Economics Bulletin

Volume 32, Issue 1

Comment on social preferences, beliefs, and the dynamics of free riding in public good experiments

Alexander Smith
Worcester Polytechnic Institute (WPI)

Abstract

Fischbacher and Gaechter (AER, 2010) find that contributions decline in repeatedly played public good games because people are imperfect conditional cooperators who match others' contributions only partly. We re-examine the data using dynamic panel data methods and find that contributions also decline because people only partially match their own contributions from previous periods. We discuss possible interpretations.

I am grateful to Robert J. Oxoby, Moritz Ritter and John Spraggon for very helpful comments and suggestions.
Citation: Alexander Smith, (2012) "Comment on social preferences, beliefs, and the dynamics of free riding in public good experiments", *Economics Bulletin*, Vol. 32 No. 1 pp. 923-931.
Contact: Alexander Smith - adksmith@wpi.edu.
Submitted: January 23, 2012. **Published:** March 21, 2012.

Economics Bulletin, 2012, Vol. 32 No. 1 pp. 923-931

1. Introduction

Urs Fischbacher and Simon Gaechter (2010) (henceforth FG) use a novel experimental design for determining why contributions decline in repeatedly played public good games. They measure preferences for cooperation with a specially designed public good game played using the strategy method and observe the same subjects in a series of ten one-shot games. In the series of one-shot games, beliefs about the contributions of others are a weighted average of what others contributed in the previous period and subjects' beliefs from the previous period. Contributions are a function of beliefs and preferences for cooperation, as measured using the strategy method. A simulation based on the estimated models of beliefs and contributions matches the data much better than models using counterfactual assumptions such as naive belief formation and perfect conditional cooperation. Preferences for imperfect conditional cooperation, in particular, are especially important for matching the data, leading FG to conclude that "contributions decline because, on average, people are imperfect conditional cooperators who match others' contributions only partly."

FG's results provide an explanation for decreasing contributions in public good games. However, by failing to account for the autoregressive nature of contributions, they falsely conclude that imperfect conditional cooperation is *the* explanation for decreasing contributions. Using the FG data (available at http://www.aeaweb.org/articles.php?doi=10.1257/aer.100.1.541), we model contributions using dynamic panel data methods. Our simulations indicate that an autoregressive model of contributions excluding preferences matches the data at least as well as FG's model based on preferences. Thus, we find that contributions also decline for reasons unrelated to imperfect conditional cooperation. Specifically, they decline because people only partially match their *own* contributions from previous periods. We discuss the potential roles of confusion and strategy.

2. Experimental Design

FG's decision-making environment is a standard linear public good game. Subjects are endowed with 20 tokens and randomly assigned to groups of four. The payoff function is:

$$\pi_i = 20 - c_i + 0.4 \sum_{j=1}^{4} c_j \tag{1}$$

where c_i is the contribution of subject i and subject i's group members are indexed by j.

There are two experiments. The first experiment (the P-experiment) measures subjects' preferences for cooperation, as a function of the cooperation of others, using the strategy method. In the second experiment (the C-experiment), subjects make contribution decisions and report their beliefs about the contributions of others in a series of ten one-shot games with random matching prior to the start of each period.[1] All subjects participate in both

[1] The belief elicitation is incentive compatible, but not strictly. Subjects receive three tokens for beliefs that are exactly correct, two tokens for deviating by one token, one token for deviating by two tokens and no tokens for deviating by more than three tokens. There is a literature about belief elicitation. Croson (2000), and Gaechter and Renner (2010) examine the effects of belief elicitation on contributions in repeated public good games. Rutstrom and Wilcox (2009) consider belief elicitation in a matching pennies game. However, addressing any weaknesses of the mechanism used by FG is beyond the scope of this comment.

experiments. In half the sessions, the P-experiment follows the C-experiment. In the other sessions, the order is reversed to control for any possible order effects, of which none are found.

3. Analysis

To identify the correlates of beliefs, FG use OLS for regressing beliefs on the average contributions of others in the previous period and beliefs from the previous period (the period is omitted because it was found to be insignificant in a prior model). The results are presented in Table 1. Unfortunately, the inclusion of lagged beliefs introduces autocorrelation that is not addressed by the "robust" least squares estimation of FG (Pearson's and Spearman's ρ between the residuals and their first lags: -0.10 and -0.11; $p < 0.01$ in both cases).

Table 1: Correlates of Beliefs

Model	(FG)	(1)	(2)
Others' contributions (t-1)	0.41***	0.46***	0.46***
	(0.02)	(0.02)	(0.03)
Belief (t-1)	0.57***	0.53***	0.45***
	(0.04)	(0.04)	(0.05)
Belief (t-2)	-	-	0.08***
			(0.03)
Constant	0.12	0.16	0.01
	(0.15)	(0.24)	(0.26)
Periods Used	2-10	2-10	3-10
Observations	1,260	1,120	980
R^2	0.64	-	-
Method	OLS	AB	AB

Notes: For the OLS regression, standard errors robust and clustered by session are reported in parentheses. For the AB regressions, the standard errors are robust.
***, **, *: Significance at 1%, 5% and 10%.

Following Wooldridge (2010, pages 371-374), we estimate the model using the method of Arellano and Bond (1991).[2] Model 1 includes the same variables as FG's model of beliefs and model 2 includes an additional lag of beliefs, which is significant at 1%. Calculating standard measures of model selection such as the Akaike and Bayesian Information Criteria (AIC and BIC) is not straightforward within the Arellano-Bond framework. However, an important requirement for the consistency of the Arellano-Bond estimator is that there is no second order correlation among residuals. For model 1, the Arellano-Bond test of the null

[2] FG use the method of Arellano and Bond for estimating a model including three lags of the average contributions of others (see Footnote 14 on page 548). They find significant evidence that the estimates are biased using the Arellano-Bond test ($p = 0.03$). However, the regression suffers from an omitted variable problem; adding one additional lag of beliefs eliminates the evidence of bias ($p = 0.37$).

hypothesis that there is no second order correlation is weakly significant ($p = 0.07$). For model 2, the test is not significant ($p = 0.17$). We favor model 2 on the basis of these tests. However, the results are similar for all three models. Neither an additional lag of the average contributions of others, nor an additional lag of beliefs is significant when added to model 2.[3]

We estimate FG's preferred model of contributions and report the results in Table 2. Contributions are regressed on predicted contributions (calculated using preferences determined in the P-experiment and current beliefs) and beliefs. Note that we use the data from periods 3-10 because in our simulations, we measure how well the models match the data from periods 3-10. FG estimate the model using the data from all ten periods, but the results are very similar to ours (see FG page 549).

Table 2: Correlates of Contributions

Model	(FG)	(1)	(2)	(3)
Predicted contribution	0.21**	-	-	0.22
	(0.07)			(0.14)
Belief	0.62***	0.60***	0.46***	0.41***
	(0.08)	(0.08)	(0.09)	(0.10)
Contribution (t-1)	-	0.23***	0.21***	0.17***
		(0.07)	(0.06)	(0.06)
Contribution (t-2)	-	-	0.10***	0.08**
			(0.04)	(0.04)
Constant	-0.21	-0.34	-0.09	-0.35
	(0.34)	(0.36)	(0.36)	(0.36)
Periods Used	3-10	2-10	3-10	3-10
Observations	1,120	1,120	980	980
R^2	0.26	-	-	-
Method	OLS	AB	AB	AB

Notes: For the OLS regression, standard errors robust and clustered by session are reported in parentheses. For the AB regressions, the standard errors are robust. ***, **, *: Significance at 1%, 5% and 10%.

Next, we examine whether lagged contributions should be included, an approach that is consistent with the work of a number of previous authors (Neugebauer et al., 2009; Ashley, Ball and Eckel, 2010; Ferraro and Vossler, 2010). The lagged contributions capture any tendency for subjects to adjust their contributions gradually over time.

Models 1 and 2 include one and two lags of contributions, but not predicted contributions (which embed preferences for conditional cooperation). In both models, a concern is that the estimated effect of beliefs is biased because beliefs are elicited and thus measured with error. Specifically, if elicited beliefs are a noisy measure of true beliefs, then the measurement error will attenuate estimates of the effect of beliefs toward zero. Within the Arellano-Bond

[3] Model 1 is based on fewer observations than the FG model even though both use the data from periods 2-10 because the Arellano-Bond method estimates models in first differences.

framework, we therefore specify beliefs as endogenous, and they are instrumented using lagged variables. There is no significant evidence of second order correlation among the residuals in models 1 or 2 ($p = 0.15$ and $p = 0.75$). Also, none of an additional of a lag of beliefs, the square of beliefs (accounting for nonlinearities), nor an additional lag of contributions are significant when added to model 2.

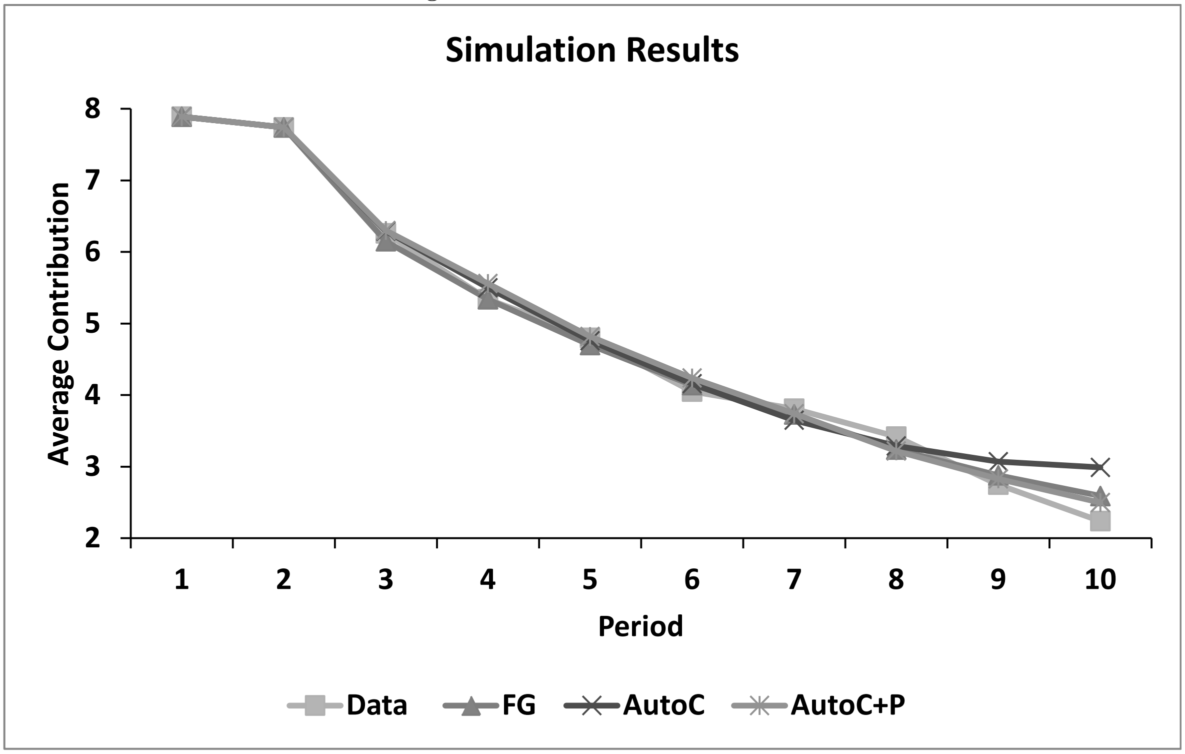
Model 3 adds predicted contributions. Here, the effects of both predicted contributions and beliefs are potentially biased by measurement error. As a result, both variables are specified as endogenous within the Arellano-Bond framework. Once again, there is no evidence of second order correlation among residuals ($p = 0.86$). Notice that the effect of predicted contributions, though similar in magnitude to in the FG model, is not statistically significant. None of a lag of predicted contributions, a lag of beliefs, the square of beliefs, nor an additional lag of contributions are significant when added to model 3. We avoid comparing the models at this stage because our interest is in judging the models by how well simulations based on their parameters match the data.

FG use simulations for testing the roles of beliefs and preferences in explaining decreasing contributions. Comparing a simulation based on the their estimated models of beliefs and contributions to models based on counterfactual assumptions including naive belief formation and perfect conditional cooperation, FG present evidence suggesting that preferences (specifically, preferences for imperfect conditional cooperation) are essential for explaining decreasing contributions. The simulations proceed in two stages. First, the simulated players form beliefs about the average contributions of others. Next, they choose contributions based partially on their beliefs.

We assume that beliefs are formed in a manner consistent with model 2 in Table 1. However, all of the comparative results are the same if we alternatively use either model 1 or the OLS estimates from the FG model. We make three different assumptions regarding the determination of contributions. The first is that they are chosen in a manner consistent with the FG model of contributions presented in Table 2. That is, we assume that predicted contributions (preferences) matter, but that contributions are not autoregressive. We call this the FG simulation model. The second scenario is that contributions are autoregressive, but that predicted contributions do not matter, as in model 2 from Table 2. This is called the AutoC model. Finally, the AutoC+P model includes predicted contributions in an autoregressive framework, consistent with model 3 in Table 2. The simulations use the exact matching structures that occurred in the real experimental sessions. Actual beliefs and contributions are used for periods 1 and 2 to accommodate the lags.

The results are shown in Figure 1. For comparison, the actual average contributions from each period are also illustrated. Visually, all three models match the data closely. The only obvious differences are that the models including predicted contributions (the FG and AutoC+P models) do better in the final two periods, suggesting that preferences are especially important in static environments. However, statistically, we compare the models based on the correlations between the simulated contributions and the data (using the observations from periods 3-10). The Pearson's and Spearman's correlations between the data and the simulated contributions from the FG model are 0.30 and 0.35 ($p < 0.01$ for all correlations). For the AutoC model, the correlations are 0.42 and 0.35, indicating that the autoregressive model excluding preferences performs at least as well as the FG model based on preferences. This is especially the case when it comes to matching extreme observations, which have a

Figure 1: Simulation Results



larger impact on the Pearson's correlation coefficient than on the Spearman's correlation coefficient.

The correlations for the AutoC+P model are 0.40 and 0.43. Both compare favorably to the FG model, but where as the AutoC+P model has a higher Spearman's correlation coefficient than the AutoC model, the AutoC model has the higher Pearson's correlation coefficient. Admittedly, the extent to which these differences are economically important is in the eye of the beholder. Nevertheless, one way of thinking about the results is to recognize that compared to the unrestricted model (AutoC+P), neither of the restricted models (eliminating either the role of preferences or the autoregressive nature of contributions) produce large losses in terms of ability to match the data.

4. Conclusion

Differently from FG, we find that imperfect conditional cooperation is not necessary for explaining decreasing contributions in repeated public good games. Our results indicate that an equally valid explanation is that contributions decline because people do not fully match their own contributions from previous periods. Two potential reasons are confusion and strategy. By confusion, we refer to the tendency for subjects to have an inaccurate understanding of the game's incentives. Repeating the game may reduce confusion, leading to lower contributions. However, though Andreoni (1995), and Houser and Kurzban (2002) report that confusion is a significant source of cooperation, Ferraro and Vossler (2010) find that confusion does not dissipate with repetition (the opportunity to learn). Thus, diminishing confusion cannot be used for explaining decreasing contributions.

Economics Bulletin, 2012, Vol. 32 No. 1 pp. 923-931

Strategy is often overlooked as an explanation for behavior in games with random rematching, but is relevant in the following way. If subjects contribute more in earlier periods trying to promote cooperation among the subjects in their sessions, then contributions will decrease as the final period approaches and the incentives for promoting cooperation decline. Neugebauer et al. (2009) present evidence that strategic incentives do not explain decreasing contributions, but their decision environment is different from FG because they examine the role of strategy by manipulating the feedback (information) given to subjects. In addition, there is an extensive literature indicating that strategic considerations are an important determinant of behavior in repeated public good games (Andreoni, 1988; see Zelmer (2002) and Andreoni and Croson (2008) for meta-analyses). As a result, it appears that strategy could be the reason that subjects fail to match their contributions from previous periods. Of course, demonstrating this conclusively requires another experiment.

References

Andreoni, James (1988) "Why Free Ride?: Strategies and Learning in Public Goods Experiments" *Journal of Public Economics* **37**, 291-304.

Andreoni, James (1995) "Cooperation in Public-Goods Experiments: Kindness or Confusion?" *American Economic Review* **85**, 891-904.

Andreoni, James and Rachel Croson (2008) "Partners versus Strangers: Random Rematching in Public Goods Experiments" in *Handbook of Experimental Economics Results* by Charles R. Plott and Vernon L. Smith, Eds., North-Holland: Amsterdam.

Arellano, Manuel and Stephen Bond (1991) "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations" *Review of Economic Studies* **58**, 277-97.

Ashley, Richard, Sheryl Ball and Catherine Eckel (2010) "Motives for Giving: A Reanalysis of Two Classic Public Goods Experiments" *Southern Economic Journal* **77**, 15-26.

Croson, Rachel (2000) "Thinking Like a Game Theorist: Factors Affecting the Frequency of Equilibrium Play" *Journal of Economic Behavior and Organization* **41**, 299-314.

Ferraro, Paul J. and Christian A. Vossler (2010) "The Source and Significance of Confusion in Public Goods Experiments" *B.E. Journal of Economic Analysis & Policy* **10** - Contributions, Article 53.

Fischbacher, Urs and Simon Gaechter (2010) "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments" *American Economic Review* **100**, 541-56.

Gaechter, Simon and Elke Renner (2010) "The Effects of (Incentivized) Belief Elicitation in Public Goods Experiments" *Experimental Economics* **13**, 364-77.

Houser, Daniel and Robert Kurzban (2002) "Revisiting Kindness and Confusion in Public Goods Experiments" *American Economic Review* **92**, 1062-69.

Neugebauer, Tibor, Javier Perote, Ulrich Schmidt and Malte Loos (2009) "Selfish-Biased Conditional Cooperation: On the Decline of Contributions in Repeated Public Goods Experiments" *Journal of Economic Psychology* **30**, 52-60.

Rutstrom, E. Elisabet and Nathaniel T. Wilcox (2009) "Stated Beliefs versus Inferred Beliefs: A Methodological Inquiry and Experimental Test" *Games and Economic Behavior* **67**, 616-32.

Wooldridge, Jeffrey M. (2010) *Econometric Analysis of Cross Section and Panel Data 2nd Ed.*, MIT Press: Cambridge.

Zelmer, Jennifer (2002) "Linear Public Goods Experiments: A Meta-Analysis" *Experimental Economics* **6**, 299-310.

Appendix

The Stata 11 code for our results (using the FG data):

*Table 1

```
reg belief l.otherscontrib l.belief, cl(idsess)
predict bhat
gen ehat = belief - bhat
pwcorr ehat l.ehat, sig
spearman ehat l.ehat
xtabond belief, lags(1) pre(l.otherscontrib) r
estat abond
xtabond belief, lags(2) pre(l.otherscontrib) r
estat abond
```

*Footnote 2

```
xtabond belief l(1/3).otherscontrib, lags(1) r
estat abond
xtabond belief l(1/3).otherscontrib, lags(2) r
estat abond
```

*Table 2

```
reg contribution predictedcontribution belief if period>2, cl(idsess)
xtabond contribution, lags(1) end(belief) r
estat abond
xtabond contribution, lags(2) end(belief) r
estat abond
xtabond contribution, lags(2) end(predictedcontribution belief) r
estat abond
```

*We conduct the simulations in the exact same manner as FG.