

# Estimation of Treatment Effects under Multiple Equilibria in Repeated Public Good Experiments \*

Jianning Kong  
Shandong University

Donggyu Sul  
University of Texas at Dallas

March 1, 2015

## Abstract

This paper shows that the conventional econometric methods fail to measure treatment effects when there exist multiple equilibria in repeated public good games. To provide an alternative method, this paper develops a new pre-test whether or not there is a single equilibrium, and proposes novel but simple statistical methods to test for treatment effects. The newly proposed tests have good finite sample properties and perform well in practice.

*Key Words: Public Good Games, Multiple Equilibria, Treatment Effects, Nonlinear Decay Model, Group Size*

*JEL Classification Number: C33, C91, C92*

---

\*We thank Ryan Greenaway-McGrevy and Jason Parker for editorial assistance. Also we are grateful to Catherine Eckel, Rachel Croson, Sherry Li, James Walker, Arlington Williams for helpful discussions, and to Joon Park and Yoosoon Chang for thoughtful comments on the previous version of this paper. Many thanks also go to the seminar participants to Indiana University and 2012 North American Economic Science Association meeting.

# 1 Introduction

Voluntary contribution experiments, or broadly public good experiments, have been popularly studied in the last two decades, showing exponential growth. Searching for ‘voluntary contribution’ in JSTOR results in more than 1,700 publications, and Google scholar, 3,000 since 2009. Public good experiments have been studied in almost all of the social sciences including economics, political sciences, finance, marketing, behavioral economics, and psychology, as well as some of the natural sciences. A typical public good experiment starts with recruiting  $G$  human subjects (i.e. group size  $G$ ). Each subject is given a fixed amount of tokens and is asked to invest either a private or group account. From a token invested in the private account the subject can earn one cent usually, and from a token invested in the group account every subject in the group can earn some amount which is called “the marginal per capita return from the group account (MPCR).” The game is repeated over a certain number of rounds, typically 10 rounds. Obviously, traditional game theory predicts that each subject invests all her tokens to the private account if MPCR is less than one, which results in the subgame perfect Nash equilibrium. However the experimental outcomes – the average contributions to the public account – are not always zero even in the long run. Usually experimentalists have studied how to improve the contribution level to the public account by controlling various factors including group size, MPCR, the number of rounds, and punishments for free riders. Econometrically, the estimation of the treatment effects becomes a key task to experimentalists.

Since human subjects are randomly recruited, it might seem easy to think that the estimation of the average treatment effect becomes straightforward. However, this is not true at all. Over rounds, subjects learn about the nature of the game and form their strategies by observing other players. The resulting outcomes are cross-sectionally and serially correlated. In fact, repeated lab experimental data generally contains rich information, but at the same time it is not straightforward to pull out meaningful results. Standard econometric methods are not useful for analyzing treatment effects with repeated experimental data. Sul (2013) points out the pitfalls of the existent econometric methods and suggests an alternative estimation method under the assumption of **one** single equilibrium. Typically, in repeated public good games before 1995, as Ledyard (1995) pointed out, the cross-sectional averages usually converge to zero as the round increases. However, this is not always the case. Ambrus and Greiner (2012) consider a binary voluntary contribution system – contribute all or nothing – and show that subjects can be classified into three types: Contributed always, sometimes, or not-at-all over successive rounds. They further show that as the degree of punishment increases, the fraction of contributing subjects increases as well.<sup>1</sup>

---

<sup>1</sup>Houser, Keane and McCabe (2004) show that subjects can be classified into three types and the cross sectional averages of the contributions across subjects for each type are quite different and divergent in 2x2 strategy games.

Even though this divergent behavior has been widely discussed, the effect of such behavior on the estimation of treatment effects remains largely unknown. This paper deals with this issue and aims to provide simple but effective estimation and testing methods for analyzing repeated public good experimental data. Specifically, we consider the case where the experimental outcomes are continuous, not binary.

By following to Ambrus and Greiner (2012) and Sul (2013), we approximate heterogeneous behaviors of subjects econometrically into three types in terms of subjects' contribution to the public account: Increasing, decreasing and confusing. The increasing group includes two types of subjects: Subjects contribute all tokens to the public account for all rounds or contribute more tokens over successive rounds. Similarly, the decreasing group includes the two types of subjects: pure free riders for all rounds and subjects contribute less tokens over rounds. The last confusing subjects are neither in the decreasing or increasing group. Under this econometric model, we can explain the "jerky" behavior of the cross sectional averages of the experimental outcomes: Under multiple equilibria, we find that the cross sectional variances are usually increasing over successive rounds, the cross sectional averages either fluctuate over rounds or converge to a non-zero positive constant.

The empirical tests for the existence of multiple equilibria in repeated public good experiments have not been developed yet.<sup>2</sup> Usually experimentalists use a rather crude regression method to test for convergence based on the coefficient on the inverse trend ( $1/t$ ) term in the regression. However such method fails when the second moments – cross sectional variance – diverge over successive rounds. We provide a formal test for the existence of single equilibrium. When the null of single equilibrium is not rejected, Sul (2013)'s trend regression can be used to estimate the average treatment effects. However if the null of single equilibrium or no-divergence is rejected, the estimation of the average treatment effects requires clustering subjects into a few groups. We develop a simple clustering mechanism by using a trend regression and provide the statistical justification of the clustering method. Finally, by utilizing the clustering results, we build a new estimator of the average treatment effects.

This paper is written for both general audiences and theoretical econometricians. Sections 2

---

Under such divergent behaviors, Wilcox (2006) shows that pooling estimators are inconsistent in 2x2 strategy games by means of Monte Carlo experiments.

<sup>2</sup>In 2x2 repeated games, the likelihood approach (for example, McKelvey and Palfrey, 1992; El-Gamal and Grether, 1995; Stahl and Wilson, 1995; Camerer and Ho, 1999; Houser, Keane and McCabe, 2004) has been used to examine the divergent behavior. This approach requires the assumption of a finite number of decision choices (usually two choices), and then estimates either (both) the type of heterogeneity or (and) heterogeneous parameter values. When the number of choices or the set of state variables is large, this approach fails due to the curse of dimensionality. For repeated public good games, there are usually a large number of choices available to each subject so that the likelihood method cannot be used.

and 4 are designed with general audiences in mind. The next section shows an empirical example of multiple equilibria and why they matter for the estimation of the treatment effects; it lays out econometric models for multiple equilibria; and lastly, it provides step-by-step procedures to estimate the average treatment effects. The econometric justification of the suggested procedures is provided in Section 3. We provide formal asymptotic theories of suggested estimators and tests. Section 4 demonstrates how to use our newly suggested methods with the actual data. Section 5 examines the finite sample performance of the suggested methods. Section 6 concludes. Technical derivation and proofs are presented in the Appendix.

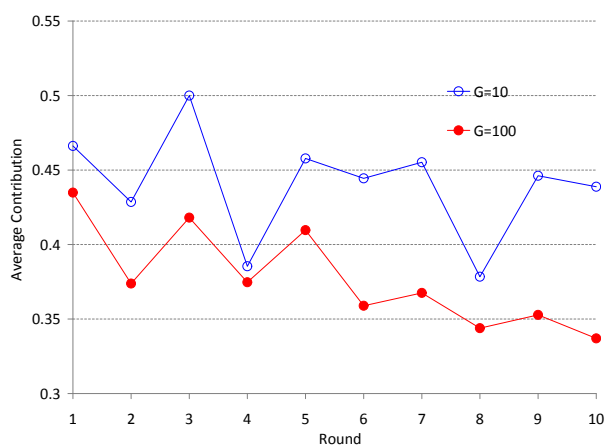
## 2 Multiple Equilibria in Repeated Experiments

First we provide the motivation of the paper by showing an empirical example where multiple equilibria are present. Next, we discuss how to test whether or not there is a single or multiple long run equilibrium(s) and how to estimate overall average effects under the presence of multiple equilibria.

### 2.1 Motivation: Canonical Empirical Example

Throughout the paper, we use the lab experimental data sets provided by Isaac, Walker and Williams (1994, hereafter IWW) which is one of the most popularly cited papers in public good experiments. The game setting is exactly the same as the earlier public good games such as Isaac and Walker (1988a, 1988b) and Andreoni (1995). IWW consider the effects of the group size  $G$  and MPCR on the contribution to the public account. They consider four group sizes ( $G=4,10,40,100$ ) and three different MPCR (0.03,0.3,0.75). Subjects earn either extra-credit points or cash.

Panel A: Average Contribution



Panel B: Cross Sectional Variance

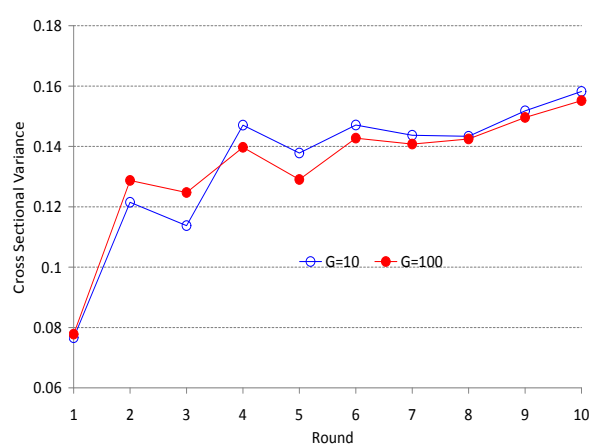


Figure 1: IWW (1994)'s experiments with group sizes of 10 and 100 (MPCR=0.75)

We choose two games as empirical examples where the MPCR is fixed at 0.75, Group Size  $G$  varies from 10 to 100 and the subjects earn extra-credit points. The total numbers of subjects ( $N$ ) in  $G = 10$  and  $G = 100$  are 100 and 300, respectively. Panel A in Figure 1 displays the cross sectional averages of subjects’ contributions to the public account over rounds. By using the so-called “inter-ocular trauma test” or eyeball examination, one may reach a conclusion that subjects tend to contribute more when group size is 10. As Ledyard (1995) and Chaudhuri (2010) point out, the initial contribution is around 50% of the optimal level. However contributions don’t decline steadily over rounds. Rather the experimental outcomes are volatile. Since the eyeball examination ignores uncertainty, a formal statistical test is required to compare the pair of experimental outcomes at each round. IWW used the standard t-test, we report the same test results in Table 1<sup>3</sup> and Wilcoxon, Mann-Whitney’s ranksum results as well. For each comparison, we report the differences of two outcomes and their t-statistics, ranksum statistics in parentheses. Only five out of ten differences are statistically different from zero at the 5% level. Based on these results, IWW concluded that the group size does not influence on the overall contributions to the public account.

Table 1: Comparison for Each Round

	Rounds									
	1	2	3	4	5	6	7	8	9	10
Difference	0.031	0.055	0.082	0.011	0.048	0.086	0.088	0.035	0.093	0.102
Z-score	(0.98)	(1.35)	(2.08)	(0.24)	(1.13)	(1.94)	(2.01)	(0.79)	(2.08)	(2.22)
Rank sum	(0.98)	(1.36)	(2.01)	(0.02)	(1.05)	(1.99)	(2.15)	(0.85)	(2.18)	(2.25)

Panel B in Figure 1 shows cross sectional variances of experimental outcomes for each round. Evidently the cross sectional variance of the two games are increasing over time. In other words the outcomes of the two experiments don’t converge. If all the subjects in one experiment are with the same type, then they will behave in a similar pattern, the cross sectional variance shouldn’t be increasing. As a result, holding difference constant, the standard  $t$  and ranksum tests become less significant as the round increases. For example, the differences in rounds 1 and 8 are similar, but the Z-score is lower in round 8 than that in round 1. Then how can we estimate the overall

<sup>3</sup>Let  $y_{c,Nt}$  and  $y_{\tau,Nt}$  be the cross sectional averages of the contributions of the public account in the controlled and treated groups, respectively. That is,  $y_{s,Nt} = N_s^{-1} \sum_{i=1}^{N_s} y_{s,it}$  where  $N_s$  is the total number of subjects and  $s = c$  or  $\tau$ . Then the simple treatment effect between the two at round  $t$  becomes  $\Pi_t = y_{c,Nt} - y_{\tau,Nt}$ , and its variance can be calculated by the sum of two variances of  $y_{c,Nt}$  and  $y_{\tau,Nt}$ . That is,  $V(\Pi_t) = N_c^{-2} \sum_{i=1}^{N_c} (y_{c,it} - N_c^{-1} \sum_{i=1}^{N_c} y_{c,it})^2 + N_\tau^{-2} \sum_{i=1}^{N_\tau} (y_{\tau,it} - N_\tau^{-1} \sum_{i=1}^{N_\tau} y_{\tau,it})^2$ . The standard z-score is, then, calculated by  $\Pi_t / \sqrt{V(\Pi_t)}$ .

treatment effects under such circumstance? We provide answers shortly.

## 2.2 Econometric Modelling of Multiple Equilibria

We utilize Sul (2013)'s decay model to explain the statistical issues regarding the estimation of treatment effects under multiple equilibria. Sul (2013) approximates the experimental outcome,  $y_{it}$ , as<sup>4</sup>

$$y_{it} = a_i + (\mu_i - a_i) \rho^{t-1} + e_{it}, \text{ for } e_{it} = u_{it} \rho^{t-1}, \quad (1)$$

where  $a_i$  is the long run value of  $y_{it}$ ,  $\mu_i$  is the unknown initial mean of  $y_{i1}$ ,  $\rho$  is the decay rate and  $e_{it}$  is the approximation error. Note that  $u_{it}$  is a bounded random variable with mean zero and variance  $\tilde{\sigma}_i^2$ , or  $u_{it} \sim iB(0, \tilde{\sigma}_i^2)_{-\mu_i}^{1-\mu_i}$ . As  $t$  increases, the unconditional mean of individual outcome,  $y_{it}$ , converges to  $a_i$  which can be interpreted as a long run equilibrium. In public good games, the long run Nash equilibrium occurs when  $a_i = 0$  and  $0 < \rho < 1$  so that as the experiment repeats, the fraction of free riders increases. Meanwhile if  $a_i = 1$  and  $0 < \rho < 1$ , then the unconditional mean of the individual outcome,  $y_{it}$ , converges to the Pareto optimum ( $a_i = 1$ ). Also except for the initial round, the experimental outcomes are cross sectionally dependent so that individual's decision depends on the overall group outcome, which supports the finding by Ashely, Ball and Eckel (2010). Lastly, if the decay rate becomes unity or  $\rho = 1$ , then the individual outcome becomes purely random.

To be specific, we consider the following simple case of triple equilibria.

$$y_{it} = \begin{cases} \mu_i \rho^{t-1} + e_{1,it} & \text{if } i \in G_1, \text{ or } a_i = 0 \text{ and } 0 < \rho < 1 \\ \mu_i + e_{2,it} & \text{if } i \in G_2, \text{ or } \rho = 1 \\ 1 - (1 - \mu_i) \rho^{t-1} + e_{3,it} & \text{if } i \in G_3, \text{ or } a_i = 1 \text{ and } 0 < \rho < 1 \end{cases}. \quad (2)$$

For groups  $G_1$  and  $G_3$ , subjects are learning each round at the same rate,  $\rho$ .<sup>5</sup> It is worthy noting that the asymptotic results in this paper continuously work when the decay function is replaced by  $1/\sqrt{t}$  or  $t$ , which has been considered for the learning effects in public good experiments.<sup>6</sup> In fact, if the decay function were known or followed a deterministic time polynomial function, the asymptotic results would hold much strongly. If a subject belongs to  $G_1$ , the contribution is decreasing over

<sup>4</sup>Here the experimental outcome  $y_{it}$  is normalized by the maximum number of tokens so that  $0 \leq y_{it} \leq 1$  always for all  $i$  and  $t$ .

<sup>5</sup>Here we impose homogeneity restriction of  $\mu_i$  and decay rate  $\rho$  across different groups for two reasons. First, there is no theoretical and empirical evidence that the unknown initial mean is different across subgroups. As Figure 1 and 2 show, regardless of time varying patterns in the cross sectional variances, the initial mean seems to be around 50% of the Pareto optimal level, which is a well known fact in this literature. Second, usually the games are symmetric so that the decay rate becomes identical in  $G_1$  and  $G_3$ .

<sup>6</sup>See Chaudhuri, Graziano and Maitra (2006), Ho and Su (2009), Ashely, Ball and Eckel (2010), Bruyn and Bolton (2008).

rounds and if she belongs to  $G_3$ , the contribution is increasing. In the long run,  $G_1$  becomes the Nash group and  $G_3$  becomes the Pareto group. The remaining subjects are classified together in a ‘confused’ group. The learning rate of the subjects in the confused group is assumed to be unity.<sup>7</sup> The latent model in (2) explains different variances between the confused and non-confused groups. For non-confused groups ( $G_1$  and  $G_3$ ), the random errors can be rewritten as  $e_{s,it} = u_{s,it}\rho^{t-1}$  for  $s = 1, 3$ . Hence as round increases, the variance of  $e_{s,it}$  decreases and eventually converges to zero. Meanwhile the random error for  $G_2$  is not time varying since  $\rho$  is always unity.

The model in (2) takes into account several important features of learning which have been the focuses in the literature: Serial and cross sectional dependence, heterogeneous patterns. First, in belief learning models (Broseta, 1995; Crawford, 1995; Fudenberg and Levine, 1995, 1998; Cheung and Friedman, 1997), players form beliefs based on their past decisions and then they choose the best-response by maximizing their expected payoffs. In reinforcement learning models (Harley, 1981; Cross, 1983; Bush and Mosteller, 1995; Roth and Erev, 1995), players are likely to choose the strategy which gave best payoff in the past. Camerer and Ho (1999) suggested the Experience-Weighted Attraction model (EWA hereafter) which combines both reinforcement learning and belief learning. EWA model depreciates previous experience and attraction using several parameters which can be think of as decay rates. The serial correlation in (2) is reflected through term  $\mu_i\rho^{t-1}$  and  $\rho$  is the decay or learning rate.

Second, subjects’ behavior is correlated with others. Hyndman, Ozbay, Schotter, and Ehrblatt (2005) consider a teaching and learning process and show that some subjects try to influence the beliefs of their opponent and lead the way to a more favorable outcome. Blume, Duffy, Franco (2009) found that in the learning behavior subjects appear to take into account the confounding effects of the simultaneous learning of others. The model suggested here imposes the cross-sectional correlation through the common factor  $\rho^{t-1}$ .

Third, subjects behave differently from each other and can be classified into several categories. See Houser, Keane, and McCabe (2004), Nathaniel and Wilcox (2006), Isaac, Walker and Thomas (1984), Gächter and Thöni (2005) for detail. In our model, we generally consider three types of behavior patterns  $G_1, G_2$  and  $G_3$ .

Let  $N_1$  be the number of subjects in  $G_1$  and  $n_1 = N_1/N$ . Then if  $n_1 = 1$ , then there is a unique

---

<sup>7</sup>Mathematically,  $G_2 \neq (G_1 \cup G_3)^c$  but  $G_2 \subset (G_1 \cup G_3)^c$  since the case where  $a_i = a$  with  $0 < a < 1$  and  $0 < \rho < 1$  is ignored. In other words, there is possibility that some subjects form a convergent group such that  $a_i = a$  for  $0 < a < 1$ . Here we don’t allow such sub-convergent behavior in public good games because of at least two reasons. First, there is no theoretical justification of why some subjects form sub-dominant strategy in repeated public good games. Second, in our limited experience, there is little empirical evidence that subjects actually form the convergence club among the confused subjects. Rather most of the confused subjects switch their contributions from zero to one or from near zero to near one, which leads to larger variance of  $e_{2,it}$  than  $e_{1,it}$  or  $e_{3,it}$ .

equilibrium. Alternatively, if  $n_1 < 1$ , then multiple equilibria exist. Similarly we can define  $N_2$  and  $N_3$  as the numbers of subjects in  $G_2$  and  $G_3$ , and  $n_2$  and  $n_3$  as their fractions, respectively. Here we need to impose additional restriction that  $n_k$  for  $k = 1, 2, 3$  is not a negligible number. Statistically we can say that  $n_k$  should be a fixed number but should not go to zero as  $N \rightarrow \infty$ .

Next we discuss the estimation issues under multiple equilibria. The sample cross sectional averages for each round under single (only  $G_1$  exists) and multiple equilibria can be written as

$$\frac{1}{N} \sum_{i=1}^N y_{it} = \begin{cases} \mu \rho^{t-1} + e_{Nt}^\dagger & \text{if all } i \in G_1 \\ \tau + \phi \rho^{t-1} + e_{Nt}^* & \text{if some } i \notin G_1 \end{cases}, \quad (3)$$

where  $\tau = n_2 \mu + n_3$ ,  $\phi = \mu - \tau$ , and

$$\begin{aligned} e_{Nt}^\dagger &= N^{-1} \sum_{i=1}^N e_{it} + \rho^{t-1} N^{-1} \sum_{i=1}^N (\mu_i - \mu), \\ e_{Nt}^* &= \rho^{t-1} N^{-1} \sum_{i=1}^N (\mu_i - \mu) + (1 - \rho^{t-1}) N^{-1} \sum_{i=1}^{N_2} (\mu_i - \mu) + e_{Nt}. \end{aligned}$$

Therefore, the time varying behavior of the cross sectional averages under a single equilibrium become very different from those under multiple equilibria. First, under this single equilibrium, the average outcome decreases each round and converges to zero. However under multiple equilibria, depending on the value of  $\phi$ , the average outcome may decrease ( $\phi > 0$ ), increase ( $\phi < 0$ ) or does not change at all ( $\phi = 0$ ) over rounds. Second, the variance of the random part of the cross sectional average,  $e_{Nt}^\dagger$ , under a single equilibrium is much smaller than that under multiple equilibria. In other words, the cross sectional averages under multiple equilibria become much more volatile than those under a single equilibrium. Third, both random errors,  $e_{Nt}^\dagger$  and  $e_{Nt}^*$ , are serially correlated. However the serial correlation of the random errors under single equilibrium,  $E e_{Nt}^\dagger e_{Nt-s}^\dagger$ , goes away quickly as  $s$  increases, meanwhile the serial correlation under multiple equilibria,  $E e_{Nt}^* e_{Nt-s}^*$  never goes to zero even when  $s \rightarrow \infty$  as long as  $n_2 \neq 0$ . This implies that the standard t-test under multiple equilibria becomes invalid because of the permanent serial dependence in the error terms. Also note that under multiple equilibria, Sul's trend regression fails since  $\tau \neq 0$ .

### 2.3 Identifying Existence of Multiple Equilibria and Estimation of Treatment Effects

Both under single and multiple equilibria, the sample cross sectional average can be written as

$$\frac{1}{N} \sum_{i=1}^N y_{it} = \tau_N + (\mu_N - \tau_N) \rho^{t-1} + e_{Nt}, \quad (4)$$

where

$$\tau_N = n_3 + n_2 \mu_{N_2}, \quad \mu_{N_2} = N_2^{-1} \sum_{i=1}^{N_2} \mu_i, \quad \mu_N = N^{-1} \sum_{i=1}^N \mu_i, \quad \text{and } e_{Nt} = N^{-1} \sum_{i=1}^N e_{it}.$$



Also the probability limit of the average outcome is given by

$$\text{AE} = \text{plim}_{N \rightarrow \infty} \frac{1}{TN} \sum_{i=1}^N \sum_{t=1}^T y_{it} = \frac{1 - \rho^T}{1 - \rho} \frac{n_1 \mu}{T} + n_2 \mu + n_3 - \frac{n_3(1 - \mu)}{T} \left( \frac{1 - \rho^T}{1 - \rho} \right).$$

Evidently, the expected outcome becomes a function of the number of rounds,  $T$ . To obtain a robust measure, as defined in Sul (2013), the asymptotic average effect (AAE) can be used by letting  $N, T \rightarrow \infty$ . That is,

$$\text{Asym AE} = \text{plim}_{N, T \rightarrow \infty} \frac{1}{TN} \sum_{i=1}^N \sum_{t=1}^T y_{it} = n_2 \mu + n_3 = \tau. \quad (5)$$

Under Nash equilibrium, AAE,  $\tau$ , becomes zero. In this case, as Sul (2013) suggests, one needs to calculate the asymptotic overall effect (OE) given by

$$\text{Asym OE} = \text{plim}_{N, T \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^T y_{it} = \frac{\mu}{1 - \rho}. \quad (6)$$

Nonetheless, testing the existence of multiple equilibria and estimating the AAE can be jointly done by utilizing the sample cross sectional average given in (4) if the decay rate,  $\rho$ , is less than unity. That is,

$$\boxed{\begin{array}{l} \mathcal{H}_0 : \text{Single Equilibrium} \\ \mathcal{H}_A : \text{Multiple Equilibria} \end{array}} \iff \boxed{\begin{array}{l} \mathcal{H}_0 : \tau = 0 \text{ or } 1 \\ \mathcal{H}_A : 0 < \tau < 1 \end{array}} \quad (7)$$

However, when  $\rho$  is equal to unity, then the above identify fails: If all subjects are confused so that  $n_2 = 1$ , then  $\rho = 1$  but  $\tau \neq 0$ . In practice all three parameters,  $\rho$ ,  $\tau$  and  $\phi (= \mu - \tau)$  are unknown. Even when  $N$  is large,  $\tau$  is impossible to estimate consistently by running nonlinear least squares in (4) due to the lack of identification. When  $\mu = \tau$ ,  $\rho$  is not identified since  $\phi = 0$ . Alternatively, when  $\rho = 1$ ,  $\phi$  is not identified from  $\tau$ . To overcome this identification issue, we provide the following sequential estimation and testing procedures.

**Step 1 (Pretesting of Single Equilibrium)** Under a single equilibrium, all subjects must be in one of groups in (2). In this case, the expectation of the cross sectional variance is either constant or decreasing over rounds. To see this, define  $H_{N,t}$  as the cross sectional variance of  $y_{it}$  at round  $t$ ,  $\hat{\sigma}_\mu^2$  as the sample variance of unknown  $\mu_i$  and  $\hat{\sigma}^2$  as the cross sectional variance of the error  $e_{it}$ . That is,  $H_{N,t} = N^{-1} \sum_{i=1}^N \left( y_{it} - N^{-1} \sum_{i=1}^N y_{it} \right)^2$ , and  $\hat{\sigma}^2$  and  $\hat{\sigma}_\mu^2$  are similarly defined. The expectation of the cross sectional variance of  $y_{it}$  becomes

$$\text{E}H_{N,t} = \begin{cases} (\sigma_\mu^2 + \sigma^2) \rho^{2t-2} & \text{if all } i \in G_1 \text{ or } G_3 \\ \sigma_\mu^2 + \sigma^2 & \text{if all } i \in G_2 \end{cases} \quad (8)$$

Hence if all subjects are either in  $G_1$  or  $G_3$ , they choose the same outcome as the round increases, which results in a decreasing pattern in the cross sectional variance. However if

there are more than single equilibrium, then the cross sectional variance is generally (not always) increasing over time as long as some subjects are altruists. See the next section for a detailed discussion. We suggest running the following trend regression with the sample cross sectional variances.

$$H_{N,t} = \alpha_T + \beta_T t + \epsilon_t. \quad (9)$$

Under a restrictive assumption, the hypotheses in (7) can be re-expressed as

$$\mathcal{H}_0 : \beta_T \leq 0, \text{ v.s. } \mathcal{H}_A : \beta_T > 0. \quad (10)$$

To test this null hypothesis, one can use a standard  $t$ -statistic. Since the regressor is a trend, the asymptotic variance can be calculated in the standard way. That is, the  $t$ -statistic can be estimated by

$$t_{\hat{\beta}} = \hat{\beta}_T / \hat{\sigma}_{\beta},$$

where  $\hat{\sigma}_{\beta}^2 = \hat{\sigma}_u^2 \left( \sum_{t=1}^T \tilde{t}^2 \right)^{-1}$  for  $\tilde{t} = t - T^{-1} \sum_{t=1}^T t$ , and  $\hat{\sigma}_u^2$  is the average of squared residuals. If  $t_{\hat{\beta}} \geq 1.65$ , then the null of no divergence is rejected at the 5% level. See Section 3.4 for the asymptotic properties of the suggested method.

**Step 2 (Clustering Subjects)** This step includes several sub-steps of estimation.

**Step 2-1: [Clustering Regression]** First sort out individuals whose outcomes are constant. If  $y_{it} = 0, 1$ , or  $d$  with  $0 < d < 1$  for all  $t$ , then cluster this individual into  $G_1, G_3$  or  $G_2$ , respectively. Collect the remaining individuals and run the following trend regression for each subject.

$$y_{it} = a + bt + w_{it}.$$

Construct the  $t$ -ratio.

**Step 2-2: [Clustering Core Groups]** Cluster subjects to core long run Nash and Pareto groups as follows

$$i \in \begin{cases} G_1^c, \text{ Core long run Nash} & \text{if } t_{\hat{\beta}} < -cv \\ G_3^c, \text{ Core long run Pareto} & \text{if } t_{\hat{\beta}} > cv \end{cases},$$

where  $cv$  is the critical value. We recommend using  $cv = 2.58$  which is the critical value at the 0.5% level. Take cross sectional averages of the subjects' outcomes in the core long run Nash and Pareto groups. Note that core long run Nash (Pareto) group includes all individuals whose outcomes are always zero (one) for all rounds. Denote them as  $y_{N,t}^{cn}$  and  $y_{N,t}^{cp}$ , respectively.

**Step 2-3: [Estimation of Decay Rate  $\rho$  and  $\mu$ ]** Further define  $\hat{\mu}$  as the cross sectional average of the first observations for all subject. That is,  $\hat{\mu} = N^{-1} \sum y_{i,1}$ . Next, run the following pooled regression

$$\begin{aligned} \ln \left( y_{N,t}^{cn} \right) - \ln \hat{\mu} &= (\log \rho) (t - 1) + v_{1t}, \\ \ln \left( 1 - y_{N,t}^{cp} \right) - \ln (1 - \hat{\mu}) &= (\log \rho) (t - 1) + v_{3t}, \end{aligned}$$

and get  $\hat{\rho} = \exp \left( \widehat{\log \rho} \right)$ .

**Step 2-4: [Estimation of  $n_2$  and  $\sigma_\mu^2$ ]** Cluster subjects to the confused group by using the following criteria

$$i \in G_2 \text{ if } -cv_T \leq t_{\hat{i}} \leq cv_T, \quad (11)$$

where  $cv_T = 1.28$  if  $T = 10$  but  $cv_T = 1.632$  if  $T = 20$ . Let  $\hat{N}_2$  be the total number of the clustered confused subjects. Then  $\hat{n}_2 = \hat{N}_2/N$ . Take the time series mean for each clustered confused subject, and then calculate the variance of  $\hat{N}_2$  time series means. This statistic becomes a consistent estimator for  $\sigma_\mu^2$ . That is,

$$\hat{\sigma}_\mu^2 = \frac{1}{\hat{N}_2} \sum_{i \in G_2} \left( \frac{1}{T} \sum_{t=1}^T y_{it} - \frac{1}{\hat{N}_2} \sum_{i \in G_2} \frac{1}{T} \sum_{t=1}^T y_{it} \right)^2.$$

**Step 3: (Estimation of AAE for Each Game)** Run (4) by using  $\hat{\rho}$ . That is,

$$y_{N,t} = \tau + \phi \hat{\rho}^{t-1} + e_{Nt}^+, \quad (12)$$

where  $y_{N,t}$  is the cross sectional average of  $y_{it}$ . The variance of  $\hat{\tau}$  is calculated by

$$V(\hat{\tau}) = \hat{n}_2 \hat{\sigma}_\mu^2 / N + \frac{\hat{n}_2}{NT^2} \sum_{t=1}^T (\hat{e}_{Nt}^+)^2,$$

and the  $t$ -statistic becomes

$$t_{\hat{\tau}} = \hat{\tau} / \sqrt{V(\hat{\tau})}.$$

**Step 4: (AATE between Two Experimental Outcomes)** The asymptotic average treatment effect between two experiments  $A$  and  $B$  can be defined as  $\Pi = \tau_A - \tau_B$ . Now consider the following two cases. First, experiment  $A$  converges to the Nash equilibrium but  $B$  diverges. Second both  $A$  and  $B$  diverge. The asymptotic average treatment effect in the first case is given by

$$\tau_A - \tau_B = \text{plim}_{N,T \rightarrow \infty} \frac{1}{TN} \sum_{i=1}^N \sum_{t=1}^T (y_{it}^A - y_{it}^B) = -\tau_B < 0,$$

since  $\tau_A$  becomes zero as  $T$  becomes a large number. Hence a divergent experiment always provides a better outcome than a convergent experiment in terms of the average number of

tokens contributed to the public account. Next consider the second case. Since both  $A$  and  $B$  diverge, the t-ratio can be constructed by

$$\hat{\Pi} = \hat{\tau}_A - \hat{\tau}_B, \quad t_{\hat{\Pi}} = \hat{\Pi} / \sqrt{V(\hat{\tau}_A) + V(\hat{\tau}_B)}.$$

If  $|t_{\hat{\Pi}}| > 1.96$ , then the difference between the two experiments becomes statistically significant at the 5% level.

In the next section, asymptotic properties of the convergence, clustering method and treatment regression are discussed.

### 3 Asymptotic Theory

There are six parameters of interest:  $n_3, n_2, \mu, \sigma_\mu^2, \sigma^2$ , and  $\rho$ . The question is whether or not there is a statistical procedure to estimate all six parameters jointly from (4). As we discussed earlier, the nonlinear LS estimator cannot identify  $\tau$ ,  $\phi$  and  $\rho$  jointly. Even when  $\rho$  is known,  $n_3$  and  $n_2$  are not identifiable by running the LS regression in (4). Hence we consider the sequential estimation strategy described in Section 2. Here we provide formal asymptotic theory for the suggested methods. We take the following assumptions.

#### Assumption A: Bounded Distributions

**(A.1)**  $\mu_i$  is independent and identically distributed with mean  $\mu$  and the variance  $\sigma_\mu^2$  but bounded between 0 and 1:  $\mu_i \sim iidB(\mu, \tilde{\sigma}_\mu^2)_0^1$

**(A.2)**  $e_{it} = u_{it}\rho^{t-1}$  for  $0 \leq \rho \leq 1$  where  $u_{it}$  is independently distributed with mean zero and variance  $\sigma_i^2$  but bounded between  $-\mu_i^*$  and  $\mu_i^*$  where  $\mu_i^* = \min[\mu_i, 1 - \mu_i]$ . That is,  $u_{it} \sim idB(0, \tilde{\sigma}_i^2)_{-\mu_i^*}^{\mu_i^*}$ .

#### Assumption B: Data Generating Process

**(B.1)** The data generating process is given by  $y_{it} = a_i + (\mu_i - a_i)\rho^{t-1} + e_{it}$ . Each group is defined as follows:  $a_i = 0$  and  $0 < \rho < 1$  for  $i \in G_1$ ,  $a_i = 1$  and  $0 < \rho < 1$  for  $i \in G_3$ , and  $\rho = 1$  for  $i \in G_2$ .

**(B.2)** The group fraction rates  $n_1, n_2$ , and  $n_3$  are independent from the size of  $N$ .

Assumption A is a standard assumption. See Sul (2013) for a detailed discussion. Here we add the symmetric distributional assumption and distinguish the latent variances of  $\mu_i$  and  $u_{it}$  ( $\tilde{\sigma}_\mu^2$

and  $\tilde{\sigma}_i^2$ ) from their actual variances ( $\sigma_\mu^2$  and  $\sigma_i^2$ ). Since both random variables are bounded, their actual variances should be less than or equal to their latent variances:  $\sigma_\mu^2 \leq \tilde{\sigma}_\mu^2$  and  $\sigma_i^2 \leq \tilde{\sigma}_i^2$ . The justification of Assumption B.1 was discussed in the previous section. In many standard public good games it is not hard to see that a few subjects contribute all tokens to the public accounts for all rounds. However the fraction of  $n_3$  is very tiny so that usually the cross sectional variance becomes decreasing over rounds. With a fixed  $N$ , however it is impossible to test whether or not such a small value of  $n_3$  or  $n_2$  is statistically negligible. Hence we take Assumption B.2.

### 3.1 Pre-Testing of Multiple Equilibria

The true data generating process of the cross sectional variance under both the null and alternative can be written as

$$H_{Nt} = \lambda_1 + n_2\lambda - 2\lambda_1\rho^{t-1} + (\lambda_1 + \lambda_2)\rho^{2t-2} + \varepsilon_t \quad (13)$$

where  $\varepsilon_t = H_{Nt} - \mathbb{E}(H_{Nt})$ ,  $\lambda_1 = n_1\tau^2 + n_2\phi^2 + n_3(1-\tau)^2$ ,  $\lambda_2 = (1-n_2)\lambda$ ,  $\lambda = \sigma_\mu^2 + \sigma^2$ ,  $\tau = n_3 + n_2\mu$ , and  $\phi = \mu - \tau$ . Note that  $\lambda = \mathbb{E}N^{-1} \sum_{i=1}^N (y_{i1} - \mathbb{E}(y_{i1}))^2$ . Even when  $\rho$  is known, it is impossible to identify  $\lambda_1$  and  $\lambda_2$  by running (13) as  $N, T \rightarrow \infty$  jointly. Let  $X_t = [1, \rho^{t-1}, \rho^{2t-2}]'$  and  $X = [X_1, \dots, X_T]'$ . Then it is easy to see that as  $t \rightarrow \infty$ ,  $H_{Nt} \rightarrow \lambda_1 + n_2\lambda$  and the matrix  $X'X$  becomes singular.

As we discussed in the previous section, under a single equilibrium  $\mathbb{E}(H_{Nt})$  does not change or decrease over rounds. Meanwhile under multiple equilibria,  $\mathbb{E}(H_{Nt})$  has a U-shape function over rounds and its minimum value is obtained at the following round  $t^*$ .

$$t^* = 1 + \frac{\ln \lambda_1 - \ln(\lambda_1 + \lambda_2)}{\ln \rho}. \quad (14)$$

Hence as  $\rho$  approaches to unity,  $t^*$  is getting larger. In other words,  $\mathbb{E}(H_{Nt})$  is decreasing over  $t$  until  $t \leq t^*$ . Of course, as  $\rho \rightarrow 1$ , group  $G_1$  and  $G_3$  subjects behave more like confused subjects so that it is getting harder to identify the existence of multiple equilibria. To exclude this case, we take the following assumption.

**Assumption C: The Condition for Asymptotic Divergence** *Let  $\tau^* = \min[\tau, 1 - \tau]$ . Then under multiple equilibria,  $\tau_*^2(1 + 2\rho) > \sigma_\mu^2 + \sigma^2$  where  $\sigma^2 = \lim_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N \sigma_i^2$ .*

Under Assumption C, the identification of multiple equilibria can be examined by utilizing the cross sectional variance. That is, if the cross sectional variance is not increasing, then this evidence becomes the existence of single equilibrium. Note that the condition of  $\mathbb{E}[H_{N1} - H_{NT}] < 0$  becomes a sufficient condition for Assumption C. Under Assumption A, B, and C, the null hypothesis in (10) becomes valid to test for the existence of multiple equilibria. More formally, we state the following Theorem 1 and 2.

**Theorem 1: (Asymptotic Properties under the Null of a Single Equilibrium)** *Under the Assumption A and B, as  $N, T \rightarrow \infty$ ,*

(i) *the probability limit of  $\hat{\beta}_T$  is given by*

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_T = \begin{cases} \frac{-6\lambda}{(1-\rho^2)(T^2-1)} + \Delta_\beta & \text{when } 0 < \rho < 1 \text{ and } i \in G_1, G_3 \\ 0 & \text{when } \rho = 1 \text{ or } i \in G_2 \end{cases}, \quad (15)$$

where  $\Delta_\beta$  is  $O(T^{-3})$  and its exact expression is given in (32) in Appendix A.

(ii) *the probability limit of  $t$ -statistic is converging to a negative constant  $\nu$  if  $0 < \rho < 1$ ,*

$$\text{plim}_{N, T \rightarrow \infty} t_{\hat{\beta}} = \text{plim}_{N, T \rightarrow \infty} \left( \hat{\beta}_T / \sqrt{V(\hat{\beta}_T)} \right) < -1.732 \text{ when } 0 < \rho < 1, \quad (16)$$

but when  $\rho = 1$ ,  $t_{\hat{\beta}}$  converges in the following limiting distribution.

$$t_{\hat{\beta}} \rightarrow^d \mathcal{N}(0, 1) \text{ when } \rho = 1. \quad (17)$$

where  $V(\hat{\beta}_T) = T^{-1} \sum_{t=1}^T \hat{\epsilon}_t^2 / \sum_{t=1}^T \hat{t}^2$  and  $\hat{\epsilon}_t$  is the regression residual from the trend regression in (9).

See Appendix A for a detailed proof of Theorem 1. As  $T \rightarrow \infty$ ,  $E\hat{\beta}_T$  converges to zero with  $|\rho| < 1$ , but its t-ratio converges to a negative constant. Meanwhile when  $\rho = 1$  (all subjects are confused), then the t-ratio follows the standard normal distribution. Hence the null hypothesis can be evaluated by use the conventional critical value for the one-sided test. Next we consider the asymptotic properties for the case of multiple equilibria.

**Theorem 2: (Asymptotic Properties under the Alternative of Multiple Equilibria)**

*Under the Assumption A, B and C, as  $N, T \rightarrow \infty$  jointly*

(i) *the probability limit of  $\hat{\beta}_T$  is given by*

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_T = 6 \frac{\lambda_1 (1 + 2\rho) - \lambda_2}{(1-\rho^2)(T^2-1)} + \Delta_\beta \rightarrow 0 \text{ as } T \rightarrow \infty, \quad (18)$$

where  $\Delta_\beta$  is the same as before,

(ii) *the probability limit of the  $t$ -statistic is given by*

$$\text{plim}_{N, T \rightarrow \infty} t_{\hat{\beta}} = \text{plim}_{N, T \rightarrow \infty} \left( \hat{\beta}_T / \sqrt{V(\hat{\beta}_T)} \right) > 1.732 \text{ when } 0 < \rho < 1, \quad (19)$$

where  $V(\hat{\beta}_T)$  is defined in Theorem 1.

See Appendix A also for a detailed proof of Theorem 2. Since the t-ratio is converging to a positive constant which is larger than 1.732, the 5% level test is suitable. That is, if the t-ratio is greater

than 1.65, the null hypothesis of a single equilibrium can be rejected at the 5% level. When the null is not rejected, the asymptotic average effects can be estimated by running Sul (2013)'s trend regression if the cross sectional averages are either increasing or decreasing over rounds.<sup>8</sup> However if the cross sectional averages are neither increasing nor decreasing, then all subjects may be in the confused group. In this case, the sample average becomes a consistent estimator for the asymptotic average effect.

Now we provide one important remark under the local to unity.

**Remark 1: (Asymptotic Theory Under the Local to Unity)** Let  $\rho_T = 1 - c/T$ . Then as  $N, T \rightarrow \infty$ , the t-ratio has the following limits

$$\text{plim}_{N,T \rightarrow \infty} t_{\hat{\beta}} = \begin{cases} -\infty & \text{if } n_1 = 1 \text{ or } n_3 = 1 \\ +\infty & \text{under multiple equilibria} \end{cases} .$$

Of course, when  $n_2 = 1$ ,  $\rho = 1$  by definition. Hence the limiting distribution of  $t_{\hat{\beta}}$  does not change. See Appendix A for a detailed proof.

### 3.2 Clustering Method

When the null of a single equilibrium is rejected, the asymptotic average effects should be estimated by running (12). We will show later the asymptotic variance of  $\hat{\tau}$  includes the value of  $\sigma_{\mu}^2$  and  $n_2$ . Hence we need to estimate  $n_2$  and  $\sigma_{\mu}^2$ . To do this, we need to identify confused subjects.

There are many clustering algorithms in the growth literature which have been developed to classify countries into a few groups. Among them, Phillips and Sul (2007)'s algorithm is known to be very effective. Their clustering method hinges on the selection of core members in one of the subgroups. They suggest using the last observation ordering as the initial guideline to select a few core members. After that, they add each individual one-at-a-time to judge whether or not this individual can be added into the same group. However, we can't apply their method directly in experimental games. In our setting, the core members are easily identified since some of individuals always choose either Pareto or Nash outcomes (all ones or zeros). Since their outcomes do not change over time, selecting such individuals as core members and comparing additional individual with them is statistically equivalent to run the following simple trend regression.

$$y_{it} = a_{k,i} + b_{k,i}t + w_{it}, \text{ for } t = 1, \dots, T, \tag{20}$$

---

<sup>8</sup>Suppose that the cross sectional averages to the contributions to the public accounts from a pair of games are increasing over rounds. Then the asymptotic average effects are unity so that the asymptotic treatment effect become zero in the long run. However the two games can be compared in the short run robustly. The overall effects for games  $A$  and  $B$  become  $T - (1 - \mu_A)/(1 - \rho_A)$  and  $T - (1 - \mu_B)/(1 - \rho_B)$ . Hence in the short run, the treatment effect can be written as  $-(1 - \mu_A)/(1 - \rho_A) + (1 - \mu_B)/(1 - \rho_B)$ .

It is important to note that the trend regression in (20) is misspecified if  $i \in G_1$  or  $G_3$ . Only when  $i \in G_2$  is the trend regression well defined. In fact, the decay function,  $\rho^{t-1}$ , converges to zero as  $t$  increases. Hence the well specified regression, replacing  $t$  in (20) by  $\rho^{t-1}$ , becomes useless since the point estimates do not have any limiting distribution.

Define  $c_{k,i} = \mu_i - \alpha_k$  and  $t_{\hat{b}_{k,i}} = \hat{b}_{k,i} / (T^{-1} \sum_{t=1}^T \hat{w}_{it}^2 / \sum_{t=1}^T \tilde{t}^2)^{1/2}$  where  $\hat{b}_{k,i}$  is the LS estimator in (20) and  $\hat{w}_{it}$  is the LS residual. Then formally we have

**Theorem 3 (Limit Theory for the Clustering Method)** *Under Assumptions A through C, as  $T \rightarrow \infty$ ,*

(i) *the expected value of  $\hat{b}_{k,i}$  becomes*

$$\mathbb{E}\hat{b}_{k,i} = \begin{cases} \mp \left( \frac{6c_{k,i}}{(1-\rho)T^2} \right) + O(T^{-3}) & \text{for } k = 1, 3, \\ 0 & \text{for } k = 2, \end{cases}, \quad (21)$$

(ii) *its limiting distribution is given by*

$$\begin{aligned} T^2 (\hat{b}_{k,i} - \mathbb{E}\hat{b}_{k,i}) &\rightarrow^d \mathcal{N} \left( 0, \frac{36\sigma_i^2}{1-\rho^2} \right) & \text{for } k = 1, 3, \\ T^{3/2} (\hat{b}_{2,i} - 0) &\rightarrow^d \mathcal{N} (0, 12\sigma_i^2) & \text{for } k = 2, \end{aligned} \quad (22)$$

(iii) *the limiting distribution of its  $t$ -statistic becomes*

$$t_{\hat{b}_{k,i}} = \frac{\hat{b}_{k,i}}{\sqrt{V(\hat{b}_{k,i})}} \rightarrow^d \begin{cases} \mathcal{N} \left( \mp \sqrt{\left( \frac{3}{1+\zeta_i} \right) \left( \frac{1+\rho}{1-\rho} \right)}, \frac{3\zeta_i}{1+\zeta_i} \right) & \text{for } k = 1, 3, \\ \mathcal{N}(0, 1) & \text{for } k = 2 \end{cases}, \quad (23)$$

where  $\zeta_i = \sigma_i^2 / c_{k,i}^2$  and  $0 \leq \zeta_i \leq 1$ .

See Appendix B for the detailed proof of Theorem 3. Here we provide intuitive explanations. For  $i \in G_2$  (confused subject), the trend regression in (20) is well specified so that  $\hat{b}_{2,i}$  becomes unbiased and consistent. Also the limiting distribution of its  $t$ -ratio becomes the standard normal distribution. However for  $i \in G_1$  or  $G_3$ , the trend regression becomes misspecified. The expected value of  $\hat{b}_{k,i}$  becomes  $O(T^{-2})$  since  $\sum_{t=1}^T \rho^{t-1}$  term becomes  $O(1)$ . This implies that  $T^2 \hat{b}_{k,i}$  for  $k \neq 2$  becomes  $O_p(1)$  and the convergence rate becomes also  $O(1)$  in (22). Also the conventional estimator for the variance of  $\hat{b}_{k,i}$  becomes also inconsistent. Hence the limiting distribution of its  $t$ -ratio becomes a non-standard normal distribution. As  $T \rightarrow \infty$ ,  $t_{\hat{b}_{k,i}}$  does not diverge to infinity but to the normal distribution in (23) which has two nuisance parameters:  $\rho$  and  $\zeta_i = \sigma_i^2 / c_{k,i}^2$ . Note that  $e_{it} = u_{it} \rho^{t-1}$  and  $u_{it}$  is bounded between  $-\mu_i$  and  $\mu_i$  for  $i \in G_1$  and between  $-(1-\mu_i)$  and  $(1-\mu_i)$  for  $i \in G_3$ . Hence  $\sigma_i^2$  should be smaller than  $\mu_i^2$  for  $i \in G_1$  and  $(1-\mu_i)^2$  for  $i \in G_3$ .



since the maximum value of  $u_{it}^2$  is either  $\mu_i^2$  for  $i \in G_1$  or  $(1 - \mu_i)^2$  for  $i \in G_3$ . This implies that  $0 \leq \zeta \leq 1$ . The relative variance ratio  $\zeta$  becomes zero when  $u_{it} = 0$  for all  $i$ , but becomes one when  $u_{it}$  has only two values of  $c_{k,i}$  and  $-c_{k,i}$ .

Nonetheless, the most important thing is whether or not the trend regression can be used for identifying each subject's type. We can use the confused case as the numeraire and consider the validity of the clustering rule in (11). Suppose that  $cv = 1.28$  is used for the boundary value. Then the t-ratio for  $i \in G_1$  should be smaller than  $-1.28$  but the t-ratio for  $i \in G_3$  should be larger than  $1.28$ . The critical value of  $t_{\hat{b}_{k,i}}$  for  $k = 1$  at the  $(1 - \alpha)\%$  level,  $x_{1-\alpha}$ , is given by

$$x_{1-\alpha} = \varkappa \sqrt{\frac{3\zeta_i}{1 + \zeta_i}} - \left( \frac{3}{(1 + \zeta_i)} \frac{1 + \rho}{1 - \rho} \right)^{1/2}, \quad (24)$$

meanwhile the critical value of  $t_{\hat{b}_{k,i}}$  for  $k = 3$  at the  $\alpha\%$  level,  $x_\alpha$ , is

$$x_\alpha = -\varkappa \sqrt{\frac{3\zeta_i}{1 + \zeta_i}} + \left( \frac{3}{(1 + \zeta_i)} \frac{1 + \rho}{1 - \rho} \right)^{1/2}, \quad (25)$$

where  $\varkappa$  is the critical value for the  $\alpha\%$  level. It is easy to see that  $x_{1-\alpha}$  becomes maximized when  $\zeta_i = 1$ . To get the maximum value, we evaluate  $x_\alpha$  and  $x_{1-\alpha}$  with  $\rho = 0.8$  and  $\zeta_i = 1$ . Then the maximum value of  $x_{1-\alpha}$  is obtained when  $\rho \uparrow 0.8$  and it becomes

$$\begin{aligned} \max x_{1-\alpha} &< \varkappa \sqrt{\frac{3}{2}} - \sqrt{\frac{27}{2}} < -1.65 \text{ if } \varkappa = 1.65, \\ \min x_\alpha &> -\varkappa \sqrt{\frac{3}{2}} + \sqrt{\frac{27}{2}} > 1.65 \text{ if } \varkappa = 1.65. \end{aligned}$$

Hence the maximum false inclusion rate to the confused group can be obtained when  $\zeta_i = 1$  for all  $i$ , and this maximum rate reaches to the 5% for each  $G_1, G_3$  subgroups. Meanwhile the false exclusion rate for the confused group becomes 20% total with  $cv = 1.28$ . This leads to under-estimation of  $n_2$ . Hence we need the selection rule for the critical value which attenuate the under-estimation

problem. See below remark 3.

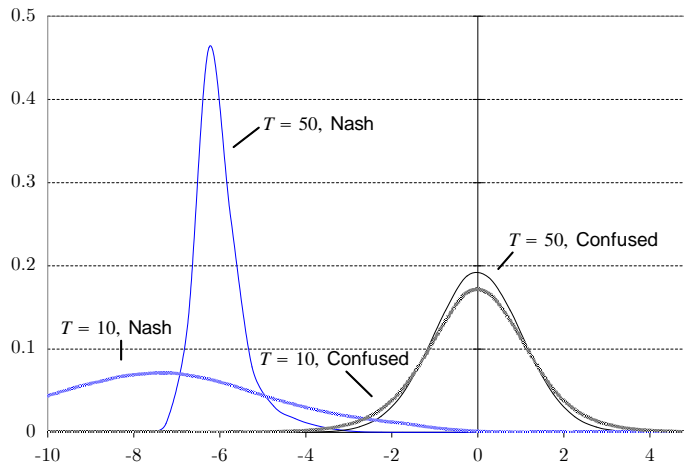


Figure 2: Empirical Distributions of t-ratios for  $i \in G_1$  and  $i \in G_2$

Figure 2 displays two empirical distributions of the t-ratios for  $i \in G_1$  and  $i \in G_2$ . We generate  $\mu_i$  and  $u_{it}$  from the bounded normal distributions of  $iidN(0.5, 0.05)_0^1$  and  $iidN(0, 0.01)_{-\mu_i}^{\mu_i}$ , respectively and set  $\rho = 0.8$  for  $G_1$  and  $G_3$ . Evidently with  $T = 10$ , both t-ratios for  $i \in G_1$  and  $i \in G_2$  have fat tails. The clustering rule in (11) seems to work well even with  $T = 10$ . When  $T = 50$ , the empirical distribution of the t-ratio for  $i \in G_1$  becomes much shrunk but that for  $i \in G_2$  becomes a standard normal so that the two empirical distributions can be distinguished well by using the clustering rule in (11).

Next, we provide some important remarks.

**Remark 2: Asymptotic Properties of the Clustering Rule under the Local to Unity**

Let  $\rho_T = 1 - c/T$ . Under Assumption A,B and C, the t-ratio for  $i \notin G_2$  can be expressed by

$$t_{\hat{b}_{k,i}} = \mp \left( \frac{3c_{k,i}^2 (2T - c)}{(c_{k,i}^2 + \sigma_i^2) c} \right)^{1/2} = \begin{cases} -\infty & \text{when } i \in G_1 \\ +\infty & \text{when } i \in G_3 \end{cases} \text{ as } t \rightarrow \infty.$$

Since  $\rho_T$  is getting persistent as  $T$  increases, as we discussed before, the minimum  $t^*$  is getting larger as  $T$  increases. However, the increase in the sample size helps more to cluster subjects correctly since its t-ratio approaches either negative or positive infinity.

**Remark 3: Selection of Critical Values** The clustering outcome is hinging on the assigned critical value. With the fixed significance level, the fraction of the confused members,  $n_2$ , becomes always under-estimated even when  $T \rightarrow \infty$ . To avoid such inconsistency problem at least asymptotically, we let the critical value be dependent on the sample size. More precisely, following Han,

Phillips and Sul (2012), we can set the significance level  $\alpha_T$  be a function of the critical value  $d_T$ .

$$\alpha_T = 2[1 - \Phi(d_T)],$$

where  $\Phi(\cdot)$  is the standard normal cdf. The following rule was found to work well in our simulations:

$$\alpha_T = \exp \left[ \ln(p) \sqrt{T/10} \right] \quad (26)$$

When  $p = 0.20$ , this choice of  $\alpha_T$  delivers a nominal size of 20%, 10%, and 6% for  $T = 10, 20$  and  $30$ , respectively. The corresponding critical values becomes 1.28, 1.65, and 1.87 respectively.

**Remark 4: Estimation of  $n_2$  and  $\sigma_\mu^2$**  Under the sample dependent clustering rule in (26), the under-estimation probability of  $n_2$  becomes reduced as  $T$  increases. With any given critical value, there is always a non-zero probability that non-confused subjects are assigned to  $G_2$ . However such probability may be ignorable due to Assumption A.2. Note that  $u_{it} \sim idB(0, \hat{\sigma}_i^2)^{-\mu_i^*}$ , so that the probability of  $T^{-1} \sum_{t=1}^T u_{it}^2$  becomes  $(\mu_i^*)^2$  goes to zero as  $T \rightarrow \infty$ . For some  $t$ , as long as there is some probability that  $-\mu_i^* < u_{it} < \mu_i^*$ ,  $\zeta_i$  cannot be equal to one. In fact, as  $T$  increases, the maximum value of  $\zeta_i$  decreases due to the boundary condition in Assumption A.2. Let  $\zeta_\alpha$  be the value for the upper  $\alpha\%$  of the  $\zeta_i$  distribution, then as  $T \rightarrow \infty$ ,  $\zeta_\alpha$  decreases as well. This implies that with given  $\rho$ ,  $x_{1-\alpha}$  in (24) increases in absolute value. Hence it may be reasonable to assume that as  $N, T \rightarrow \infty$

$$\Pr(\text{subject } i \text{ is assigned to } G_2) \rightarrow n_2. \quad (27)$$

Under this additional assumption, define an indicator function for each subject

$$Z_i = \begin{cases} 1 & \text{if } i \text{ is assigned to } G_2 \\ 0 & \text{otherwise} \end{cases}.$$

Then  $Z_i$  follows a Bernoulli distribution with mean  $n_2$  and variance  $n_2(1 - n_2)$ . That is,

$$\hat{n}_2 = \frac{\hat{N}_2}{N} = \frac{\sum_{i=1}^N Z_i}{N} \rightarrow^d \mathcal{N} \left( n_2, \frac{n_2(1 - n_2)}{N} \right).$$

Meanwhile as long as the false inclusion rate to the confused group becomes zero, the variance of  $\sigma_\mu^2$  can be consistently estimated. Let  $\hat{G}_2$  be the estimated subset,  $G_2$ , from the clustering. As  $T \rightarrow \infty$ ,  $\Pr(\hat{G}_2 \subseteq G_2) \rightarrow 1$  under (27). Then  $\hat{\mu}_i$  can be obtained by taking time series average of  $y_{it}$  so that  $\hat{\sigma}_\mu^2$  can be estimated from  $\hat{\mu}_i$ . It is easy to show that  $\hat{\sigma}_\mu^2 = \sigma_\mu^2 + O_p(N^{-1/2})$ .

**Remark 5 (Estimation of  $\rho$  and  $\mu$ )** We can estimate  $\mu$  consistently by using the first observation of  $y_{it}$ . That is,  $\hat{\mu} = N^{-1} \sum_{i=1}^N y_{i1} = \mu + O_p(N^{-1/2})$ . Next use a higher critical value in (20) to

identify core long run Nash and Pareto groups. Note that a subject  $i$  must be in  $G_1$  or  $G_3$  group if  $y_{it} = 0$  or 1 for all  $t$ , respectively. And then run the following pooled OLS to estimate  $\log \rho$ .

$$\begin{aligned} \ln \left( \frac{1}{N_1^c} \sum_{i \in G_1^c} y_{it} \right) - \ln \hat{\mu} &= \gamma (t-1) + v_{1t}, \\ \ln \left( 1 - \frac{1}{N_3^c} \sum_{i \in G_3^c} y_{it} \right) - \ln (1 - \hat{\mu}) &= \gamma (t-1) + v_{3t}, \end{aligned}$$

where  $\gamma = \log \rho$ ,  $G_1^c$  and  $G_3^c$  are core long run Nash and Pareto groups, respectively.  $N_1^c$  and  $N_3^c$  are the total numbers of the core long run Nash and Pareto groups, respectively.

As  $N \rightarrow \infty$  with a fixed  $T$ , the limiting distribution of  $\hat{\gamma}$  is approximated by

$$\sqrt{N} (\hat{\gamma} - \gamma) \rightarrow^d \mathcal{N} (0, w_\gamma^2),$$

where  $w_\gamma^2 = 9 (\sigma_\mu^2 + n_1 (\sigma^2 - \sigma_\mu^2)) / (16 \mu^2 n_1 T^2) + 9 (\sigma_\mu^2 + n_3 (\sigma^2 - \sigma_\mu^2)) / (16 (1 - \mu)^2 n_3 T^2)$ . The estimate of  $\hat{\rho}$  becomes  $\exp(\hat{\gamma})$ . See Appendix B for a detailed derivation.

### 3.3 Estimation of Asymptotic Average Effects

Here we consider a rather simple but efficient method for estimating  $\tau$  in (5). First, we consider the infeasible case where the decay rate,  $\rho$ , and the fraction of confused subjects are assumed to be known. When  $\rho$  is known, the treatment regression in (4) can be run directly. Here we rewrite (4) as

$$y_{N,t} = \tau_N + \phi_N \rho^{t-1} + e_{Nt} \quad (28)$$

where  $\phi_N = \mu_N - \tau_N$ . The limit theory in this case can be formally stated as

**Theorem 4 (Limit Theory under Infeasible Estimation)** *Under Assumption A and B, as  $N, T \rightarrow \infty$  jointly, the limiting distribution of  $\hat{\tau}_N$  is given by*

$$\begin{bmatrix} \sqrt{N} (\hat{\tau}_N - \tau) \\ \sqrt{N} (\hat{\phi}_N - \phi) \end{bmatrix} \rightarrow^d \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} n_2 \sigma_\mu^2 & 0 \\ 0 & \sigma_\alpha^2 \end{bmatrix} \right). \quad (29)$$

where  $\sigma_\alpha^2 = (1 - n_2)^2 \sigma_\mu^2 + (1 - \rho^2) (1 + n_2 \rho^2) (1 + \rho^2)^{-1} \sigma^2$ .

See Appendix C for a detailed proof. Here we briefly discuss why the convergence rate is  $\sqrt{N}$  rather than  $\sqrt{NT}$ . Note that the parameters of interest are  $\tau$  and  $\phi$ , but the time series regression in (28) estimates  $\tau_N$  and  $\phi_N$  with given  $N$ . Since  $\tau_N = \tau + n_2 N_2^{-1} \sum_{i \in G_2}^{N_2} (\mu_i - \mu)$  and  $\phi_N = N^{-1} \sum_{i=1}^N (\mu_i - \mu) - N^{-1} \sum_{i \in G_2}^{N_2} (\mu_i - \mu) + \mu - \tau$ , the convergence rate is mainly determined by  $\sqrt{N}$ . Next, the decay rate function,  $\rho^{t-1}$ , is  $o(1)$  and  $\sum_{t=1}^T \rho^{t-1}$  is  $O(1)$ . This means that  $\hat{\phi}_N - \phi$  is  $O_p(1)$  as  $T \rightarrow \infty$  for a fixed  $N$ .

From Remark 4 and 5,  $\hat{\rho} - \rho = O_p(N^{-1/2})$ . Hence we can run  $y_{N,t}$  on a constant and  $\hat{\rho}^{t-1}$ . That is,

$$y_{N,t} = \tau_N + \phi_N \hat{\rho}^{t-1} + e_{Nt}^+, \quad (30)$$

where  $e_{Nt}^+ = e_{Nt} + \phi_N (\rho^{t-1} - \hat{\rho}^{t-1})$ . Since  $\hat{\rho} - \rho = O_p(N^{-1/2}T^{-1})$ , the limiting distributions of  $\hat{\tau}_N$  and  $\hat{\phi}_N$  in (30) follow (29). That is,

**Theorem 5 (Limit Theory under Feasible Estimation)** *Under Assumption A and B,*

(i) *as  $N, T \rightarrow \infty$  jointly, the limiting distributions of  $\hat{\tau}_N$  and  $\hat{\phi}_N$  in (30) becomes*

$$\begin{bmatrix} \sqrt{N}(\hat{\tau}_N - \tau) \\ \sqrt{N}(\hat{\phi}_N - \phi) \end{bmatrix} \rightarrow^d \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} n_2 \sigma_\mu^2 & 0 \\ 0 & \sigma_\alpha^2 \end{bmatrix} \right).$$

(ii) *If  $N \rightarrow \infty$  with a fixed  $T$ , the limiting distribution of  $\tau_N$  becomes*

$$\sqrt{N}(\hat{\tau}_N - \tau) \rightarrow^d \mathcal{N}(0, n_2 \sigma_\mu^2 + n_2 \sigma^2/T).$$

See Appendix C for a detailed proof of Theorem 5.

## 4 Return to Empirical Examples and Monte Carlo Simulation

### 4.1 Return to Empirical Examples

In this subsection, we apply our pre-testing of no-divergence and the AATE summation method to the two empirical examples used in Section 2. First we examine whether or not there are multiple equilibria by running the trend regression with the cross sectional variance in (9). The point estimates for  $\hat{\beta}$  are 0.007 and 0.006 and  $t_{\hat{\beta}}$  are 4.653 and 4.287 for  $G = 10$  and  $G = 100$ , respectively. Hence the null hypothesis of a single equilibrium is rejected even at the 1% level. This result confirms what Panel B in Figure 1 showed before, that there is clear divergence in the two games.

Table 2: Estimation of Asymptotic Average Effects

	$\hat{\mu}$	$\hat{\rho}$	$\hat{\tau}$	$\hat{\sigma}_\tau^2 \times 10^{-4}$	$\hat{\sigma}_\mu^2$	$\hat{\sigma}^2 \times 10^{-1}$	$\hat{n}_1$	$\hat{n}_2$	$\hat{n}_3$
$N = 100, G = 10$	0.47	0.76	0.43	2.50	0.035	0.010	0.30	0.41	0.29
$N = 300, G = 100$	0.43	0.89	0.30	1.08	0.052	0.003	0.40	0.43	0.16

Based on these results, we run the clustering regression in (20) and sort each individual into one of the three groups using the critical value of 1.28. The results are showed in Table 2. From

the confused group, we calculate  $\hat{\sigma}_\mu^2$  and  $\hat{n}_2$ . See Remark 4 and 5 for the estimation methods. The estimated mean of initial outcome  $\hat{\mu}$ , are near to 0.5 for the two games. The decay rates are around 0.8. The estimates of the variances of  $e_{it}$ ,  $\hat{\sigma}^2$  in the two games are very small after controlling out the decay pattern. The fractions of Nash and Pareto (or decreasing and increasing) groups are all different over two experiments. As the group size increases, more subjects become free riders and fewer subjects become pure altruists even in the long run. The fractions of confused group are around the same level.

Next, we run the feasible treatment regression in (30) and get the point estimates of  $\tau$ . The AAE in the game with group size 10 is higher than that in the game with group size 100. The asymptotic average treatment effect which is the difference between the two AAE is 0.13 ( $= 0.43 - 0.30$ ), and the t-statistic is 6.87 ( $= 0.13 / \sqrt{(0.25 + 1.08) 10^{-4}}$ ) which is large enough for us to reject the null of no difference even at the 1% level. The asymptotic average outcome of the game with group size 10 is significantly higher. Comparing these results with those in Table 1 where we couldn't find any meaningful results due to invalid statistical inference, our method provides an accurate and robust measure of the treatment effect.

## 4.2 Monte Carlo Study

This section reports finite sample performances of the suggested tests and estimators. The data generating process is given in (2). We generate  $\mu_i \sim iidN(0.5, \tilde{\sigma}_\mu^2)_0^1$ , and  $u_{it} \sim iidN(0, \tilde{\sigma}_i^2)_{-\mu_i^*}^{\mu_i^*}$ . Since  $\mu_i$  is bounded, the corresponding true variances for  $\mu_i$ ,  $\sigma_\mu^2$ , become approximately 0.03 and 0.048 for  $\tilde{\sigma}_\mu^2 \in [0.03, 0.05]$ , respectively. For all cases, we set  $\tilde{\sigma}_\mu^2 \in [0.03, 0.05]$  but we change the values of  $\rho$  and  $\tilde{\sigma}_i^2$  depending on cases. For all simulations, we consider the total number of repetition of 2,000,  $T \in [10, 20, 30]$  and  $N \in [25, 50, 100, 200]$ . Various cases were considered, but to save space, the following few cases – which highlight the findings – are reported. The remaining simulation results are available at the authors' webpage.<sup>9</sup>

**Case 1 (Testing Multiple Equilibria)** For this case, we set  $\tilde{\sigma}_i^2 = \tilde{\sigma}^2 \in [0.005, 0.01]$  to calculate  $t^*$  in (14) and  $\rho \in [0.97, 0.98, 0.99, 1.00]$  under the null of single equilibrium. Note that  $n_1 = 1$  with  $\rho < 1$  but  $n_2 = 1$  for  $\rho = 1$  under the null. Meanwhile under the alternative, we consider two subcases:  $n_3 = 0.1$ ,  $n_2 = 0.2$  and  $n_3 = 0.2$ ,  $n_2 = 0.1$ . In the first and second subcases,  $\tau$  becomes 0.2 and 0.25, respectively. Note that under the alternative, the expected cross sectional variance,  $E(H_{Nt})$ , becomes minimized at  $t^*$  as we discussed in (14). Moreover,  $t^*$  increases as  $\sigma_\mu^2$ ,  $\sigma^2$ , and  $\rho$  increase but  $t^*$  decreases as  $\tau$  increases. We set  $\rho \in [0.8, 0.9]$ . For all cases, Assumption C holds.

Table 3 reports the size of the test – the rejection rate when the null is true. The nominal size is

---

<sup>9</sup> All results are available in MS excel format at [http://www.utdallas.edu/~d.sul/papers/Exp2\\_MC.xlsx](http://www.utdallas.edu/~d.sul/papers/Exp2_MC.xlsx)

5%. When all subjects are free riders ( $n_1 = 1$ ), the null hypothesis is rarely rejected even with very persistent  $\rho$ . Only when all subjects are confused ones ( $\rho = 1$ ), the rejection rate reaches at the nominal size of the 5% rate. This implies that the designed test does not reject the null hypothesis at all if all subjects are free riders and their learning speed is moderately fast. The designed test is rejected more or less at the nominal rate only when all subjects are confused. When  $T = 10$ , there seems to be a mild size distortion regardless values of  $\tilde{\sigma}_\mu^2$  and  $\tilde{\sigma}^2$ . However as  $T$  increases to 20, such size distortion seems to go away. This evidence supports that our theoretical claim in Theorem 1 works fairly well even with a finite sample.

Table 4 shows the power of the test. We consider two alternatives of  $\tau = 0.2$  and 0.25. Note that  $\tau$  can be 0.2 when  $n_3 = 0.2$  but  $n_2 = 0$ . In this case, the finite sample performance is much better than the reported case in Table 4. We report the ratio of  $t^*/T$  which depends on  $\tau$ ,  $\tilde{\sigma}_\mu^2$ ,  $\tilde{\sigma}^2$ , and  $\rho$ . When  $\rho$  is moderately persistent ( $\rho = 0.8$ ), the designed test achieves almost perfect power regardless values of  $\tilde{\sigma}_\mu^2$  and  $\tilde{\sigma}^2$ . The power loss happens when  $\rho$  is very persistent and  $T$  is small. When  $\tau = 0.2$ ,  $t^*/T$  becomes around 0.5 with  $\rho = 0.9$ ,  $\tilde{\sigma}_\mu^2 = 0.05$  and  $T = 10$ . We find that in this case  $H_{Nt}$  decreases rapidly until around  $t = 5$  and then slowly increases. Hence the power of the test becomes lower as  $N$  increases. Except for this case, the power of the test usually increases as  $N$  increases. That is, when either  $\tau$  or  $T$  increases, such abnormal behavior disappears quickly. In our empirical examples,  $\tau$  is much higher than 0.2, hence the power of the test must be very high. In fact, we set  $N = 100, 300$  but  $T = 10$ , and mimic our empirical examples via simulation where we set  $\tilde{\sigma}_i^2 \sim iidU [0, \tilde{\sigma}_{\max}^2]$ ,  $\tau \in [0.3, 0.4]$  and consider the cases of  $\tilde{\sigma}_{\max}^2 \in [0.03, 0.05]$ . For all cases, we get almost perfect power of the test.

**Case 2 (Clustering Mechanism)** In contrast to case 1, we don't impose the homogeneity restriction on  $\tilde{\sigma}_i^2$ . It is because the clustering performance is dependent on the maximum value of  $\tilde{\sigma}_i^2$ . We set  $\tilde{\sigma}_i^2 \sim iidU [0, \tilde{\sigma}_{\max}^2]$  for  $\tilde{\sigma}_{\max}^2 \in [0.03, 0.05]$ . We consider three values of  $p$  in (26):  $p = 0.2, 0.1$  and 0.05, meanwhile set  $n_1 = 0.4, n_2 = 0.4, n_3 = 0.2$ .  $N$  is set to be 100 since the estimator for  $n_2$  is independent of the size of  $N$ . Note that  $\sigma_\mu^2$  is calculated via simulation from generated  $\mu_i \sim iidN (0.5, \tilde{\sigma}_\mu^2)_{\mu_i}^{-\mu_i}$ . When  $\tilde{\sigma}_\mu^2$  is 0.03,  $\sigma_\mu^2$  becomes 0.03. That is,  $\mu_i$  is almost always generated between zero and one. However when  $\tilde{\sigma}_\mu^2$  becomes 0.05,  $\sigma_\mu^2$  becomes 0.048, which implies that  $\mu_i$  for some  $i$  becomes either zero and one.

Table 5 reports the simulation results: Averages of  $\hat{\sigma}_\mu^2$ ,  $\hat{n}_2$  and false inclusion rates with various  $p$  values. Overall the results support our theoretical claims in Theorem 3 and 4. When  $T = 10$ , the estimators of  $\hat{\sigma}_\mu^2$  are slightly biased upward and its bias is getting larger as  $p$  and  $\rho$  increase. Meanwhile  $\hat{n}_2$  becomes bias downward with  $\rho = 0.8$  or with  $p = 0.2$ . As  $p$  increases,  $\hat{n}_2$  becomes smaller. Hence the downward bias becomes more serious with  $\rho = 0.8$ . The estimated false inclusion rate is also increasing as  $\rho$ ,  $\sigma_\mu^2$  and  $\sigma_{\max}^2$  increase. However, such bias and the false inclusion rate

become much milder as  $T$  increases.

**Case 3 (Size of Test for Asymptotic Average Treatment Effects: Infeasible Estimation)**

Before we investigate the finite sample performance in the feasible estimation and testing, we consider the size properties of infeasible estimation first. We set  $\tilde{\sigma}_{\max}^2 = \tilde{\sigma}_{\mu}^2 \in [0.03, 0.05]$  and  $\rho \in [0.8, 0.9]$ . Two samples are generated but their underlying data generating processes are exactly the same. Subject's types and the decay rate are all known in this case. Table 6 reports the size of the test with the nominal size of 5%. Interestingly, there is somewhat mild size distortion with small  $N$  regardless the size of  $T$  and the size distortion is getting larger as  $\rho$  increases. Also as Theorem 1 states, the size of the test converges to the nominal 5% level as  $N$  increases.

**Case 4 (Asymptotic Average Treatment Effects: Feasible Estimation)**

In this case, we don't know any parameter value. Hence the clustering is required to estimate key parameters. We construct various simulation results but report here only a few highlight cases where  $\tilde{\sigma}_{\max}^2 \in [0.03, 0.05]$  and  $\tilde{\sigma}_{\mu}^2 = 0.05$ . Under the null of the same treatment effects, we set  $n_3 = 0.3$ ,  $n_2 = 0.3$  so that  $\tau = 0.45$ .

First we report the size of the test in Table 7. Note that the estimates of  $\tau_N$  are very accurate for all cases and show little bias. So we don't report them in Table 7. The nominal size is 5%. Here we also consider the impact of the clustering  $p$ -values on the size of the test so that we set  $p \in [0.2, 0.1, 0.05]$  also. The size of the test with the clustering  $p$ -value of 0.2 becomes very reasonable compared to Table 6. When the clustering  $p$ -value is less than 0.2, there are somewhat significant size distortion with a large  $N$ . The clustering  $p$ -value with 0.2 works reasonable well.

We consider various alternative but here report only  $\tau_A - \tau_B = 0.05$  case. The first and second samples are generated with  $(n_3 = 0.2, n_2 = 0.5)$  and  $(n_3 = 0.2, n_2 = 0.4)$ , respectively. Here we report the power of the test only with  $p = 0.2$ . The nominal size is 5%. The results are shown in Table 8. Overall, the power of the test decreases as  $\tilde{\sigma}_{\mu}^2$ ,  $\tilde{\sigma}_{\max}^2$  and  $\rho$  increase. However the power of the test increases definitely as  $N$  increases. When  $\rho$  is persistent, the power of the test is also dependent on  $T$  naturally. Note that when the asymptotic average treatment effect  $(\tau_A - \tau_B)$  is larger than 0.05, the power of the test becomes perfect for almost cases, which is not reported here.

## 5 Conclusion

When there exist multiple equilibria in repeated public good games, estimating the treatment effects is not straightforward. Usually conventional methods including standard z-score, Tobit regression, and Sul (2013)'s trend regression fail to estimate the treatment effects under multiple equilibria. We propose a new pre-test to find out the existence of multiple equilibria, and then show how



to estimate the treatment effects consistently and effectively. This paper showed that the newly suggested methods have good finite sample performances. We applied our methods to IWW (1994)'s data where two samples differ only in terms of group size. Contrast to what IWW (1994) claimed, we found that the average contribution to the public account with the group size of 10 ( $G = 10$ ) and MPCR of 0.75 is significantly higher than that with  $G = 100$  and MPCR=0.75.

The newly suggested method can be widely used in various settings of public good games. From a simple experiment where experimentalists want to examine the average treatment effects to a complicated experiment where experimentalists want to see the evolution of multiple equilibria, the proposed test can provide a simple but accurate empirical result. For an example, the most interesting research question in the public good experiments with punishment is the optimal punishment level for a Pareto equilibrium. By using the proposed tests, one can examine how the subgame perfect Nash equilibrium changes to multiple equilibria, and merges into a single Pareto equilibrium by changing the degree of punishment.

Table 3: Rejection Rate of the Null of Single Equilibrium under the Null  
 ( $n_1 = 1$  for  $\rho < 1$ ,  $n_2 = 1$  for  $\rho = 1$ ; Nominal Size=5%)

$\tilde{\sigma}_\mu^2$	$\tilde{\sigma}^2$	$N$	$\rho (T = 10)$				$\rho (T = 20)$			
			0.97	0.98	0.99	1	0.97	0.98	0.99	1
0.03	0.005	25	0	0	0.005	0.064	0	0	0	0.044
		50	0	0	0	0.055	0	0	0	0.036
		100	0	0	0	0.064	0	0	0	0.039
		200	0	0	0	0.072	0	0	0	0.045
0.03	0.01	25	0	0	0.007	0.068	0	0	0	0.038
		50	0	0	0	0.053	0	0	0	0.038
		100	0	0	0	0.065	0	0	0	0.036
		200	0	0	0	0.07	0	0	0	0.042
0.05	0.005	25	0	0	0.001	0.062	0	0	0	0.039
		50	0	0	0	0.059	0	0	0	0.035
		100	0	0	0	0.054	0	0	0	0.038
		200	0	0	0	0.068	0	0	0	0.043
0.05	0.01	25	0	0	0.004	0.066	0	0	0	0.037
		50	0	0	0	0.054	0	0	0	0.034
		100	0	0	0	0.054	0	0	0	0.039
		200	0	0	0	0.067	0	0	0	0.043

Table 4: Rejection Rate of the Null of Single Equilibrium under the Alternative  
 ( $n_3 = 0.1, n_2 = 0.2$  for  $\tau = 0.2$  &  $n_3 = 0.2, n_2 = 0.1$  for  $\tau = 0.25$ )

			$\rho = 0.8$					$\rho = 0.9$				
$T = 10$			$N$									
$\tau = 0.2$	$\tilde{\sigma}_\mu^2$	$\tilde{\sigma}^2$	$t^*/T$	25	50	100	200	$t^*/T$	25	50	100	200
	0.03	0.005	0.20	1	1	1	1	0.32	0.81	0.94	0.97	1
	0.03	0.01	0.21	1	1	1	1	0.34	0.66	0.83	0.91	0.98
	0.05	0.005	0.25	0.98	1	1	1	0.42	0.45	0.50	0.48	0.48
	0.05	0.01	0.26	0.96	1	1	1	0.44	0.34	0.35	0.30	0.26
$\tau = 0.25$	0.03	0.005	0.18	1	1	1	1	0.27	0.99	1	1	1
	0.03	0.01	0.19	1	1	1	1	0.29	0.96	0.99	1	1
	0.05	0.005	0.22	1	1	1	1	0.35	0.84	0.88	0.94	0.99
	0.05	0.01	0.23	1	1	1	1	0.37	0.75	0.80	0.87	0.95
$T = 20$			$N$									
$\tau = 0.2$	$\tilde{\sigma}_\mu^2$	$\tilde{\sigma}^2$	$t^*/T$	25	50	100	200	$t^*/T$	25	50	100	200
	0.03	0.005	0.10	1	1	1	1	0.16	1	1	1	1
	0.03	0.01	0.11	1	1	1	1	0.17	1	1	1	1
	0.05	0.005	0.13	1	1	1	1	0.21	1	1	1	1
	0.05	0.01	0.13	1	1	1	1	0.22	0.99	1	1	1
$\tau = 0.25$	0.03	0.005	0.09	1	1	1	1	0.13	1	1	1	1
	0.03	0.01	0.01	1	1	1	1	0.14	1	1	1	1
	0.05	0.005	0.11	1	1	1	1	0.18	1	1	1	1
	0.05	0.01	0.11	1	1	1	1	0.19	1	1	1	1

Table 5: Finite Sample Performance of the Clustering Results

$\hat{\sigma}_{\max}^2$	$\sigma_{\mu}^2$	$\hat{\sigma}_{\mu}^2$			$\hat{n}_2 (n_2 = 0.4)$			False inclusion rate		
		p=0.05	p=0.1	p=0.2	p=0.05	p=0.1	p=0.2	p=0.05	p=0.1	p=0.2
$T = 10, \rho = 0.8$										
0.03	0.030	0.037	0.035	0.034	0.375	0.347	0.302	0.016	0.011	0.006
0.05	0.030	0.041	0.038	0.036	0.389	0.357	0.308	0.030	0.021	0.012
0.03	0.048	0.056	0.054	0.052	0.384	0.354	0.307	0.025	0.018	0.011
0.05	0.048	0.059	0.056	0.054	0.399	0.365	0.314	0.040	0.028	0.018
$T = 10, \rho = 0.9$										
0.03	0.030	0.042	0.041	0.039	0.435	0.392	0.333	0.076	0.056	0.038
0.05	0.030	0.044	0.043	0.042	0.484	0.431	0.361	0.124	0.095	0.065
0.03	0.048	0.060	0.059	0.058	0.446	0.403	0.342	0.087	0.067	0.046
0.05	0.048	0.060	0.060	0.059	0.490	0.438	0.368	0.130	0.102	0.072
$T = 20, \rho = 0.8$										
0.03	0.030	0.031	0.031	0.030	0.391	0.378	0.349	0.002	0.001	0.001
0.05	0.030	0.031	0.030	0.030	0.392	0.378	0.349	0.004	0.002	0.001
0.03	0.048	0.049	0.048	0.048	0.396	0.381	0.351	0.007	0.005	0.003
0.05	0.048	0.049	0.048	0.047	0.398	0.382	0.352	0.009	0.006	0.004
$T = 20, \rho = 0.9$										
0.03	0.030	0.033	0.032	0.031	0.397	0.382	0.351	0.009	0.006	0.003
0.05	0.030	0.035	0.033	0.032	0.405	0.387	0.353	0.017	0.010	0.005
0.03	0.048	0.053	0.051	0.050	0.406	0.389	0.356	0.018	0.013	0.008
0.05	0.048	0.054	0.052	0.050	0.416	0.395	0.359	0.027	0.018	0.011
$T = 30, \rho = 0.8$										
0.03	0.030	0.030	0.029	0.029	0.397	0.389	0.369	0.001	0.000	0.000
0.05	0.030	0.029	0.029	0.029	0.397	0.389	0.369	0.002	0.000	0.000
0.03	0.048	0.047	0.046	0.045	0.400	0.389	0.369	0.005	0.001	0.000
0.05	0.048	0.046	0.044	0.044	0.401	0.390	0.369	0.006	0.001	0.000
$T = 30, \rho = 0.9$										
0.03	0.030	0.030	0.029	0.029	0.397	0.389	0.369	0.002	0.001	0.000
0.05	0.030	0.030	0.029	0.029	0.399	0.390	0.370	0.003	0.001	0.000
0.03	0.048	0.048	0.046	0.045	0.402	0.391	0.370	0.006	0.002	0.001
0.05	0.048	0.047	0.045	0.044	0.404	0.392	0.370	0.008	0.003	0.001

Table 6: Size of Test for Infeasible Estimation  
 ( $n_3 = 0.3, n_2 = 0.3$ )

$\tilde{\sigma}_\mu^2 = \tilde{\sigma}_{\max}^2$	$N$	$T = 10$		$T = 20$	
		$\rho=0.8$	$\rho=0.9$	$\rho=0.8$	$\rho=0.9$
0.03	25	0.065	0.059	0.075	0.073
	50	0.058	0.059	0.055	0.058
	100	0.056	0.059	0.052	0.053
	200	0.051	0.058	0.048	0.049
0.05	25	0.067	0.057	0.074	0.071
	50	0.056	0.065	0.055	0.057
	100	0.057	0.061	0.051	0.051
	200	0.053	0.058	0.045	0.046

Table 7: Size of Test for Asymptotic Average Treatment Effects  
 ( $\tilde{\sigma}_\mu^2 = 0.05$ ,  $n_2 = 0.3$ ,  $n_3 = 0.3$ , 5% Test)

$T$	$\tilde{\sigma}_{\max}^2$	$N$	$\rho = 0.8$			$\rho = 0.9$		
			p=0.05	p=0.1	p=0.2	p=0.05	p=0.1	p=0.2
10	0.03	25	0.043	0.054	0.071	0.043	0.056	0.072
		50	0.038	0.05	0.074	0.038	0.045	0.061
		100	0.026	0.034	0.054	0.029	0.039	0.049
		200	0.034	0.043	0.059	0.038	0.042	0.061
10	0.05	25	0.033	0.046	0.058	0.053	0.062	0.076
		50	0.03	0.038	0.054	0.034	0.041	0.054
		100	0.018	0.025	0.042	0.032	0.04	0.052
		200	0.027	0.034	0.049	0.041	0.049	0.061
20	0.03	25	0.067	0.074	0.083	0.048	0.059	0.074
		50	0.053	0.058	0.069	0.039	0.047	0.059
		100	0.042	0.051	0.06	0.027	0.036	0.046
		200	0.043	0.051	0.061	0.031	0.04	0.051
20	0.05	25	0.055	0.066	0.079	0.034	0.049	0.071
		50	0.046	0.054	0.066	0.029	0.039	0.051
		100	0.036	0.043	0.056	0.019	0.028	0.044
		200	0.038	0.046	0.056	0.021	0.033	0.047

Table 8: Power of the Test  
(AATE=0.05, 5% Test)

$\tilde{\sigma}_\mu^2$	$\tilde{\sigma}_{\max}^2$	$N$	$\rho = 0.8$		$\rho = 0.9$	
			$T = 10$	$T = 20$	$T = 10$	$T = 20$
0.03	0.03	25	0.452	0.468	0.308	0.435
		50	0.562	0.623	0.384	0.584
		100	0.834	0.879	0.629	0.850
		200	0.979	0.993	0.881	0.988
0.05	0.03	25	0.402	0.456	0.252	0.410
		50	0.497	0.611	0.311	0.552
		100	0.782	0.869	0.516	0.829
		200	0.969	0.993	0.774	0.983
0.03	0.05	25	0.349	0.341	0.287	0.325
		50	0.446	0.467	0.351	0.435
		100	0.716	0.720	0.553	0.692
		200	0.935	0.937	0.815	0.930
0.05	0.05	25	0.327	0.338	0.244	0.312
		50	0.402	0.470	0.305	0.417
		100	0.667	0.723	0.476	0.679
		200	0.914	0.943	0.733	0.924

## References

- [1] Ambrus, A., and B. Greiner (2012). Imperfect Public Monitoring with Costly Punishment: An Experimental Study. *American Economic Review*, 102(7), 3317-3332.
- [2] Andreoni, J. (1995). Warm-Glow Versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments. *Quarterly Journal of Economics*, 110(1), 1-21.
- [3] Ashley, R., S. Ball, and C. Eckel (2010). Motives for Giving: A Reanalysis of Two Classic Public Goods Experiments. *Southern Economic Journal*, 77(1), 15-26.
- [4] Bruyn, A.D. and G.E. Bolton (2008) Estimating the Influence of Fairness on Bargaining Behavior. *Management Science*, 1774–1791.
- [5] Blume, A., J. Duffy, and A. M. Franco (2009). Decentralized organizational learning: An experimental investigation. *American Economic Review*, 99(4), 1178-1205.
- [6] Broseta, B. (2000). Adaptive Learning and Equilibrium Selection in Experimental Coordination Games: An ARCH(1) Approach. *Games and Economic Behavior*, 32(1), 25–50.
- [7] Bush, R., and F. Mosteller (1995). Stochastic Models for Learning. New York: John Wiley & Sons.
- [8] Camerer, C. and T. Ho (1999). Experience-Weighted Attraction Learning in Normal Form Games. *Econometrica*, 67, 827-874.
- [9] Chaudhuri, A. (2011). Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature. *Experimental Economics*, 14(1), 47–83.
- [10] Cheung, Y. W., and D. Friedman (1997). Individual Learning in Normal Form Games: Some Laboratory Results. *Games and Economic Behavior*, 19, 46-76
- [11] Crawford, V. P. (1995). Adaptive Dynamics in Coordination Games. *Econometrica*, 63, 103-143
- [12] Cross, J. G. (1983). A Theory of Adaptive Economic Behavior. New York/London: Cambridge University Press
- [13] El-Gamal, M. A., and T.R. Palfrey (1995). Vertigo: Comparing Structural Models of Imperfect Behavior in Experimental Games. *Games and Economic Behavior*, 8, 322-348.
- [14] Fudenberg, D., and D. K. Levine (1995). Consistency and Cautious Fictitious Play. *Journal of Economics Dynamics and Control*, 19, 1065-1090.



- [15] Gächter, S., and C. Thöni (2005). Social Learning and Voluntary Cooperation Among Like-minded People. *Journal of the European Economic Association*, 3(2-3), 303–314.
- [16] Greenaway-McGrevy, R., C. Han and D. Sul (2012). Estimating the Number of Common Factors in Serially Dependent Factor Models. *Economic Letters*, 116, 531-534.
- [17] Harley, C. B. (1981). Learning the Evolutionary Stable Strategies. *Journal of Theoretical Biology*, 89, 611-633.
- [18] Ho, T. and X. Su (2009). Peer-Induced Fairness in Games. *American Economic Review*, 2022-2049
- [19] Houser, D., M. Keane and K. McCabe (2004), Behavior in a Dynamic Decision Problem: An Analysis of Experimental Evidence Using a Bayesian Type Classification Algorithm. *Econometrica*, 72(3), 781-822.
- [20] Hyndman, K., E. Y. Ozbay, A. Schotter, and W. Ehrblatt (2012). Convergence: An Experimental Study of Teaching and Learning in Repeated Games. *Journal of the European Economic Association*, 10(3), 573–604.
- [21] Isaac, R. M., J. M. Walker, and S. H. Thomas (1984). Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations. *Public Choice*, 43(2), 113-149.
- [22] Isaac, R. M., and J. Walker (1988a). Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism. *Quarterly Journal of Economics*, 103(1), 179-199.
- [23] Isaac, R. M., and J. Walker (1988b). Communication and Free Riding behavior: the Voluntary Contribution Mechanism. *Economic Inquiry*, 26(4), 585–608.
- [24] Isaac, R. M., J. Walker and A. W. Williams (1994). Group Size and the Voluntary Provision of Public Goods. *Journal of Public Economics*, 54, 1–36.
- [25] Ledyard, J. (1995). Public Goods: A Survey of Experimental Research. in *The Handbook of Experimental Economics*. Roth and Kagel, eds. Princeton, NJ: Princeton University Press.
- [26] McKelvey, D. and T.R. Palfrey (1992). An Experimental Study of the Centipede Game. *Econometrica*, 4, 803-836.
- [27] Meredith W. and J. Tisak (1990). Latent Curve Analysis. *Psychometrika*, 55, 107-122.
- [28] Phillips, P. C. B. and T. Magdalinos (2007). Limit Theory for Moderate Deviations from a Unit Root. *Journal of Econometrics*, 136, 115–130.

- [29] Phillips, P. C. B. and D. Sul (2007). Transition modeling and econometric convergence tests. *Econometrica*, 75(6), 1771-1855.
- [30] Roth, A., and I. Erev (1995). Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term. *Games and Economic Behavior*, 8, 164-212
- [31] Stahl, D. O. and P.W. Wilson (1995). On Players' Models of Other Players: Theory and Experimental Evidence. *Games and Economic Behavior*, 10, 218-254.
- [32] Sul, D. (2012). Estimation of Treatment Effects in Repeated Public Good Experiments. mimeo, University of Texas at Dallas.
- [33] Wilcox, N. (2006). Theories of learning in games and heterogeneity bias. *Econometrica*, 74, 1271-1292.

## Appendix A: Proofs of Theorem 1 and 2

The following lemmas are helpful to prove Theorem 1.

**Lemma 1** The probability limit of the cross sectional variance of  $y_{it}$  is given by

$$\text{plim}_{N \rightarrow \infty} H_{N,t} = \lambda_1 + n_2 \lambda - 2\lambda_1 \rho^{t-1} + (\lambda_1 + \lambda_2) \rho^{2t-2}$$

where  $\phi = \mu - \tau$ ,  $\tau = n_3 + n_2 \mu$ ,  $\lambda_1 = n_1 \tau^2 + n_2 \phi^2 + n_3 (1 - \tau)^2$ ,  $\lambda = \sigma_\mu^2 + \sigma^2$  and  $\lambda_2 = (1 - n_2) \lambda$ .

**Lemma 2** Let  $f_t(a, b) = -2\lambda_1 \rho^{t-1} + (\lambda_1 + \lambda_2) \rho^{2t-2}$  and  $\xi = \sum_{t=1}^T \left( t - T^{-1} \sum_{t=1}^T t \right) \left( f_t - T^{-1} \sum_{t=1}^T f_t \right)$  where  $\lambda_1 > 0$  and  $\lambda_2 > 0$ . Then as  $T \rightarrow \infty$ ,  $\xi > 0$  as long as  $\lambda_1 (1 + 2\rho) > \lambda_2$  with any  $0 < \rho < 1$ .

**Lemma 3** Let  $\tau_* = \min(\tau, 1 - \tau)$ , if  $\tau_*^2 (1 + 2\rho) > \sigma_\mu^2 + \sigma^2$ , then  $\lambda_1 (1 + 2\rho) > \lambda_2$ .

**Proof of Lemma 1.** The cross sectional average of  $y_{it}$  can be written as

$$\begin{aligned} y_{Nt} &= \frac{1}{N} \sum_{i=1}^N y_{it} = \frac{1}{N} \sum_{i \in G_1}^{N_1} (\mu_i \rho^{t-1} + e_{it}) + \frac{1}{N} \sum_{i \in G_2}^{N_2} (\mu_i + e_{it}) \\ &\quad + \frac{1}{N} \sum_{i \in G_3}^{N_3} (1 - (1 - \mu_i) \rho^{t-1} + e_{it}) \\ &= n_3 + n_2 \mu_{N_2} + (\mu_N - n_2 \mu_{N_2} - n_3) \rho^{t-1} + e_{Nt}. \end{aligned}$$

where  $\mu_{N_2} = \frac{1}{N_2} \sum^{N_2} \mu_i$ ,  $\mu_N = \frac{1}{N} \sum^N \mu_i$ . As  $N \rightarrow \infty$ ,  $\text{plim}_{N \rightarrow \infty} \mu_{N_2} = \text{plim}_{N \rightarrow \infty} \mu_N = \mu$ . Hence we have

$$\text{plim}_{N \rightarrow \infty} y_{Nt} = n_3 + n_2 \mu + (\mu - n_2 \mu - n_3) \rho^{t-1}.$$

From the direct calculation, the probability limit of the cross sectional variance is given as

$$\text{plim}_{N \rightarrow \infty} H_{N,t} = \lambda_1 + n_2 \lambda - 2\lambda_1 \rho^{t-1} + (\lambda_1 + \lambda_2) \rho^{2t-2}$$

where  $\phi = \mu - \tau$ ,  $\tau = n_3 + n_2 \mu$ ,  $\lambda_1 = n_3 (1 - \mu) + \tau \phi$ ,  $\lambda = \sigma_\mu^2 + \sigma^2$  and  $\lambda_2 = (1 - n_2) \lambda$ . Note that  $\lambda_1 = n_3 (1 - \mu) + \tau \phi = n_1 \tau^2 + n_2 \phi^2 + n_3 (1 - \tau)^2 \geq 0$  and  $\lambda_2 \geq 0$ . ■

**Proof of Lemma 2.** From the direct calculation, with any fixed  $|\rho| < 1$ , we have

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{T} \xi &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \left( t - T^{-1} \sum_{t=1}^T t \right) (-2\lambda_1 \rho^{t-1} + (\lambda_1 + \lambda_2) \rho^{2t-2}) \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \left( \frac{\lambda_1}{1 - \rho} T + \frac{\lambda_1 + \lambda_2}{2(1 - \rho^2)} T + O(1) \right) = \frac{(1 + 2\rho) \lambda_1 - \lambda_2}{2(1 - \rho^2)} \end{aligned}$$

Since  $\phi = \mu - \tau$ ,  $\tau = n_3 + n_2 \mu$ ,  $\lambda_1 = n_1 \tau^2 + n_2 \phi^2 + n_3 (1 - \tau)^2$ ,  $\lambda = \sigma_\mu^2 + \sigma^2$  and  $\lambda_2 = (1 - n_2) \lambda$ ,

$$\lim_{T \rightarrow \infty} \frac{1}{T} \xi = \left\{ \begin{array}{ll} -(\sigma_\mu^2 + \sigma^2) / [2(1 - \rho^2)] < 0 & \text{if } n_2 = n_3 = 0 \text{ or if } n_1 = n_2 = 0 \\ 0 & \text{if } n_1 = n_3 = 0 \\ [\lambda_1 - \lambda_2 + 2\rho\lambda_1] / [2(1 - \rho^2)] & \text{if } n_k \neq 0 \text{ for } k = 1, 2, 3. \end{array} \right\}$$

The first two cases are the convergent cases while the last case is the divergent one. Therefore if  $(1 + 2\rho)\lambda_1 > \lambda_2$ , then  $\lim_{T \rightarrow \infty} \frac{1}{T}\xi > 0$ . ■

**Proof of Lemma 3.** Let  $\tau_* = \min(\tau, 1 - \tau)$ , then the following inequality holds.

$$\begin{aligned} n_1\tau^2 + n_3(1 - \tau)^2 &\geq \tau_*^2(n_1 + n_3), \\ \left(n_1\tau^2 + n_3(1 - \tau)^2\right)(1 + 2\rho) &\geq \tau_*^2(n_1 + n_3)(1 + 2\rho) \text{ with } \rho \geq 0. \end{aligned}$$

Suppose that  $\sigma_\mu^2 + \sigma^2 < \tau_*^2(1 + 2\rho)$ , then

$$\begin{aligned} \lambda_2 &= (n_1 + n_3)(\sigma_\mu^2 + \sigma^2) < \tau_*^2(n_1 + n_3)(1 + 2\rho) \leq \left(n_1\tau^2 + n_3(1 - \tau)^2\right)(1 + 2\rho) \\ &< \left(n_1\tau^2 + n_3(1 - \tau)^2 + n_2\phi^2\right)(1 + 2\rho) = \lambda_1(1 + 2\rho). \end{aligned}$$

■

**$N$ -Probability Limits of the OLS estimator and its t-Statistic.** Let

$$H_{N,t} = \lambda_1 + n_2\lambda - 2\lambda_1\rho^{t-1} + (\lambda_1 + \lambda_2)\rho^{2t-2} + \varepsilon_t \quad (31)$$

where  $\varepsilon_t = H_{N,t} - \mathbb{E}(H_{N,t}) = O_p(N^{-1/2})$ . Now consider the following trend regression.

$$H_{N,t} = \alpha_T + \beta_T t + \epsilon_t.$$

The OLS estimators are defined as

$$\begin{aligned} \begin{bmatrix} \hat{\alpha}_T \\ \hat{\beta}_T \end{bmatrix} &= \begin{bmatrix} T & \sum_{t=1}^T t \\ \sum_{t=1}^T t & \sum_{t=1}^T t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^T H_{N,t} \\ \sum_{t=1}^T t H_{N,t} \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 + n_2\lambda + \Delta_\alpha \\ 6 \frac{\lambda_1(1 + 2\rho) - \lambda_2}{(1 - \rho^2)(T^2 - 1)} + \Delta_\beta \end{bmatrix} + \begin{bmatrix} T & \sum_{t=1}^T t \\ \sum_{t=1}^T t & \sum_{t=1}^T t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^T \varepsilon_t \\ \sum_{t=1}^T t \varepsilon_t \end{bmatrix} \end{aligned}$$

where  $\Delta_\alpha = O(T^{-1})$ , and

$$\begin{aligned} \Delta_\beta &= 6\rho^T \frac{2\lambda_1(1 + \rho) - (\lambda_1 + \lambda_2)\rho^T}{(T^2 - 1)(1 - \rho^2)} - 6 \frac{\lambda_1(1 + 5\rho^2 + 6\rho + 2\rho^3) - \lambda_2(1 + \rho^2)}{T(T^2 - 1)(\rho^2 - 1)^2} \\ &\quad + 12 \frac{\lambda_1(1 + \rho)}{(1 - \rho)^2(T^2 - 1)T} \rho^T - 6 \frac{(1 + \rho^2)(\lambda_1 + \lambda_2)}{T(T^2 - 1)(1 - \rho^2)^2} \rho^{2T}. \end{aligned} \quad (32)$$

The OLS residuals,  $\hat{\epsilon}_t$ , is given by

$$\hat{\epsilon}_t = \left(H_{N,t} - \frac{1}{T} \sum_{t=1}^T H_{N,t}\right) - \hat{\beta}_T \left(t - \frac{1}{T} \sum_{t=1}^T t\right) = \tilde{H}_{N,t} - \hat{\beta}_T \tilde{t}, \text{ let say,}$$

so that

$$\frac{1}{T} \sum_{t=1}^T \hat{\epsilon}_t^2 = \frac{1}{T} \sum_{t=1}^T \left(\tilde{H}_{N,t}^2 + \hat{\beta}_T^2 \tilde{t}^2 - 2\hat{\beta}_T \tilde{H}_{N,t} \tilde{t}\right).$$

Define  $\tilde{\rho}^{t-1} = \rho^{t-1} - T^{-1} (1 - \rho^T) (1 - \rho)^{-1}$ ,  $\tilde{\rho}^{2t-2} = \rho^{2t-2} - T^{-1} (1 - \rho^{2T}) (1 - \rho^2)^{-1}$  and  $\tilde{\varepsilon}_t = \varepsilon_t - T^{-1} \sum_{t=1}^T \varepsilon_t$ . Then

$$\frac{1}{T} \sum_{t=1}^T \tilde{H}_{N,t}^2 = \frac{1}{T} \sum_{t=1}^T [-2\lambda_1 \tilde{\rho}^{t-1} + (\lambda_1 + \lambda_2) \tilde{\rho}^{2t-2} + \tilde{\varepsilon}_t]^2 = I_A + II_A,$$

where

$$\begin{aligned} I_A &= 4\lambda_1^2 \frac{1}{T} \sum_{t=1}^T (\tilde{\rho}^{t-1})^2 + (\lambda_1 + \lambda_2)^2 \frac{1}{T} \sum_{t=1}^T (\tilde{\rho}^{2t-2})^2 - 4\lambda_1 (\lambda_1 + \lambda_2) \frac{1}{T} \sum_{t=1}^T \tilde{\rho}^{t-1} \tilde{\rho}^{2t-2} \\ &\quad + \frac{1}{T} \sum_{t=1}^T \tilde{\varepsilon}_t^2, \\ II_A &= -4\lambda_1 \frac{1}{T} \sum_{t=1}^T \tilde{\varepsilon}_t \tilde{\rho}^{t-1} + 2(\lambda_1 + \lambda_2) \frac{1}{T} \sum_{t=1}^T \tilde{\varepsilon}_t \tilde{\rho}^{2t-2}. \end{aligned}$$

Consider the  $N$ -probability limit of  $t_{\hat{\beta}}$  given by

$$\text{plim}_{N \rightarrow \infty} t_{\hat{\beta}} = \frac{\text{plim}_{N \rightarrow \infty} \hat{\beta}_T}{\sqrt{\text{plim}_{N \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_t^2 / \sum_{t=1}^T \tilde{t}^2}},$$

where

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_T = 6 \frac{\lambda_1 (1 + 2\rho) - \lambda_2}{(1 - \rho^2) (T^2 - 1)} + \Delta_\beta = \beta_T, \text{ let say,}$$

and

$$\text{plim}_{N \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_t^2 = \text{plim}_{N \rightarrow \infty} I_A + \beta_T^2 \frac{1}{T} \sum_{t=1}^T \tilde{t}^2 - 2\beta_T \frac{1}{T} \sum_{t=1}^T (-2\lambda_1 \tilde{\rho}^{t-1} + \lambda_2 \tilde{\rho}^{2t-2}) \tilde{t},$$

so that

$$\text{plim}_{N \rightarrow \infty} I_A = 4\lambda_1^2 \frac{1}{T} \sum_{t=1}^T (\tilde{\rho}^{t-1})^2 + (\lambda_1 + \lambda_2)^2 \frac{1}{T} \sum_{t=1}^T (\tilde{\rho}^{2t-2})^2 - 4\lambda_1 (\lambda_1 + \lambda_2) \frac{1}{T} \sum_{t=1}^T \tilde{\rho}^{t-1} \tilde{\rho}^{2t-2},$$

since  $\text{plim}_{N \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \tilde{\varepsilon}_t^2 = 0$ .

Under the local to unity,  $\rho = 1 - c/T$ , the probability limit of  $\hat{\beta}_T$  is given as

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_T = 6 \frac{T^2 (3\lambda_1 - \lambda_2) + 2Tc\lambda_1}{c(2T - c)(T^2 - 1)} + \Delta_\beta.$$

## Proof of Theorem 1

The global  $N$ -asymptotic convergence holds either if  $n_k = 1$  for  $k = 1, 2, 3$ . When  $n_2 = 1$ ,  $H_{Nt}$  does not either converge or diverge. Meanwhile when  $n_1 = 1$  or  $n_3 = 1$ ,  $H_{Nt}$  converges to zero.

**(i) Case of  $n_1 = 1$  or  $n_3 = 1$  with a fixed  $0 < \rho < 1$ .** In this case, the probability limit of  $\hat{\beta}_T$  can be rewritten as

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_T = \frac{-6(\sigma_\mu^2 + \sigma^2)}{(1 - \rho^2)(T^2 - 1)} + \Delta_\beta = \beta_T,$$

and

$$\text{plim}_{N \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \hat{\epsilon}_t^2 = (\sigma_\mu^2 + \sigma^2)^2 \frac{1}{T} \sum_{t=1}^T (\tilde{\rho}^{2t-2})^2 + \beta_T^2 \frac{1}{T} \sum_{t=1}^T \tilde{t}^2 - 2\beta_T \frac{1}{T} \sum_{t=1}^T (\sigma_\mu^2 + \sigma^2) \tilde{\rho}^{2t-2} \tilde{t}.$$

As  $N, T \rightarrow \infty$  jointly, it is easy to show

$$\text{plim}_{N, T \rightarrow \infty} t_{\hat{\beta}} = -\sqrt{3} \sqrt{1 + \frac{2\rho^2}{1-\rho^2}} < -1.732 \text{ for } |\rho| < 1. \quad (33)$$

**(ii) Case of  $n_1 = 1$  or  $n_3 = 1$  with  $\rho = 1 - c/T$ .** As  $N, T \rightarrow \infty$  jointly, the probability limit of  $t_{\hat{\beta}}$  statistic is given by

$$\text{plim}_{N, T \rightarrow \infty} t_{\hat{\beta}} = -\lim_{T \rightarrow \infty} \sqrt{3} \sqrt{1 + \frac{2(T^2 - 2cT + c^2)}{2cT - c^2}} = -\lim_{T \rightarrow \infty} \sqrt{3 \frac{T}{c}} = -\infty \text{ for a fixed } c.$$

**(iii) Case of  $n_2 = 1$**  In this case,  $H_{N,t}$  is given by

$$H_{N,t} = \sigma_{N,\mu}^2 + \sigma_{N,t}^2 + 2\omega_{N,t},$$

where  $\sigma_{N,\mu}^2 = N^{-1} \sum_{i=1}^N (\mu_i - N^{-1} \sum_{i=1}^N \mu_i)^2$ ,  $\sigma_{N,t}^2 = N^{-1} \sum_{i=1}^N (u_{it} - N^{-1} \sum_{i=1}^N u_{it})^2$  and  $\omega_{N,t} = \frac{1}{N} \sum_{i=1}^N \mu_i u_{it} - \frac{1}{N^2} \sum_{i=1}^N u_{it} \sum_{i=1}^N \mu_i$ . Hence the OLS estimator can be rewritten as

$$\hat{\beta}_T = \frac{\sum_{t=1}^T \tilde{t} (\sigma_{N,\mu}^2 + \sigma_{N,t}^2 + 2\omega_{N,t})}{\sum_{t=1}^T \tilde{t}^2}.$$

From Assumption B, we have

$$\sqrt{N} \sum_{t=1}^T \tilde{t} (\sigma_{N,t}^2 - \sigma^2) = \frac{1}{\sqrt{N}} \sum_{t=1}^T \tilde{t} \sum_{i=1}^N (u_{it}^2 - \sigma^2) + O_p\left(\frac{1}{\sqrt{N}}\right) \rightarrow^d \mathcal{N}\left(0, \sum_{t=1}^T \tilde{t}^2 v_4\right),$$

where  $\sigma^2 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \sigma_i^2$  and  $v_4 = \lim_{N \rightarrow \infty} \mathbb{E} \frac{1}{N} \sum_{i=1}^N (u_{it}^2 - \sigma^2)^2$ . Next,

$$\sqrt{N} \omega_{N,t} = \frac{1}{\sqrt{N}} \sum_{i=1}^N \mu_i u_{it} - \frac{1}{N^{3/2}} \sum_{i=1}^N u_{it} \sum_{i=1}^N \mu_i \rightarrow^d \mathcal{N}(0, \sigma^2 \sigma_\mu^2) \text{ for each } t.$$

so that

$$\sqrt{N} \sum_{t=1}^T 2\tilde{t} \omega_{N,t} \rightarrow^d \mathcal{N}\left(0, 4 \sum_{t=1}^T \tilde{t}^2 \sigma^2 \sigma_\mu^2\right).$$

This implies that as  $N, T \rightarrow \infty$  jointly,

$$\sqrt{NT}^{3/2} \hat{\beta}_T \rightarrow^d \mathcal{N}(0, 12 [v_4 + 4\sigma^2 \sigma_\mu^2]).$$

Next, the sum of the residual squares times  $N$  is given by

$$N \sum_{t=1}^T \hat{\epsilon}_t^2 = I + II + III + IV + V + VI,$$

where

$$\begin{aligned}
I &= N\hat{\beta}_T^2 \sum_{t=1}^T (t - \bar{t})^2, \quad II = N \sum_{t=1}^T \left( \sigma_{N,t}^2 - \frac{1}{T} \sum_{t=1}^T \sigma_{N,t}^2 \right)^2, \\
III &= N4 \sum_{t=1}^T \left( \omega_{N,t} - \frac{1}{T} \sum_{t=1}^T \omega_{N,t} \right)^2, \\
IV &= -2N \sum_{t=1}^T \hat{\beta}_T (t - \bar{t}) \left( \sigma_{N,t}^2 - \frac{1}{T} \sum_{t=1}^T \sigma_{N,t}^2 \right), \\
V &= -2N \sum_{t=1}^T \hat{\beta}_T (t - \bar{t}) 2 \left( \omega_{N,t} - \frac{1}{T} \sum_{t=1}^T \omega_{N,t} \right), \\
VI &= 4N \sum_{t=1}^T \left( \sigma_{N,t}^2 - \frac{1}{T} \sum_{t=1}^T \sigma_{N,t}^2 \right) \left( \omega_{N,t} - \frac{1}{T} \sum_{t=1}^T \omega_{N,t} \right).
\end{aligned}$$

The probability limit for each term is as follow.

$$\begin{aligned}
\text{plim}_{N \rightarrow \infty} I &= v_4 + 4\sigma^2\sigma_\mu^2, & \text{plim}_{N \rightarrow \infty} II &= Tv_4 - v_4, & \text{plim}_{N \rightarrow \infty} III &= 4T\sigma^2\sigma_\mu^2 - 4\sigma^2\sigma_\mu^2, \\
\text{plim}_{N \rightarrow \infty} IV &= 0, & \text{plim}_{N \rightarrow \infty} V &= 0, & \text{plim}_{N \rightarrow \infty} VI &= 0.
\end{aligned}$$

Hence the probability limit of the sum of the residual squares times  $N$  is

$$\text{plim}_{N \rightarrow \infty} \frac{N}{T} \sum \hat{\epsilon}_t^2 = v_4 + 4\sigma^2\sigma_\mu^2.$$

Therefore, as  $N, T \rightarrow \infty$  jointly, the t-statistic has the following limiting distribution

$$t_{\hat{\beta}} \rightarrow^d \mathcal{N}(0, 1).$$

■

## Proof of Theorem 2

First, the probability limits of  $\hat{\beta}_T$  and the average of the residual squares are given by

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_T = 6 \frac{\lambda_1(1+2\rho) - \lambda_2}{(1-\rho^2)(T^2-1)} + \Delta_\beta = \beta_T,$$

$$\text{plim}_{N \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \hat{\epsilon}_t^2 = \text{plim}_{N \rightarrow \infty} I_A + \beta_T^2 \frac{1}{T} \sum_{t=1}^T \tilde{t}^2 - 2\beta_T \frac{1}{T} \sum_{t=1}^T (-2\lambda_1 \tilde{\rho}^{t-1} + \lambda_2 \tilde{\rho}^{2t-2}) \tilde{t},$$

where

$$\text{plim}_{N \rightarrow \infty} I_A = 4\lambda_1^2 \frac{1}{T} \sum_{t=1}^T (\tilde{\rho}^{t-1})^2 + (\lambda_1 + \lambda_2)^2 \frac{1}{T} \sum_{t=1}^T (\tilde{\rho}^{2t-2})^2 - 4\lambda_1(\lambda_1 + \lambda_2) \frac{1}{T} \sum_{t=1}^T \tilde{\rho}^{t-1} \tilde{\rho}^{2t-2}.$$

Next as  $T \rightarrow \infty$ , we have

$$\lim_{T \rightarrow \infty} \left( \text{plim}_{N \rightarrow \infty} t_{\hat{\beta}} \right) = \nu = \sqrt{3} \left( \frac{L_1}{L_2} \right)^{1/2}$$

where

$$\begin{aligned} L_1 &= (\lambda_1 (1 + 2\rho) - \lambda_2)^2 (1 + \rho^2) (1 + \rho + \rho^2), \\ L_2 &= 4\lambda_1^2 (1 - \rho^4) (1 + \rho + \rho^2) + (\lambda_1 + \lambda_2)^2 (1 + \rho) (1 - \rho^3) - 4\lambda_1 (\lambda_1 + \lambda_2) (1 - \rho) (1 - \rho^4). \end{aligned}$$

Note that

$$\begin{aligned} L &= L_1 - L_2 = 2\rho\lambda_1^2 (4\rho^5 + 4\rho^4 + 7\rho^3 + 7\rho^2 + 3\rho + 2) \\ &\quad + 2\rho\lambda_2^2 (\rho^3 + \rho^2 + \rho) - 2\rho\lambda_1\lambda_2 (4\rho^4 + 4\rho^3 + 4\rho^2 + 4\rho + 2). \end{aligned}$$

If  $L > 0$ , then  $\nu > 1.732$ . Since  $\lambda_1 > 0$  and the function  $L$  is continuous, and  $L$  is an increasing function of  $\lambda_1$ . That is,

$$\partial L / \partial \lambda_1 = 2\rho (\lambda_1 (1 + 2\rho) - \lambda_2) (4\rho^4 + 4\rho^3 + 4\rho^2 + 4\rho + 2) + 4\rho\lambda_1 (1 - \rho) (1 + \rho) (1 - \rho + 2\rho^2) > 0,$$

This implies that the minimum value of  $L$  is achieved as  $\lambda_1 \rightarrow 0$ . That is,

$$L_{\min} = 2\lambda_2^2 (\rho^4 + \rho^3 + \rho^2) > 0$$

regardless of values of  $\lambda_2$  and  $\rho$ . Therefore

$$\lim_{T \rightarrow \infty} \left( \text{plim}_{N \rightarrow \infty} t_{\hat{\beta}} \right) = \nu > 1.732.$$

Under the local to unity, the limit  $\nu$  becomes

$$\nu = \lim_{T \rightarrow \infty} \left( \frac{9(\lambda_2 - 3\lambda_1)(2T\lambda_2 - 6T\lambda_1 + 23c\lambda_1 - 5c\lambda_2)}{22c\lambda_1^2 - 20c\lambda_1\lambda_2 + 6c\lambda_2^2} \right)^{1/2} = +\infty.$$

■

## Appendix B: Proofs of Theorem 3 and Some Remarks

### Appendix B.1: Proof of Theorem 3

Rewrite the data generating process as

$$y_{k,it} = \alpha_k + (\mu_i - \alpha_k) \rho_k^{t-1} + e_{k,it} \text{ for } k = 1, 2, 3.$$

Note that  $\alpha_1 = \alpha_2 = 0$  but  $\alpha_3 = 1$ ,  $\rho_1 = \rho_3 = \rho$ , and  $\rho_2 = 1$ . When  $k = 1, 2$ , and  $3$ , subjects become free riders, confused subjects and altruists, respectively. The individual trend regression, which is a misspecified time series regression, is given by

$$y_{k,it} = a_{k,i} + b_{k,i}t + w_{k,it}.$$



The OLS estimator for  $b_{k,i}$  can be written as

$$\hat{b}_{k,i} = \frac{\sum \tilde{t} \tilde{y}_{k,it}}{\sum \tilde{t}^2} = c_{k,i} \frac{\sum \tilde{t} \tilde{\rho}_k^{t-1}}{\sum \tilde{t}^2} + \frac{\sum \tilde{t} \tilde{e}_{k,it}}{\sum \tilde{t}^2}, \quad (34)$$

where  $c_{k,i} = (\mu_i - \alpha_k)$  and  $\tilde{\rho}_k^{t-1} = \rho_k^{t-1} - T^{-1} \sum \rho_k^{t-1}$ . The limiting distributions of  $\hat{b}_{k,i}$  and its  $t$ -ratio are depending on the value of  $\rho$ . The first term is given by

$$\frac{\sum \tilde{t} \tilde{\rho}_k^{t-1}}{\sum \tilde{t}^2} = \begin{cases} -\frac{6}{(1-\rho)T^2} + O(T^{-3}) & \text{for } k = 1, 3, \\ 0 & \text{for } k = 2 \end{cases} = \Delta_{k,i}, \text{ let say.}$$

The limiting distribution of the second term is depending on the value of  $\rho$ . Here we consider the following three cases.

**(i) Case of  $k = 1$  or  $k = 3$  with a fixed  $|\rho| < 1$ .** The regression residual is given by

$$\hat{w}_{k,it} = \tilde{y}_{k,it} - \hat{b}_{k,i} \tilde{t} = c_{k,i} \tilde{\rho}_k^{t-1} - \hat{b}_{k,i} \tilde{t} + \tilde{e}_{k,it}$$

but the expectation of its variance is given by

$$\lim_{T \rightarrow \infty} \mathbb{E} T^{-1} \sum_{t=1}^T \hat{w}_{it}^2 = \frac{c_{k,i}^2 + \sigma_i^2}{(1-\rho^2)T} + O(T^{-2}),$$

since

$$\begin{aligned} \mathbb{E} \sum_{t=1}^T \hat{w}_{it}^2 &= c_{k,i}^2 \sum \rho^{2t-2} + \sigma_i^2 \sum \rho^{2t-2} + O(T^{-1}) \\ &= \frac{c_{k,i}^2 + \sigma_i^2}{1-\rho^2} + O(T^{-1}) \neq \mathbb{E} \sum_{t=1}^T \tilde{e}_{it}^2 = \frac{\sigma_i^2}{1-\rho^2} + O(T^{-1}). \end{aligned}$$

Due to the use of the misspecified regression, the conventional  $t$ -ratio becomes also inconsistent. To see this, first define the  $t$ -ratio for  $\hat{b}_{k,i}$  as

$$t_{\hat{b}_{k,i}} = \frac{\hat{b}_{k,i}}{\sqrt{\frac{1}{T} \sum \hat{w}_{it}^2 / \sum \tilde{t}^2}}.$$

Next, let's derive the limiting distribution of  $\hat{b}_{k,i}$ . Note that the second term in (34) has the following limiting distribution.

$$T^2 \frac{\sum \tilde{t} \tilde{e}_{k,it}}{\sum \tilde{t}^2} = T^2 \left( \hat{b}_{k,i} - c_{k,i} \Delta_{k,i} \right) \rightarrow^d \mathcal{N} \left( 0, \frac{36\sigma_i^2}{(1-\rho^2)} \right),$$

Hence we have

$$T^2 \hat{b}_{k,i} \sqrt{\frac{(1-\rho^2)}{36\sigma_i^2}} = \mp \left( \frac{c_{k,i}^2}{\sigma_i^2} \frac{1+\rho}{1-\rho} \right)^{1/2} + T^2 \sqrt{\frac{(1-\rho^2)}{36\sigma_i^2} \frac{\sum_{t=1}^T \tilde{t} e_{it}}{\sum_{t=1}^T \tilde{t}^2}} + O_p(T^{-1}) \rightarrow^d \mathcal{N}(b_i^*, 1),$$

where

$$b_i^* = \mp \left( \frac{c_{k,i}^2}{\sigma_i^2} \frac{1+\rho}{1-\rho} \right)^{1/2}.$$

Meanwhile

$$\begin{aligned} t_{\hat{b}_{k,i}} &= \frac{\hat{b}_{k,i}}{\sqrt{\frac{1}{T} \sum \hat{w}_{it}^2 / \sum \tilde{t}^2}} = \mp \left( \frac{3c_{k,i}^2 (1+\rho)}{(c_{k,i}^2 + \sigma_i^2) (1-\rho)} \right)^{1/2} + T^2 \sqrt{\frac{1-\rho^2}{12(c_{k,i}^2 + \sigma_i^2)} \frac{\sum_{t=1}^T \tilde{t} e_{it}}{\sum_{t=1}^T \tilde{t}^2}} \\ &\rightarrow^d \mathcal{N}\left(b_i^+, \frac{3\sigma_i^2}{c_{k,i}^2 + \sigma_i^2}\right) \end{aligned}$$

where

$$b_i^+ = \mp \left( \frac{3c_{k,i}^2 (1+\rho)}{(c_{k,i}^2 + \sigma_i^2) (1-\rho)} \right)^{1/2}.$$

Let  $\zeta_i = \sigma_i^2/c_{k,i}^2$  where  $0 \leq \zeta_i \leq 1$  since

$$\sigma_i^2 = \mathbb{E} \left( \frac{1}{T} \sum_{t=1}^T u_{it}^2 \right) \leq \max u_{it}^2 = c_{k,i}^2.$$

Hence

$$b_i^+ = \mp \left( \frac{3(1+\rho)}{(1+\zeta_i)(1-\rho)} \right)^{1/2}, \quad 0 \leq \frac{3\sigma_i^2}{c_{k,i}^2 + \sigma_i^2} = \frac{3\zeta_i}{1+\zeta_i} \leq 1.5$$

To evaluate the critical value, we consider the  $(1-\alpha)\%$  level for  $k=1$  but the  $\alpha\%$  for  $k=3$ . Let  $\varkappa$  be the  $\alpha\%$  critical value for  $N(0,1)$ . Since

$$t_{\hat{b}_{k,i}} \sqrt{\frac{1+\zeta_i}{3\zeta_i}} \pm \left( \frac{1+\zeta_i}{3\zeta_i} \frac{3(1+\rho)}{(1+\zeta_i)(1-\rho)} \right)^{1/2} \rightarrow^d \mathcal{N}(0,1),$$

the critical value of  $t_{\hat{b}_{k,i}}$  for  $k=1$  at the  $(1-\alpha)\%$  level,  $x_{1-\alpha}$ , is given by

$$x_{1-\alpha} = \varkappa \sqrt{\frac{3\zeta_i}{1+\zeta_i}} - \left( \frac{3}{(1+\zeta_i)} \frac{1+\rho}{1-\rho} \right)^{1/2},$$

meanwhile the critical value of  $t_{\hat{b}_{k,i}}$  for  $k=3$  at the  $\alpha\%$  level,  $x_\alpha$ , is

$$x_\alpha = -\varkappa \sqrt{\frac{3\zeta_i}{1+\zeta_i}} + \left( \frac{3}{(1+\zeta_i)} \frac{1+\rho}{1-\rho} \right)^{1/2}.$$

It is easy to see that  $x_{1-\alpha}$  becomes minimized when  $\zeta_i = 0$  but maximized when  $\zeta_i = 1$ . To get the maximum value, we consider  $\rho \geq 0.8$  which is the boundary value for almost all empirical studies. Then the maximum value of  $x_{1-\alpha}$  is obtained when  $\rho \uparrow 0.8$  and it becomes

$$\max x_{1-\alpha} < \varkappa \sqrt{\frac{3}{2}} - \sqrt{\frac{27}{2}}.$$

When  $\varkappa = 1.65$  (critical value at the 5% level for the one size test),

$$\max x_{1-\alpha} < -1.6534.$$

Meanwhile the minimum value of  $x_\alpha$  becomes

$$\min x_\alpha > 1.6534.$$

**(ii) Case of  $k = 1$  or  $3$  with the local to unity,  $\rho = 1 - c/T$**  In this case, as  $T \rightarrow \infty$ ,

$$t_{\hat{b}_{k,i}} = \mp \left( \frac{3c_{k,i}^2 (2T - c)}{(c_{k,i}^2 + \sigma_i^2) c} \right)^{1/2} = \begin{cases} -\infty & \text{when } i \in G_1 \\ +\infty & \text{when } i \in G_3 \end{cases}.$$

**(iii) Case of  $k = 2$**  The trend regression becomes well specified in this case. It is straightforward to show that

$$T^{3/2} \hat{b}_{2,i} \rightarrow^d \mathcal{N}(0, 12\sigma_i^2),$$

and

$$t_{\hat{b}_{2,i}} \rightarrow^d \mathcal{N}(0, 1).$$

■

## Appendix B.2: Proof of Remark 4

When we identify the club membership, we assign each subject into groups using the following rule:

$$i \in G_1 \text{ if } t_{\hat{b}_1} < -c_0, \quad i \in G_2 \text{ if } -c_0 \leq t_{\hat{b}_1} \leq c_0, \quad \text{and } i \in G_3 \text{ if } t_{\hat{\beta}_1} > c_0,$$

As shown before, as  $T \rightarrow \infty$

$$t_{\hat{b}_1} \rightarrow -\infty \text{ if } i \in G_1, \quad t_{\hat{b}_3} \rightarrow \infty \text{ if } i \in G_3, \quad \text{and } t_{\hat{b}_2} \rightarrow \mathcal{N}(0, 1) \text{ if } i \in G_2.$$

As we set higher critical value of  $c_0$ ,

$$\Pr(\text{subject } i \text{ is assigned to } G_2) \rightarrow n_2.$$

Define an indicator function for each subject

$$Z_i = \begin{cases} 1 & \text{if } i \text{ is assigned to } G_2 \\ 0 & \text{otherwise} \end{cases}.$$

Then  $Z_i$  follows Bernoulli distribution with mean  $n_2$  and variance  $n_2(1 - n_2)$ . That is,

$$\hat{n}_2 = \frac{\hat{N}_2}{N} = \frac{\sum_{i=1}^N Z_i}{N} \rightarrow^d \mathcal{N}\left(n_2, \frac{n_2(1 - n_2)}{N}\right).$$

■

### Appendix B.3: Proof of Remark 5

Let

$$\begin{aligned} y_{N_1t}^c &= \mu_{N_1} \rho^{t-1} + u_{N_1t}^c \rho^{t-1}, \\ y_{N_3t}^c &= 1 - (1 - \mu_{N_3}) \rho^{t-1} + u_{N_3t}^c \rho^{t-1}. \end{aligned}$$

Next, define  $\hat{\mu}_N = N^{-1} \sum_{i=1}^N y_{i1} = N^{-1} \sum_{i=1}^N (\mu_i + u_{i1})$ . Then we have

$$\begin{aligned} y_{N_1t}^c &= \hat{\mu}_N \rho^{t-1} + u_{N_1t}^c \rho^{t-1} + (\mu_{N_1} - \hat{\mu}_N) \rho^{t-1}, \\ 1 - y_{N_3t}^c &= (1 - \hat{\mu}_N) \rho^{t-1} - u_{N_3t}^c \rho^{t-1} - (\mu_{N_3} - \hat{\mu}_N) \rho^{t-1}. \end{aligned}$$

Taking logarithm yields

$$\begin{aligned} \log y_{N_1t}^c &= \log \hat{\mu}_N + (t-1) \log \rho + \frac{u_{N_1t}^c}{\hat{\mu}_N} + \frac{\mu_{N_1} - \hat{\mu}_N}{\hat{\mu}_N} + O_p(N^{-1}), \\ \log(1 - y_{N_3t}^c) &= \log(1 - \hat{\mu}_N) + (t-1) \log \rho - \frac{u_{N_3t}^c}{(1 - \hat{\mu}_N)} - \frac{\mu_{N_3} - \hat{\mu}_N}{1 - \hat{\mu}_N} + O_p(N^{-1}). \end{aligned}$$

The pooled OLS estimator is given by

$$\hat{\gamma} - \gamma = \frac{\sum_{t=1}^T (t-1) [u_{N_1t}^c + (\mu_{N_1} - \hat{\mu}_N)]}{2\hat{\mu}_N \sum_{t=1}^T (t-1)^2} - \frac{\sum_{t=1}^T (t-1) [u_{N_3t}^c + (\mu_{N_3} - \hat{\mu}_N)]}{2(1 - \hat{\mu}_N) \sum_{t=1}^T (t-1)^2}.$$

Note that for a large  $T$ , as  $N \rightarrow \infty$ , the probability limit of the first denominator term becomes

$$\text{plim}_{N \rightarrow \infty} \hat{\mu}_N \sum_{t=1}^T (t-1)^2 = \frac{\mu}{3} T^3 + O(T^2),$$

meanwhile the expectation of the variance of the numerator terms are given by

$$\begin{aligned} & \mathbb{E} \left\{ \sum_{t=1}^T (t-1) u_{N_s t}^c + (\mu_{N_s} - \hat{\mu}_N) \sum_{t=1}^T (t-1) \right\}^2 \\ &= \frac{\sigma^2}{N_s} \sum_{t=1}^T (t-1)^2 + \left( \frac{1}{N_s} \sigma_\mu^2 + \frac{1}{N} (\sigma^2 - \sigma_\mu^2) \right) \frac{1}{4} (T^2 - T)^2, \text{ for } s = 1, 3 \end{aligned}$$

since

$$\begin{aligned} E(\mu_{N_s} - \hat{\mu}_N)^2 &= E\left(N_s^{-1} \sum_{i=1}^{N_s} \mu_i - N^{-1} \sum_{i=1}^N (\mu_i + u_{i1})\right)^2 \\ &= \frac{1}{N_s} \sigma_\mu^2 + \frac{1}{N} (\sigma^2 - \sigma_\mu^2) \text{ for } s = 1, 3. \end{aligned}$$

Then

$$E(\hat{\gamma} - \gamma)^2 = \frac{9(\sigma_\mu^2 + n_1(\sigma^2 - \sigma_\mu^2))}{16\mu^2 n_1 T^2 N} + \frac{9(\sigma_\mu^2 + n_3(\sigma^2 - \sigma_\mu^2))}{16(1-\mu)^2 n_3 T^2 N} + O(T^{-3} N^{-1})$$

Therefore as  $N \rightarrow \infty$  with a fixed  $T$ ,

$$\sqrt{N}(\hat{\gamma} - \gamma) \rightarrow^d \mathcal{N}(0, \varpi_\gamma^2 / T^2),$$

where

$$\varpi_\gamma^2 = \frac{9(\sigma_\mu^2 + n_1(\sigma^2 - \sigma_\mu^2))}{16\mu^2 n_1} + \frac{9(\sigma_\mu^2 + n_3(\sigma^2 - \sigma_\mu^2))}{16(1-\mu)^2 n_3}$$

and as  $N, T \rightarrow \infty$ , jointly,

$$\sqrt{NT}(\hat{\gamma} - \gamma) \rightarrow^d \mathcal{N}(0, \varpi_\gamma^2).$$

Since  $\rho = \exp(\gamma)$ , finally the limiting distribution of  $\hat{\rho}$  can be obtained by using delta method. That is,

$$\sqrt{NT}(\hat{\rho} - \rho) \rightarrow^d \mathcal{N}(0, \rho^2 \varpi_\gamma^2).$$

■

## Appendix C: Proofs of Theorem 4 and 5

### Appendix C.1: Proof of Theorem 4

Rewrite the regression of interest again.

$$y_{N,t} = \tau_N + \phi_N \rho^{t-1} + e_{Nt}.$$

The OLS estimators are defined as

$$\begin{bmatrix} \hat{\tau}_N \\ \hat{\phi}_N \end{bmatrix} = \begin{bmatrix} \tau_N \\ \phi_N \end{bmatrix} + \begin{bmatrix} T & \sum_{t=1}^T \rho^{t-1} \\ \sum_{t=1}^T \rho^{t-1} & \sum_{t=1}^T \rho^{2(t-1)} \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^T e_{Nt} \\ \sum_{t=1}^T \rho^{t-1} e_{Nt} \end{bmatrix},$$

and let the covariance and variance matrix be

$$E \begin{bmatrix} \left(\sum_{t=1}^T e_{Nt}\right)^2 & \left(\sum_{t=1}^T e_{Nt}\right) \left(\sum_{t=1}^T \rho^{t-1} e_{Nt}\right) \\ \left(\sum_{t=1}^T e_{Nt}\right) \left(\sum_{t=1}^T \rho^{t-1} e_{Nt}\right) & \left(\sum_{t=1}^T \rho^{t-1} e_{Nt}\right)^2 \end{bmatrix} = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} \\ \sigma_{12} & \sigma_{22}^2 \end{bmatrix}, \text{ let say.}$$

The joint limiting distributions of  $(\hat{\tau}_N - \tau)$  and  $(\hat{\phi}_N - \phi)$  can be decomposed into

$$\begin{bmatrix} \sqrt{N}(\hat{\tau}_N - \tau) \\ \sqrt{N}(\hat{\phi}_N - \phi) \end{bmatrix} = \begin{bmatrix} \sqrt{N}(\tau_N - \tau) \\ \sqrt{N}(\phi_N - \phi) \end{bmatrix} + \begin{bmatrix} \sqrt{N}(\hat{\tau}_N - \tau_N) \\ \sqrt{N}(\hat{\phi}_N - \phi_N) \end{bmatrix}.$$

We derive the limiting distribution of the second term first. Note that

$$e_{Nt} = \frac{1}{N} \sum_{i \in G_1, G_3} u_{it} \rho^{t-1} + \frac{1}{N} \sum_{i \in G_2} u_{it} = \rho^{t-1} \frac{1}{N} \sum u_{it} + (1 - \rho^{t-1}) \frac{1}{N} \sum_{i \in G_2} u_{it}.$$

Hence we have

$$\begin{aligned} \sigma_{11}^2 &= \mathbb{E} \left( \sum^T e_{Nt} \right)^2 = \frac{\sigma^2}{N} \left[ (1 - n_2) \frac{1 - \rho^{2T}}{1 - \rho^2} + n_2 T \right], \\ \sigma_{12} &= \mathbb{E} \left( \sum^T e_{Nt} \right) \left( \sum^T \rho^{t-1} e_{Nt} \right) = \frac{\sigma^2}{N} \left[ (1 - n_2) \frac{1 - \rho^{3T}}{1 - \rho^3} + n_2 \frac{1 - \rho^T}{1 - \rho} \right], \\ \sigma_{22}^2 &= \mathbb{E} \left( \sum \rho^{t-1} e_{Nt} \right)^2 = \frac{\sigma^2}{N} \left[ (1 - n_2) \frac{1 - \rho^{4T}}{1 - \rho^4} + n_2 \frac{1 - \rho^{2T}}{1 - \rho^2} \right]. \end{aligned}$$

Therefore as  $N, T \rightarrow \infty$ ,

$$\begin{bmatrix} \sqrt{NT}(\hat{\tau}_N - \tau_N) \\ \sqrt{N}(\hat{\phi}_N - \phi_N) \end{bmatrix} \rightarrow^d \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} n_2 & 0 \\ 0 & \frac{1 - \rho^2}{1 + \rho^2} (1 + n_2 \rho^2) \end{bmatrix} \sigma^2 \right).$$

Next, the limiting distribution of the first term is given by

$$\begin{bmatrix} \sqrt{N}(\tau_N - \tau) \\ \sqrt{N}(\phi_N - \phi) \end{bmatrix} \rightarrow^d \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} n_2 & 0 \\ 0 & (1 - n_2)^2 \end{bmatrix} \sigma_\mu^2 \right).$$

where we use the fact that the covariance of  $\hat{\tau}_N$  and  $\hat{\alpha}_N$  becomes zero since

$$\begin{aligned} E(\tau_N - \tau)(\phi_N - \phi) &= E \left( N^{-1} \sum^{N_2} (\mu_i - \mu) \right) \left( N^{-1} \sum^N (\mu_i - \mu) - N^{-1} \sum^{N_2} (\mu_i - \mu) \right) \\ &= n_2 N^{-1} \sigma_\mu^2 - n_2 N^{-1} \sigma_\mu^2 = 0. \end{aligned}$$

Therefore

$$\begin{bmatrix} \sqrt{N}(\hat{\tau}_N - \tau) \\ \sqrt{N}(\hat{\phi}_N - \phi) \end{bmatrix} \rightarrow^d \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} n_2 \sigma_\mu^2 & 0 \\ 0 & \sigma_\alpha^2 \end{bmatrix} \right),$$

where  $\sigma_\alpha^2 = (1 - n_2)^2 \sigma_\mu^2 + (1 - \rho^2) (1 + n_2 \rho^2) (1 + \rho^2)^{-1} \sigma^2$ .

■

## Appendix C.2: Proof of Theorem 5

Rewrite  $y_{Nt}$  as

$$y_{Nt} = \tau_N + \phi_N \hat{\rho}^{t-1} + \phi_N (\rho^{t-1} - \hat{\rho}^{t-1}) + e_{Nt} = \tau_N + \phi_N \hat{\rho}^{t-1} + e_{Nt}^+,$$

where

$$e_{Nt}^+ = \phi_N (\rho^{t-1} - \hat{\rho}^{t-1}) + e_{Nt}.$$

Note that

$$\hat{\rho}^{t-1} = \rho^{t-1} + (t-1)\rho^{t-2}(\hat{\rho} - \rho) + O_p(N^{-1}).$$

where we have that  $\sqrt{NT}(\hat{\rho} - \rho) \xrightarrow{d} N(0, \varpi_\rho)$ , which we have shown before. Hence the regression error can be written as

$$e_{Nt}^+ = -\phi_N (t-1)\rho^{t-2}(\hat{\rho} - \rho) + e_{Nt} + O_p(N^{-1}),$$

and its variance is given by

$$E(e_{Nt}^+)^2 = \phi^2 (t-1)^2 \rho^{2(t-2)} \frac{\varpi_\rho}{N} + \frac{\sigma^2}{N} \left( n_2 + (1-n_2)\rho^{2(t-2)} \right) + O(N^{-2}),$$

meanwhile its autocovariance is written as

$$Ee_{Nt}^+ e_{Ns}^+ = \phi^2 (t-1)(s-1)\rho^{t+s-4} \frac{\varpi_\rho}{N} + O(N^{-2}).$$

From the direct calculation, the covariance matrix is derived as

$$\begin{aligned} \Sigma &= E \begin{bmatrix} \left( \sum^T e_{Nt}^+ \right)^2 & \left( \sum^T e_{Nt}^+ \right) \left( \sum^T \hat{\rho}^{t-1} e_{Nt}^+ \right) \\ \left( \sum^T e_{Nt}^+ \right) \left( \sum^T \hat{\rho}^{t-1} e_{Nt}^+ \right) & \left( \sum^T \hat{\rho}^{t-1} e_{Nt}^+ \right)^2 \end{bmatrix} \\ &= \begin{bmatrix} \frac{\sigma^2}{N} (1-n_2) \frac{1-\rho^{2T}}{1-\rho^2} + \frac{T\sigma^2}{N} n_2 & \frac{\sigma^2}{N} (1-n_2) \frac{1-\rho^{3T}}{1-\rho^3} + \frac{\sigma^2}{N} n_2 \frac{1-\rho^T}{1-\rho} \\ \frac{\sigma^2}{N} (1-n_2) \frac{1-\rho^{3T}}{1-\rho^3} + \frac{\sigma^2}{N} n_2 \frac{1-\rho^T}{1-\rho} & \frac{\sigma^2}{N} (1-n_2) \frac{1-\rho^{4T}}{1-\rho^4} + \frac{\sigma^2}{N} n_2 \frac{1-\rho^{2T}}{1-\rho^2} \end{bmatrix} + O(T^{-1}N^{-1}). \end{aligned}$$

Hence as  $N, T \rightarrow \infty$ , the limiting distributions become

$$\begin{bmatrix} \sqrt{N}(\hat{\tau}_N - \tau) \\ \sqrt{N}(\hat{\phi}_N - \phi) \end{bmatrix} \rightarrow^d \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} n_2 \sigma_\mu^2 & 0 \\ 0 & \sigma_\alpha^2 \end{bmatrix} \right),$$

which is the same as the infeasible case.

Meanwhile if  $N \rightarrow \infty$  with a fixed  $T$ , the limiting distribution of  $\hat{\tau}_N$  becomes

$$\sqrt{N}(\hat{\tau}_N - \tau) \rightarrow^d \mathcal{N}(0, n_2 \sigma_\mu^2 + n_2 \sigma^2 / T).$$

■