agreement would be difficult to achieve ... and ... some of these states may have only a small probability of realization. Therefore, extending a protocol to cover all of these (low-probability) states is cost inefficient, and bargainers often fail to cover some of them, with the result that they are compelled to bargain de novo whenever one of these improbable states emerges" (p. 871). This is precisely the explanation given by Thompson and Cairnie (1973) for the 1971–72 episode:

> [T]he contrast between the negotiation results of 1971 and 1972 arose from experience with a new method of determining salary increases. This new method consisted of a formula based on weighted averages of wage and salary changes in the private sector. Originally proposed by the BCTF, the formula was accepted by school trustees for the 1971 negotiations, resulting in the virtual absence of arbitration. After their initial experience with the formula

> and the relatively generous settlements it produced, however, the trustees refused to accept it as a basis for a second round of negotiations, and bargaining for 1972 salaries was bitter and generally fruitless. (p. 13)

So in 1971, the parties adopted a very explicit protocol covering wage increases. At the time of the agreement, the parties, or at least the school boards, apparently considered the circumstances that arose in 1971 to be improbable. The breakdown of the protocol in 1972 made coordination harder, and dispute rates rose.

In sum, the most important implication of our work is that focusing on the mechanics of bargainer learning may lead to a new and fruitful understanding of conflict resolution procedures. Because the issues are closely related, such a study may yield important insights into the question of why bargaining sometimes ends in dispute in the first place. Some speculation: It is not apparent to us that outside party intervention need slow bargainer learning. The right sort of intervention might promote faster learning—truly effective conflict resolution.

APPENDIX: WRITTEN INSTRUCTIONS PROVIDED TO SUBJECTS

General

Please read the instructions carefully. If at any time you have questions or problems, raise your hand and the monitor will be happy to assist you. From now until the end of the session, unauthorized communication of any nature with other participants is prohibited.

At the end of the session, you will be paid a $5 cash show-up fee. During the session, you will play a series of bargaining games with another participant. Each game gives you an opportunity to earn additional cash.

Description of the Bargaining Game

The game involves two bargainers, Player A and Player B. They must decide how to divide 100 (abstract) chips. The game is played in two stages:

Stage 1. Each bargainer proposes a division of the chips. The computer will display both proposals simultaneously, meaning that neither bargainer will be able to see the other bargainer's proposal before making their own.

Stage 2. After reviewing both proposals, each bargainer decides whether to ‘‘accept’’ the other bargainer's proposal or to ‘‘maintain’’ their own proposal. The computer will display both decisions simultaneously, meaning that neither bargainer will be able to see the other bargainer's decision before making their own.

Chip Division Rules. The decisions made in Stages 1 and 2 determine how the chips are divided:

—If one bargainer ‘‘accepts’’ and the other ‘‘maintains,’’ then the maintaining bargainer’s proposal determines the number of chips each bargainer receives.

—If both bargainers ‘‘accept’’ then, using a process that is equivalent to a coin flip, the computer will randomly choose one bargainer’s proposal.

—If both bargainers ‘‘maintain’’ then each bargainer receives zero (0) chips. (*For ELA*45 *this was replaced with*: If both bargainers ‘‘maintain’’ then an arbitrator will determine how many chips each bargainer receives. Your folder includes a list of the arbitrators’ last 100 decisions. Please review this list. You should expect similar decisions in the future.)

Please note: Chip division rules may or may not be changed for some of the games. We cannot tell you at this time what the potential changes are. However, any change will be fully explained before it is put into effect.

Role Assignments

You will have the same role, Player A or Player B, for all games. Your role is determined by the ‘‘A’’ or ‘‘B’’ that precedes the cubicle number you drew when you entered the room. Your role will also appear on your computer screen during games.

Pairing Procedure

You will be paired with the same person for all games. This person will be selected at random from the group of participants in the room who have the opposite role that you have. All pairings are anonymous: you will not know the identity of the person you are playing, nor will they know yours, nor will these identities be revealed after the session is completed.

Bargaining Record

Several blank ‘‘Bargaining Records’’ are provided in your folder. At the conclusion of each game fill out one of these forms. Completed forms provide you with a history of your past games, and you may reference them at any time during the session.

Chip Values

For every game, each chip has a value of $0.24. So the value of obtaining Z chips in a game is $0.24 \times Z$ dollars. For your convenience, the computer will automatically calculate the value of Stage 1 proposals. Scratch paper and a pen have been provided (in your folder) for any further calculations you might wish to perform or if you wish to make private notes.

Money Earnings

You will be paid the money you make for one game. We will play more than one game. The one that you are paid for will be selected by a lottery at the conclusion of the session. Each game played has an equal chance of being selected, so it is in your interest to make as much money as you can in each and every game. You will be paid your earnings in cash, immediately upon conclusion of the session. Earnings are confidential: only you and the monitor will know the amount you make.

Consent Forms

If you wish to participate in this study, please read and sign the accompanying consent form. Please note: In order to collect your earnings from the game, you must stay until the end of the session, which will last about 90 minutes.

Practice Games

You should now play some practice games. Practice until you feel comfortable with the game and its rules. Be aware that no money will be paid for the practice games (so you may experiment freely). Second, your bargaining partner for the practice games will be the computer. It has been programmed to make proposals and decisions completely at random, so do not be concerned if its bargaining appears irrational. Remember, the point of practice is to familiarize you with the game and its rules before you are paired with another person.

(*The following was given to bargainers in the arbitration treatments after round* 6.)

Notice of Change in Division Rules

The chip division rules will be changed to the following for all of the remaining games in the session:

Chip Division Rules. The decisions made in Stages 1 and 2 determine how the chips are divided:

—If one bargainer "accepts" and the other "maintains," then the maintaining bargainer's proposal determines the number of chips each bargainer receives.

—If both bargainers "accept" then, using a process that is equivalent to a coin flip, the computer will randomly choose one bargainer's proposal.

The above two rules are the same as before, but the last rule is different:

—If both bargainers "maintain" then an arbitrator will determine how many chips each bargainer receives.

Below are the arbitrators' last 100 decisions. You should expect similar decisions in the future.

ACKNOWLEDGMENT

Many thanks to Keith Ord for the statistical consultations.

REFERENCES

Allison, J. R. (1990). "Five Ways to Keep Disputes Out of Court," *Harvard Business Rev.*, January–February, 166–177.

Anderson, J. C. (1981). "The Impact of Arbitration: A Methodological Assessment," *Indust. Relations* **20**, 129–148.

Ashenfelter, O. (1987). "A Model of Arbitrator Behavior," *Amer. Econom. Rev.* **77**, 342–346.

Ashenfelter, O., and Bloom, D. E. (1984). "Models of Arbitrator Behavior: Theory and Evidence," *Amer. Econ. Rev.* **74**, 111–124.

Ashenfelter, O., Currie, J., Farber, H. S., and Spiegel, M. (1992). "An Experimental Comparison of Dispute Rates in Alternative Arbitration Systems," *Econometrica* **60**, 1407–1433.

Bazerman, M. H., and Neale, M. A. (1982). "Improving Negotiation Effectiveness Under Final Offer Arbitration: The Role of Selection and Training," *J. Appl. Psychol.* **67**, 543–548.

Bloom, D. E. (1981). "Is Arbitration *Really* Compatible with Bargaining?" *Indust. Relations* **20**, 233–244.

Bolton, G. E. (1997). "The Rationality of Splitting Equally," *J. Econom. Behav. Organization* **32**, 365–381.

Brams, S. J., Kilgour, D. M., and Merrill, S. (1991). "Arbitration Procedures," in *Negotiation Analysis* (Peyton Young, Ed.), pp. 47–65. Ann Arbor: Univ. of Michigan Press.

Butler, R. J., and Ehrenberg, R. G. (1981). "Estimating the Narcotic Effect of Public Sector Impasse Procedures," *Indust. Labor Relations Rev.* **35**, 3–20.

Chamberlain, G. (1980). "Analysis of Covariance with Qualitative Data," *Rev. Econom. Stud.* **146**, 225–238.

Charness, G. (1998). "Bargaining Efficiency and Screening: An Experimental Investigation," working paper, Universitat Pompeu Fabra, Barcelona.

Crawford, V. P. (1982). "Compulsory Arbitration, Arbitral Risk and Negotiated Settlements: A Case Study in Bargaining under Imperfect Information," *Rev. Econom. Stud.* **49**, 69–82.

Currie, J. (1989). "Who Uses Interest Arbitration? The Case of British Columbia Teachers, 1947–1981," *Indust. Labor Relations Rev.* **42**, 363–379.

Currie, J., and McConnell, S. (1991). "Collective Bargaining in the Public Sector: The Effect of Legal Structure on Dispute Costs and Wages," *Amer. Econom. Rev.* **81**, 693–718.

Currie, J., and McConnell, S. (1996). "Collective Bargaining in the Public Sector: Reply," *Amer. Econom. Rev.* **86**, 327–328.

Farber, H. S., and Bazerman, M. H. (1986). "The General Bias of Arbitrator Behavior: An Empirical Analysis of Conventional and Final-Offer Arbitration," *Econometrica* **54**, 1503–1528.

Gale, J., Binmore, K., and Samuelson, L. (1995). "Learning to be Imperfect: The Ultimatum Game," *Games Econom. Behav.* **8**, 56–90.

Gunderson, M., Hebdon, R., and Hyatt, D. (1996). "Collective Bargaining in the Public Sector: Comment," *Amer. Econom. Rev.*, **86**, 315–326.

Harsanyi, J. C. (1977). *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge, England: Cambridge Univ. Press.

Kochan, T. A., and Baderschneider, J. (1978). "Dependence on Impasse Procedures: Police and Firefighters in New York State," *Indust. Labor Relations Rev.* **31**, 431–449.

Lester, R. A. (1984). *Labor Arbitration in State and Local Government*. Princeton, NJ: Princeton Univ. Industrial Relations Section.

Lester, R. A. (1989). "Analysis of Experience Under New Jersey's Flexible Arbitration System," *Arbitration J.* **44**, 14–21.

Mundlak, Y. (1978). "On the Pooling of Time-Series and Cross-Section Data," *Econometrica* **46**, 69–85.

Olson, C. A., and Rau, B. L. (1997). "Learning from the Interest Arbitration Experience: The Next Round," *Indust. Labor Relations Rev.* **50**, 237–251.

Perkins, D. N., and Salomon, G. (1989). "Are Cognitive Skills Context-Bound?" *Educational Researcher* **18**, 16–25.

Reder, M. W., and Neumann, G. R. (1980). "Conflict and Contract: The Case of Strikes," *J. Political Economy* **88**, 867–886.

Roth, A. E. (1995). "Bargaining Experiments," in *Handbook of Experimental Economics* (J. Kagel and A. E. Roth, Eds.). Princeton, NJ: Princeton Univ. Press.

Roth, A. E., and Erev, I. (1995). "Learning in Extensive Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games Econom. Behav.* **8**, 164–212.

Roth, A. E., and Schoumaker, F. (1983). "Expectations and Reputations in Bargaining: An Experimental Study," *Amer. Econom. Rev.* **78**, 362–372.

Schelling, T. C. (1963). *The Strategy of Conflict*. New York: Oxford Univ. Press.

Thompson, M., and Cairnie, J. (1973). "Compulsory Arbitration: The Case of British Columbia Teachers," *Indust. Labor Relations Rev.* **27**, 3–17.

Wirtz, W. (1963). "Address Before National Academy of Arbitrators," *Daily Labor Rep.* **23**, F1–F4.