Erratum for Bargaining and Reputation: An Experiment on Bargaining in the Presence of Behavioral Types*

Matthew Embrey
University of Sussex

Guillaume R. Fréchette
New York University

Steven F. Lehrer
Queen's University and NBER

July 16, 2019

Abstract

This note corrects an error in the units of the delay variable in *Embrey et al.* (2015).

1 Introduction

This note corrects an error in the units of the delay variable reported in *Embrey et al.* (2015). It reports the results of implementing a two-stage bilateral bargaining game. In the first stage, subjects make announcements for an amount out of 30 that they would like for themselves. If these announcements are incompatible—sum to more than 30—then bargaining proceeds to a second stage. The second stage is a concession stage, where subject can either continue to wait for the other to accept their offer, or concede and accept the other's offer.

*We wish to thank Samreen Malik for discovering and alerting us to the discrepancy in the delay statistics that we are correcting with this Erratum. We are responsible for all errors.

The computer interface implemented the concession stage with a refresh rate of 1/100th of a second (i.e. the clock proceeds in ticks of 1/100th of a second). As is common in many programming frameworks, the software records the time in the concession stage in milliseconds. However, this resulted in the units of delay being ten times smaller than the smallest time interval that was measured. Not recognizing this discrepancy meant that the actual average delay statists are 1/10 of those reported in the original paper; equivalently, the units should be 1/10th of a second rather than one second.

The key finding of *Embrey et al.* (2015) is the observation that subjects mimic behavioral types. This finding is unaffected by the change in units since it is based on first-stage announcement data. The more detailed analysis that did explore second-stage delay did not focus heavily on the point estimates for average delay, not least because average delay by subgame appeared too long compared to a simple theoretical upper bound.[1] After noting that delays appeared significantly above these upper bounds, we focussed on a comparative static exercise that compared average delay in *explicit* subgames against other subgames.[2] The main result of this analysis was that average delay is closer to the upper bound for explicit subgames, in particular in sessions with both a restricted design and induced types.

The first result no longer holds with the corrected delay units as average delay is typically within the simple upper bounds. The second result, being based on a comparative static, is unaffected by the change in units. Average delays remain smaller compared to respective delay bounds (i.e. there is less delay on a subgame by subgame basis) for explicit subgames than non-explicit subgames. Section 2 provides the corrected tables and statistical tests for the average delay analysis, which was covered in Subsection 5.2 of the original paper. Whether subjects engaged in overly long conflicts in the second stage, in either explicit or non-explicit subgames, requires a more detailed analysis that is given below in Section 3.[3]

[1]This upper bound does not depend on the finer details of the equilibrium predictions. It is found by ignoring the possibility of initial concession and that the support for concession by rational players is a strict subset of $[0, \infty)$.

[2]Explicit subgames were those that involved both announcements being for either a demand that mimicked some computer player, or a demand for a 50-50 split (i.e. 15 out of 30).

[3]That average delay was within its upper bound does not in itself lead to the conclusion that observed delays are consistent with theoretical predictions. How tight the upper bound is to the actual predicted average delay depends on both how long concession should last—something that is decreasing in the likelihood of the other player being irrational—

It is important to emphasize that the majority of the results are not affected by the correction to the delay units. Subsection 5.1, which covered first-stage announcements and outcomes based on average points earned, is completely unaffected by the units of the delay variable. In particular, the number of points earned in a round was computed directly and not derived using the delay variable. The concession analysis in Subsection 5.3 is also not affected by the delay units, with the exception that the time cut-off for initial concession reported in footnote 39 and in Figure 2 should read 2/10th of a second (with a robustness analysis in the online appendix that looks at 1/10, 1/2 and 1 second) rather than 2 seconds.[4]

Indeed this correction leaves unaffected the analyses for the bulk of the paper. The main results, as stated in the conclusions of the original paper, are:

1. Subjects recognize the role of reputation and mimic induced obstinate types.

2. Evidence for the presence of complementary types that acquiesce to credible obstinate demands.

3. There is a tendency to make more demanding announcements too often and to remain in the concession stage too long.

4. Second-stage behavior is closer to the predicted pattern in sessions with a restricted design and induced types.

---

and whether a rational player should concede initially or not. This makes the bound tighter for symmetric subgames—or for all subgames when the set of behavioural types includes only aggressive types, such as in treatments *R0* and *R3*—because initial concession is not predicted, and for subgames where the probability of meeting a behavioural type is smaller. It is worth noting in this latter case, however, that overall delay (i.e. aggregating over the likelihood of being in each subgame) is predicted to be small, as the probability of being matched to a behavioural type becomes small, so long as the Nash Bargaining demand is in the set of behavoural types—here that means a demand of 15 out of 30 (see *Abreu et al.* 2015). Aggregate delay is predicted to be smaller, despite subgame by subgame delay potentially getting longer, because rational players should make the Nash bargaining demand (50-50) more often, leading to less second stages. As reported in the original paper, there is evidence (not affected by the delay units) that subjects are making more demanding announcements, such as 20 out of 30, too often.

[4]Given initial concession should be instantaneous, the analysis is based on a threshold that is much closer to instant reaction times.

Only the second part of point 3 is affected by the correction to the delay units, and the correction does not eliminate the evidence that there are quantitative deviations from the finer predictions of the sequential equilibrium. Overall, the corrected, and more reasonable, second-stage delay statistics add further weight to the assertion in the conclusion of the original paper that, despite the complexity of the equilibrium predictions, the two-sided reputation model performs well in predicting bargaining behavior, especially when the possibilities for behavioral play are transparent (i.e. in explicit subgames).

2 Corrected Delay Analysis

This section presents the corrected table and statistical tests reported in Section 5.2 of the original paper. Table 1 presents the summary information on delay in the second stage, originally reported in Table 6 (page 623) of *Embrey et al.* (2015); columns 6 and 8 have been corrected. As can be seen, and contrary to footnote 36 of the original paper, mean delay is consistently shorter than the theoretical upper bound, both overall and for individual subgames with at least 15 observations. The observed delays are significantly shorter than their respective upper bounds for all treatments $[p, p_n < 0.001]$.[5]

Contrasting the performance across designs, mean delay is further below the bounds dictated by the model in the restricted design: the delay to bound ratio is smaller at 0.47, compared to 0.66 in the unrestricted design. As before, the difference between the designs is particularly notable for treatments with computer players – that is, *U1* and *U2* versus *R3* and *R4* – and is statistically significant $[p, p_n < 0.05]$. Also as before, this difference is statistically insignificant if only explicit subgames are used – that is, subgames involving announcements that might have been made by a computer or announcements of 15 $[p, p_n > 0.1]$.[6] Finally, as before, comparing treatments within the restricted design provides further evidence that explicit subgames have shorter delays $[p, p_n < 0.01]$.

In summary, the comparisons across designs are unaffected by the correction to the units of the delay variable. However, given that in all cases

[5]Statistical tests as in the original paper.

[6]For *U1*, only the subgames 15-20 and 20-20 are explicit; for *U2*, 12-20, 15-20 and 20-20. For *R3* and *R4* all subgames are explicit by design. For *C0* and *R0*, none of the subgames are explicit since there are no computer players, and an announcement pair of 15-15 does not result in a second-stage.

average delay is below the theoretical bound, it is no longer possible to interpret the reduction in average delay by subgame as necessarily an indication of improved performance of the model's prediction.

	Subgame[§]		Obs		Delay (Seconds)		
Treatment	α_L	α_H	Freq	%	Mean	Bound	Ratio[§§]
C0	*All*		180		24.8		0.58
	15	20	17	9.4	16.3	20.0	
U1	*All*		162		20.3		0.66
	15	20	32	19.8	5.2	20.0	
	20	20	24	14.8	37.8	50.0	
U2	*All*		183		13.2		0.75
	15	20	29	15.8	9.4	20.0	
	20	20	27	14.8	9.2	50.0	
R0	*All*		245		15.9		0.58
	15	18	38	15.5	6.4	11.1	
	15	20	139	56.7	10.3	20.0	
	18	20	27	11.0	21.0	36.4	
	20	20	40	16.3	40.6	50.0	
R3	*All*		193		6.9		0.45
	15	20	119	61.7	5.3	20.0	
	18	20	16	8.3	13.6	36.4	
	20	20	49	25.4	8.7	50.0	
R4	*All*		160		6.1		0.32
	15	18	15	9.4	2.5	11.1	
	15	20	73	45.6	2.6	20.0	
	18	20	17	10.6	9.3	36.4	
	20	20	51	31.9	11.2	50.0	

Table 1: Second-Stage Delay

[§] Only subgames with at least 15 observations reported.
[§§] Weighted (frequency) average of mean delay divided by bound.

3 Second-Stage Delay and Equilibrium Predictions Revisited

A strong prediction from the stylized model of *Abreu and Gul* (2000) concerns the interval of time over which rational types might concede. Since neither player will continue to hold out once the prior probability of their opponent being irrational has hit one, there is a strictly finite upper bound for longest time it could take for an agreement to be reached (i.e. conditional on the game not ending in outright disagreement, which only happens if two irrational types meet). The rate at which an opponent is predicted to concede (unconditional on knowing whether they are rational or not) is determined by the two initial demands, which must be incompatible to reach the second stage, and the discount rate. Consequently the probability that this opponent is rational, which must start strictly less than one, will decrease until it becomes zero. At this point, neither player should still be in the concession game unless both are irrational; but then the game will end in disagreement.

As shown in Section 1.6.2 of the online appendix, the maximum time before concession can be bounded using the (unconditional) probability of meeting an irrational type, and the smaller that probability the longer the bound. This can be inverted so that the longest observed delay, for each second-stage announcement combination, can be used to construct an upper bound on the probability of a player being irrational. For treatments with induced types, this upper bound can be compared to the lower bound implied by the presence of pre-programmed robot players. This analysis is done in Table 2.[7] For all treatments with induced types, except *R4* there is evidence of excessive delays. In *U1*, the longest delay in the $15-20$ subgame is inconsistent with there being at least a $2/15$ chance of being matched with a 20-type computer player; similarly, for the *U2* treatment, the longest delays observed in the $14-20$ and $15-20$ subgames are inconsistent with a $1/15$ chance of being matched with the 20-type computer player. In *R3*, the longest delays observed in the $15-18$, $18-18$, $18-20$ and $20-20$ subgames all provide evidence of excessive delay given the induced probability of meeting either of the 15, 18, or 20-type computer players (a probability of $1/15$ for each).

To complete the analysis, Table 3 shows this analysis for the two treatments that do not include induced types. For *C0*, the longest delays in the $15-19$ and $15-23$ subgames suggest that the probability of being matched with a subject 15-type cannot be so large. Similarly, in *R0*, the longest delays observed in the $15-18$ and $15-20$ suggest that the probability of subject types, of any of the three possible demands, also cannot be large. While these observations do not in themselves contradict the predictions of the

[7] This prediction is not dependent on both players being rational; it is only necessary that at least one player is rational so that there is eventually an agreement at some point. Consequently the analysis includes all subgames except those where two robot players were matched.

Subgame		Obs		Delay	Upper Bounds§	
α_L	α_H	Freq	%	T_{max}	z_L	z_H
U1						
11	20	5	1.3	21.3	1.7	11.9
14	20	8	2.0	54.6	11.2	25.5
15	20	52	13.1	221.1	0.1	1.2
15	25	7	1.8	31.8	62.1	85.3
20	20	51	12.9	180.2	16.5	16.5
20	21	6	1.5	205.3	15.5	18.6
20	25	7	1.8	79.4	58.9	76.7
U2						
12	20	27	6.8	27.4	8.5	25.4
14	20	10	2.5	682.7	0.0	0.0
15	18	6	1.5	8.7	64.7	70.6
15	20	37	9.3	233.2	0.1	0.9
16	20	9	2.3	84.7	13.9	24.4
18	20	10	2.5	70.5	34.7	41.4
20	20	47	11.8	92.1	39.8	39.8
20	25	14	3.5	128.3	42.5	65.2
20	29	7	1.8	246.2	27.4	87.8
R3						
15	18	33	8.4	366.0	0.0	0.0
15	20	158	40.4	85.8	7.6	18.0
18	18	5	1.3	150.2	5.0	5.0
18	20	41	10.5	211.2	4.2	7.1
20	20	67	17.1	555.2	0.4	0.4
R4						
15	18	37	9.7	39.3	14.0	20.8
15	20	102	26.7	62.9	15.1	28.4
18	18	9	2.4	80.6	20.0	20.0
18	20	38	9.9	111.8	18.7	24.7
20	20	75	19.6	186.3	15.5	15.5

Table 2: Further Analysis of Delay in Treatments with Induced Types

For *U1* and *U2*, only subgames with at least 5 observations reported.

§ Upper bounds on z (in %) implied by T_{max}.

model, it provides further evidence that subjects are likely to be announcing 20 too often in the first stage. This is because rational players should make the Nash bargaining demand (15 out of 30) more often as the probability of being matched with an irrational player gets small, so long as this Nash bargaining demand is in the set of behavioural types, (see *Abreu et al.* 2015).

Subgame		Obs		Delay	Upper Bounds[§]	
α_L	α_H	Freq	%	T_{max}	z_L	z_H
C0						
14	20	4	1.4	6.4	77.4	85.2
14	22	5	1.7	6.1	85.1	92.2
15	16.5	4	1.4	0.3	97.0	97.3
15	17	14	4.8	31.0	9.8	13.3
15	18	4	1.4	40.1	13.5	20.1
15	19	7	2.4	84.5	4.2	9.8
15	20	17	5.9	92.1	6.3	15.9
15	21	3	1.0	7.0	84.0	90.1
15	22	6	2.1	46.8	36.7	58.6
15	23	3	1.0	466.2	0.0	1.7
15	26	4	1.4	71.6	37.7	77.1
16	18	5	1.7	26.7	39.3	44.9
16	20	4	1.4	10.7	77.9	83.6
17	20	5	1.7	87.7	19.6	28.6
17	22	4	1.4	14.4	81.2	88.0
18	20	3	1.0	40.4	54.6	60.4
20	22	5	1.7	170.6	24.1	32.1
20	25	3	1.0	323.7	11.6	34.0
R0						
15	18	38	10.9	88.2	1.2	2.9
15	20	139	39.7	156.3	0.9	4.4
18	18	1	0.3	30.5	54.3	54.3
18	20	27	7.7	120.4	16.4	22.2
20	20	40	11.4	190.1	14.9	14.9

Table 3: Further Analysis of Delay in Treatments without Induced Types

For *C0*, only subgames with at least 3 observations reported.

[§] Upper bounds on z (in %) implied by T_{max}.

4 Conclusion

This erratum reports an error in how the units of delay were treated in *Embrey et al.* (2015). The correction, however, does not affect the main finding that concerned first-stage announcement behavior, nor comparative-static results on second-stage concession behaviour. This change in units implies that some behavior that used to appear inconsistent with the *Abreu and Gul* (2000) model, cannot be explicitly interpreted as such anymore. Overall, the corrected, and more reasonable, second-stage delay statistics add further weight to the assertion in the conclusion of the original paper that, despite the complexity of the equilibrium predictions, the two-sided reputation model performs well in predicting bargaining behavior, especially when the possibilities for behavioral play are transparent (i.e. in explicit subgames).

References

Abreu, D., and F. Gul, Bargaining and reputation, *Econometrica*, *68*, 85–117, 2000.

Abreu, D., D. Pearce, and E. Stacchetti, One-sided uncertainty and delay in reputational bargaining, *Theoretical Economics*, *10*(3), 719–773, 2015.

Embrey, M., G. R. Fréchette, and S. F. Lehrer, Bargaining and reputation: An experiment on bargaining in the presence of behavioural types, *The Review of Economic Studies*, *82*(2), 608–631, 2015.