Journal of Economic Behavior and Organization 152 (2018) 124–146

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Estimating heterogeneous contributing strategies in threshold public goods provision: A structural analysis☆

Yonghong An[a,*], Yingyao Hu[b], Pengfei Liu[c]

[a] Department of Economics, Texas A&M University, USA
[b] Department of Economics, Johns Hopkins University, USA
[c] Department of Aquaculture and Fisheries, University of Arkansas at Pine Bluff, USA

ARTICLE INFO

Article history:
Received 18 September 2017
Revised 14 May 2018
Accepted 2 June 2018

Keywords:
Public goods
Private provision
Measurement errors
Unobserved heterogeneity
Nonparametric identification and estimation
Lab experiment

ABSTRACT

We analyze a structural model of threshold public goods contributions, where the public good is provided only if the aggregated contributions reach or surpass the predetermined cost; otherwise contributions will be returned to individuals. Based on individual contributions to a public good in multiple periods, we are able to identify the number of contributing strategies, functional form for each strategy and the transition probabilities of contributing strategies conditional on the previous provision outcomes. The result of the constructive identification suggests a multi-step procedure to estimate the model primitives. Monte Carlo results illustrate that the procedure works well in practice. We apply the methodology to the experimental data we collected and show that subjects strategically respond to provision history by making an adjustment based on their preceding contributing strategies. We also find that subjects are more likely to adjust contribution strategies upon provision failures.

1. Introduction

Private provision of public goods is important for governments or organizations to seek support to cover project costs partially or entirely. Prominent examples include the newly emerged crowd-funding industry, annual fundraising of non-profit organizations such as Wikipedia and National Public Radio (NPR).[1] The prevalence and rapid development of private provision of public goods call for better understanding of individuals' contributing behavior, which could shed light on some policy-related issues such as setting appropriate mechanisms for the provision. In this paper, we analyze a structural model of threshold public goods contributions by allowing heterogeneity and learning among individuals.

A large body of literature has been devoted to the study of the private provision of public goods with a focus on individual behavior. It has been documented that individuals do not always reveal their true values toward the public good, (e.g.,

☆ We are grateful to Alex Brown, Catherine Eckel, and Daniel Stephenson for their helpful comments and discussions. The authors would like to thank Enago (www.enago.com) for the English language review. This paper is a revised version of the Centre for Microdata Methods and Practice (Cemmap) working paper under the title "Estimating Private Provision of Public Goods with heterogeneous Participants: A Structural Analysis". All errors are our own.

* Corresponding author.
E-mail addresses: y.an@tamu.edu (Y. An), yhu@jhu.edu (Y. Hu), pengfei.liu@uconn.edu (P. Liu).

[1] The Crowd-funding Industry Report's data indicating the overall crowd-funding industry has raised \$2.7 billion in 2012, across more than 1 million individual campaigns globally. In 2013 the industry is projected to grow to \$5.1 billion. Wikipedia organizes an annual fundraising campaign to support its operations, which usually lasts from mid-November to mid-January. The total money raised increases from \$94,000 in 2005 to \$25 million in 2012.

https://doi.org/10.1016/j.jebo.2018.06.003

see Andreoni (1988); Weimann (1994) and Olson (1965)) and that they exhibit strategic and heterogeneous contributing behavior (Oliveira et al. (2014) and Fischbacher and Gächter (2010)). However, existing studies of individuals' behavior mainly rely on behavioral assumptions on their beliefs or preferences, without explicitly modeling individuals' strategic behavior and interactions with a microeconomic foundation. To fill the gap, we propose a structural model of private provision of public goods that allows individuals' contributing behavior to be heterogeneous and evolve over time. The model primitives including the number of different contributing strategies, functional form for each strategy, and the transition probabilities among all possible strategies are shown to be identifiable and estimable from the revealed contribution choices of individuals. Based on the data collected in a threshold public good experiment, we find that subjects strategically respond to provision history by making an adjustment based on their preceding contributions; such response is heterogeneous and dependent on subjects' current contributing strategies.

This paper focuses on threshold public good games with the money back guarantee (Bergstrom et al., 1986; Cadsby and Maynes, 1999; Croson and Marks, 2000; Liu et al., 2016), where the public good is provided only if the aggregated contributions reach or surpass the predetermined cost (or the provision point); otherwise contributions will be returned to individuals.[2] We collect individual contributions data from a threshold public good experiment. Subjects in a group of fixed membership make contributions toward a public good with predetermined cost across 10 periods, with their induced values being randomly drawn from a uniform distribution. Subjects observe the outcome of the game, i.e., whether the public good is provided, but not other group members' contributions after each period. The data pattern suggests that subjects contribute using heterogeneous strategies and they also adjust their contributing strategies based on the outcome as well as their own strategies in the preceding period.

Based on the observed pattern of our data, we propose a structural model describing individuals' behavior in public good provision and estimate the model using the experimental data. Our model allows the individuals to employ heterogeneous contributing strategies (we label all individuals employing the same strategy as a "type") and we do not assume their strategies constitute a Bayesian Nash equilibrium. In line with Fischbacher and Gächter (2010), who focus on "linear" public goods,[3] we assume that the individual heterogeneity originates in their beliefs about other contributors' behavior. We allow subjects to adjust their contributing strategies as the belief can change based on the provision outcomes. Without specifying the number of contributing strategies, the functional form of the strategies, and the adjustment process *ex ante*, we demonstrate that all these objectives can be directly recovered from individuals' contributions. The main requirement of our approach is that each individual participates in three public good provision games (makes three contributions). The sequence of contributions from each individual is used to identify and estimate the structural econometric model based on the recently developed results in nonclassical measurement errors (Hu, 2008). We treat the unobserved type of an individual as the latent variable, while using the revealed contributions as the corresponding measurements. In this way, contributions are used as instrumental variables for the unobserved type of individuals, which enable us to identify a distinct pattern for each contributing strategy.

We employ a two-stage procedure for estimation. First, we back out the number of types as well as the contributing strategy for each type using a fully nonparametric approach. Second, we use maximum likelihood estimation to recover the adjustment process among different types based on provision history using multi-periods data. A Monte Carlo experiment demonstrates that our proposed method performs very well for samples with a similar size to our experimental data. Our empirical estimations based on the experimental data suggest that subjects can be classified into three types ranked by average contribution, from low to high. If we average all the periods, we estimate the proportions of three types to be 23.6%, 36.1%, and 40.3% for type 1, 2, and 3, respectively. The estimated contributing strategies of all three types are significantly different and all highly nonlinear. Our estimation results show that type 1 contributes substantially less than type 2 and 3. Nevertheless, we find that type 1 still contributes a significant proportion of their induced values rather than employs a complete free-riding strategy.[4]

We illustrate the adjustment among types by estimating the transition matrices with each element being a probability of type k ($k = 1, 2, 3$) in the current period conditional on one's type j ($j = 1, 2, 3$) and the outcome of the provision in the preceding period. We find that subjects maintain their proceeding contributing strategies with a probability greater than 67% in response to a successful provision. By contrast, both type 1 and 2 would adjust to higher types with a substantial probability to respond to an unsuccessful provision. The majority of type 3 would stick to their contributing strategy regardless of the outcome. We also estimate the model separately using the first and last five periods of data and compare the results with that estimated from all the 10 periods. The comparison suggests that in the last five periods subjects are less reluctant to adjust their contributing strategies. Moreover, we do not find subjects' contributing strategies converge during the ten periods' of experiments.

[2] Bagnoli and McKee (1991) find that provision point mechanism together with money back guarantee (MBG) can potentially induce Pareto efficient outcome in a single unit provision environment.

[3] In a standard linear public good game, subjects are asked to allocate their tokens between a private fund that benefits only the individual investor and a group fund that generates profits for everyone. The private fund yields a higher rate of return than the public fund for the private investor, but the public fund provides the group with a higher total return. The marginal return for the group fund is normally set such that the social optimum occurs when individuals give everything to the group fund, while the individuals' optimum occurs when one keeps all tokens in their private fund.

[4] Heterogeneous behavior in "linear" public good games indicate the existence of a substantial portion of free-rider (Fischbacher and Gächter, 2010; Fischbacher et al., 2001).

The main contribution of our paper is to provide some new findings on individuals' contributing strategies and how they are affected by provision outcomes. There is a fast-growing experimental literature on private provision of public goods focuses on investigating the heterogeneous behavior of individuals (Fischbacher and Gächter, 2010; Oliveira et al., 2014). Ours is the first paper that explicitly models and estimates individuals' heterogeneous strategies and the transition probabilities in private provision of a threshold public good without imposing behavioral assumptions on individuals' beliefs or preferences. The estimates in our paper constitute systematic evidence on individuals' heterogeneity and their strategic response to others' behavior as well as their own.

The novelty of our paper is that the analysis of individuals' heterogeneous behavior is grounded on revealed contribution choices without any behavioral assumptions. Existing studies of individuals' heterogeneous behavior (e.g., Fischbacher and Gächter (2010); Fischbacher et al. (2001), and Fischbacher et al. (2014)) mainly rely on experimental controls to classify different contributing strategies, which is either not applicable or costly to data contexts in fields. By contrast, our approach does not specify strategies *ex ante*. We rather identify strategies from the revealed contributions by exploring the structural connection between individuals' strategies and their multiple contributions. Therefore, our approach can be applied to more general data contexts.

Another main contribution of our paper is that contributing strategies can be estimated without imposing a functional form or solving equilibria explicitly. This is a great advantage since the existing studies on the threshold public good provision lack detailed analyses on individuals' contributing strategies, partly due to the difficulty of deriving an analytical solution. There are several attempts to characterize the Bayesian-Nash equilibrium for a two-players threshold public good provision game (e.g., Alboth et al. (2001); Barbieri and Malueg (2008), and Laussel and Palfrey (2003)). However, once the group size grows to three and above, an analytical solution is almost impossible without much more stringent assumptions. The possible non-equilibrium and heterogeneous contributing strategies of individuals might be rationalized by various behavioral models, e.g., level-k thinking (Crawford and Iriberri, 2007) or cognitive hierarchy (Camerer et al., 2004). Therefore, our paper sheds some lights on the analysis of non-equilibrium behavior without imposing too many structural restrictions. For instance, we can test the validity of a model by comparing its prediction of subjects' contributing strategies with our estimates. Furthermore, our paper contributes to the public good learning literature (Clemens and Riechmann, 2002; Healy, 2006). Our learning results show individuals will adjust their contributions based on the history of outcome and their own strategies, and such learning adjustments are contingent on unobserved individual types.

Our methodology relates to recent studies of unobserved heterogeneity in environments of strategic interactions using the results of measurement errors (e.g. Hu (2008); Hu and Schennach (2008)). For example, Krasnokutskaya (2011); Li et al. (2000) and Hu et al. (2013b) consider auction models unobserved heterogeneity. Hu et al. (2013a) use bandit experiments to nonparametrically estimate the learning rule using auxiliary measurements of beliefs. Xiao (2018) considers multiple equilibria in static and dynamic games. The connection between the unobserved heterogeneity and observables in these studies is similar to our paper. Nevertheless, our paper is the first study of the private provision of public goods with rigorous identification and estimation applying measurement errors method.

The remaining of this paper is organized as follow. Section 2 provides an overview of the experiment and the data. Section 3 proposes a structural model of threshold public goods with heterogeneous contributing strategies and shows the model is nonparametrically identifiable and estimable. Section 4 conducts Monte Carlo experiments to illustrate our method. Section 5 presents the empirical estimation results using experimental data. Section 6 concludes. Proofs, tables, figures and experiment instructions are collected in the Appendix.

2. The experiment and data

We conducted six experiment sessions in the College of Agriculture and Natural Resources (CANR) Lab, University of Connecticut (UConn). Subjects were recruited primarily through UConn Daily Digest where we advertised requesting volunteer participation in economic experiments. Our subject pool consists mostly undergraduates and a few graduate students from various academic majors who have indicated a willingness to participate in economic experiments. We check the participants' names and email addresses, before confirming their attendance, to ensure each subject participated only once in this sequence of experiments. We conducted experiments through networked computer terminals using z-Tree (Fischbacher (2007)). Inter-participant communications during the experiment were prohibited and subjects could not observe each other's choices. Experiment instructions were distributed to participants and were read aloud. Subjects were told that they had already earned a \$5 show-up fee before we proceeded to the instructions and were paid in cash after the experiment was finished. Each experiment session consisted of two groups with group size fixed at five. Group memberships were kept the same during the 10 decision periods, i.e., subjects knew that they would play with the same people during the 10 -period experiment. There is no time limit for each period, although the first two periods are usually longer than the remaining ones. It takes about 30–45 minutes to run ten periods of experiments.

Our experiment uses the provision point mechanism with the money back guarantee where each subject was asked to contribute to a threshold public good (Cadsby and Maynes, 1999). At the beginning of each decision period, subjects were told their induced values, which simulate the valuations for the public goods. Induced values followed a uniform distribution on the interval [8,20] and were rounded to one decimal place. Subjects knew the value distribution and their own induced values, however, they did not observe the induced value of the others. Subjects were informed the potential

Table 1
Summary statistics by period.

Variable	Period 1	2	3	4	5	6	7	8	9	10
Sample size	60	60	60	60	60	60	60	60	60	60
Provided	0.583	0.333	0.333	0.333	0.417	0.25	0.75	0.833	0.5	0.5
	(0.497)	(0.475)	(0.475)	(0.475)	(0.497)	(0.437)	(0.437)	(0.376)	(0.504)	(0.504)
Value	14.863	13.233	14.373	13.038	14.152	13.567	13.745	13.837	14.952	13.958
	(3.280)	(3.606)	(3.365)	(3.186)	(3.311)	(3.350)	(3.793)	(3.525)	(3.195)	(3.391)
Contribution	8.821	7.775	8.178	7.997	8.323	8.323	8.941	8.992	8.605	7.937
	(4.068)	(4.311)	(3.621)	(3.945)	(2.626)	(3.115)	(3.776)	(3.373)	(3.470)	(3.339)
Subj. Prob.	0.662	0.555	0.594	0.534	0.484	0.513	0.488	0.517	0.521	0.519
	(0.240)	(0.291)	(0.281)	(0.300)	(0.300)	(0.326)	(0.314)	(0.332)	(0.292)	(0.337)
Contr./Value	0.609	0.586	0.569	0.615	0.598	0.615	0.646	0.661	0.583	0.580
	(0.326)	(0.299)	(0.232)	(0.263)	(0.169)	(0.202)	(0.188)	(0.255)	(0.213)	(0.212)

Standard deviation in parentheses.

loss of contributing higher than their induced values.[5] The provision cost or the threshold c was public information. If the total contribution from the group is equal to or higher than the provision cost, the public good will be provided and one's profit equals her induced value minus contribution; otherwise, the public good was not provided and each subject received zero profits as contributions were fully returned (with the money back guarantee mechanism). After each decision period, subjects would be informed of the provision result and their own profits in the last period, but not others' profit. We set the provision cost for one unit equal to 60% of the expected induced value for an individual times the number of all individuals in a session; thus, the cost is $60\% \times 14 \times 5 = 42$. A total of 60 subjects participated in the experiment, producing 600 individual level observations. Individuals received an average earning of about \$20, including the \$5 show-up fee. The earning was based on the cumulative profits that one subject earned in all 10 decision periods. Actual earnings vary across individuals and sessions.

Table 1 presents simple summary statistics of the data. In each column (period), the variable "Provided" is a binary variable indicating the outcome of the public good game: provided=1 if provided and 0 otherwise.; "Contr./Value" is defined as the ratio of contribution over the induced value (such a ratio can be used to approximate subjects' linear contributing strategies). The table demonstrates that the proportion of groups that successfully provide the public good varies a lot across period with the minimum 0.333 and the maximum 0.833 even though the change of value and contribution is relatively small. The average contribution ratio per period ranges from 0.569 to 0.661, which is consistent with results reported in Croson and Marks (2000) for threshold public good games. To visualize the pattern of contribution for groups, we further illustrate the group contribution (the sum of individual contributions in one group) for the twelve experimental groups as well as the average group contribution relative to the provision cost in Fig. 1. We observe that the average group contribution follows the provision cost closely, which is consistent with the theoretical predictions for threshold public game with incomplete information (Bagnoli and Lipman, 1989; Bagnoli and McKee, 1991; Croson and Marks, 2000). Nevertheless, the individual group's contributions display frequent and significant adjustments across the period. Most of the groups increase their contributions to respond to lower contributions in the preceding period and decrease their contributions upon a provision failure.

We check whether the groups' contributing behavior converge to an equilibrium in the ten periods. Due to the complicated nature of the threshold public good game, the analytical solution under various equilibrium concepts including the BNE often does not exist when the group size is larger than two (Alboth et al., 2001; Barbieri and Malueg, 2008). Therefore, it is hard for individuals to figure out the best responses during a short period of time. We compute the variance of contributions among groups for each period, then regress the variance on the period number. The result indicates that the variance does not decrease in period significantly ($variance_t = 7.475 - 0.04906t$, p-value=0.884), implying that the contributing behavior does not converge (at least) to a single equilibrium. This is can also be seen from Fig. 1, where the dispersion of group contributions does not shrink across periods. Our finding is consistent with the literature (e.g., see Cason and Zubrickas (2017)) which suggest that more than 30 periods are often needed to examine equilibrium behaviors.

Next, we present some interesting observations regarding individual subjects' contributing behavior. As subjects are informed the provision outcome after each decision period, the dependence of contributing behavior on the outcomes may exist. For this purpose, we further illustrate in Fig. 2 the relationship between subjects' value and contribution conditional on the outcome in the preceding period. The blue and red markers are for successful provision and provision failure, respectively. There are two important observations regarding this figure. First, the relationship between value and contribution varies across periods for a given outcome. For example, the blue markers are concentrated in period 5, however, they are scattered in period 9, implying that contributing behavior in those two periods is distinct. Second, the relationship between

[5] Our data shows about 2.8% of contributions are higher than the induced value. Subjects were not allowed to contribute more than 20 (in experimental currency) in any single period.

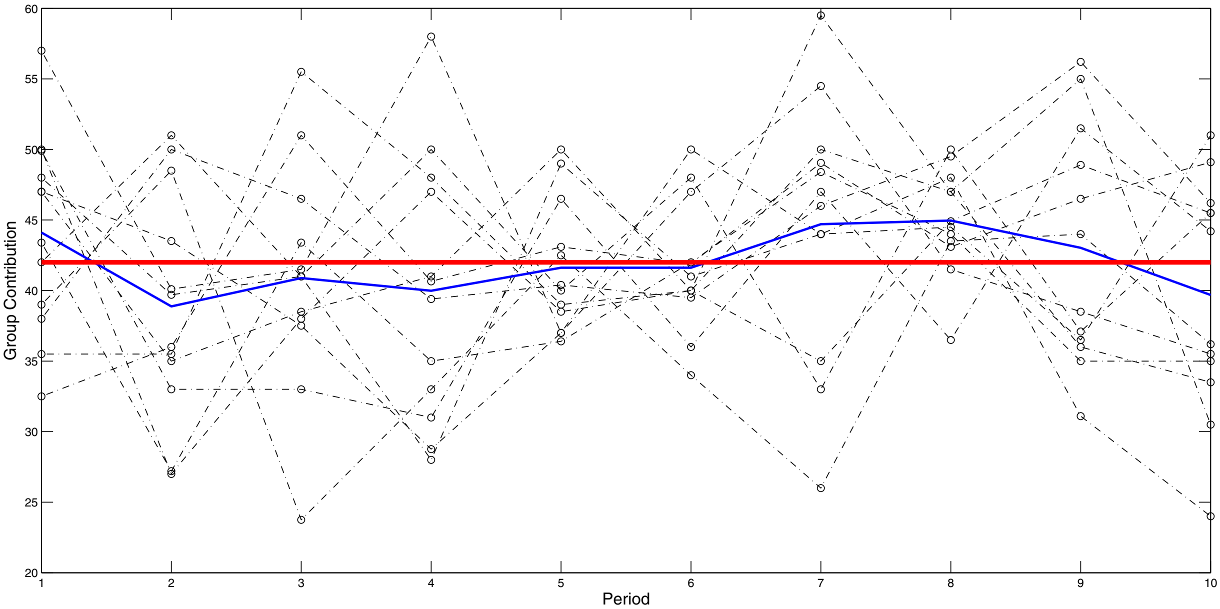

Fig. 1. Group contribution by period.

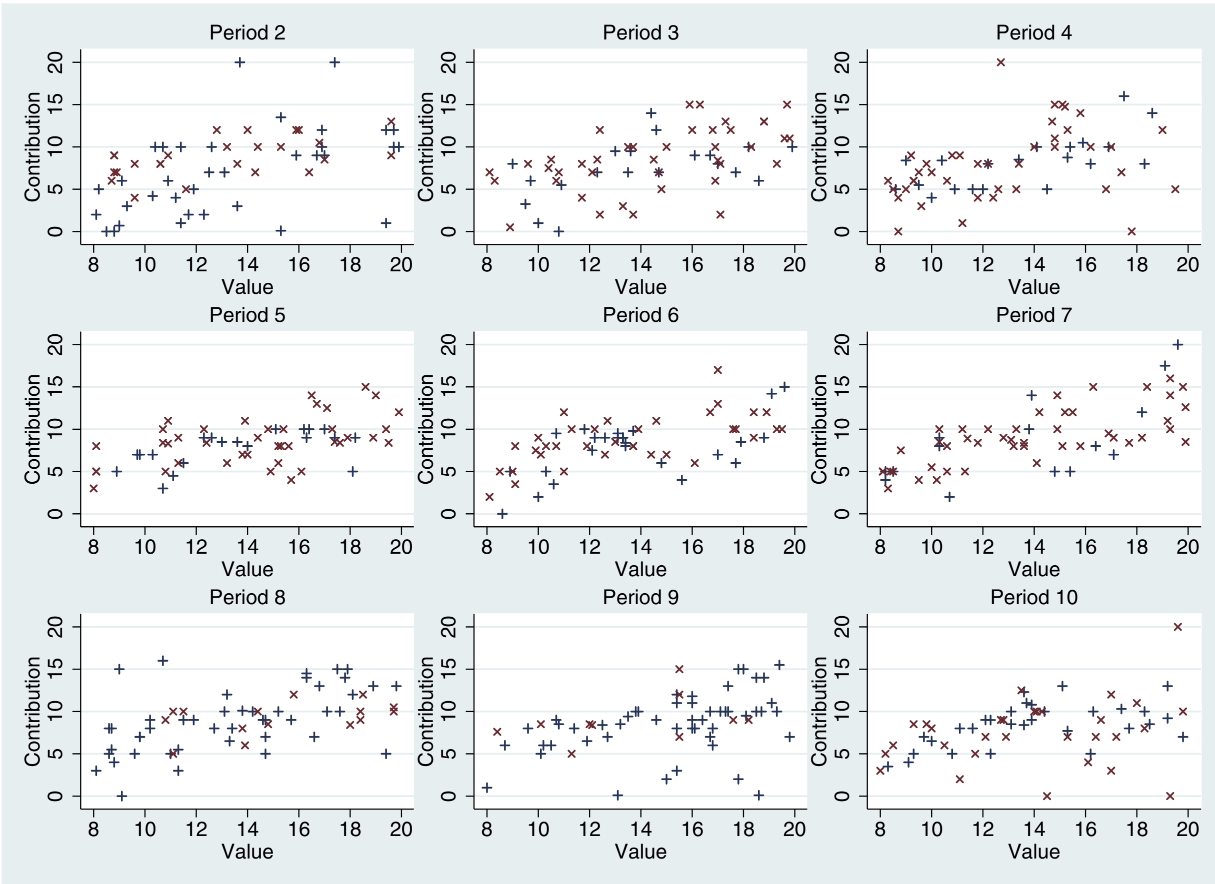

Fig. 2. Value-contribution conditioning on outcome.

value and contribution is different across outcomes; e.g., in period 2 those subjects who had a successful outcome in period 1 contribute relatively less: for a given value, contributions indicated by the blue markers are below the red ones. However, such a pattern is less obvious for period 3. A possible interpretation is that subjects respond to the preceding period's different outcomes differently and such response could also be distinctive across subjects and/or periods.

In summary, the pattern in Figs. 1 and 2 of the experimental data reveals that (1) subjects may be heterogeneous in their contributing behavior; (2) subjects may adjust their contributing behavior based on the provision history and such adjustment can be heterogeneous; (3) the contributing behavior does not converge to an equilibrium in the ten periods' of the game. We present a structural model in the next section to quantitatively assess these findings.

3. A structural model with heterogeneous subjects

In this section, we propose a structural model of public goods provision with heterogeneous subjects to rationalize the findings in Section 2. The main components of the model are shown to be nonparametrically identifiable and estimable under mild conditions.

3.1. The model

A group of $I \geq 2$ risk-neutral subjects contribute to a public good across $T \geq 3$ periods.[6] The private values of subjects $v_{it} \in [\underline{v}, \overline{v}]$, $i = 1, 2, \ldots, I; t = 1, 2, \cdots T$ are i.i.d. draws across i and t from a cumulative distribution function $G(\cdot)$ with density $g(\cdot)$. At period t, subject i makes a contribution $b_{it} \in [0, v_{it}]$ and the public good is provided only if the total contribution of all the I subjects exceeds the cost (threshold) $c > 0$, i.e., $\Sigma_i b_{it} \geq c$ where c is a known constant over periods. We maintain that $\overline{v} < c$ such that it is impossible for one subject to provide the good. Subject i obtains a payoff $v_{it} - b_{it}$ if the public good is provided and zero otherwise. The common knowledge among subjects at the beginning of period t includes the value distribution $G(\cdot)$, group size I, cost c and the outcome of previous periods $w_{-t} \equiv \{w_1, w_2, \cdots, w_{t-1}\}$, which are binary variables with $w_s = 1$ indicating a successful provision in period s, and $w_s = 0$ otherwise. In summary, subject i solves the following maximization problem in period t:

$$\max_{b_{it}} \ (v_{it} - b_{it}) \Pr\left(\sum_{j=1}^{I} b_{jt} \geq c \middle| \mathcal{I}_{-it} \right), \tag{1}$$

where $\mathcal{I}_{-it} \equiv \{(w_s, b_{is}, v_{is}),\ s = 1, 2, \ldots, t-1\}$ is a set of information available to subject i prior to period t.

The probability in (1) summarizes a subject's belief about others' behavior. To model the heterogeneous contributing behavior, we follow the previous findings in the literature and assume that the probability may be different across subjects. Given a subject's value v_{it}, each possible probability implies a corresponding b_{it} as the optimal solution to problem (1). Let all the subjects be one of the K ($K \geq 2$) discrete private types with each type corresponding to a specific contributing strategy, which is defined as follows.[7] Subject i's type is denoted as $\tau_i \in \{1, 2, \ldots, K\}$, then her contributing strategy (we only consider those strategies monotone in value) is a mapping from her private value and type to her contribution, i.e.,

$$s_i(\cdot, \cdot) : [\underline{v}, \overline{v}] \times \{1, 2, \ldots, K\} \to [0, \overline{v}].$$

For ease of notation, we rewrite $s_i(v_i, \tau_i = k)$ as $s_k(v_i)$. This strategy also depends on the group size I, value distribution $G(\cdot)$, and the threshold c; however, we suppress the argument I, G and c in $s_k(\cdot)$ to simplify the notation. We assume that the number of types K does not vary across time and each subject's type is private information. Each subject potentially adjusts her type over time based on the outcome in previous periods. We denote such adjustments by a transition matrix $\Pr(\tau_{t'} | \tau_t, w_{-t})$. Given a vector of outcome history w_{-t}, $\Pr(\tau_{t'} | \tau_t, w_{-t})$ is a $K \times K$ matrix with its (i, j)-th element being the probability for a type j in period t that changes to type i in period t'.

Note that we are not requiring or prohibiting subjects' contributing strategies constitute a Nash equilibrium in our model. The type τ can be understood as a "reduced-form" description of subjects' contributing behavior. For the possible non-equilibrium behavior, the heterogeneous contributing strategies of subjects might be rationalized by various behavioral models, e.g., level-k thinking (Crawford and Iriberri, 2007) or cognitive hierarchy (Camerer et al., 2004).

Using data from three consecutive periods, the joint distribution of b_1, b_2, b_3 can be directly identified. Our goal of identification is to uniquely determine the number of type K, the proportion and bidding strategy for each type as well as the transition matrix $\Pr(\tau_{t'} | \tau_t, w_{-t})$. Let $F(\cdot)$ be the distribution of subjects' contributions and $F(\cdot | \tau = k), k = 1, 2, \cdots, K$ be the distribution for subjects of type k. Then using the data for a given period t (we drop the index of period for simplicity), the model provides a finite mixture of distributions for all the types:

$$F(b_1, b_2, b_3) = \sum_{k=1}^{K} F(b_1, b_2, b_3 | \tau = k) p_k, \tag{2}$$

[6] Subjects are assumed to be risk-neutral, as did in most of the existing literature (see e.g., Cadsby and Maynes (1999); Croson and Marks (2000) and Fischbacher and Gächter (2010)). Given that the average earning is only \$20, we believe that risk-neutrality is a reasonable assumption.

[7] The assumption that subjects are of K types is relevant in environments where members are known a priori to belong to a small number of distinct groups. In other cases where type is in fact continuously distributed, our method below should be interpreted as showing identification for a coarser, discretized version of the model.

where p_k is the proportion of type k. To explore the dependence of model primitives on the relationship above, we consider a similar equation for one period,

$$F(b)=\sum_{k=1}^{K}F(b|\tau=k)p_k=\sum_{k=1}^{K}G\big(s^{-1}_{\tau=k}(b)\big)p_k. \tag{3}$$

The (Eq. (3)) holds because $F(b|\tau=k)=\Pr(B\le b|\tau=k)=\Pr(s_{\tau=k}(V)\le b)=\Pr(V\le s^{-1}_{\tau=k}(b))=G\big(s^{-1}_{\tau=k}(b)\big)$, where B is a random variable following distribution $F(\cdot|\tau=k)$ with $s_{\tau=k}(V)$ denote the contribution strategy for type k. We have two observations from (3) regarding identification of the model. First, cross-sectional observations of subjects' values and contributions are insufficient for identification. The cross-sectional data allow us to recover a relationship between values and contributions, which is the combined contributing strategies for all the types. Without prior information about the number of types as well as proportion and functional form of the strategy for each type, it is impossible to back out the contributing strategy for each type. Second, we have a short panel with multiple observations for each subject and the subject's identity; however, identification still requires a novel method. With a panel data of values and contributions being observed, a possible approach is to apply the method in Athey and Haile (2002) to recover a subject's contributing strategy using her multiple values and contributions under the assumption of invariant type. Such an approach requires a large number of observations for each subject. Nevertheless, it is unlikely that we have both type invariance and a long panel in both experimental and field data.

To solve the identification issues, we apply the recent development in the literature of measurement error, namely Hu (2008) to identify the model based upon (2) and (3). It is worth noting that our methodology of identification only requires researchers to observe the distribution of induced values but not individuals' values. This allows us to accommodate more flexible data structures, e.g., in many field data individuals' values are unknown but researchers may have prior information about the distribution of values.

3.2. Identification

We consider the case where M groups of subjects sequentially participate in T games of provision for the public good. The cost of the public good or the threshold is fixed for all the game. Similarly, the group size and the group members remain the same. As subjects' contributing strategies may depend on group size and cost, maintaining them fixed allows us to control for their effects when we conduct our analysis. Suppose we observe an i.i.d. sample $\{b^m_{it}, w^m_t\}, i=1,2,\cdots,I; m=1,2,\cdots,M; t=1,2,\cdots,T,$ where i, m and t indicate individual, group, and time period, respectively, and we use $N\equiv M\cdot I$ to denote the sample size or the total number of individuals. We assume individuals' values are unknown to researchers but the distribution is known. For ease of notation, we suppress the superscript m and subscript i. As will be shown, three periods of data ($T=3$) are sufficient for identification, hence the sample is denoted as $\{b_1, w_1; b_2, w_2; b_3; w_3\}$. As discussed previously, subject i's type in period t is denoted as $\tau_{it}\in\{1,2,\ldots,K\}$, where K is unknown, and the type may evolve across periods. The objectives of interest are: (1) number of type, (2) contributing strategy for each of the type, (3) the proportion of each type in the first period and (4) the transition matrix of type across period, or the "learning rules".

We start our identification strategy from a joint distribution of subjects' contributions and the provision outcome, b_1, b_2 and b_3 and w_2. By the law of total probability, we have

$$\begin{aligned} f_{b_3,w_2,b_2,b_1} &= \sum_{\tau_3}\sum_{\tau_2} f_{b_3,\tau_3,w_2,b_2,\tau_2,b_1} \\ &= \sum_{\tau_3}\sum_{\tau_2} f_{b_3|\tau_3,w_2,b_2,\tau_2,b_1} f_{\tau_3|w_2,b_2,\tau_2,b_1} f_{w_2|b_2,\tau_2,b_1} f_{b_2|\tau_2,b_1} f_{\tau_2,b_1}, \end{aligned} \tag{4}$$

where f_{R_1,R_2} and $f_{R_1|R_2}$ denote the joint and conditional densities R_1 and R_2 respectively. For simplicity of exposition, we still use the notation of $f(\cdot)$ when R_1 and/or R_2 are discrete whenever there is no ambiguity. Let $\Omega_{-t}\equiv\{(w_s, b_s, \tau_s)\text{for } s=1,\ldots,t-1\}$ be a set of information available for subjects prior to period t, where τ_s contains all the subjects' types from period 1 to $t-1$. Our first assumption specifies the dependence of subjects' contributions on the information set Ω_{-t}.

Assumption 1. A subject's contribution in each period is only determined by her induced value and her current type, i.e., $b_{it}=s_{it}(v_{it},\tau_{it})$, which is $s_k(v_{it})$ for $\tau_{it}=k$, $k\in\{1, 2, \cdots, K\}$.

This assumption excludes the dependence of the current contribution on the preceding information Ω_{-t}. It states all the information available to a subject is absorbed into her current type. That is, a subject sufficiently utilizes the history of outcomes, her contributions and strategies to determine the strategy at the current period, which implies that the type is a "sufficient statistic" of the information set Ω_{-t}. This assumption simplifies the conditional density $f_{b_3|\tau_3,w_2,b_2,\tau_2,b_1}$ as $f_{b_3|\tau_3}$ and $f_{b_2|\tau_2,b_1}$ as $f_{b_2|\tau_2}$. Accordingly, (4) can be rewritten as

$$f_{b_3,w_2,b_2,b_1}=\sum_{\tau_3}\sum_{\tau_2} f_{b_3|\tau_3} f_{\tau_3|w_2,b_2,\tau_2,b_1} f_{w_2|b_2,\tau_2,b_1} f_{b_2|\tau_2} f_{\tau_2,b_1}. \tag{5}$$

In the equation above, $f_{w_2|b_2,\tau_2,b_1}$ is the probability that the public good is provided successfully in period 2 for $w_2 = 1$. Recall that $w_2 = 1$ only if the summation of all the contributions in this period exceeds the cost, $\sum_{j=1}^{I} b_{j2} \geq c$ hence the probability $f_{w_2|b_2,\tau_2,b_1}$ is independent of τ_2 conditional on b_2, i.e., $f_{w_2|b_2,\tau_2,b_1} = f_{w_2|b_2,b_1}$. The conditional probability $f_{\tau_3|w_2,b_2,\tau_2,b_1}$ captures the transition process of subjects' type from period $t = 2$ to $t = 3$. Similar to Assumption 1, we impose some restrictions on how subjects' type evolves.

Assumption 2. The contributing strategy in the next period for a subject only depends on the outcome of provision and her contributing strategy in the current period.

Under this assumption, the transition of types $\Pr(\tau_{t+1}|w_t, b_t, \tau_t, \Omega_{-t})$ can be simplified to $\Pr(\tau_{t+1}|w_t, \tau_t)$. The restriction imposed by this assumption is twofold: first, the history Ω_{-t}, especially outcomes before period t play no role in subjects' learning rule given the current period's information. We do not rule out the possibility that subjects consider the information Ω_{-t}, however, it's irrelevant under Assumption 1 since the current type τ_t absorbs the history Ω_{-t}. This leaves us with the transition probability being $\Pr(\tau_{t+1}|w_t, b_t, \tau_t)$. This part of the assumption is also supported by the data pattern in the proceeding section, where w_{t-2} has little impact on b_{it} after controlling w_{t-1}. Second, a subject's contribution in the preceding period has no impact on her strategy for this period given the previous outcome and her previous strategy. This restriction is a natural consequence of the independence of subjects' values across periods: since values are independent, a subject can only learn from the provision outcome and her type in the last period. Intuitively, the contribution b_t contains no additional information other than τ_t for subjects with independent values across period. Nevertheless, it is worth noting that the independence of type τ_{t+1} and the information set Ω_{-t} is an assumption of first-order Markov process, which is widely used in the literature, and it can be relaxed when more periods of data are available for each subject.

Under Assumption 2 we further simplify (5) as

$$f_{b_3|w_2,b_2,b_1} f_{b_1,b_2} = \sum_{\tau_2} f_{b_3|w_2,\tau_2} f_{b_2|\tau_2} f_{\tau_2,b_1}. \tag{6}$$

Integrating out b_2 on both sides of the equation above, we obtain

$$\int f_{b_3|w_2,b_2,b_1}(\cdot|\cdot,\cdot,u) f_{b_1,b_2}(\cdot,u) du = \sum_{\tau_2} f_{b_3|w_2,\tau_2} f_{\tau_2,b_1}. \tag{7}$$

The two equations above provide a structural link between directly observed objectives on the L.H.S. and unknowns on the R.H.S. Following Hu (2008), we adopt a matrix form of Eqs. (6) and (7) for the purpose of identification. Specifically, we discretize the contributions b_1 and b_3, which are both continuous variables, as L values and denote the discretized contributions as d_1 and d_3, respectively.[8]

For a given outcome $w_2 \in \{0, 1\}$, and discretized contributions d_1 and d_3, we define the following matrices:

$$\begin{aligned} A_{ij} &\equiv \Pr(d_3 = i|w_2, b_2, d_1 = j) f(d_1 = j, b_2), \\ E_{ij} &\equiv \int \Pr(d_3 = i|w_2, b_2, d_1 = j) f(d_1 = j, b_2) db_2, \\ (B_{d_3|w_2,\tau_2})_{i,k} &\equiv [\Pr(d_3 = i|w_2, \tau_2 = k)]_{ik}, \\ (C_{\tau_2,d_1})_{k,j} &\equiv [\Pr(\tau_2 = k, d_1 = j)]_{kj}, \\ D_{b_2|\tau_2} &\equiv \mathrm{diag}[f(b_2|\tau_2 = 1) \quad f(b_2|\tau_2 = 2) \quad \cdots \quad f(b_2|\tau_2 = K)]. \end{aligned} \tag{8}$$

All the matrices are pointwise in b_2, where A and E are of dimension $L \times L$, B, C and D are of dimension $L \times K$, $K \times L$ and $K \times K$, respectively, where the number of types K is still unknown. Similar to their continuous counter-parts, the matrices defined above describe the distributions of observed and unobserved variables. For example, the (i, k)-th element in $B_{d_3|w_2=0,\tau_2}$ is the probability that the discretized contributions of the third period for those subjects who are of type k is in the i-th segment given the second period's outcome is "not provided". The k-th element of $D_{b_2|\tau_2}$ is the density $f_{b_2|\tau_2}$ for the type $\tau_2 = k$ evaluated at b_2.

The matrices defined above allow us to express (6) and (7) in a matrix form as follows:

$$\begin{aligned} A &\equiv B_{d_3|w_2,\tau_2} D_{b_2|\tau_2} C_{\tau_2,d_1}, \\ E &\equiv B_{d_3|w_2,\tau_2} C_{\tau_2,d_1}. \end{aligned} \tag{9}$$

[8] The discrete contribution d_t is determined by the following method of discretization.

$$d_t = \begin{cases} 1 & \text{if } b_t \in [\underline{b}, b_t(1)], \\ 2 & \text{if } b_t \in (b_t(1), b_t(2)], \\ \ldots & \\ L & \text{if } b_t \in (b_t(L-1), \overline{b}], \end{cases}$$

where the support of contribution, $[\underline{b}, \overline{b}]$ is divided into L segments by the $L-1$ cutoff points $b(1), b(2), \ldots, b(L-1)$, $\underline{b} < b(1) < b(2) < \ldots < b(L-1) < \overline{b}$, and $d_t \in \{1, 2, \ldots, L\} (L \geq 2)$ is the discretized contribution. Both d_1 and d_3 take values from $\{1, 2, \ldots, L\}$, however, the cutoff points for discretizing b_1 and b_3 can be different. Then $\Pr(d_t = l) \equiv \int_{b_t(l-1)}^{b_t(l)} f_{b_t}(u) du$.

For a given value of w_2, the matrix E describes the joint distribution of two discretized contributions d_1 and d_3. As argued in An (2017), the rank of this matrix can be used to identify the number of types under two conditions: first, the support of τ_t does not change along with t; second, contribution distribution of any type is not a linear combination of those for other types. We employ this insight here and make the following assumption.

Assumption 3. The inverse contributing functions $s_k^{-1}(\cdot)$ for $k = 1, 2, \cdots, K$ are linearly independent. Formally, there does not exist some $c_k \in \mathbb{R}, k = 1, 2, \cdots, K$ not all zero such that $\sum_{k=1}^{K} c_k s_k^{-1}(b) = 0$ for all $b \in [0, \bar{v}]$.

The restrictions imposed by this assumption on the inverse contributing strategies $s_k^{-1}(\cdot)$ can be described as a nonzero Wronskian if $s_k^{-1}(\cdot)$ has $(K-1)$-th continuous derivatives (see e.g., chapter 2 in Shilov (2013) for details). Recall that the distribution of contributions for type k, $F(b|\tau = k)$ is equal to $G\big(s_k^{-1}(b)\big)$, which can be further simplified as $s_k^{-1}(b)/(\bar{v} - \underline{v})$ because the induced values are uniformly distributed in our experiment. Thus Assumption 3 implies that the distributions of contributions for different types are linearly independent. We require the linear independence holds regardless of the conditioning on the outcome. The unconditional linear independence implies that the row rank of C_{τ_2, d_1} is equal to K, the number of types. Similarly, the linear independence conditional on the outcome w_2 guarantees that the column rank of $B_{d_3|w_2,\tau_2}$ for any $w_2 \in \{0, 1\}$ is also K. The essential restriction of this assumption is that there are enough variations of contributing strategies across type. Recall that the values of subjects who are of different types are i.i.d. It is unlikely that two different mappings from values to contributions (contributing strategies) lead to linearly dependent distributions of contributions. Similar assumptions of full rank have been widely imposed to identify structural models in econometrics. For example, in Newey and Powell (2003) and Chernozhukov et al. (2007) the full rank condition is essential for the identification of nonparametric instrumental variable models.

Lemma 1. *Under Assumptions 1–3, the number of types $K = rank(E)$.*

The proof of this lemma is similar to that of lemma 2 in An (2017). This lemma provides an important guidance to choose L, the number of values the discretized b_1 and b_3 take. Specifically, $L = \min\{l : det(E_{l\times l}) = 0\} - 1$. More details of the discretization will be discussed in estimation.

The assumption of invertibility implies $E^{-1} = C_{\tau_2,d_1}^{-1} B_{d_3|w_2,\tau_2}^{-1}$. Combining (9) with the relationship above, we obtain

$$A \times E^{-1} = B_{d_3|w_2,\tau_2} D_{b_2|\tau_2} B_{d_3|w_2,\tau_2}^{-1}, \tag{10}$$

where $D_{b_2|\tau_2}$ and $B_{b_3|w_2,\tau_2}$ are matrices of eigenvalues and eigenvectors, respectively for the observed matrix $A \times E^{-1}$. Especially, each of the diagonal element of $D_{b_2|\tau_2=k}, k \in \{1, 2, \ldots, K\}$ is the density of contributions for subjects of type k evaluated at b_2. Employing the strategies of identification proposed in Hu (2008), if the matrix decomposition in (10) is unique, then both $B_{d_3|w_2,\tau_2}$ and $D_{b_2|\tau_2}$ are identified since the L.H.S of the equation can be recovered from data.

To achieve the uniqueness of the decomposition, it is necessary to normalize the eigenvector matrix $B_{d_3|w_2,\tau_2}$ and make the eigenvector unique for each given eigenvalue. Considering that for a given outcome $w_2 \in \{0, 1\}$, each element in the eigenvector matrix $B_{d_3|w_2,\tau_2}$ is a conditional probability, hence each column of the matrix sums up to one, i.e., $\sum_{d_3} B_{d_3|w_2,\tau_2} = 1$. Then a plausible method of normalization is to divide each column by the corresponding column sum. To achieve the uniqueness of eigenvector for each eigenvalue, it is necessary for the eigenvalues to be distinctive, which is guaranteed by the following lemma.

Lemma 2. *If subjects' values are uniformly distributed, then the distributions of contributions for any two different types of subjects are distinct, i.e., for any two different types $k, j \in \{1, 2, \cdots, K\}$, the density $f_{b|\tau}(b|\tau = k)$ is different from $f_{b|\tau}(b|\tau = j)$, i.e., the set $\{b : f_{b|\tau}(b|\tau = k) \neq f_{b|\tau}(b|\tau = j)\}$ has nonzero Lebesgue measure.*

It is straightforward to prove Lemma 2 because of the link between the cdf of b and the inverse contributing function,

$$F(b|\tau = k) = \frac{s_k^{-1}(b) - \underline{v}}{\bar{v} - \underline{v}}, \quad k = 1, \cdots, K. \tag{11}$$

The equation above allows us to recover the contributing strategies for $\tau_2 = 1, 2, \cdots, K$, i.e.,

$$s_k^{-1}(b) = (\bar{v} - \underline{v})F(b|\tau = k) + \underline{v}, \quad k = 1, 2, \cdots, K.$$

The result in Lemma 2 is testable from (10) because once we obtain all the eigenvalues for each contribution b_2, it is straightforward to verify whether the result is violated, i.e., whether there exist at least two types whose distributions of contributions are always the same for any b_2.

Assumption 4. There exists a functional $\mathcal{F}$ such that $\mathcal{F}(s_k^{-1}(b)), k = 1, 2, \cdots, K$ can be completely ordered.

This assumption guarantees that the contributing strategies of K types can be strictly ordered. The choice of $\mathcal{F}$ is flexible: it can be a known quantile of $b \in [0, \bar{v}]$ or the mean. For example, let $b_{0.5}$ be the median of the contribution b, then a possible condition to order $s_k^{-1}(\cdot)$ is $s_1^{-1}(b_{0.5}) > s_2^{-1}(b_{0.5}) > \cdots > s_K^{-1}(b_{0.5})$, which implies subjects of type 1 would have the largest value to contribute $b_{0.5}$ and type K have the smallest value, i.e., type 1 contributes the least. The restriction of this assumption is flexible and in the following identification we assume that the average contributions for different types can be ordered. Recall that $s_k^{-1}(b)/(\bar{v}-\underline{v}) = F(b|\tau = k)$, Assumption 4 implies that we can distinguish different types according to their average contribution. Without loss of generality, we always label types in an ascending order according to expected contribution, i.e., on average type 1 contributes the least while type K contributes the most. The approach to label the types is consistent with the findings in the literature of public good. For example, in Fischbacher and Gächter (2010) the three types free riders, learners and contributors are classified according to how much they contribute. By imposing Assumption 4, the ordering of eigenvalues (eigenvectors) is fixed and the eigenvector matrix $B_{d_3|w_2,\tau_2}$ is uniquely determined from the eigenvalue-eigenvector decomposition of the observed matrix $A \times E^{-1}$.

For each period, the observed distribution of contributions is a weighted average of distributions for all the possible types, i.e.,

$$f(b) = \sum_{\tau} f(b|\tau)\Pr(\tau). \tag{12}$$

This relationship allows us to identify the proportion of each type $\Pr(\tau_2)$ in period 2 once the distribution for each type $f(b_2|\tau_2)$ is identified from the eigenvalue-eigenvector decomposition. In summary, all the important components of the model are identified from (10) and the results are summered as follows.

Proposition 1. *Under Assumptions 1–4, the distribution of contributions in period 3 conditioning on the outcome and type in the last period* $\Pr(d_3|w_2,\tau_2)$, *the distribution of contributions* ($f_{b_2|\tau_2}$) *and the proportion for each type* ($\Pr(\tau_2)$) *in period 2 are uniquely determined by the joint distribution of outcome in period 2 and contributions in three periods,* f_{b_3,w_2,b_2,b_1}. *Furthermore, if the distribution of values is known, the contributing strategy of each type* $s_k(\cdot)$ *is also identified.*

Based on the results of identification in Proposition 1, we show next that the two learning rules $\Pr(\tau_3|w_2,\tau_2)$ and $\Pr(\tau_2|w_1,\tau_1)$ are identified, too. First of all, the identified distribution of period 3 conditional on the outcome and type in period 2, $\Pr(d_3|w_2,\tau_2)$ is associated with the learning rule $\Pr(\tau_3|w_2,\tau_2)$ as

$$\begin{aligned}\Pr(d_3|w_2,\tau_2) &= \sum_{\tau_3}\Pr(d_3|\tau_3,w_2,\tau_2)\Pr(\tau_3|w_2,\tau_2)\\ &= \sum_{\tau_3}\Pr(d_3|\tau_3)\Pr(\tau_3|w_2,\tau_2).\end{aligned} \tag{13}$$

It is necessary to utilize an important implication of our model: the distribution of subjects' contributions for a certain type is invariant across periods, i.e., $f_{b_3|\tau_3} = f_{b_2|\tau_2} = f_{b_1|\tau_1}$. This conclusion is due to the fact that the provision game is homogeneous and subjects' values are i.i.d. in each period, therefore, the distribution of contributions must remain the same for each type in different periods. Using this property, $\Pr(d_3|\tau_3)$ can be obtained from the identified conditional density $f_{b_3|\tau_3} = f_{b_2|\tau_2}$, and the learning rule $\Pr(\tau_3|w_2,\tau_2)$ is identified from (13). We exemplify the procedure by assuming subjects are of two types, and correspondingly the discretized contribution d_3 takes two values. Then the aforementioned equation can be expressed in a matrix form:

$$\begin{aligned}&\begin{bmatrix}\Pr(d_3=1|w_2,\tau_2=1) & \Pr(d_3=1|w_2,\tau_2=2)\\ \Pr(d_3=2|w_2,\tau_2=1) & \Pr(d_3=2|w_2,\tau_2=2)\end{bmatrix}\\ &= \begin{bmatrix}\Pr(d_3=1|\tau_3=1) & \Pr(d_3=1|\tau_3=2)\\ \Pr(d_3=2|\tau_3=1) & \Pr(d_3=2|\tau_3=2)\end{bmatrix} \times \begin{bmatrix}\Pr(\tau_3=1|w_2,\tau_2=1) & \Pr(\tau_3=1|w_2,\tau_2=2)\\ \Pr(\tau_3=2|w_2,\tau_2=1) & \Pr(\tau_3=2|w_2,\tau_2=2)\end{bmatrix},\end{aligned} \tag{14}$$

where $w_2 \in \{0, 1\}$. This is a linear system and the learning rule $\Pr(\tau_3|w_2,\tau_2)$ can be uniquely solved from it only if the first matrix on the R.H.S. is full rank, which is guaranteed under Assumption 3. A similar argument can be applied to identify the learning rule of subjects from the first to the second period $\Pr(\tau_2|w_1,\tau_1)$. Alternatively, we might identify the learning rule as follows. Considering the observed joint density of contribution b_2, b_1 and the outcome w_1, f_{b_2,w_1,b_1}, we employ the law of total probability to obtain

$$\begin{aligned}f_{b_2,w_1,b_1} &= \sum_{\tau_2}\sum_{\tau_1} f_{b_2,\tau_2,w_1,b_1,\tau_1}\\ &= \sum_{\tau_2}\sum_{\tau_1} f_{b_2|\tau_2,w_1,b_1,\tau_1} f_{\tau_2|w_1,b_1,\tau_1} f_{w_1|b_1,\tau_1} f_{b_1|\tau_1} f_{\tau_1}\\ &= \sum_{\tau_2}\sum_{\tau_1} f_{b_2|\tau_2}\Pr(\tau_2|w_1,\tau_1) f_{w_1|b_1} f_{b_1|\tau_1} f_{\tau_1},\end{aligned} \tag{15}$$

where the first two equalities hold without any assumption and the third equality is due to Assumptions 1 and 2. In the equation above, the L.H.S. as well as $f_{w_1|b_1}$ are directly observed from the data. The distribution for each type $f_{b_2|\tau_2} = f_{b_1|\tau_1}$ and f_{τ_1} are identified using Proposition 1.

Proposition 2. *Under Assumptions 1–4, the learning rules regarding how subjects adjust their contributing strategies* $f_{\tau_2|w_1,\tau_1}$ *and* $f_{\tau_3|w_2,\tau_2}$ *are uniquely determined by the joint distribution of outcomes and contributions in three periods,* f_{b_3,w_2,b_2,w_1,b_1}.

The results of identification in Propositions 1 and 2 are constructive and they suggest a convenient multi-step procedure for estimation. We discuss the procedure briefly and leave the technical details of estimation in Appendix B. The first step of estimation is to determine the number of types by testing the rank of the matrix E. Next, by the eigenvalue-eigenvector decomposition in (10), we obtain the eigenvector matrix as well as the conditional distribution of contributions for each type in the second period, where the L.H.S. of (10) is estimated nonparametrically by kernel estimation. Consequently, the corresponding probability of each type can be estimated from (12). Lastly, based on (15) the learning rules are estimated by maximum likelihood estimation (MLE) since the learning rule only contains $K \times K$ parameters, where K is the number of types.

4. Monte Carlo experiments

In this section, we present some Monte Carlo evidence to demonstrate the performance of our estimator. We consider a game of public good provision similar to the experimental setting in Section 2. The game is played by groups with size $m = 5$ for three periods ($T = 3$). Values V_{it} are drawn from a standard uniform distribution and independent across individuals and over the three periods. The cost of the public good is set to be $c = 0.6 \times E[V_{it}] \times m = 1.5$. Individuals are of three types with their contributing strategies respectively being as follows:[9]

$$s_1(v) = \sqrt{v+1} - 1, \quad s_2(v) = \frac{2v}{3}, \quad s_3(v) = \Phi^{-1}\Big((\Phi(1) - \Phi(0))v + \Phi(0)\Big), \tag{16}$$

where $\Phi(\cdot)$ is the cumulative distribution function for the standard normal distribution. Notice that all the three strategies are strictly increasing in value on the support [0, 1].

Starting from period $t = 1$, we randomly draw N values from a standard uniform distribution $U[0, 1]$, then assign one of the three types to the N individuals according to the probability $\Pr(\tau_1 = 1) = 0.4, \Pr(\tau_1 = 2) = 0.3$ and $\Pr(\tau_1 = 3) = 0.3$. After we simulate the contributions for all the individuals based on their values and the contributing strategies in (16), the indicator of outcome w_1 is generated as $w_1 = 1(\sum_{m=1}^{5} b_{1m} \geq c = 1.5)$. Conditioning on w_1 and individuals' type τ_1 in period $t = 1$, we simulate their type τ_2 in period $t = 2$ according to the following transition matrix of types:

$$f(\tau'|\tau, w = 1) = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.2 & 0.6 & 0.4 \\ 0.3 & 0.1 & 0.4 \end{bmatrix}, f(\tau'|\tau, w = 0) = \begin{bmatrix} 0.8 & 0.1 & 0.2 \\ 0.1 & 0.7 & 0.6 \\ 0.1 & 0.2 & 0.2 \end{bmatrix},$$

where τ' indicates the type in the next period. For simplicity, it is assumed that the two transition matrices are invariant across periods. For example, if an individual was type 1 in a certain period, and her group successfully provides the public good, then she will be type 1, 2, and 3 with probabilities 0.5, 0.3 and 0.2, respectively in the next period. By applying this procedure repeatedly, we simulate a sample of contributions and outcomes $\{b_{i1}, w_1, b_{i2}, w_2, b_{i3}\}, i = 1, 2, \cdots, N$ for 1000 replications.

We first estimate the number of types through the rank of E defined previously using two alternative approaches. First, we use statistics of the condition number and determinant for the matrix E under the hypotheses of different number of types.[10] The condition number is a measure of how close a matrix is singular: a matrix with large condition number is nearly singular, whereas a matrix with condition number close to 1 is far from being singular. Alternatively, we may implement a more complicated but rigorous method to test the rank of E (e.g., Robin and Smith (2000)). In the simulation, we discretize $b_{it}, t = 1, 3$ into 2–6 segments, and compute the condition number and the determinant of the matrix E for each segment. Tables 2 presents the results for $w = 0, N = 500$ and $w = 1, N = 1000$, in the top and bottom panel, respectively, and the results for $w = 1, N = 500$ and $w = 0, N = 1000$ are similar, hence omitted for brevity.[11] As the results show, both the condition number and the determinant jump between 3 and 4 at different quantiles. For instance, the median of condition number for $N = 500$ jumps more than three-fold from 53 to 167.45 and a similar pattern is also observed for the case with $N = 1000$. The pattern of determinants is consistent with the condition number, and this offers some statistical confirmation for the rank of E, i.e., the number of types being three. We also use the method in Robin and Smith (2000) to test the rank of E and the result also supports 3 types. Table 3 presents the testing result, where b_1 and b_3 are discretized into 2 to 6 values. The null hypothesis H_0 is that $rank(E) = r$ and the alternative hypothesis H_1 is that $rank(E) > r$. For both $N = 500$ and $N = 1000$, we see that there are straight rejections at 5% significance level across all discretization methods for $r = 1$ and $r = 2$. For $r = 3$, we fail to reject the null hypothesis at 5% significance level when $M > 3$. As a result, we are confident to determine the number of types to be 3. Note that in the table, the test is invalid whenever $M \leq r$ since a $r \times r$ has a rank of at most r.

[9] For simplicity, we assume away the dependence of contributing strategies on group size. Nevertheless, the estimation still relies on group size because the winning indicator w is determined by contributions and the cost c, which is a linear function of group size m.

[10] Condition number of a matrix A is defined as $||A|| \cdot ||A^{-1}||$, where $||\cdot||$ is a matrix norm. We adopt the Euclidean norm, i.e., $||A||_2$, which is defined as the largest eigenvalue of the matrix $A'A$.

[11] The results are obtained by discretizing the subjects' contributions equally on the support. A different approach of discretization might change the reported numbers but the pattern that both condition number and determinant jump from 3 to 4 does not change with discretizations. We use the Epanechnikov kernel function and choose the bandwidth to be $2N^{-0.2}$.

Table 2
Identification of number of types.

Discretize Level	2	3	4	5	6
		$w = 0, N = 500$			
Condition Number					
Mean	190.03	126.14	792.29	2124.35	1164.67
25 percentile	18.40	26.86	101.93	93.31	138.51
Median	32.03	53.00	167.45	199.17	247.45
75 percentile	63.20	93.47	425.37	490.10	590.13
Determinant					
Mean	4.83E-03	-2.75E-02	2.16E-04	1.96E-07	-3.49E-05
25 percentile	-4.66E-04	-3.01E-05	-6.54E-08	-8.15E-11	-9.75E-13
Median	1.08E-03	-1.02E-06	-2.99E-10	0.00E+00	-7.24E-18
75 percentile	5.05E-03	4.48E-06	8.59E-09	1.92E-10	2.14E-13
		$w = 1, N = 1000$			
Mean	115.89	145.99	1307.84	774.56	849.63
25 percentile	20.54	32.48	104.16	135.55	142.82
Median	35.87	60.07	196.40	220.76	253.25
75 percentile	77.65	116.02	440.30	461.40	476.87
Determinant					
Mean	1.50E-04	9.20E-05	4.24E-07	5.91E-09	5.54E-13
25 percentile	-2.72E-04	-4.61E-06	-8.23E-09	-4.79E-11	-6.84E-14
Median	8.53E-04	2.41E-07	1.60E-10	-8.34E-14	0.00E+00
75 percentile	2.95E-03	1.09E-05	1.12E-08	1.52E-11	6.16E-14

Table 3
Monte Carlo: results of Rank test.

	p-value	$M = 2$	$M = 3$	$M = 4$	$M = 5$
$N = 500$	$r = 1$	0.000	0.000	0.000	0.001
	$r = 2$	N/A	0.024	0.018	0.014
	$r = 3$	N/A	N/A	0.074	0.254
$N = 1000$	$r = 1$	0.000	0.000	0.000	0.000
	$r = 2$	N/A	0.001	0.001	0.017
	$r = 3$	N/A	N/A	0.135	0.219

Table 4
Estimate of type probability.

	type 1	type 2	type 3
True value	0.40	0.30	0.30
$N = 500$	0.400***	0.301***	0.299***
	(0.047)	(0.059)	(0.045)
$N = 1000$	0.401***	0.297***	0.301***
	(0.032)	(0.041)	(0.031)

Standard errors in parentheses, * $p < .10$, ** $p < .05$, *** $p < .01$.

The estimates of contributing strategies for three types, together with the corresponding [10%, 90%] point-wise confidence intervals are illustrated in Figs. 3. The estimates perform well for modest sample-size datasets of $N = 500, 1000$. For the three types, the estimated contributing strategies track the actual ones very closely. Notice that for types 1 and 2 the estimated contribution is larger than the true one when the value is close to its upper bound. This is because the estimate of contribution distribution is less accurate when a contribution is close to the upper bound due to the sparse observations.

We provide the estimation of the initial type probabilities, i.e., at period $t = 1$ in Table 4. The probabilities are accurately estimated for both sample size. Tables 5 presents the estimated transition matrices of types. Notice that the estimate of $f(\tau'|\tau, w = 1)$ performs very well while that of $f(\tau'|\tau, w = 0)$ is a bit noisy, and this is because the observations for $w = 0$ is smaller than $w = 1$ due to the setting of our transition matrix of type. Nevertheless, both estimated matrices are accurate enough to capture the transition pattern of the type.

In summary, the Monte Carlo evidence illustrates that our procedure of estimation performs well for modest-sized samples.

Robustness check. In the simulation, we assume individuals' induced values are independent across periods in the data generating process; such assumption is automatically satisfied as induced values are randomly re-assigned to each subject every period in the experiment. However, this assumption might be violated for some field data where values may be correlated. As a robustness check, we allow values to be correlated across periods and then estimate the model by the proposed method

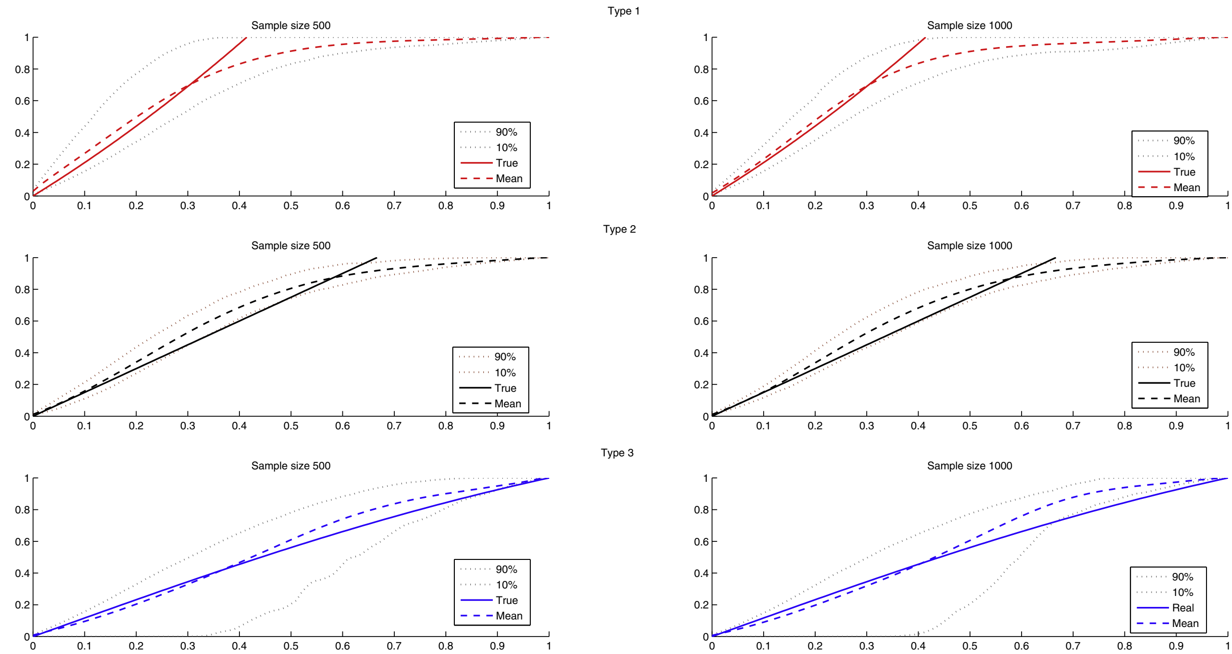

Fig. 3. Estimate of contributing strategies.

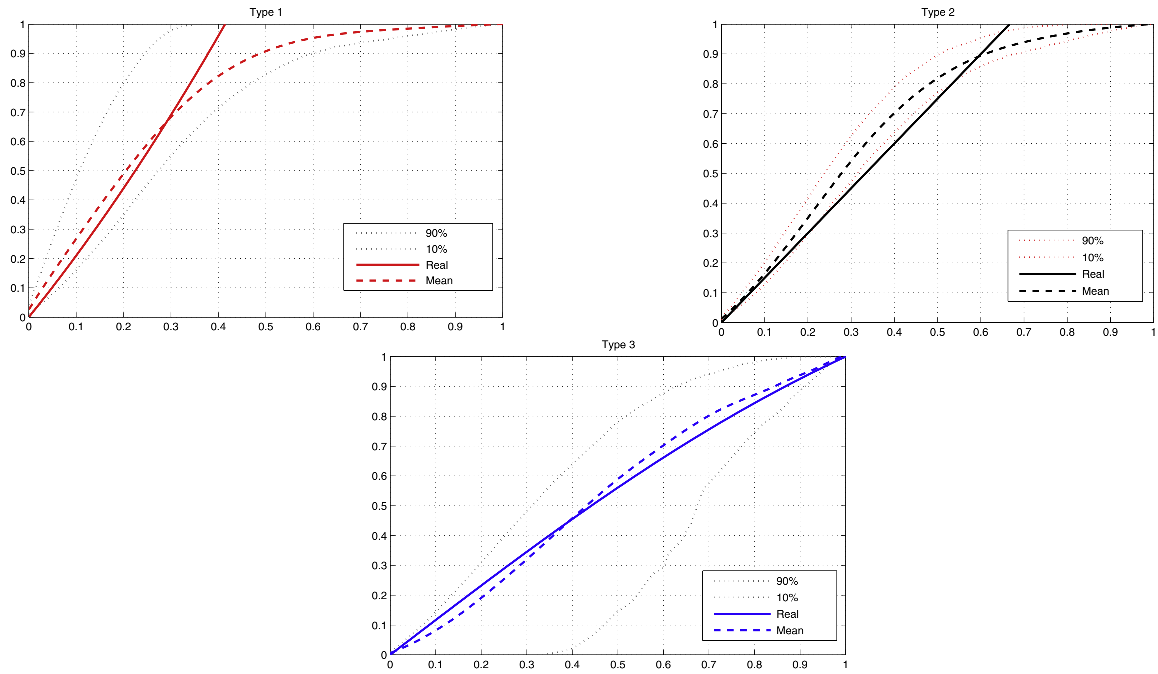

Fig. 4. Estimated contributing strategies with correlated values, $N = 500$.

assuming independence of the values.[12] Fig. 4 and the bottom panel in Table 5 present the estimated contributing strategies and the transition matrices, respectively for sample size 500 when the values of two consecutive periods are correlated with a coefficient 0.2. The estimate of strategies is very close to that in Fig. 3. Similarly, the probabilities of transition in Table 5 also closely track the corresponding elements in the true transition matrices. A comparison of the estimated results

[12] To generate the uniformly distributed values with Pearson correlation, we first generate normally distributed draws with Spearman correlations then apply the uniform transformation to those random draws. Please see Embrechts et al. (2003) for details.

Table 5
Estimated transition matrix of type.

	Type1	Type2	Type3
$\Pr(\tau_2 \mid w=1, \tau_1): N = 500$			
Type1	0.494***	0.308*	0.206
	(0.170)	(0.240)	(0.170)
Type2	0.226	0.543**	0.423**
	(0.193)	(0.285)	(0.225)
Type3	0.279***	0.150	0.370**
	(0.143)	(0.167)	(0.172)
$\Pr(\tau_2 \mid w=0, \tau_1): N = 500$			
	Type1	Type2	Type3
Type1	0.735***	0.218	0.213
	(0.192)	(0.279)	(0.288)
Type2	0.160	0.555*	0.618*
	(0.188)	(0.358)	(0.401)
Type3	0.105	0.227	0.169
	(0.110)	(0.234)	(0.250)
$\Pr(\tau_2 \mid w=1, \tau_1): N = 1000$			
	Type1	Type2	Type3
Type1	0.504***	0.317	0.176
	(0.125)	(0.203)	(0.136)
Type2	0.206	0.556**	0.441**
	(0.146)	(0.230)	(0.179)
Type3	0.291***	0.127	0.382***
	(0.102)	(0.115)	(0.130)
$\Pr(\tau_2 \mid w=0, \tau_1): N = 1000$			
	Type1	Type2	Type3
Type1	0.766***	0.192	0.168
	(0.153)	(0.236)	(0.239)
Type2	0.140	0.608**	0.615**
	(0.138)	(0.320)	(0.362)
Type3	0.099	0.204	0.215
	(0.084)	(0.190)	(0.244)
$\Pr(\tau_2 \mid w=1, \tau_1)$: corr. values$(N = 500)$			
	Type1	Type2	Type3
Type1	0.590***	0.258	0.104
	(0.152)	(0.216)	(0.134)
Type2	0.174	0.607**	0.408**
	(0.167)	(0.275)	(0.214)
Type3	0.236**	0.135	0.488***
	(0.135)	(0.175)	(0.193)
$\Pr(\tau_2 \mid w=0, \tau_1)$: corr. values$(N = 500)$			
	Type1	Type2	Type3
Type1	0.843***	0.207	0.196
	(0.157)	(0.295)	(0.314)
Type2	0.087	0.593**	0.607*
	(0.144)	(0.359)	(0.394)
Type3	0.069	0.200	0.197
	(0.092)	(0.248)	(0.293)

in three panels of Tables 5 illustrates that the correlation of values negatively affects the accuracy of estimates. For example, the probability $\Pr(\tau_2 = 1 \mid w = 1, \tau_1 = 1)$ is estimated to be 0.494 in the top panel, which is close to the true value 0.50, while the estimate is 0.590 in the bottom panel. Nevertheless, with the modest correlation of values, our proposed method based on independence of values still performs well in estimating the model.

5. Empirical results

In this section, we use the experimental data described in Section 2 to identify heterogeneous contributing strategies and learning in the private provision of the threshold public good game.[13] To maximize the number of observations and explore

[13] As discussed in Section 3.2, three periods of data are sufficient for identification. Under the assumption that the learning rule is invariant across periods, each of the three consecutive periods of the 10 periods in our sample can be employed to estimate the learning rule. We reorganize 10 periods of data into {1, 2, 3}, {2, 3, 4}, ⋯{8, 9, 10} to increase the number of observations. Nevertheless, we keep the original data structure for the estimation of the initial type distribution.

Table 6
Estimation of number of types.

Discretize level	2	3	4	5	6
		$w = 0$			
Condition Number					
Original Sample	13.84	15.36	3.34E+03	9.62E+02	2.03E+03
Mean	21.58	289.23	3.12E+04	1.03E+05	7.63E+15
25 percentile	7.65	18.79	1.60E+03	1.02E+03	1.32E+03
Median	11.16	24.24	4.35E+03	2.28E+03	3.71E+03
75 percentile	13.90	38.30	1.36E+04	1.03E+04	9.99E+03
Determinant					
Original Sample	0.0050	4.12E-05	-1.27E-09	8.65E-13	4.87E-16
Mean	0.0071	2.40E-05	7.58E-10	8.14E-13	1.09E-16
25 percentile	0.0052	1.34E-05	-1.69E-10	-2.10E-14	-1.18E-17
Median	0.0065	2.15E-05	1.34E-10	1.64E-13	6.86E-18
75 percentile	0.0091	3.22E-05	1.71E-09	5.89E-13	6.16E-17
		$w = 1$			
Condition Number					
Original Sample	26.65	34.77	3.16E+03	8.95E+02	1.16E+03
Mean	20.95	65.52	1.33E+04	1.56E+16	1.38E+16
25 percentile	8.035	19.10	1.43E+03	1.15E+03	1.54E+03
Median	11.35	24.99	3.16E+03	2.49E+03	3.92E+03
75 percentile	15.81	35.20	9.65E+03	8.69E+03	1.03E+04
Determinant					
Original Sample	0.0026	1.53E-05	6.27E-10	2.33E-13	1.26E-16
Mean	0.0066	2.39E-05	9.60E-10	3.01E-13	3.84E-17
25 percentile	0.0044	1.35E-05	-2.10E-10	-3.97E-14	-1.33E-17
Median	0.0063	2.01E-05	1.20E-10	6.20E-14	6.43E-18
75 percentile	0.0090	2.85E-05	1.22E-09	4.20E-13	7.80E-17

Table 7
Results of Rank Test.

p-value	$M = 2$	$M = 3$	$M = 4$	$M = 5$
$r = 1$	0.000	0.000	0.028	0.016
$r = 2$	N/A	0.001	0.001	0.023
$r = 3$	N/A	N/A	0.147	0.220

the possibility that subjects may change their learning rules across periods, we consider the following three approaches to aggregate the observations: (a) pool all the 10 periods' data, where learning rule $\Pr(\tau'|w, \tau)$ is invariant between any two periods, (b) use only the first five periods' data and (c) use the last five periods' data, where we assume that subjects' learning behavior is invariant in the first and the last five periods, respectively. The results of (a) can be considered as the baseline and (b) and (c) are used for robustness checks.

5.1. Estimation results

The first set of results are condition numbers and determinants of matrix E, which are used to determine the number of types. Tables 6 presents the results when we pool the 10 periods' data together. The first and last five periods' of data both lead to very similar results and are hence omitted. The top panel of the table is conditional on the outcome that the public good is not provided, $w_2 = 0$ and the bottom panel is for the outcome $w_2 = 1$. Both panels reveal a clear pattern that the condition number and determinant jump when the number of discretization changes from 3 to 4, and this identifies the number of types to be 3. We also use the method in Robin and Smith (2000) to test the rank of E. The result in Table 7 also supports 3 types.

Next, the procedure of identification using matrix decomposition enables us to recover the conditional density $f_{b_2|\tau_2}$.[14] It is used to obtain the contributing strategies for three types and the initial type probability through f_{b_1} and $f_{b_2|\tau_2} = f_{b_1|\tau_1}$. The estimates of probability and contributing strategy for each type are provided in Table 8 and Fig. 5, respectively. We label the three types such that on average type 3 contributes the most, whereas type 1 the least. The results in Table 8 indicate that the proportion of each type is significantly positive. The first two rows are both estimates for the first period using different samples of data. Both rows reveal that the proportion of type 3 is the largest (about 40%), and that of type 1 and type 2 are slightly different. By contrast, the type probability of period 5 (using the data of the last five periods) displays a different pattern: the proportion of type 2 is the largest (about 38%) whereas the other two types have a proportion 25.3%

[14] In the first and this step of estimation, we discretize subjects' contributions b_1 and b_3 into equal intervals on their supports. In estimating the kernel density $\widehat{f}(b_2, d_3|w_2)$, we choose the Epanechnikov kernel function and the bandwidth to be $2N^{-0.2}$.

Table 8
Estimate of type probability in the first period.

	Type1	Type2	Type3
pooled data	0.236***	0.361***	0.403***
	(0.049)	(0.054)	(0.041)
the first 5 periods	0.334***	0.263***	0.403***
	(0.051)	(0.058)	(0.046)
the last 5 periods	0.253***	0.381***	0.365***
	(0.041)	(0.035)	(0.051)

Standard errors in parentheses, * $p < .10$, ** $p < .05$, *** $p < .01$.

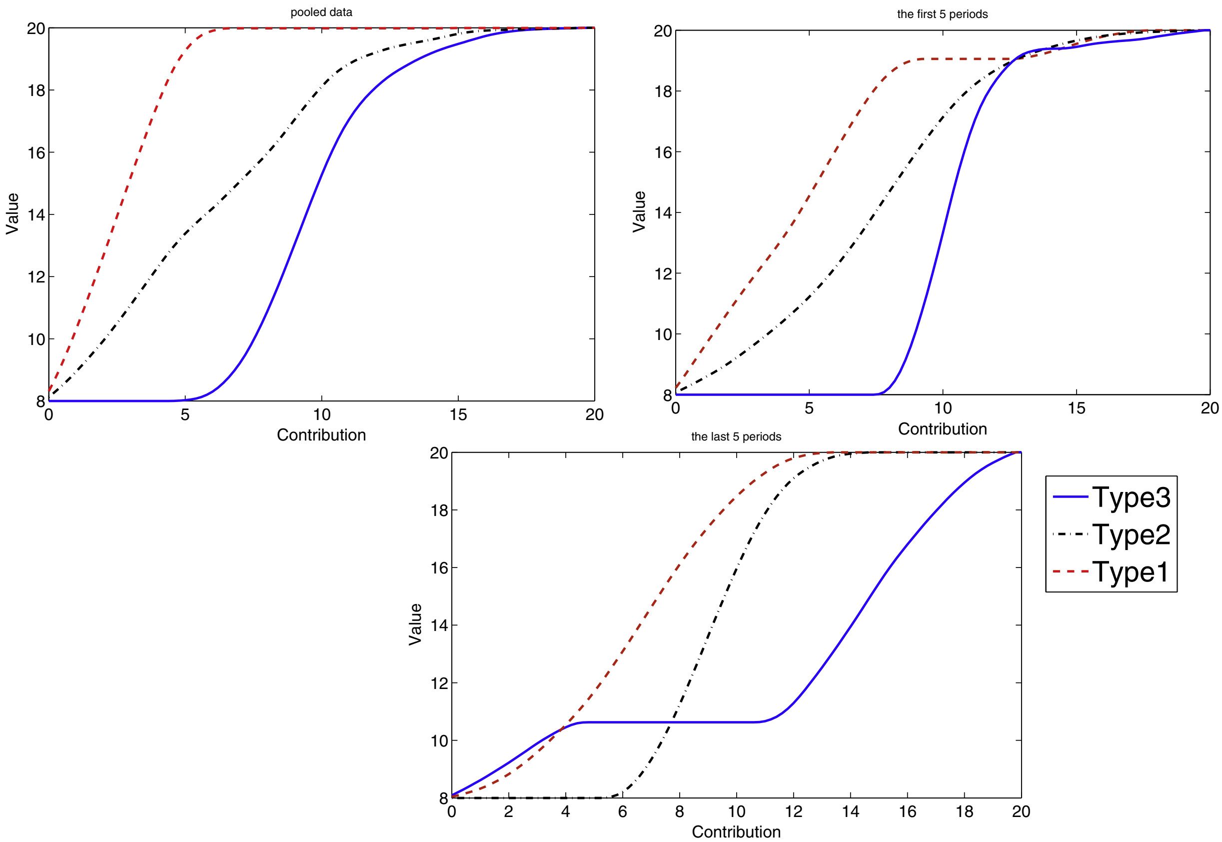

Fig. 5. Heterogeneous contributing strategies.

(type 1) and 36.5%. The results imply that subjects tend to contribute more at the beginning (i.e., the larger proportion of types 1 and 2). As they continue to participate in the game, more subjects may learn to contribute strategically and adapt the contributing strategies of lower types.

The three subplots in Fig. 5 illustrate the contribution as a function of the value for subjects of three types.[15] A formal test indicates that the contributing strategy of each type is significantly different from the other two at the 5% significance level.[16] This pattern holds across all the three subplots, which implies that during the 10 periods of the experiment, subjects' contributing strategies do not "converge" through learning. This result provides strong evidence that subjects contribute using heterogenous strategies.

The first subplot using the 10 periods' of data illustrates that type 1 and type 2 contribute significantly less than type 3, especially when a subject's value is small. When their values are near the lower bound 8, type 1 and type 2 behave as free-riders and contribute nothing; whereas type 3 contribute up to 5. However, as the value increases the three types contribute

[15] Note that some of the contributing strategies are not strictly increasing. This is due to our small sample size, and the strategies would be strictly increasing asymptotically.

[16] According to Lemma 2, testing the difference between two bidding strategies is equivalent to that of two distributions of contributions. Therefore, we conduct Kolmogorov-Smirnov tests on the distributions of contributions for any of the two types.

a significant proportion of their values, e.g., for a value 14 the contributions are 2.5, 5 and 9, which is about 17.9%, 35.7% and 64.3% of the value, respectively for type 1, 2, and 3. Furthermore, it is worth noting that the difference between type 2 and type 3 diminishes as the value increases. Especially, when a value is greater than 18, which is near the upper bound (the upper bound of value is 20), the contributions of the two types are very close to each other. However, type 1 only contributes 5 when his value is 20. The pattern revealed by the first subplot is comparable to the results in Fischbacher and Gächter (2010), where the three types are free riders, learners, and contributors. Nevertheless, we do not find free riders in our experiment: the contribution of type 1 is significantly less than the other two types but greater than zero.

The second and the third subplots present the results using the first and last five periods of data, respectively. A comparison between the two subplots illustrates how the three types of subjects may adjust their contributing strategies across period. First of all, in the last five periods, we observe that type 2 contributes a significant proportion of his value but type 3 contributes a similar proportion to type 1 when their values are small. This contrasts with the pattern in the first five periods. An implication is that type 2 and type 3 adjust their strategies differently based on the history of provision outcomes. Secondly, in the first five periods, the behavior of the three types is similar when their values are high, while in the last five periods only the discrepancy between type 1 and type 2 diminishes at a high value. These results show that (1) the three types contribute heterogeneously; (2) the three types adjust their contributing strategies heterogeneously, too. In addition, we do not find the evidence that subjects' strategies converge after they spend more time with other group members.

The last set of results are on the learning rule of different types and they are presented in Table 9, where each 3×3 matrix represents transition probabilities of types conditional on provision outcomes: each column contains the probabilities that a certain type is being adjusted to three types in the next period, hence the column sum is one. For example, in Table 9 (a) the second column implies that if the public good is provided successfully, then the subjects of type 2 in the current period would continue to be type 2 with probability 67.2% and adjust to be type 1 (contributes less) with probability 32.8% in the next period. However, they never transit to type 3 and contribute more.

Across all the three data scenarios, the matrices $\Pr(\tau_2|w=1, \tau_1)$ are close to upper diagonal or diagonally dominant, whereas $\Pr(\tau_2|w=0, \tau_1)$ are close to lower diagonal except the probabilities of type 3 in 9 (b) and 9 (f). Such a difference reveals that subjects' contributing strategies are negatively affected by the outcome, i.e., they will maintain their strategies or contribute less if the public good is provided successfully but contribute more or at least the same for a failure provision. In Table 9 (a) 73.2% of type 1, 67.2% of types 2 and 89.5% of type 3 would keep their own type in the next period conditioning on a successful outcome. Moreover, a substantial proportion of type 2 would change to type 1 and contribute less in the next period. By contrast, if the good is not provided in the current period, significant proportions of type 1 and type 2 transit to type 2 and type 3, respectively, and contribute more in the next period. In summary, the learning rules of subjects are heterogeneous and crucially rely on the experiment outcome in the preceding session. Nevertheless, the behavior of type 3 is a bit noisy: they may contribute less upon observing an unsuccessful outcome.

Another important observation is that the estimates in the middle and bottom panels of Tables 9 indicate different patterns of subjects' learning. Subjects in the first five periods are more reluctant to adjust their contributing strategies than in the last five periods. For instance, conditioning on an unsuccessful provision, with probability 54.8% of type 2 transits to type 3 in the last five periods. By contrast, this probability is only 46.4% for the first five periods. In response to a successful provision, almost all the subjects of type 1 remain as type 1 in the first five periods, whereas this probability is 82.9% in the last five periods. The difference between the first and the last five periods suggest that the learning can also be dynamic, which is out of the primary focus of our paper but provides an interesting venue for future research.

It may take one or two periods for the subjects to learn how exactly the game works, even though detailed instructions are given to the subjects before the experiment. To address this concern, we reestimate the model by using only the last 8 and 9 periods, respectively. Our results are qualitatively consistent with the results from using all 10 periods data, thus are omitted in the paper. Specifically, the percentage of type 1 is at least 20% and the result based on the last 8 periods suggests that type 3 has the highest percentage. The estimated learning rules also suggest that conditional on a positive outcome, subjects are more likely to stay on current strategy. Also, subjects are more likely to adjust contribution strategies upon a provision failure.

5.2. Implications of estimated results

The estimated number of types, contributing strategy of each type, and the learning rule are important for us to understand individuals' strategic behavior, predict the outcome of public good provision, and provide policy implications regarding the provision of public goods.

First of all, the estimated contributing strategies enable us to test the validity of various models that can be potentially used to explain subjects' behavior. Our analysis of contributing strategies does not rely on any equilibrium or behavioral assumptions of subjects' behavior: if the number of type $K = 1$, then subjects might follow BNE or BNE with risk aversion; if $K > 1$, the strategies are not at equilibrium and level-k thinking (Crawford and Iriberri, 2007) or cognitive hierarchy (Camerer et al., 2004) might be appropriate to interpret subjects' behavior. In either scenario, we can test which model explains the data better. This can be done by comparing our estimate of contributing strategies with that predicted by the model we try to test. The prediction of subjects' contributing strategies is possible because their value distribution is known. Such a testing procedure is similar to Bajari and Hortacsu (2005), where several alternative explanations of bidders' behavior

Table 9
Learning rules.

$\Pr(\tau_2 \| w=1, \tau_1)$: pooled data			
	Type1	Type2	Type3
Type1	0.732***	0.328*	0.105
	(0.087)	(0.198)	(0.191)
Type2	0.268	0.672***	0.000
	(0.144)	(0.155)	(0.030)
Type3	0.000	0.000	0.895***
	(0.165)	(0.135)	(0.195)
$\Pr(\tau_2 \| w=0, \tau_1)$: pooled data			
	Type1	Type2	Type3
Type1	0.661***	0.066	0.104
	(0.127)	(0.244)	(0.146)
Type2	0.291**	0.532**	0.251***
	(0.105)	(0.269)	(0.055)
Type3	0.048	0.402**	0.645***
	(0.158)	(0.195)	(0.141)
$\Pr(\tau_2 \| w=1, \tau_1)$: first 5 periods			
	Type1	Type2	Type3
Type1	0.942***	0.000	0.399
	(0.309)	(0.349)	(0.339)
Type2	0.000	1.000**	0.000
	(0.222)	(0.400)	(0.179)
Type3	0.058	0.000	0.601**
	(0.222)	(0.363)	(0.333)
$\Pr(\tau_2 \| w=0, \tau_1)$: first 5 period			
	Type1	Type2	Type3
Type1	0.616***	0.000	0.000
	(0.217)	(0.192)	(0.132)
Type2	0.384	0.536*	0.063
	(0.445)	(0.326)	(0.184)
Type3	0.000	0.464*	0.937***
	(0.432)	(0.245)	(0.211)
$\Pr(\tau_2 \| w=1, \tau_1)$: last 5 periods			
	Type1	Type2	Type3
Type1	0.829***	0.000	0.099
	(0.147)	(0.102)	(0.093)
Type2	0.171	0.849***	0.390***
	(0.141)	(0.080)	(0.057)
Type3	0.000	0.151	0.511***
	(0.057)	(0.105)	(0.080)
$\Pr(\tau_2 \| w=0, \tau_1)$: last 5 periods			
	Type1	Type2	Type3
Type1	0.502***	0.179	0.000
	(0.163)	(0.140)	(0.068)
Type2	0.381**	0.273***	0.220**
	(0.187)	(0.093)	(0.087)
Type3	0.117	0.548***	0.779***
	(0.215)	(0.103)	(0.114)

in first-price auctions are tested by comparing the known value distribution to those identified from different models. The testing procedure is beyond the scope of this paper and we leave the details in future research.

Second, our results can be used to predict group contribution dynamic across multiple periods, which can not be done using previous reduced form analyses. Such dynamic outcomes may provide important policy implications for the provision of public goods. Particularly, we consider how to predict provision outcomes for a group of individuals (group size is fixed at 5 and the cost is also 42) with known but different characteristics from our subjects. Suppose that there are three types in the population, the initial type distribution depends on observed characteristics of individuals, and the transition probabilities do not depend on the initial type distribution. The estimate of initial type distribution and subjects' characteristics in our experiments can be used to predict the initial type distribution of the new group of individuals by assuming e.g., type probabilities are logit function of characteristics. Based on the predicted initial type distribution, together with the learning rule and contributing strategies, we are able to simulate the contribution for each individual in every period.

Fig. 6 shows the simulation results and illustrates how initial type distribution affects provision outcome. In each of the seven subfigures, we plot the average contribution of 100 groups in dotted line and the horizontal line is the cost. We

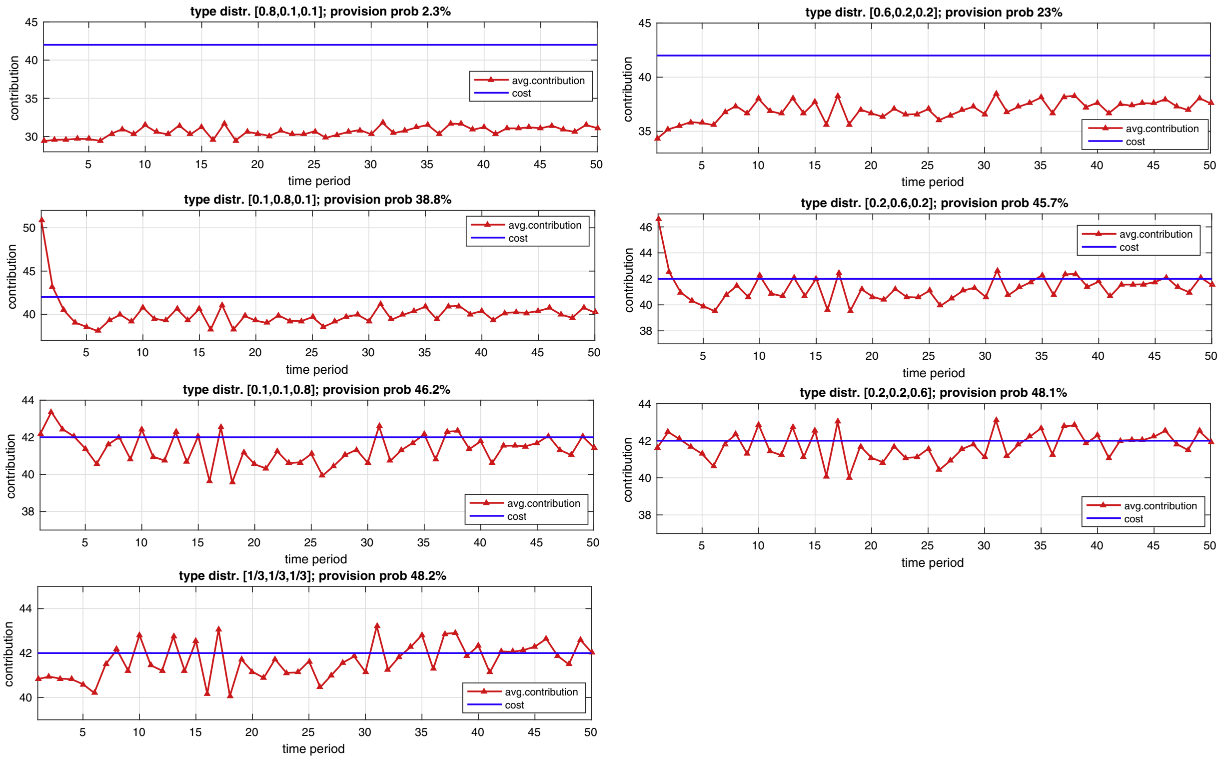

Fig. 6. Simulated group contribution outcomes.

also explicitly state the type distribution (e.g., [0.8, 0.1, 01] means the proportion of type 1, 2, and 3 is 0.8, 0.1, and 0.1, respectively) and the percentage of successful periods in the title of each subfigure. We find from the figure that when the low type dominates (the top panel), the average group contribution falls significantly below the provision cost and the group seldom provides the public good (the provision rate is 2.3% and 24% respectively). By contrast, when three types are equally distributed or type 3 dominates (the third and the last row, respectively), the average group contribution closely follow the provision cost across the periods and the group successfully provides the public good at a much higher frequency (48.2% for equally distributed type, and 46.2% and 48.1% for the case of type 3 dominating). The two medium cases fall in between in terms of group contribution and provision outcomes.

The simulation results could be useful for a policymaker who plans to conduct a public good provision. For example, if the predicted initial type distribution is close to [0.8, 0.1, 0.1], then given the very low provision rate simulated the policymaker should consider alternative strategies to improve the rate, e.g., decrease the cost or increase group size, etc.

6. Conclusion

We have studied the identification and estimation of a structural model for the private provision of a threshold public good allowing heterogeneous contributing strategies. The main motivation for our paper is the need to explain individuals' heterogeneous contributing behavior and possible adjustments of their strategies based on provision history. Individual heterogeneity and learning have been documented in previous experimental studies and also observed in our experimental data of threshold public good games. Our structural model allows for individuals to employ heterogeneous contributing strategies, which may be adjusted upon the outcome of the provision history. A prominent advantage of our approach over the existing studies is that we are able to recover the number of different strategies, function form of each strategy and the transition probability among the strategies without imposing any parametric assumptions on these objectives from individual revealed contributions.

The structural estimates of our experimental data suggest that subjects can be classified into three types with each type employing a different contributing strategy. A subject of type 1 contributes a much smaller share of her value than types 2 and 3, and type 2 and 3 exhibit distinct learning processes. The estimates of learning indicate that subjects tend to maintain their strategies in response to a successful provision in the last period. By contrast, they tend to contribute more if the good is not provided in the last period. Also, the three types display different patterns of learning. Particularly, type 1 makes the relative smaller adjustment than the other two types. By dividing the data into the first and last five periods, we find that subjects in the first five periods are more reluctant to adjust their contributing strategies than in the last five periods.

There are a few directions for future research. First, our methodology for threshold public good could be applied to linear public good provision, where the contributing strategy is a mapping from the endowment to the ratio of contributions made to public good over their own account. Furthermore, even though we allow individuals change their learning behavior across 10 periods, we leave out their forward-looking behavior. It will be interesting to incorporate dynamics into our model and explore individuals' dynamic learning behavior. Last but not least, the estimated contributing strategies could be used to test the validity of various equilibrium or non-equilibrium models that used to describe public goods provision games.

Appendix A. Estimation

In this section, we propose a procedure to estimate the objectives that are identified nonparametrically in Section 3 . The procedure follows directly from the argument of identification, and a similar approach is also applied in An et al. (2010). We estimate all the objectives in multiple steps.

Step one: Estimation of the conditional distribution $f(b_2|\tau_2)$. Recall our identification is mainly based on (Eq. (10)), which holds for all b_2. To improve the performance of our estimator, we take integral of this equation with respect to b_2 and use the aggregated version for estimation:

$$\int_{b_2} b_2 A \times E^{-1} db_2 = B_{d_3|w_2,\tau_2} D_{Eb_2|\tau_2} B^{-1}_{d_3|w_2,\tau_2} Eb_2, \tag{A.1}$$

where $D_{Eb_2|\tau_2} \equiv \int_{b_2} b_2 D_{b_2|\tau_2} db_2$. The L.H.S. of the equation above is estimable from data, then both $B_{b_3|w_2,\tau_2}$ and $D_{Eb_2|\tau_2}$ can be estimated by the eigenvalue-eigenvector decomposition described in (10). The details can be found in An et al. (2010) and An (2017), and thus omitted here. To construct the matrices A, B and E, we need to discretize b_1 and b_3 as discussed in footnote [8]. There are infinitely many different ways to discretize b_1 and b_3. Nevertheless, the method of discretization does not affect the result of decomposition asymptotically given Assumptions 1–3 hold. In practice, when sample is small, we may vary methods of discretization to ensure robustness of results.

Let $\widehat{B}_{b_3|w_2,\tau_2}$ be the estimated eigenvector matrix, we estimate the conditional density $f(b_2|\tau_2)$ from the joint density $f(b_2, \tau_2)$ and the probability distribution $\Pr(\tau_2)$. First we consider the relationship

$$f(b_2,\tau_2) = \sum_{w_2\in\{0,1\}} f(b_2,\tau_2|w_2)\Pr(w_2),$$

where $\Pr(w_2)$ can be directly recovered from data and the joint distribution of b_2, τ_2 conditional on the outcome w_2, $f(b_2,\tau_2|w_2)$ is determined by the following equation:

$$\begin{aligned} f(b_2,d_3|w_2) &= \sum_{\tau_2} f(b_2,d_3,\tau_2|w_2) = \sum_{\tau_2} f(d_3|w_2,b_2,\tau_2) f(b_2,\tau_2|w_2) \\ &= \sum_{\tau_2} f(d_3|w_2,\tau_2) f(b_2,\tau_2|w_2). \end{aligned}$$

The L.H.S. of the equation above is estimable from data, and $f(d_3|w_2,b_2,\tau_2)$ is obtained from the eigenvalue-eigenvector decomposition. Thus, we get an estimator of $f(b_2,d_3|w_2)$. We exemplify the estimation for $w_2=0$:

$$f(b_2,d_3|0) = B_{d_3|0,\tau_2} f(b_2,\tau_2|0) \Rightarrow \widehat{f}(b_2,\tau_2|0) = \widehat{B}^{-1}_{d_3|0,\tau_2}\widehat{f}(b_2,d_3|0),$$

where $\widehat{B}_{d_3|0,\tau_2}$ is invertible by construction, and $\widehat{f}(b_2,d_3|0)$ is a kernel estimator defined as:

$$\widehat{f}(b_2,d_3=j|0) = \frac{1}{Nh}\sum_{i=1}^{N} K\left(\frac{b_2-b_{2i}}{h}\right)\mathbf{1}(b_{3i}=j).$$

Consequently we have the estimator of the joint distribution (b_2, τ_2),

$$\widehat{f}(b_2,\tau_2) = \widehat{f}(b_2,\tau_2|0)\widehat{\Pr}(w_2=0) + \widehat{f}(b_2,\tau_2|1)\widehat{\Pr}(w_2=1) \tag{A.2}$$

Similarly, the type distribution $\Pr(\tau_2)$ can be estimated from

$$\Pr(\tau_2) = \sum_{w_2\in\{0,1\}} \Pr(\tau_2|w_2)\Pr(w_2),$$

where $\Pr(\tau_2|w_2)$ is associated with estimable $\Pr(d_3|w_2)$ and estimated $\Pr(d_3|\tau_2,w_2)$.

$$\begin{aligned} \Pr(d_3|w_2) &= \sum_{\tau_2}\Pr(d_3,\tau_2|w_2) = \sum_{\tau_2}\frac{\Pr(d_3|\tau_2,w_2)\Pr(\tau_2,w_2)}{\Pr(w_2)} \\ &= \sum_{\tau_2}\Pr(d_3|\tau_2,w_2)\Pr(\tau_2|w_2). \end{aligned}$$

We again illustrate our estimator for $w_2 = 0$. Let $\overrightarrow{\Pr}(d_3|0)$ denote a column vector with three elements $[\Pr(d_3 = 1|w_2 = 0)\ \Pr(d_3 = 2|w_2 = 0)\ \Pr(d_3 = 3|w_2 = 0)]^T$, and $\overrightarrow{\Pr}(\tau_2|0)$ is similarly defined. Then the last equation can be rewritten as

$$\overrightarrow{\Pr}(d_3|0) = B_{d_3|0,\tau_2}\overrightarrow{\Pr}(\tau_2|0),$$

which implies an estimator $\widehat{\overrightarrow{\Pr}}(\tau_2|0) = \widehat{B}^{-1}_{d_3|0,\tau_2}\widehat{\overrightarrow{\Pr}}(d_3|0)$. Then the type probabilities are estimated as

$$\widehat{\overrightarrow{\Pr}}(\tau_2) = \widehat{\overrightarrow{\Pr}}(\tau_2|0)\widehat{\Pr}(w_2 = 0) + \widehat{\overrightarrow{\Pr}}(\tau_2|1)\widehat{\Pr}(w_2 = 1).$$

Step two: Estimation of contributing strategies. In our paper, as in most of the experiments, the distribution of subjects' values is known to the researcher. Combining this distribution with the estimated conditional density $\widehat{f}(b_2|\tau_2)$ allows us to recover the contributing strategies for $\tau_2 = 1, 2, 3$. Let $F_{B|\tau}$ denote the conditional cdf of the observed contributions for a type τ, then lemma 2 states that

$$F_{B|\tau}(b|\tau = k) = \frac{s_k^{-1}(b) - \underline{v}}{\overline{v} - \underline{v}},\ \ k = 1, 2, 3,$$

The relationship above implies that

$$s_k^{-1}(b) = (\overline{v} - \underline{v})F_{B|\tau}(b|\tau = k) + \underline{v},\ \ k = 1, 2, 3.$$

Then, our estimate of $s_k^{-1}(b)$ is

$$\widehat{s}_k^{-1}(b) = (\overline{v} - \underline{v})\widehat{F}_{B|\tau}(b|\tau = k) + \underline{v},\ \ k = 1, 2, 3. \tag{A.3}$$

Step three: Estimation of transition matrices of types. The learning rule $f_{\tau_2|w_1,\tau_1}$ is estimated from (15), which is repeated as follows.

$$f_{b_2,w_1,b_1} = \sum_{\tau_2}\sum_{\tau_1} f_{b_2|\tau_2} f_{\tau_2|w_1,\tau_1} f_{w_1|b_1} f_{b_1|\tau_1} f_{\tau_1}.$$

Based on the equation above, we maximize the likelihood function of the L.H.S. to estimate the learning rule on the R.H.S. Specifically, suppose we fix $w_1 = 1$ then the log likelihood function is expressed as:

$$\begin{aligned}\log\ \mathcal{L} &= \sum_{i=1}^{N}\log\sum_{\tau_2=1}^{3}\left(f_{b_{2i}|\tau_2}\sum_{\tau_1=1}^{3}\Pr(\tau_2|w_1 = 1, \tau_1)f_{w=1|b_{1i}}f_{b_{1i}|\tau_1}\Pr(\tau_1)\right)\\ &= \sum_{i=1}^{N}\log\sum_{\tau_2=1}^{3}\left(f_{b_{2i}|\tau_2}\frac{f_{w=1,b_{1i}}}{f_{b_{1i}}}\sum_{\tau_1=1}^{3}\Pr(\tau_2|w_1 = 1, \tau_1)f_{b_{1i}|\tau_1}\Pr(\tau_1)\right)\\ &= \sum_{i=1}^{N}\log\sum_{\tau_2=1}^{3}\left(f_{b_{2i}|\tau_2}\frac{f_{b_{1i}|w_1=1}\Pr(w_1 = 1)}{f_{b_{1i}}}\sum_{\tau_1=1}^{3}\Pr(\tau_2|w_1 = 1, \tau_1)f_{b_{1i}|\tau_1}\Pr(\tau_1)\right).\end{aligned} \tag{A.4}$$

Recall that the unknown transition matrix $\Pr(\tau_2|w_1 = 1, \tau_1)$ contains six independent parameters (denoted by θ). Given the estimated results in the proceeding steps, MLE of $\Pr(\tau_2|w_1 = 1, \tau_1; \theta)$ is

$$\widehat{\Pr}(\tau_2|w_1 = 1, \tau_1; \theta) \equiv \max_{\theta\in[0,1]^6}\log\mathcal{M}, \tag{A.5}$$

where $\log\mathcal{M}$ is the log-likelihood function $\log\mathcal{L}$ with all the terms but the transition matrix being replaced by their corresponding estimates. Especially, we employ the relationship $f_{b_1|\tau_1} = f_{b_2|\tau_2}$ and $f(b_1) = \sum_{\tau_1} f_{b_1|\tau_1}\Pr(\tau_1)$ in estimating $f_{b_1|\tau_1}$ and $\Pr(\tau_1)$, respectively.

Properties of the estimators can be proved by standard methods and we refer interested reader to An (2017) and An et al. (2010) for details.

Appendix B. Experiment instruction

In this experiment, you will be divided into different groups where each group can provide one unit of public good. If the sum of contributions from your group exceeds the cost, the public good is provided, and your profit is your value minus your contribution; otherwise your profit is zero. Your value is randomly drawn from 8 to 20; that is, someone may have a value as low as 8, and someone may have a value as high as 20, while for the most of the time, your value is between 8 and 20. Your value will vary across periods.

Your goal is to maximize your profit. In order to make better decisions, you may need to guess how much other people would contribute in your group. In each period, you need to enter 1) your guess on how likely your group will provide the public good (subjective probability, between 0 and 1); 2) your contribution to the public good.

What you need to do? Once the program is activated, please enter your guess on how likely your group will provide the public good and then make an offer to the public good.

How is your profit calculated?

- Your profit= Your benefit - Your cost.
- Your benefit= your value, if the public good if provided; Your benefit= 0, if the public good if not provided. Suppose that your value is $10, if the public good is provided, you benefit equals your value, which is $10; if the public good is not provided, you benefit is 0.
- Your cost= your offer, if the public good if provided; Your cost= 0, if the public good if not provided. Suppose that you make an offer of $5, if the public good is provided, you cost is $5; if the public good is not provided, you cost is 0.
- Under this situation, your profit=$10-$5=$5 if the public good is provided; your profit=$0 if the public good is not provided.

All the numbers used in examples serve only illustrative purpose; please do not try to use these examples to guess what would actually happen in the experiment.

How to decide if the public good can be provided? We will compare the total offer of your group with the cost of the public good. If the group's total offer is higher or equal to the cost for the public good, we will provide the public good, otherwise not.

Quiz (4 mins):

1. If your offer on the public good is $10, you value $20, what's if your profit if the public good is provided /not provided?
2. If the total offers of your group is $50, the cost of the public good is $40, is the public good provided? What if the cost of the public good is $60?

Instructions At-A-Glance

- You will be asked to decide how much money to offer towards the cost of the public good.
- The administrator will use the offers of everyone in your group to determine if we can provide the public good.
- If you offer more, in exchange for incurring some of the costs, you may get a higher profit by increasing the probability of the public good being provided.
- If you offer less, you may decrease the probability of the public good being provided; however, you may get a higher profit since you pay less if the public good is provided.

At the end of the experiment, your earnings will be totaled across all periods and converted from experimental dollars to real dollars. You will be paid as you leave.

Now please make your decisions!

References

Alboth, D., Lerner, A., Shalev, J., 2001. Profit maximizing in auctions of public goods. J. Public Econ. Theory 3 (4), 501–525.
An, Y., 2017. Identification of first-price auctions with non-equilibrium beliefs: a measurement error approach. J. Econom. 200, 326–343.
An, Y., Hu, Y., Shum, M., 2010. Estimating first-price auctions with an unknown number of bidders: a misclassification approach. J. Econom. 157, 328–341.
Andreoni, J., 1988. Why free ride?: Strategies and learning in public goods experiments. J. Public Econ. 37 (3), 291–304.
Athey, S., Haile, P.A., 2002. Identification of standard auction models. Econometrica 70 (6), 2107–2140.
Bagnoli, M., Lipman, B.L., 1989. Provision of public goods: fully implementing the core through private contributions. Rev. Econ. Stud. 56 (4), 583–601.
Bagnoli, M., McKee, M., 1991. Voluntary contribution games: efficient private provision of public goods. Econ. Inq. 29 (2), 351–366.
Bajari, P., Hortacsu, A., 2005. Are structural estimates of auction models reasonable? evidence from experimental data. J. Polit. Econ. 113 (4), 703–741.
Barbieri, S., Malueg, D.A., 2008. Private provision of a discrete public good: efficient equilibria in the private-information contribution game. Econ. Theory 37 (1), 51–80.
Bergstrom, T., Blume, L., Varian, H., 1986. On the private provision of public goods. J. Public Econ. 29 (1), 25–49.
Cadsby, C.B., Maynes, E., 1999. Voluntary provision of threshold public goods with continuous contributions: experimental evidence. J. Public Econ. 71 (1), 53–73.
Camerer, C.F., Ho, T.-H., Chong, J.-K., 2004. A cognitive hierarchy model of games. Q. J. Econ. 861–898.
Cason, T.N., Zubrickas, R., 2017. Enhancing fund raising with refund bonuses. Games Econ. Behav. 101, 218–233.
Chernozhukov, V., Imbens, G.W., Newey, W.K., 2007. Instrumental variable estimation of nonseparable models. J. Econom. 139 (1), 4–14.
Clemens, C., Riechmann, T., 2002. Discrete public goods: contribution levels and learning as outcomes of an evolutionary game. Technical Report. Diskussionspapiere der Wirtschaftswissenschaftlichen Fakultät, Universität Hannover.
Crawford, V.P., Iriberri, N., 2007. Level-k auctions: can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private–value auctions? Econometrica 75 (6), 1721–1770.
Croson, R.T., Marks, M.B., 2000. Step returns in threshold public goods: a meta-and experimental analysis. Exp. Econ. 2 (3), 239–259.
Embrechts, P., Lindskog, F., McNeil, A., 2003. Modelling dependence with copulas and applications to risk management. In: Handbook of Heavy Tailed Distributions in Finance, 8, pp. 329–384.
Fischbacher, U., 2007. Z-tree: zurich toolbox for ready-made economic experiments. Exp. Econ. 10 (2), 171–178.
Fischbacher, U., Gächter, S., 2010. Social preferences, beliefs, and the dynamics of free riding in public goods experiments. Am. Econ. Rev. 541–556.
Fischbacher, U., Gächter, S., Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. Econ. Lett. 71 (3), 397–404.
Fischbacher, U., Schudy, S., Teyssier, S., 2014. Heterogeneous reactions to heterogeneity in returns from public goods. Soc. Choice Welfare 43 (1), 195–217.
Healy, P.J., 2006. Learning dynamics for mechanism design: an experimental comparison of public goods mechanisms. J. Econ. Theory 129 (1), 114–149.
Hu, Y., 2008. Identification and estimation of nonlinear models with misclassification error using instrumental variables: a general solution. J. Econom. 144, 27–61.
Hu, Y., Kayaba, Y., Shum, M., 2013. Nonparametric learning rules from bandit experiments: the eyes have it!. Games Econ. Behav. 81, 215–231.
Hu, Y., McAdams, D., Shum, M., 2013. Identification of first-price auctions with non-separable unobserved heterogeneity. J. Econom. 174 (2), 186–193.
Hu, Y., Schennach, S., 2008. Instrumental variable treatment of nonclassical measurement error models. Econometrica 76, 195–216.

Krasnokutskaya, E., 2011. Identification and estimation of auction models with unobserved heterogeneity. Rev. Econ. Stud. 78 (1), 293–327.
Laussel, D., Palfrey, T.R., 2003. Efficient equilibria in the voluntary contributions mechanism with private information. J. Public Econ. Theory 5 (3), 449–478.
Li, T., Perrigne, I., Vuong, Q., 2000. Conditionally independent private information in OCS wildcat auctions. J. Econom. 98 (1), 129–161.
Liu, P., Swallow, S.K., Anderson, C.M., 2016. Threshold-level public goods provision with multiple units: experimental effects of disaggregated groups with rebates. Land Econ. 92 (3), 515–533.
Newey, W., Powell, J., 2003. Instrumental variable estimation of nonparametric models. Econometrica 71 (5), 1565–1578.
Oliveira, A., Croson, R., Eckel, C., 2014. One bad apple? Heterogeneity and information in public good provision. Exp. Econ. 1–20.
Olson, M., 1965. Logic of collective action: public goods and the theory of groups.. Harvard Economic Studies.
Robin, J.-M., Smith, R.J., 2000. Tests of rank. Econ. Theory 16 (02), 151–175.
Shilov, G.E., 2013. Elementary Functional Analysis. Courier Corporation.
Weimann, J., 1994. Individual behaviour in a free riding experiment. J. Public Econ. 54 (2), 185–200.
Xiao, R., 2018. Identification and estimation of incomplete information games with multiple equilibria. J. Econom.. forthcoming