

TRUST AND TRANSFORMATION IN REVERSE-CROWDING OUT:
SOME EVIDENCE FROM LABORATORY EXPERIMENTS

R. Mark Isaac and Douglas A. Norton

Florida State University

Prepared for the APEE Annual Meetings

Las Vegas, Nevada, 2008

\

We thank Austin Boyle and Glenn Dutcher for programming the experiments reported here. All sessions were run using Z-Tree (Fischbacher [2007]). We are grateful to the National Science Foundation (grant #0451970) and to the John and Hallie Quinn Eminent Scholar Chair for support, and to the participants in the 2008 Social Dilemmas Conference for many helpful comments.

I. INTRODUCTION: LOOKING WHERE THE LIGHT IS BETTER.

This paper is about what we have taken to calling “reverse crowding out,” or the possibility that a renewed commitment on the part of a citizenry to private institutions of charity and compassion could “crowd out”, that is “replace” the decades-old predominance of government, especially the federal government, as the primary engine of social welfare. It is quite likely that this paper will seem strange to many, if not most, of its audience for three reasons.

First, it has come to be taken as received wisdom that this size and scope of the government in these areas is a ratchet process in which the government footprint can only enlarge, and never diminish. Even to raise the possibility of “reverse crowding out” may appear to be an oxymoron. As some of the students in our class on the economics of compassion remarked to us when we discussed early private, religious institutions for the care of unwed mothers, it is simply unbelievable to them that such institutions could ever again play a similar role.

Secondly, our research comes from a series of laboratory economics experiments. Although the time has apparently passed in which the use of experimental economics needs to be justified *de novo* in every research project, it may strike some as odd that such a broad social force as reverse crowding out could be studied in the laboratory. We respond by noting that our goals are limited, and our results tentative. But, if our presentation of these experiments motivates others using different economic techniques to join the inquiry, that would be a good thing, not a bad thing.

Finally, although our course of experiments are still in progress, our early conclusions take us in directions that may not be anticipated by many neo-classical economists, such as the role of trust, the importance of endogenously formed common expectations in dealing with public goods, and the possibility that individual references are fluid or even transformed rather than a part of the *ceteris paribus* conditions. (Of course, there's nothing in neoclassical economics that rules out changes in preferences, it's just not the usual sandbox in which we tend to dig. As the old saying goes, we seem to be looking for our lost keys over here, because the light is better).

II. BACKGROUND: THE NURSE RATCHET EFFECT

It is beyond controversy to note that the United States federal government's historical role in funding what have been called programs of charity or compassion exploded in the early part of the 20th century in association with the arrival of the New Deal. What is sometimes overlooked is that the seeds of the political demand for federalization evolved long before. During the period of the Third Great Awakening, a Christian religious revival that swept both the United States and Great Britain in the 1860s – 1920s era, the religious mobilization of a generation yielded numerous, perhaps seemingly paradoxical consequences. On the one hand, the Third Great Awakening fostered the creation of numerous private institutions of charity, such as the Salvation Army, originally in Great Britain. In an undergraduate class that we teach on the economics of compassion, we highlight the role of the Free Methodist Church in founding homes for unwed mothers, such as the Home of Redeeming Love on the outskirts of Oklahoma City. A website for a successor institution notes that the women of

the church not only ran the home, they actually farmed the land on which the home was located in order to raise funds.

In terms of political activity, however, the Third Great Awakening took a different turn, away from individual action towards governmental, specifically federal governmental, activity. Of course, one of the most famous examples is the 18th Amendment to the U.S. Constitution, imposing by government decree the social experiment of Prohibition on the sale of alcoholic beverages. But the Third Great Awakening's eyes on the power of federal programs did not stop with Demon Rum. In terms of economic policy a prominent face of religious revival in this period was the so-called "Social Gospel" in which Christian revivalism merged with calls for government regulation, taxation, and expenditures championed by the Populist and Progressive movements. Even the American Economic Association, founded in 1886, was described by one of its founders, Richard Ely, as a vehicle for achieving "practical Christianity."¹ Walter Rauschenbusch, a leader in the theology of the Social Gospel, wrote the following:

"In the last resort the only hope is in the moral forces which can be summoned to the rescue. If there are statesmen, prophets, and apostles who set truth and justice above selfish advancement; if their call finds a response in the great body of the people; if a new tide of religious faith and moral enthusiasm creates new standards of duty and a new capacity for self-sacrifice; if the strong learn to direct their love of power to the uplifting of the people and see the highest self-assertion in self-sacrifice-then the entrenchments of

¹ Ely quoted in *Florence Kelly and the Nation's Work: the Rise of Women's Political Culture*, by Kathryn Sklar (Yale University Press, 1995) as available on Google Books.

vested wrong will melt away; the stifled energy of the people will leap forward; the atrophied members of the social body will be filled with a fresh flow of blood; and a regenerate nation will look with the eyes of youth across the fields of the future.”²

There are then two great ironies of the New Deal and its relationship to the Third Great Awakening, which was itself rapidly receding by the 1930s. First, it was in the Roosevelt Administration that one great Christian social experiment in federal power (Prohibition) was ended, even as the New Deal brought fruition to the economic dreams of the disciples of the Social Gospel. The second irony is that the New Deal appears to have done great damage to the fabric of the private religious social safety net that has been fostered during the Revival. Specifically, Gruber and Hungerman [2007] estimate that New Deal social spending crowded out around thirty percent of existing religious charitable spending. It is not too much a stretch to say that this damage to the private religious social safety net was an important part of the trend of secularization and governmental absorption of a range of social concerns from the sick (the fading of explicitly religious hospitals) to orphans (the receding of religious orphanages and the growth of the governmental foster care system) to education (the secularization of numerous historically religious colleges and universities).

This scenario admits of many questions; the one that concerns us here is the following: is the evolution of the social safety net into an almost exclusively governmental enterprise unidirectional and irreversible? Is it possible, and under what conditions is it possible, that private institutions could achieve a “reverse crowd out” and

² Walter Rauschenbusch, *The Social Gospel* (New York: MacMillan, 1908), excerpted on “Modern History Sourcebook,” <http://www.fordham.edu/halsall/mod/rausch-socialgospel.html>

begin to recover from the federal government some of the responsibility of compassionate care?

Our term “reverse crowding out” is a reference to the neoclassical concept of “crowding out” by government of the private provision of public goods. Standard models of equilibria in individual decision-making predict that increases in taxed provision of public goods crowd out an equivalent amount of private provision.

The theoretical channel by which crowding out may occur has been developed for many years (Bergstrom, Blume, and Varian [1986]). There seems little doubt that, as an empirical matter, crowding out exists at some level, but there is widely different empirical evidence as to its scale. Field studies include Payne [1998] Hungerman [2005] and Hungerman [2007]. Laboratory economics experiments (beginning with Andreoni [1993] and including Chan, et al. [1998] and Eckel, et al. [2004]) also have detected intermediate levels of crowding out.

Andreoni’s paper, in particular, was motivated by his theory of “warm glow” altruism in which individuals receive utility from the actual act of generosity. The intermediate levels of crowding out found in Andreoni’s experiments are consistent with a warm-glow model, although that need not be the only explanation.

The empirical evidence on crowding out has taken on an important new dimension with the Gruber and Hungerman study. Private, religious contributions for social services remained steady through the worst early years of the Great Depression, but began to tumble in the period of 1933-1934, eventually settling in at about 30 percent less than their earlier level. Gruber and Hungerman estimate that, even though the rate of crowding out was not large, the massive increase in federal social service expenditures

can account for nearly all of the reduction in private religious transfers. If accurate, the Gruber and Hungerman results indicate that crowding out occurred with significant social consequences.

We investigate the possibility of reverse crowding out through a series of laboratory experiments.

III. EXPERIMENTAL DESIGN AND HYPOTHESES

As noted, Andreoni's initial experimental design was constructed to focus on his investigation of warm glow and the possibility of incomplete crowding out. He used a relatively small group (three subjects) with a concurrent "payoff chart" technology incorporating an interior optimum and equilibrium. Indeed, the term "crowding out" in Andreoni's sense is an exercise in comparing equilibrium levels of contribution. In fact, it requires that there be a non-cooperative equilibrium in which each individual has a positive contribution to the public good. Furthermore, in the standard environment, there is no difference in the productivity of private and government provision. Government provision is observationally equivalent to a required "minimum contribution" in the voluntary contributions mechanism (indeed this is how it was explained to the subjects). Many other researchers, for purposes of comparability, have used this same structure.

Because we are asking different questions, we begin down a new path. First, we want a larger group dynamic, so our experiments consist of seven persons. Secondly, we are very much interested in a world of extensive free riding in equilibrium, and thus we drop the interior payoff structure and return to the equally traditional linear public goods model.

Thirdly, we model government provision as being less efficient than private provision. Thus, in our experiments, provision through government taxation is made to be moderately inefficient (twenty percent of tax receipts are lost). This could represent actual physical and organizational inefficiency in the government provision of social services, or population preferences that value government provision less. Either story fits our design.

Fourthly, the role of government and the levels of taxation are exogenous in most of the existing literature, while we are interested in endogenous mechanism choice.³

Finally, we are interested in how endogenous institutions respond to changes in individual preferences. Presumably, religious institutions may be most likely to reverse crowd out government provision if individuals in the community/society become more receptive to personal action.

Putting all of these features together, our research question is different from the standard crowding-out paradigm. In the traditional literature, the non-governmental state yields an interior equilibrium of private provision. As governments impose these minimum contributions, total provision of the public good (from both government and private sources) can be benchmarked against the private equilibrium.

In our research, the voluntary contributions equilibrium is defined by complete free riding. The empirical reality, as repeatedly demonstrated in laboratory experiments, is that of positive, but non-optimal, disequilibrium levels of public goods provision. In such an environment, a standard research question has been to examine those mechanism features that can maximize the efficiency of the system, (a classic example being face to

³ One partial exception is Sutter and Weck-Hannemann [2004] in which subjects are allowed to vote yes/no by majority vote on a exogenously suggested tax scheme.

face non-binding communication [Isaac and Walker, 1988]). We look at a specific mechanism sequence: purely voluntary contributions (VCM), exogenously determined forced government provision (taxation) combined with VCM, and then endogenously determined taxation combined with VCM.⁴

To implement this, we begin by using an [A, B, C, D] experimental design as follows. For the first seven periods (Stage 1), seven subjects participate in a standard VCM design. Each subject has the task of investing 100 tokens in either an individual or a group exchange. Investment in the individual exchange returns a one token payment with certainty, while the group exchange operates as a VCM process with $MPCR = .5$.

In the second round of seven periods (Stage 2), we come the closest to matching the Andreoni design. We impose a tax of 30 tokens upon every participant to fund a minimum level of the public good. The tax efficiency of 80 percent means that 24 tokens are actually transferred to the group exchange for each person; six tokens disappear.⁵ Voluntary contributions are allowed beyond the minimum through the VCM process.

In the third treatment (Stage 3), the individuals are now allowed to choose the tax level through a voting mechanism, again with VCM operating on their residual tokens. This experiment could be run with several different voting mechanisms, but in this experiment subjects will cast votes using the median voter process. Each participant will submit a number from 0 – 100 (the per-period endowment of tokens) and the experimenter chooses the median tax.⁶

⁴ Another stream of research into crowding-out has been to look at individual decisions in dictator game experiments (see Eckel, et al. [2004]).

⁵ The 20 percent inefficiency figure is not intended to be an advertisement for an actual field scale. It is one that is neither so small as to be trivial nor so large as to seem to be overwhelming to the subjects. However, it is interesting to note that prior to the 1986 tax returns, ..., estimated at 49 percent.

⁶ See Holcombe and Kenney [2007] for information on actual field implementations of this voting mechanism in Florida school finance elections.

In our fourth treatment (Stage 4) we introduce our “agents of grace”. Two of the seven participants are randomly chosen in advance to have their MPCRs changed to 1.43 (accomplished by lowering their return on their individual exchange).⁷ The fact that a change in incentives is possible, but not the exact form nor the distribution, is publicly announced to the individuals.

We note that our baseline experiments contain only this sequencing because it is of specific interest in the historical evolution of government involvement in charitable activity. We will return later to how we addressed sequencing.

The underlying contributions game by itself admits of the standard theoretical properties. In the stage game, full investment in the individual exchange is the dominant strategy, while the group income maximizing outcome is for each participant to invest all tokens in the group exchange. Nothing changes when an exogenous tax is enacted. However, with the median voter tax mechanism, there are a multiplicity of equilibria.

Consider that the choice of a subject can be represented as the pair $\langle t_i, p_i | T \rangle$, meaning that each subject chooses a vote for a tax, t_i , and, conditional on the tax rate actually chosen, T , a proportion, p_i , of his untaxed funds to contribute to the group good. In the stage game, whatever tax is chosen, it will remain a dominant strategy in the subgame to set p_i equal to zero. The only remaining question is what vector of tax choices support an equilibrium? The following provide a partial characterization of the outcomes in Stage 3:⁸

⁷ In the parlance of some public goods research, this would be known as raising only the “inside” MPCR.

⁸ The Stage 4 outcomes need only be adjusted by the change in the dominant strategies in the stage game.

Let V be any vector $(\langle t_1, 0 \rangle, \dots, \langle t_7, 0 \rangle)$ with the properties that a) Each person contributes no remaining tokens to the group exchange regardless of the value of T , and b) that no one person could have altered the median vote. Then V is an equilibrium.

The following is one category of outcomes that are not equilibria: For any vector $V = (\langle t_1, p_1(T) \rangle, \dots, \langle t_7, p_7(T) \rangle)$, if V has the property that the t_i s are strongly distinct, then V is not an equilibrium.

The first statement has the following corollary. The set of equilibrium vectors so characterized includes both the full tax outcome through the vector $(\langle 100, 0 \rangle, \dots, \langle 100, 0 \rangle)$ and the full free riding outcome through the vector $(\langle 0, 0 \rangle, \dots, \langle 0, 0 \rangle)$. The former Pareto dominates the latter. The idea of finding the maximum tax seems very focal (even though the 20 percent inefficiency means that it is not the group optimum), and the resulting stability might mean that the group would do little exploring over other alternatives. We call this scenario “full government lock-in.” The first statement also suggests the manner in which this mechanism could become “stuck” at intermediate levels of taxation, perhaps at our initial level of 30 tokens per individual.

The second (disequilibrium) statement is important because it suggests a way in which the full tax outcome could be achieved. If intermediate levels of taxation occur with distinct votes, then one person can change the level of the tax. By assumption $T < 100$, and a single unit increase in the tax will make that pivotal person better off. This might provide for a ratcheting up of the initial, exogenously imposed tax.

What, then, is the basis of what we call “reverse crowd out?” In Stage 3, the change in rules means that individuals have the capability of eliminating the tax. In Stage 4, the two agents of grace will have a dominant strategy to contribute all their tokens to the group exchange, and this gives the possibility of reverse crowd an extra impetus. Logically, they, like everyone else, would prefer that whatever contributions are made be made through the VCM process because it is more efficient. The most complete manifestation of reverse crowd out would occur if tax rates fell and total contributions to the group exchange increased. An outcome such as this, with low or zero tax rates and significant contributions to the group good, would clearly fail as an equilibrium in the usual sense, but previous laboratory experiments in public goods provision have shown that non-equilibrium outcomes are the rule, rather than the exception.

IV. PRELIMINARY RESULTS

The following are three preliminary conclusions based upon the aggregate results from our initial four baseline sessions (noted as Sessions 1-4) and three control sessions (noted as Sessions 5-8).

Result: In our initial four baseline experiments we observed no session that achieved the full tax lock-in (one session came close). In our first three control experiments, two sessions converged to the full tax lock-in. Figures 1 - 7 display the total tokens invested in the group exchange, together with the tax component of the group exchange. (As an aside, none of the groups stayed locked-in at the thirty percent exogenously imposed tax level).

Result: We observed only one baseline or control session that could even partially be interpreted as reverse crowding out. In Session 2, a baseline session, (Figure 2) taxes dropped to zero in Stage 3 and remained there for the remainder of the experiment. Furthermore, total contributions increase during Stage 4. They settle in the neighborhood of 300 tokens, far below the social optimum but above the Stage 4 dominant strategy prediction of 200 tokens.

Result: Overall, the level of the provision of the public good in the baseline experiments appears to be different than typically expected from a VCM process. In three of the four baseline experiments there is a reversal of the usual tendency for a decay in contributions in the final periods (specifically, in Stage 4).

The problem with the previous result is that the data for the baseline experiments include the contributions of the agents of grace. But we have adjusted the agents of grace incentives toward full contribution to the group good in Stage 4, so perhaps it is not surprising that there is no decay in contributions. A more focused question is: “What effect did the agents of grace have on the other five participants?” To address this question, we present in Figures 8 - 11 the contributions to the group exchange with the contributions of the agents of grace removed from all 28 periods. In particular, these figures plot the actual data together with the estimates from a regression with the following functional form:

$$\text{Tokens} = \beta_0 + \beta_1\text{PER} + \beta_2\text{MVTAX} + \beta_3\text{STG2} + \beta_4\text{STG3} + \beta_5\text{STG4} + \beta_6\text{STG2*PER} + \beta_7\text{STG3*PER} + \beta_8\text{STG4*PER}$$

where

PER = Period

MVTAX = median voter tax chosen that period

STGX = a fixed effect for Stage X (Stage 1 excluded)

STGX*PER = an interactive dummy for Stage X and Period.

The full regression results are presented in Table 1. The F statistic displayed in each of Figures 8-11 tests our complete model against a restricted model with only the constant and the period (a standard VCM formulation). Notice that even with the agents of grace excluded, the slope of contributions flattens in Stage 4. In three of the four cases this model is statistically significant. These results indicate that the non-standard behavior of these sessions extends beyond the contributions behavior of the agents of grace themselves.

However, at this point our analysis is still not complete. The agents of grace may still be having an effect in Stage 4 through a number of channels: by influencing others contributions or by influencing others tax levels, or both. Moreover, the effects of the agents of grace may be through their own tax choices as well as through their own contributions. We cannot fully distinguish these effects because we have not yet had a chance to model individual voting and contribution behavior. However, when we conducted an out-of-sample simulation of Stage 4 using Stage 3 parameters, in some of

the sessions the choices of which tax levels to assume make a large difference, and in others they don't.

The second standard of control for examining the influence of the agents of grace is to look not just at previous VCM experiments but also at our control series, in which no subject preferences are changed in Stage 4.⁹ The equivalent regression results are contained in Figures 12- 14. Two of the three control experiments failed to exhibit the “Stage 4 Effect” of a weaker decay in Stage 4, but this observation is problematic because two of the control sessions had reached the full tax level during Stage 3, and effectively locked into that level in Stage 4. Thus while Sessions 7 and 8 are structurally “control” experiments, they are behaviorally not useful for examining the effects of the agents of grace in the baseline experiment.¹⁰ At this point, we are left only with Session 9, which like our baseline experiments did not reach the 100 percent tax level in Stage 3. What is distinct about Session 9 (in stark contrast to our baseline experiments) is that the total decay rate for Stage 4 is negative, and the interactive dummy variable on decay for Stage 4 is negative and statistically lower than the relevant parameter for Stage 3. Although Session 9 is our only useful control experiment at this time, the pattern of Session 9 is one in which there was no indication of a mediating influence of the Stage 4 intervention itself, absent the additional influence of agents of grace.

⁹ For purposes of isolating the effect of the agents of grace, we maintained the other aspects of Stage 4: the announcement of a new stage and the distribution of envelopes.

¹⁰ Sessions 7 and 8 are not useful as controls for several reasons. First, the 100 percent tax level was reached in Stage 3, with stability at that level in Stage 4. Thus there is little useful statistical information in the coefficients of the Stage 4 dummy variables. Also, it could be argued that these two experiments stabilized at the full tax outcome precisely because there were no agents of grace. Perhaps the AOGs would have voted for zero tax, and perhaps changed the outcome. That is indeed a possibility, but the counterfactual (“if there had been agents of grace they might have moved away from the tax”) is difficult to evaluate. We can get an idea of how much “wobble room” there is in the data by looking at the 6th highest tax vote (that is, assume that the AOGs were critical and that they voted for zero tax, but no one else changed). For the last five periods, the taxes would have been (23, 80, 100, 100, 22) in Session 7 and (70, 40, 30, 25, and 30) for Session 8. So, the AOGs could have, at least potentially, made some difference.

IV.1. Results From a Special Subject Pool (Or “Rauschenbusch” Vs. “Redeeming Love”)

The results from the four baseline sessions were somewhat different than we had expected. The pattern was more idiosyncratic, suggesting that group composition and/or path dependence were determining factors. In particular, we did not see convincing evidence of any of three categories of boundary results: the full taxation outcome, the reverse crowd-out, or even a simple replication of the familiar VCM pattern of decay towards the free-riding equilibrium. Even when considering the control experiments without the agents of grace, we found little support for reverse crowding out.

Because of the issues of equilibrium versus non-equilibrium behavior inherent in the basic public goods design, plus the added effects of equilibrium selection derived from the addition of the voting mechanism, we became convinced that issues of the level of group expectations might be an important additional factor. Because these expectations could be different when home-grown among naturally-occurring groups rather than groups created randomly in the laboratory, we decided to investigate our design with subjects drawn from a naturally occurring intentional community with incentives generating strong common values and expectations.

To do this, we recruited enough subjects for two sessions from the FSU Wesley Foundation campus ministry community. The experiments were conducted as already noted, but there was the following exception. We announced (in a public announcement available in the Appendix) that at the end of the experiment we would match each person's earnings with a contribution to the New Life Children's Home, an orphanage in Guatemala. The FSU Wesley Foundation has been active in supporting the NLCH, and it

is likely that every subject in the experiment either had participated in a mission project at the orphanage, knew someone who had participated in a mission project at the orphanage, or had plans to go on a mission project to the orphanage.

It is important to distinguish what we did in this experiment from what has been done in similar uses of non-standard subject pools in laboratory experiments. Specifically, we emphasize that we did not tie payments to the orphanage to the provision of the experimental group exchange, but rather to the total earnings of the subjects. This is because in these sessions we were *not* interested in homegrown preferences *per se*; rather, we were interested in homegrown expectations.

If we had tied payments to the orphanage to the provision of only the group good, then we would have substituted the subjects' home grown preferences for the orphanage for the induced preference structure of our experimental design. Suppose we had found, for example, that contributions to the public good had increased. This could quite logically have been interpreted as deriving from a *de facto* increase in the "own marginal per capita return" for the public good. But the fact that increases in the MPCR drive behavioral increases in provision of the public good has been known since the work of Isaac and Walker in the 1980s.

Instead, by tying payments to the orphanage to total earnings, at one level the incentive structure of our design is preserved. A subject wanting to provide more of his funds to the orphanage would have to consider the tension between the stage game dominant strategy and the social optimum of full group contributions.¹¹ But our design changed the incentive structure of the game in other ways.

¹¹ In fact, casual comments by the subjects after the experiment indicated that they were aware of this tension.

Consider the choice of an individual in the stage game. If he increases his contribution to the group exchange by one token, the following happens:

His personal token earnings decrease by .5 tokens

His personal contributions to the orphanage decrease by .5 tokens

At this point, incentives from the contributions to the orphanage exactly mirror those of the incentives in cash from the existing game. However, this person should also consider how his additional token affects contributions to the orphanage via other subjects' earnings. Suppose we adopt the natural conjectural variation that this subject believes that the *token allocation of other subjects does not change*. Then, in addition, this subject calculates that contributions to the orphanage via others' earnings increase by 3 tokens.¹²

So, there is a well defined trade-off of a decrease in personal earnings of .5 tokens with an increase in token earnings for the orphanage of 2.5 tokens, assuming that the individual values tokens contributed to the earnings the same regardless of whether they come from him or from the other individuals. As long as this person values the contributions to the orphanage “enough” the stage game looks like it admits a dominant strategy for contributions.¹³ But what could go wrong with this scenario?

One, maybe the individual does not value contributions to the orphanage “enough.”

¹² Note that if these conditions are met, then we are, in effect, witnessing a Vickrey/Groves/Clarke mechanism created through a combination of induced and naturally occurring preferences and expectations.

¹³ Recall that because these are the home grown preferences of the subjects, they need not be additively separable, so “enough” may be a more complex relationship than a ratio of one value to another.

Two, the dominant strategy in the stage game could fail if individuals value contributions that they make much more than contributions that others make. (This is essentially the same distinction as the “Basic Charity” vs. “See and Be Seen” preferences in Isaac, Pevnitskaya, and Salmon [2008]).

Three, there could be intertemporal effects in this set-up. Subjects in this experiment carry over a portfolio of cash earnings and contribution earnings from period to period, unlike the case in the standard VCM. What if these preferences are not separable? What if I think that your preferences for the contributions to the orphanage are subject to some type of crowding-out? That is, if I add one token to the public good, I conjecture that you will decrease your contributions in future periods. In an extreme sense, this might approach a very different conjectural variation: I assume that you keep your *contributions to the orphanage* constant, and not your token allocations.

What we valued in these subjects was not so much their home-grown preferences for contributions to the orphanage but rather their common expectations. Everyone in the room knew that everyone else in the room also had strong reasons to want to maximize the total earnings of the group, and they had strong reasons to believe that everyone else knew this to be true, and so forth. Essentially, we believed that all of the conditions of a dominant strategy in the stage game above were likely to be met, and we believed that this was known with something approaching common expectations.¹⁴

The incentives that we induced were strong, but these kinds of incentives are ones that are likely overlooked as we consider how naturally occurring groups actually make

¹⁴ In fact, the Isaac, Pevnitskaya, and Salmon paper uses the same subject pool and concludes that their choices in charity auctions as not well explained by the “See and Be Seen” model.

decisions about public goods compared to our standard models. When there have been unexpectedly large amounts of cooperation in a standard VCM experiment, research conjectures have centered on either models that essentially retain the concept of the public good as an instrument to individual maximization (reciprocity) or on models that presume that these strangers care about each others' earnings (altruism). What we propose here is a model in which individuals' preferences and expectations are aligned in a way so that everyone recognizes that it is in everyone's interests to provide the public good. While this does not rule out altruism or reciprocity, we believe that it is a distinct social phenomenon.

We conjectured that if we were ever to see something approaching a reverse crowding out scenario, it ought to be here. Nevertheless, the scenario is falsifiable. Perhaps, contrary to our expectations, these individuals did not value contributions to the orphanage "enough." Perhaps they value only their own contributions, but not those of other members of the group. Perhaps they are concerned that other people will adjust their decisions in ways that offset their own generosity.

The results of these two sessions are striking. We observed both of the boundary outcomes that we had failed to see in the first four sessions. In Session 5, the tax system was "crowded out" and replaced by a purely voluntary regime of substantial, stable, although not quite optimal, levels of contribution. In Session 6, on the other hand, the subjects embraced the tax system and ratcheted the tax level up to the full tax regime. (Recall that the assignment of the two groups was made randomly by the computer from the 14 participants). Charts of tokens to the group and of the tax levels are presented in Figures 15 and 16.

Using the same decomposition techniques as in the first four sessions, we modeled the behavior of the five stable-preference participants. In Session 5, the level of public goods provision is best explained by the constant term of 483 tokens. In Session 6, on the other hand, including the dummy variables for the stages significantly improves the predictive power of the model. Setting aside the issue of the size of the confidence intervals, in both sessions the period decay in Stage 4 was approximately zero. (See Figures 15 and 16.)

One homogenous subject pool randomly divided into two groups produced two very different boundary outcomes (albeit two outcomes that each have the property of substantial provision of the public good). We suspect that part of the difference can be explained by the different behavior in Stage 1. In Session 5, the coefficient on the Stage 1 period decay in contributions the group good is only -2.7 tokens per period, only 25 percent of the next-smallest decay of any of the groups. In Session 6, however, the measured decay in Stage 1 is -34.3 tokens, the largest decay of any of the six sessions¹⁵. It should not be surprising, therefore, that Session 6 used the tax as a means of funding the public good. (As in the previous sessions, the coefficient on MVTAX suggests that the tax was used as a commitment device and not as a signal).

How did these two very different stable patterns differ in how much they provided to the orphanage? This information is provided in Figure 19. Both groups did very well, with the full tax group slightly beating out the reverse crowd-out group. The irony of this conclusion is that the two agents of grace in Session 5, the no tax session, diverged, to different extents, from the dominant strategy. One agent of grace, in particular, played the new dominant strategy for two periods, and then switched to a choice of voting for a 100

¹⁵ Recall that these numbers are from the five-person subgroup.

percent tax, but choosing zero contributions when the tax was revealed to be zero. If either of these two agents of grace had chosen the dominant strategy, the payment to NLCH would have been virtually identical, if not reversed.

IV. 2. On Trust and Taxes.

What role do taxes play in these groups? One possibility is that the participants are choosing taxes for the standard reasons of risk reduction: enact a binding mechanism to insure an increase in group profits. But, just from a visual examination of the total results from the baseline and control experiments in Figures 1 – 7 and 15-16, it seems possible that there may be a correlation between tax levels and contributions indicating that taxes serve as a coordination device, a signal of residual cooperation, or perhaps even a transformation process over preferences (any of these could be considered a “Social Gospel” type effect in which individuals are argued to be more likely to be cooperative in a high-tax environment).¹⁶ If the later effect is true, then we should see that the cooperativeness of individuals over their residual tokens (that is, tokens remaining after any tax) is positively correlated with the tax. To examine this conjecture, we consider two parts of the data. First, in *none* of the regressions discussed above are the coefficients on MVTAX (the tax rates in stages 3 and 4) greater than 1.00, the direct effect of a one unit increase in the tax rate.¹⁷ Furthermore, as a check on this result we have conducted a regression, using the same structure as discussed above, on Sessions 1-4 in which we restricted the dependent variable to “percent of after-tax tokens invested by

¹⁶ Dorothy Day is quoted as saying, “We want a society where it is easier for people to be good.” (<http://sp-usa.org/handbook/ecdemocracy.html>; Socialist Party USA.)

¹⁷ The tax measure was scaled to take account of the efficiency loss. The coefficients in the four experiments were .6975, -.05873, .6626, and .5527. The upper limits of the 95% confidence intervals were 1.09, .77, 1.24, and .92 respectively.

participants other than the agents of grace” in all 28 periods. In none of the four experiments was the coefficient on the tax variable statistically significantly different from zero, leading to the following:

Conjecture: the tax appears to be serving as a commitment device, not a signaling mechanism.

What this result does not tell us is why some groups rely more on the commitment function of the tax than do others. What we have discovered in these experiments is a result that we did not anticipate when we began the research, namely, that there *appears to be a systematic, negative relationship between the level of underlying cooperativeness in the groups and their use of the tax mechanism*. To be more specific, groups that experience a greater degree of free riding in the institution-less periods of Stage 1 tend to converge to a greater reliance on the tax mechanism when given the choice. To see this effect, consider Figure 20, which shows the scatterplot of the average of Stage 3 tax levels against Period 7 contributions to the group exchange (the other line is a quick-and-dirty OLS regression line on the data). Although we don’t know how this relationship will hold as we conduct more control experiments, at this point we offer the following conjecture:

Conjecture: Groups that establish greater degrees of trust regarding mutual cooperation will be more likely to be successful in crowding out government taxation.

V . CONCLUSION: IF TRUST AND TRANSFORMATION MATTER, THEN MAN DOES NOT LIVE BY ECONOMICS ALONE¹⁸

The results reported here are part of an ongoing investigation. We intend to conduct additional control experiments to go with this series, to analyze the individual data, and to look forward to a further research plan that includes possibilities such as varying the level of inefficiency of the tax sector or allowing for Tiebout-style mobility between groups. Nevertheless, there are two stylized facts from the current results that are worth final emphasis. First, economic systems respond to the transformation of individual behavior (which in this case we can connect directly to a change in the induced preferences of the experiment).¹⁹

There is nothing in the observations that preferences can change that is inconsistent with neo-classical microeconomic theory; virtually every principles of microeconomics class will describe a demand shift that comes from a change in consumer tastes. Nevertheless, as a practical matter most economic policy analysis takes stability of consumer preferences as a given and emphasizes instead the alteration of economic incentives that accompany changes in various economic institutions. In a session at the Southern Economic Association meetings in 2007, we were only two of three attendees polled in a session who favored changing individual preferences (rather than creating markets or switching to a government imposed opt-out program) as a more desirable next step in addressing the shortage of donor organs.²⁰ Yet the fascinating story

¹⁸ The latter statement is from an inscription in a book about University of Minnesota football written by economist James P. Quirk.

¹⁹ Doug and I have another paper on the drawing boards considering the extent to which the neo-classical economics concept of a utility function is useful in describing Jesus' teachings of the Kingdom of God.

²⁰ The polling at the session did not consider two other attractive options: organ exchange programs such as the kidney programs that Alvin Roth at Harvard and others have designed (which we view as a market-like solution) and decentralized affinity-group matching (which is as a vehicle by which a new technology

from our experiments is the extent to which altered behavior of two of seven individuals also led to a change in a behavior of the other, “stable preference” participants.²¹

Economists *qua* economists should be just as leery of preference engineering as they should be of Pigouvian-Samuelsonian social engineering. But, society is replete with institutions other than economics institutions. Just as an endogenous evolution of institutions of markets and exchange can be welcomed by economists who would never favor top-down social engineering, so should those economists who favor non-coercive solutions to social problems welcome religious and other movements whose impact is to effect a change in preferences that reduces the derived demand for government intervention.

Secondly, our experiments highlighted the role of trust in a society as a vehicle for solving resource allocation problems outside the realm of coercive institutions. So far, our groups' tax choices can be significantly predicted by their level of cooperation in the pre-institutional condition. Even with twenty percent inefficiency, the coercive tax was utilized by several groups who experienced high levels of free riding; groups which experienced greater degrees of cooperation relied much less on the tax.

We should be careful not to associate all manifestations of the kind of commitment system evidenced by our laboratory tax as being represented only by a traditional notion of “government.” Groups of individuals may evolve punishment strategies that mimic the commitment value of our tax system and still fall short of the archetypal government. Voluntarily-acting individuals may attempt to construct

– the internet – allows for both a changes of preferences and a greater reliance on private solutions to the problem).

²¹ Compare with Luke 13:20-22.

enforceable contracts. Quasi-government institutions such as homeowners associations may resemble “private governments” with taxing power enacted by land covenants and resulting bylaws, but with much lower costs of entry and exit than with a traditional government.

Our results raise the general issue as to whether less-trusting groups of individuals will be more likely to search for some kind of coercive institution. The natural complement to this statement is that seemingly non-economic institutions play an economic role in creating conditions in which coercive action is not chosen. If so, such “non-economic” institutions have very real economic consequences, and they should include in economists’ broader inquiries into social organizations.

REFERENCES

- Ahn, T.K., R. Mark Isaac, and Tim Salmon. "Endogenous Group Formation," Forthcoming, *Journal of Public Economic Theory*, 2007.
- Andreoni, James. "An Experimental Test of the Public Goods Crowding-Out Hypothesis." *American Economic Review* 83 (1993): 1317-1327.
- Bergstrom, Theodore, Lawrence Blume, and Hal Varian. "On the Private Provision of Public Goods." *Journal of Public Economics* 29 (1986): 25-49.
- Chan, Kenneth, Rob Godby, Stuart Mestelman, and R. Andrew Muller. "Crowding-out Voluntary Contributions to Public Goods." *Journal of Economic Behavior and Organization* 48 (2002): 305-317.
- Eckel, Catherine, Philip Grossman, and Rachel Johnson. "An Experimental Test of the Crowding out Hypothesis." *Journal of Public Economics* 89 (2005): 1543-1560.
- Espinola-Arrendondo, Ana and Fleix Munoz-Garcia. "The Importance of Foregone Options: Generalizing Social Comparisons in Sequential-Move Games." Working Paper. Department of Economics, University of Pittsburgh. (2007).
- Fischbacher, Urs. "z-Tree: Zurich Toolbox for Ready-Made Economics Experiments." *Experimental Economics* 10 (2007): 171-178.
- Gruber, Jonathan and Daniel Hungerman. "Faith Based Charity and Crowd Out During the Great Depression." *Journal of Public Economics*. 91 (2007):1043-1069.
- Hansman, Henry. "The Role of Non-Profit Enterprise." Working Paper 806, Institution for Social and Policy Studies, Yale University. (1978)
- Holcombe, Randall, and Lawrence Kenny. "Evidence on Voter Preferences From Unrestricted Choice Referendums." *Public Choice* 131.1-2 (2007): 197-215.
- Hungerman, Daniel. "Are Church and State Substitutes? Evidence From the 1996 Welfare Reform." *Journal of Public Economics* 89 (2005): 2245-2267.
- Hungerman, Daniel. "Diversity and Crowd-Out: A Theory of Cold-Glow Giving." NBER Working Paper No. 13348. (2007).
- Isaac, R. Mark, Svetlana Pevnitskaya, and Tim Salmon, ""Individual Behavior In Auctions with Price Proportional Benefits," Florida State University, 2008.
- Isaac, R. Mark, James Walker, and Susan Thomas. "Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations." *Public Choice* 43 (1984): 113-149.

Isaac, R. Mark, and James Walker. "Communication and Free-Riding Behavior: The Voluntary Contribution Mechanism." Economic Inquiry 26 (1988): 585-608.

Lakdawalla, Darius, and Tomas Philipson. "The Non-Profit Sector and Industry Performance." Journal of Public Economics 90 (2006): 1681-1698.

Norton, Douglas. "Revolutions Begin in Coffee Houses." Florida State University, Department of Economics, Master's Thesis in Progress.

Payne, A. Abigail. "Does the Government Crowd Out Private Donations? New Evidence from a Sample of Non-Profit Firms." Journal of Public Economics. 69 (1998):323-345.

Sutter, Matthias, and Hannelore Weck-Hannemann. " An Experimental Test of the Public-goods Crowding-out Hypothesis When Taxation is Endogenous." FinanzArchiv 60 (2004): 94-110.

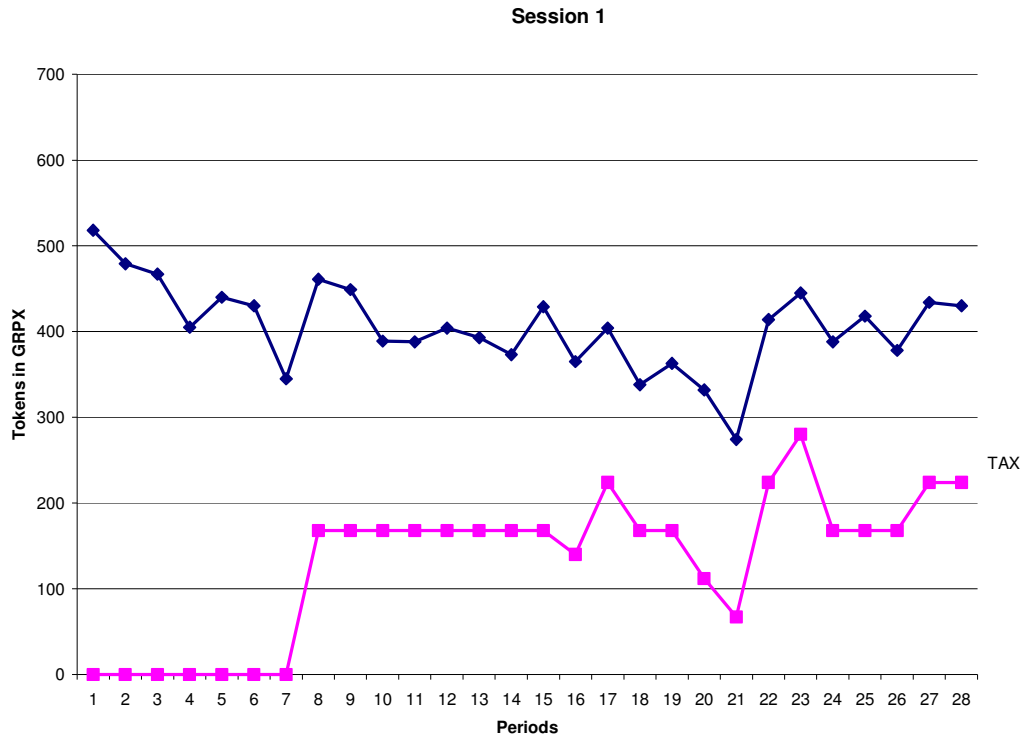


FIGURE 1 (BASELINE)

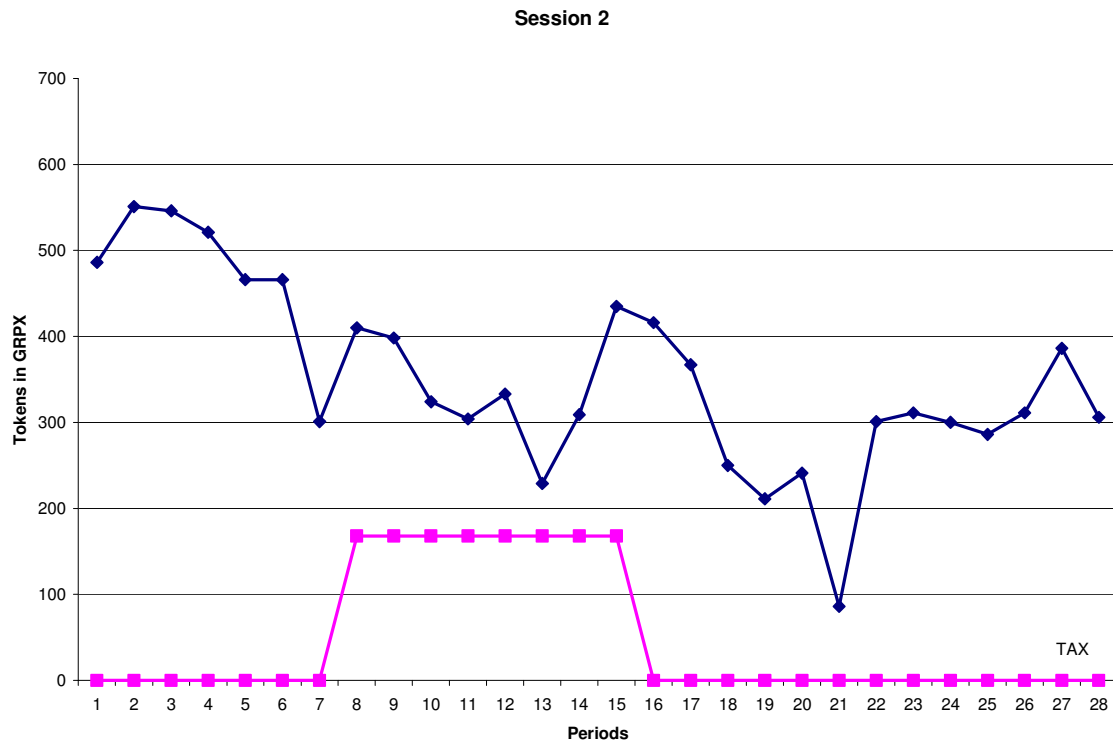


FIGURE 2 (BASELINE)

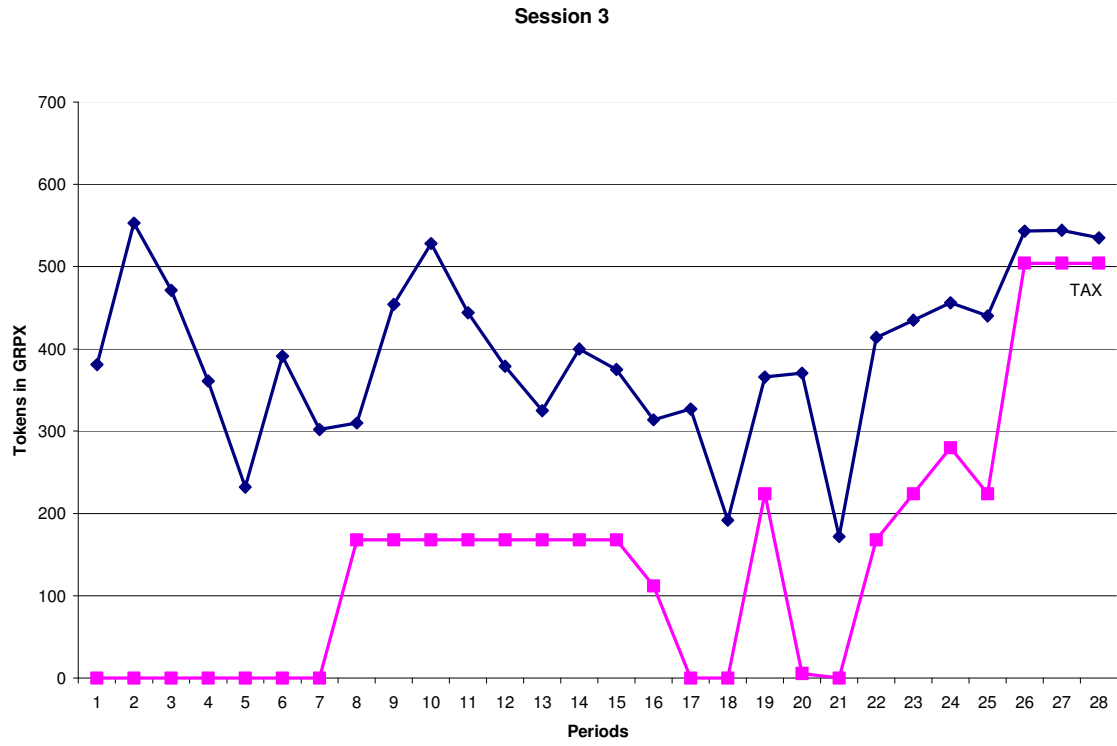


FIGURE 3 (BASELINE)

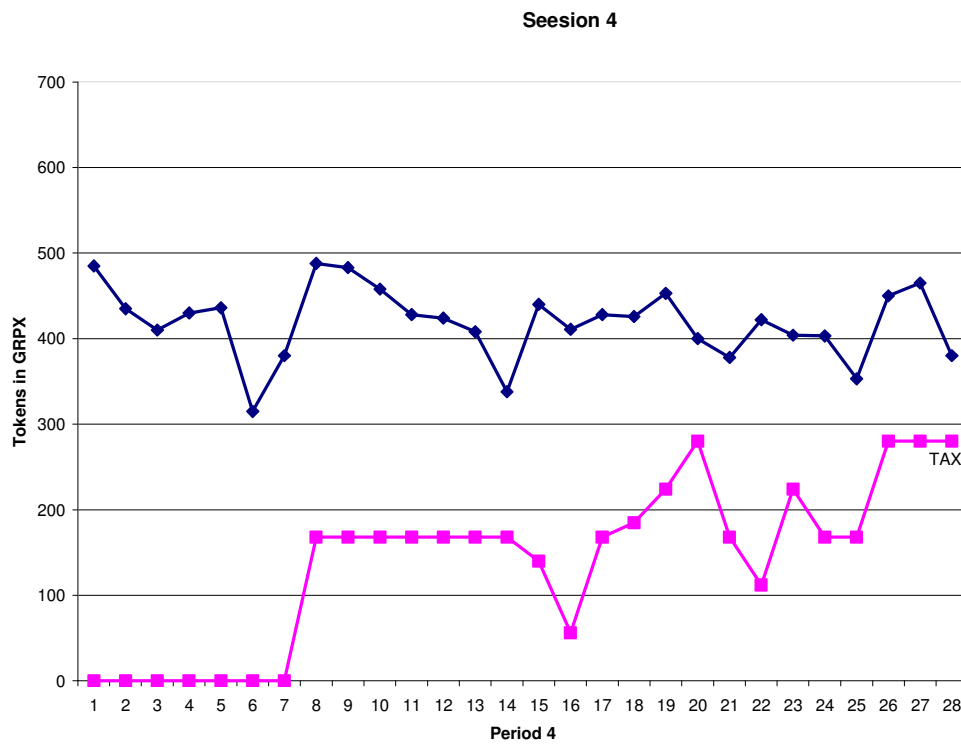


FIGURE 4 (BASELINE)

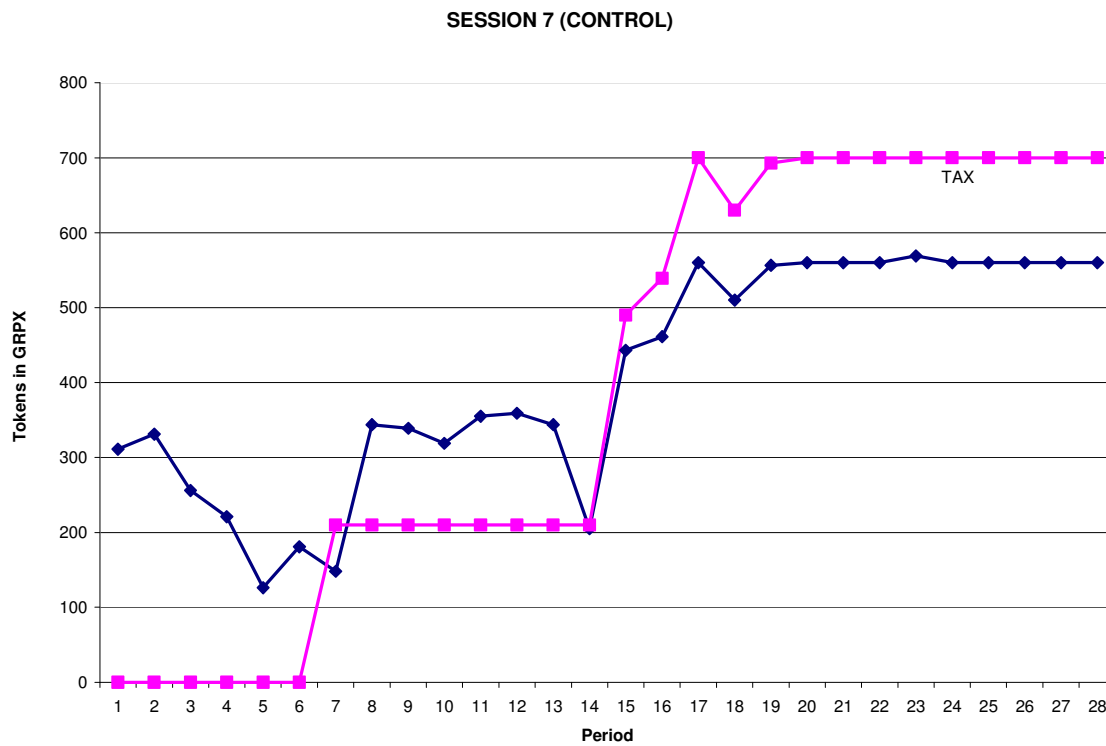


FIGURE 5 (CONTROL)

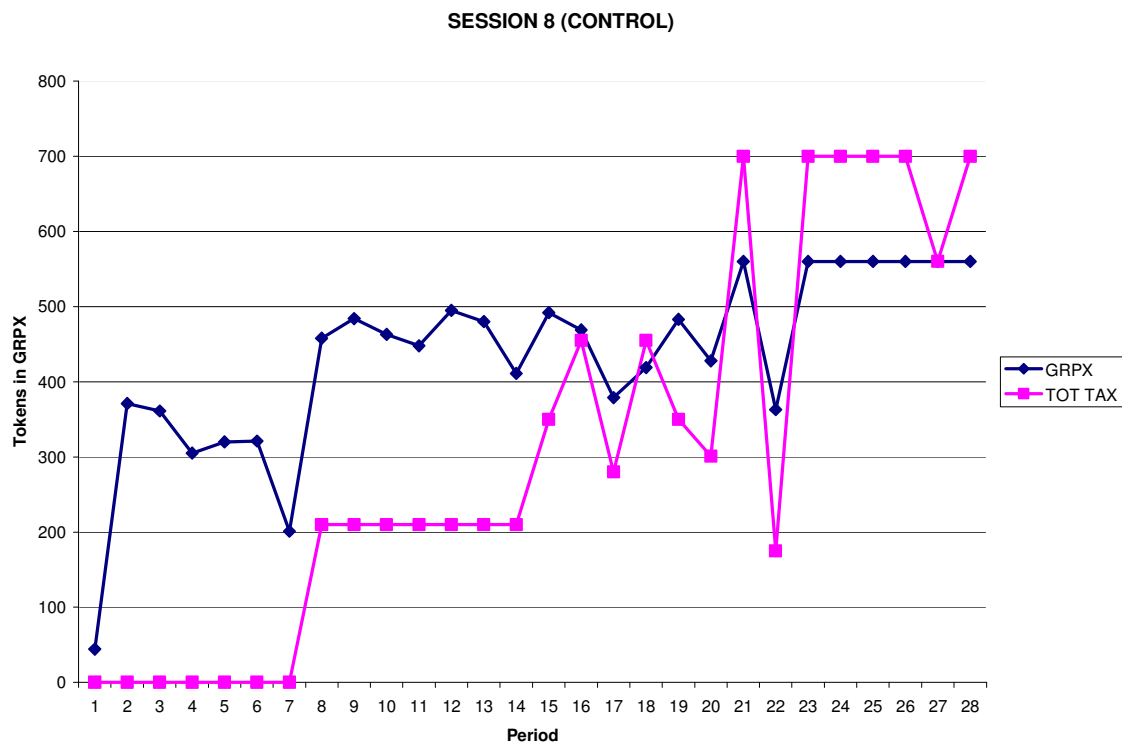


FIGURE 6 (CONTROL)

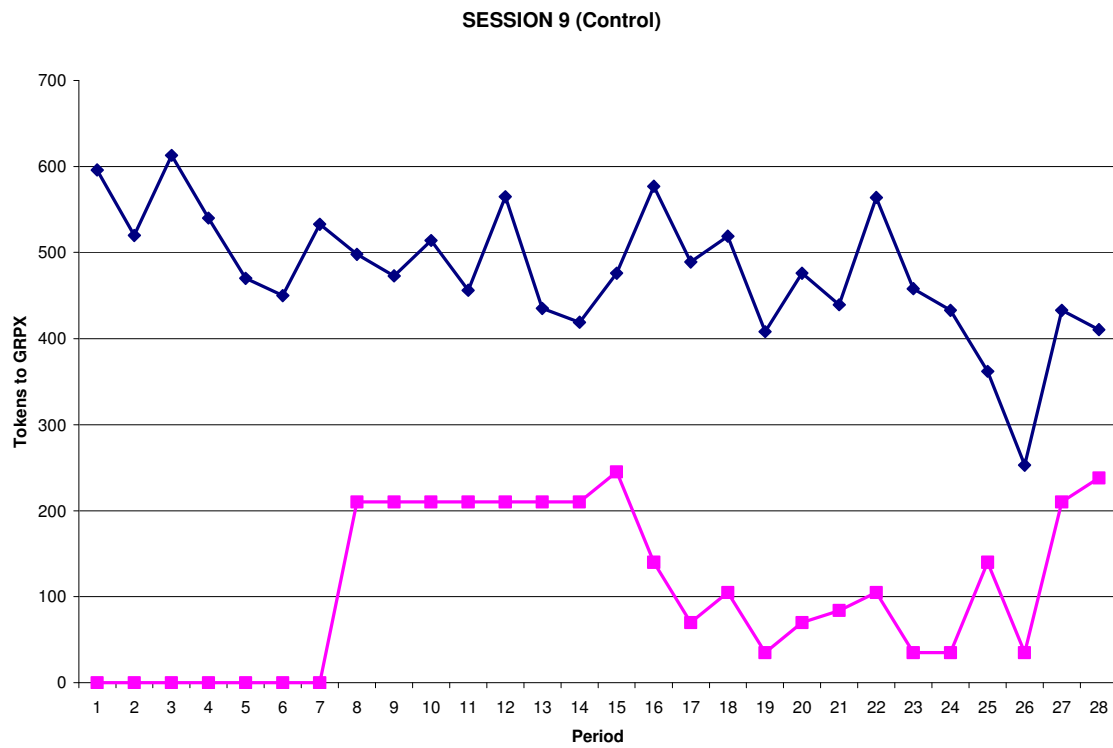


FIGURE 7 (CONTROL)

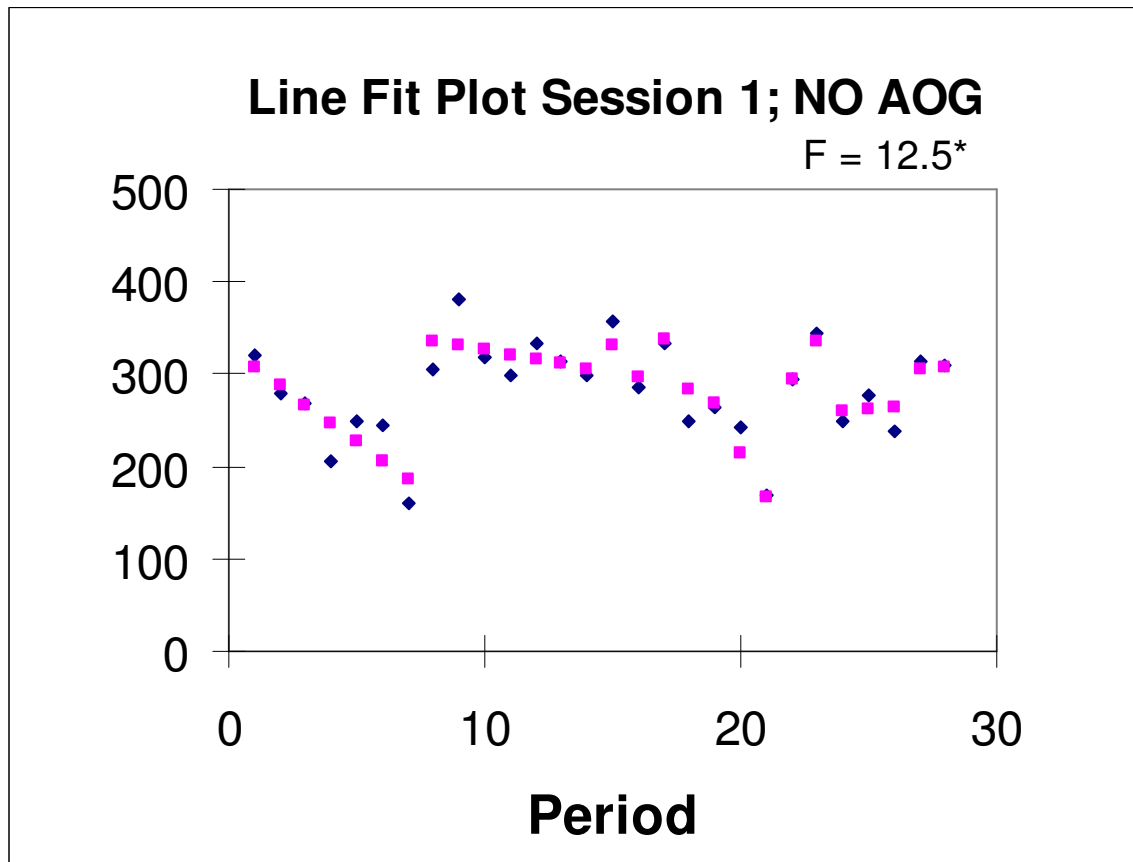


FIGURE 8 (BASELINE; AGENTS OF GRACE REMOVED)

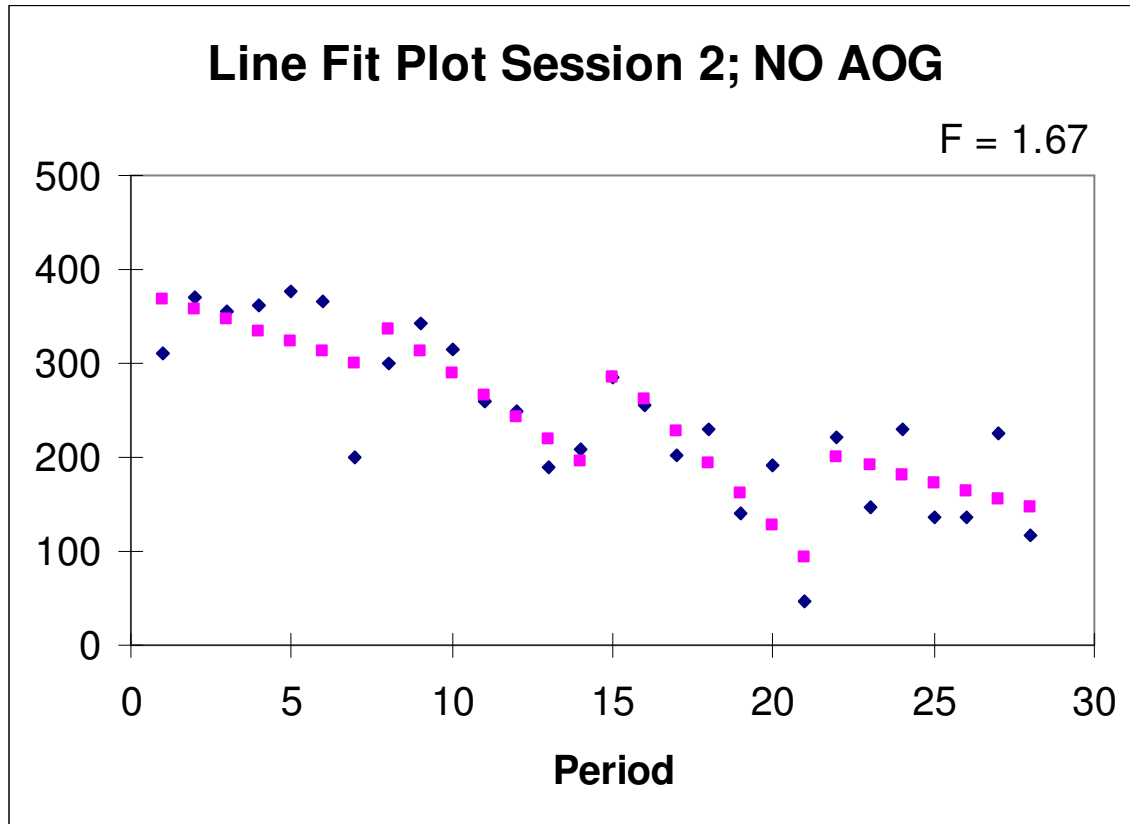


FIGURE 9 (BASELINE: AGENTS OF GRACE REMOVED)

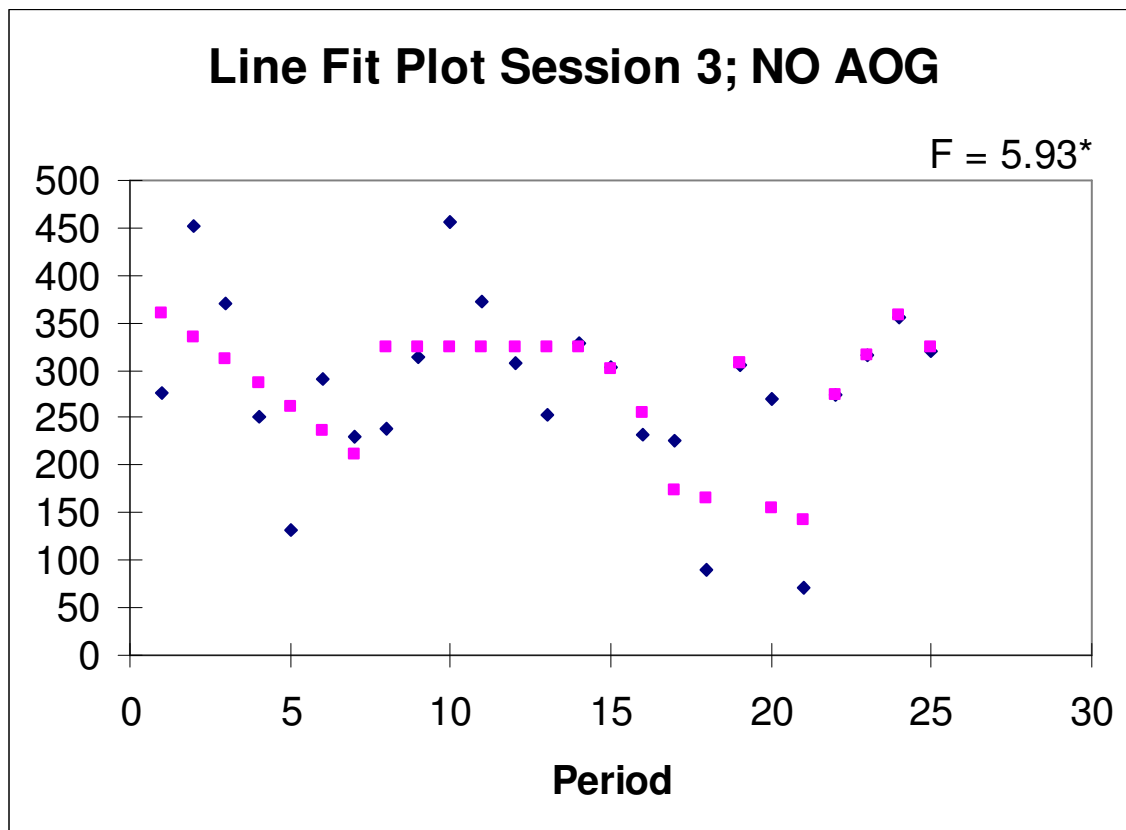


FIGURE 10 (BASELINE; AGENTS OF GRACE REMOVED)

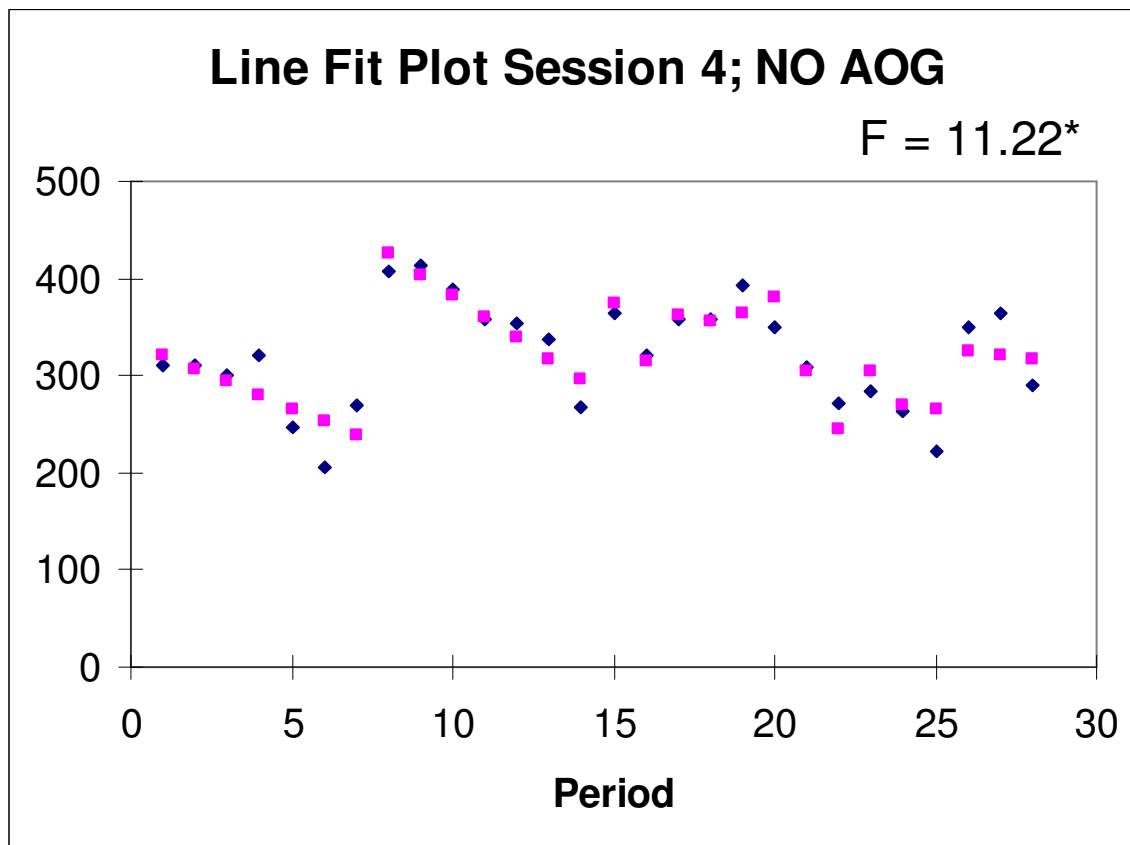


FIGURE 11 (BASELINE; AGENTS OF GRACE REMOVED)

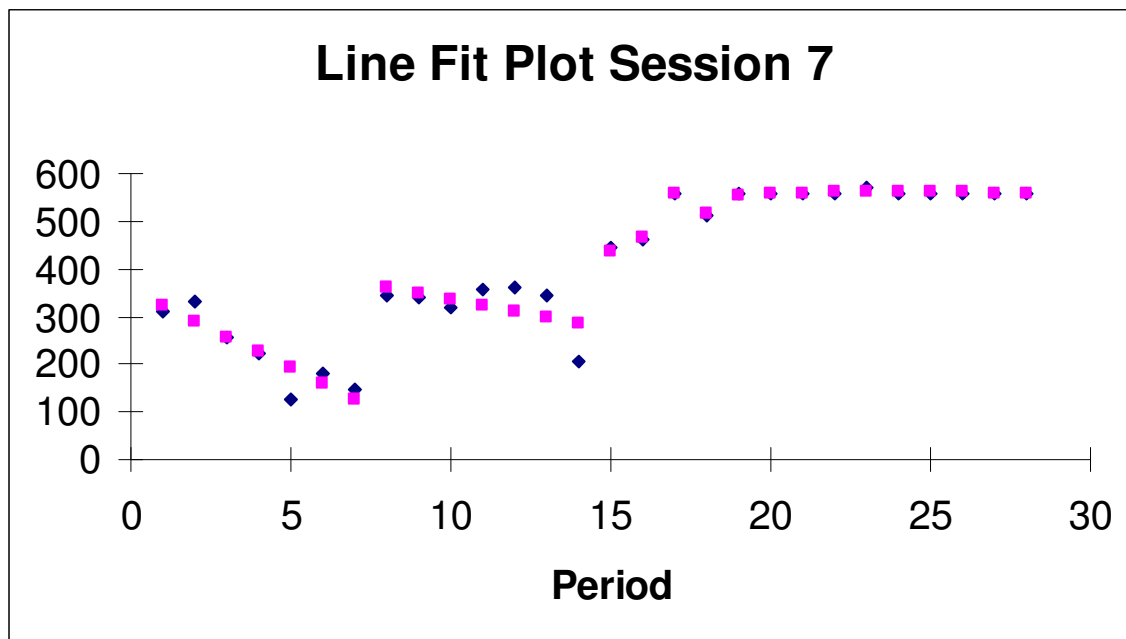


FIGURE 12 (CONTROL)

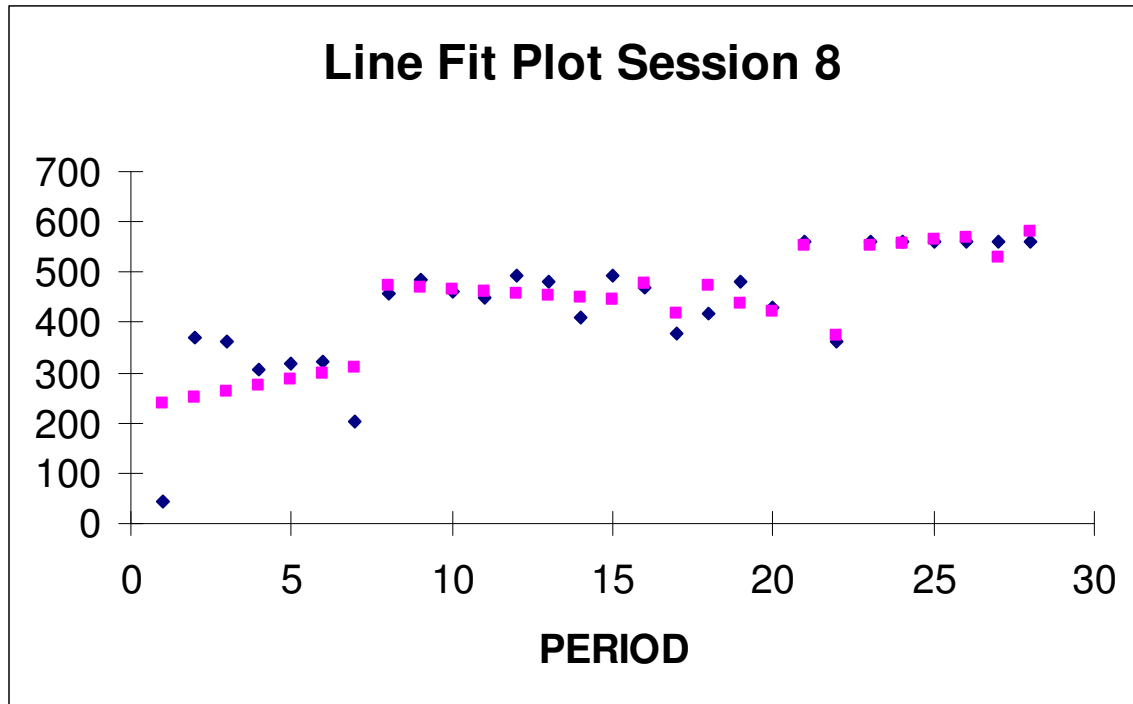


FIGURE 13 (CONTROL)

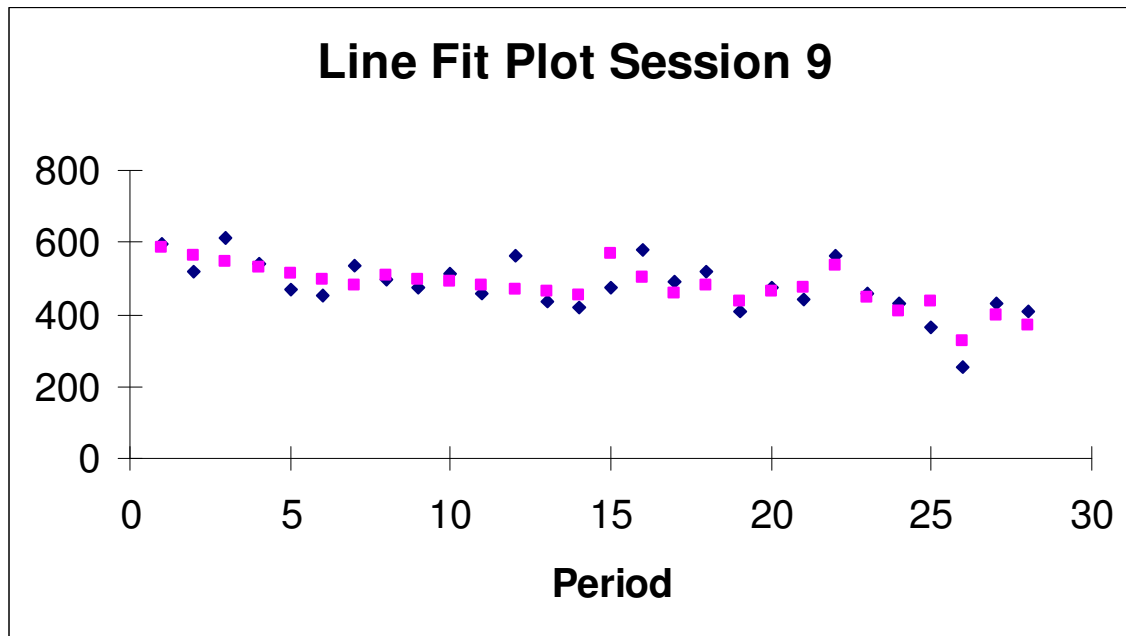


FIGURE 14 (CONTROL)

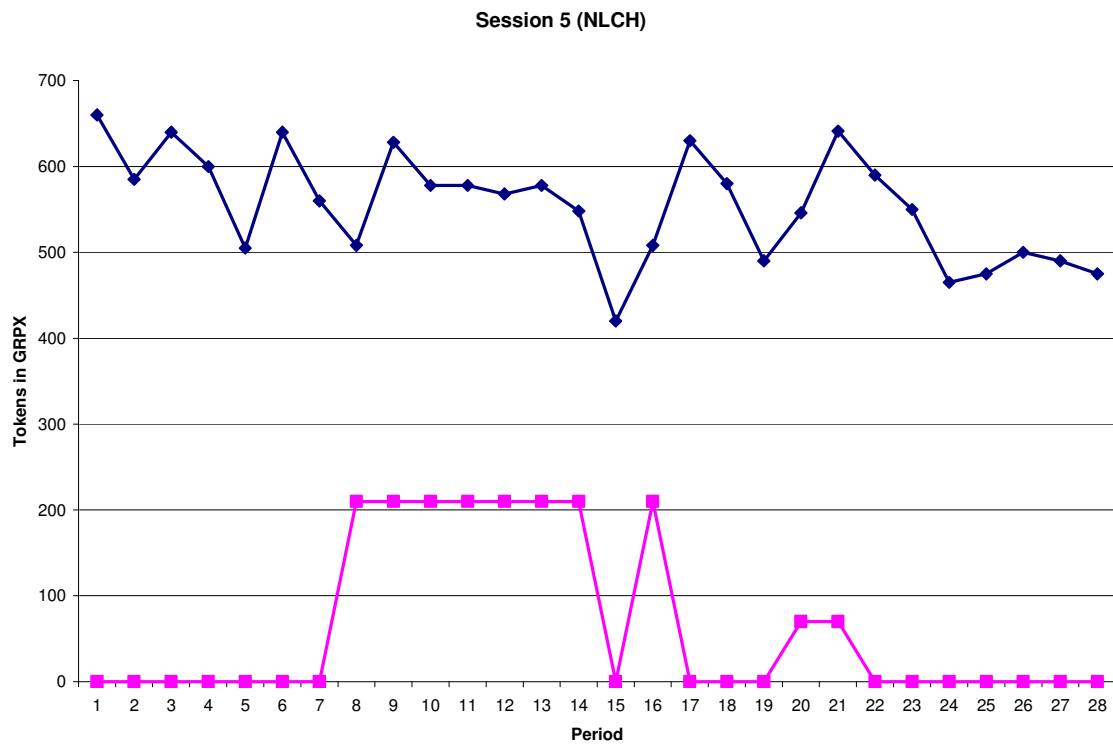


FIGURE 15 (BASELINE; SPECIAL SUBJECT POOL)

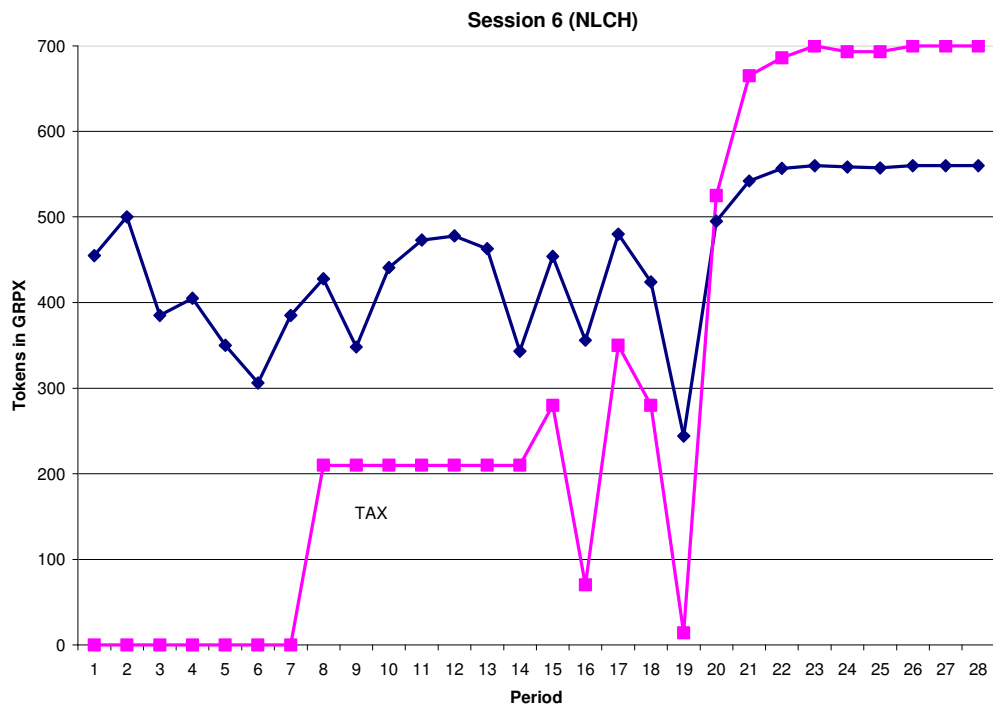


FIGURE 16 (BASELINE: SPECIAL SUBJECT POOL)

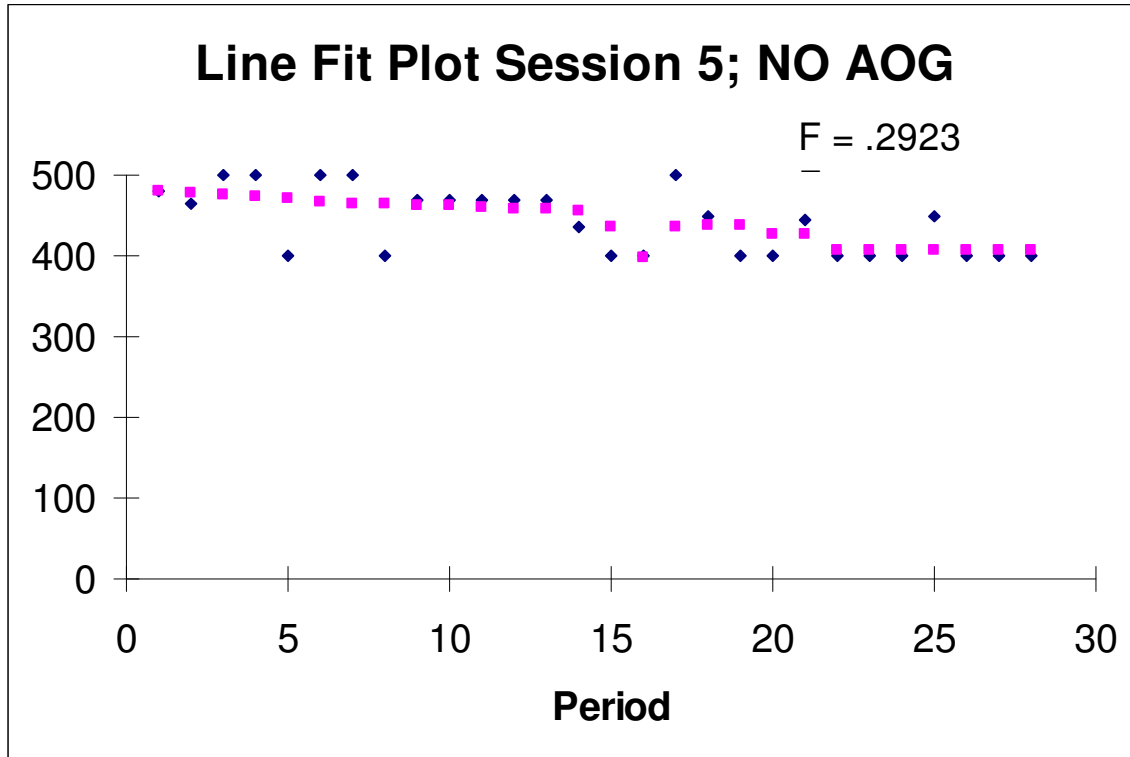


FIGURE 17 (BASELINE; SPECIAL SUBJECT POOL)

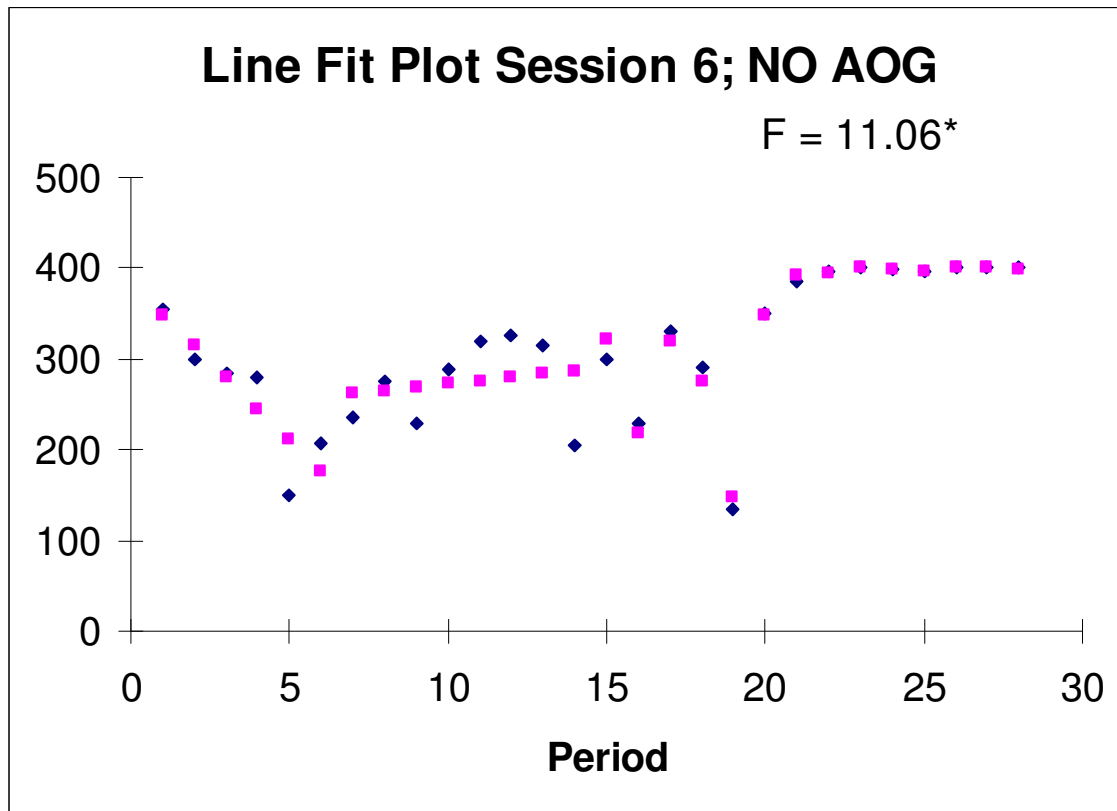


FIGURE 18 (BASELINE; SPECIAL SUBJECT POOL)

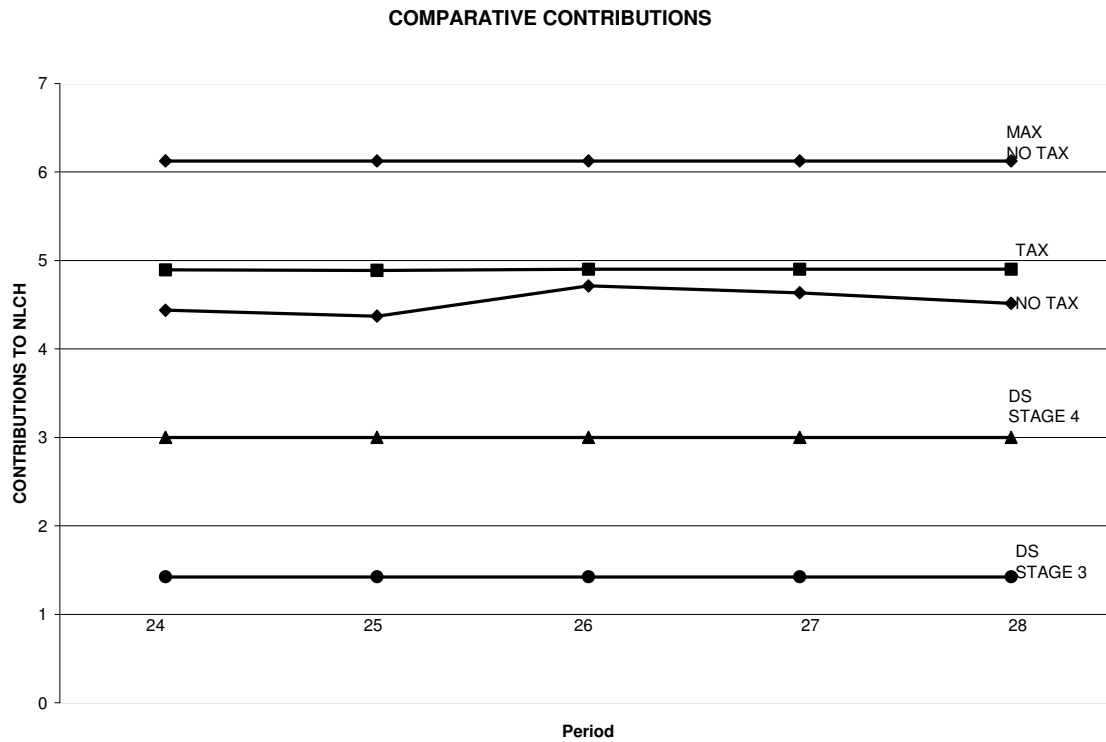


FIGURE 19

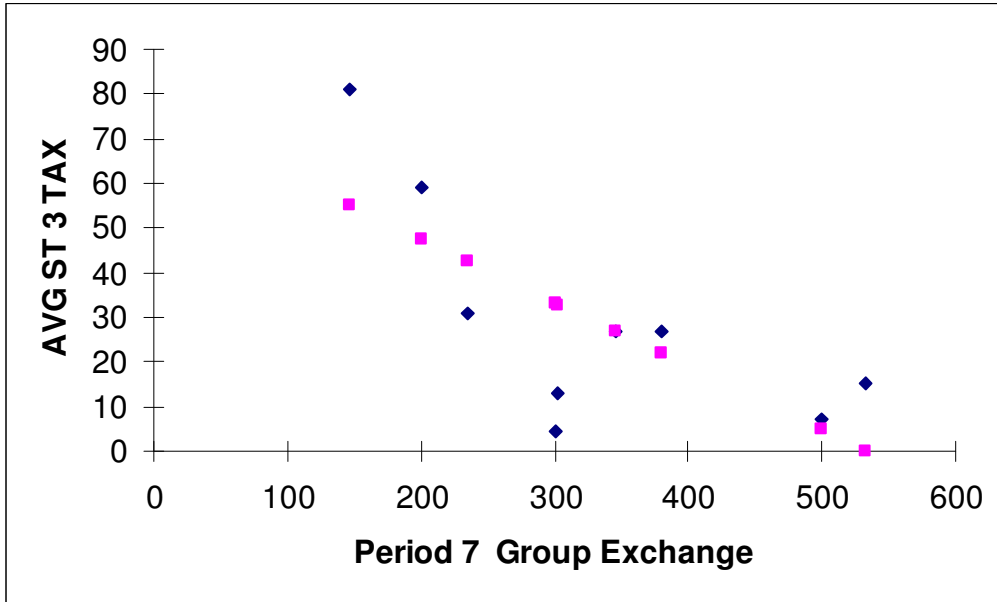


FIGURE 20

REGRESSION TABLES

| Variable | EXP 1 | P- value | EXP 2 | P- value | EXP 3 | P- value | EXP 4 | P- value | EXP 5 | P- value | EXP 6 | P- value |
|---------------------|---------|-------------|---------|-------------|---------|-------------|---------|-------------|---------|-------------|---------|-------------|
| Intercept | 328 | 8E-12 | 380.286 | 2E-08 | 385.571 | 7E-06 | 335 | 2E-11 | 483.667 | 8E-12 | 382.667 | 2E-10 |
| Period | -20.286 | 6E-04 | -11.429 | 0.232 | -24.929 | 0.093 | -13.714 | 0.02 | -2.7143 | 0.752 | -34.286 | 4E-04 |
| MVTAX | 0.69754 | 0.002 | -0.0587 | 0.884 | 0.66264 | 0.029 | 0.55269 | 0.006 | -.2200 | 0.345 | .5237 | 1E-05 |
| DUMST2 | 49.0357 | 0.423 | 139.893 | 0.224 | -61.107 | 0.723 | 263.286 | 7E-04 | -9.9167 | 0.884 | -146.92 | 0.033 |
| DUMST3 | 119.965 | 0.339 | 416.867 | 0.075 | -79.134 | 0.784 | 178.213 | 0.113 | -61.231 | 0.634 | 48.0153 | 0.721 |
| DUMST4 | -231.84 | 0.114 | 15.0357 | 0.95 | -317.3 | 0.536 | -73.24 | 0.649 | -76.524 | 0.659 | -269.73 | 0.127 |
| DUMST2*PER | 15.1786 | 0.044 | -11.679 | 0.383 | 24.9643 | 0.226 | -7.8571 | 0.317 | 1.46429 | 0.886 | 37.9405 | 8E-04 |
| DUMST3*PER | 4.6665 | 0.546 | -22.057 | 0.155 | 17.1338 | 0.415 | -0.6779 | 0.937 | 3.51484 | 0.748 | 19.1285 | 0.103 |
| DUMST4*PER | 22.2187 | 0.005 | 2.53571 | 0.848 | 29.2342 | 0.289 | 10.1658 | 0.27 | 2.71429 | 0.804 | 34.0525 | 0.004 |
| <u>ADJR Squared</u> | .75 | | .69 | | .58 | | .72 | | .23 | | .82 | |

TABLE 1

REGRESSION TABLES (CONT'D)

| Variable | EXP 7 | p- value | EXP 8 | p- value | EXP 9 | p- value |
|---------------|----------|-------------|----------|-------------|----------|-------------|
| Intercept | 356.1429 | .00 | 227.5714 | .00 | 599.1429 | .00 |
| Period | -32.8214 | .00 | 11.7857 | .38 | -16.8571 | .13 |
| MVTAX | 0.7261 | .04 | 0.4085 | .02 | 0.8127 | .04 |
| DUMST2 | 111.6071 | .16 | 281.1071 | .09 | -16.2143 | .90 |
| DUMST3 | -201.54 | .09 | 124.4187 | .62 | -207.36 | .42 |
| DUMST4 | -185.378 | .46 | -43.2572 | .90 | 794.5078 | .02 |
| DUMST2*PER | 19.7143 | .03 | -15.9643 | .40 | 7.5 | .62 |
| DUMST3*PER | 32.6822 | .02 | -13.2051 | .49 | 18.0854 | .28 |
| DUMST4*PER | 32..1786 | .01 | -5.7920 | .77 | -25.2055 | .15 |
| ADJ R Squared | .95 | | .67 | | .45 | |

TABLE 2

