



Essays on Social Network and the Role of Information

Juni Singh

► To cite this version:

Juni Singh. Essays on Social Network and the Role of Information. Political science. École des hautes études en sciences sociales (EHESS), 2020. English. NNT : 2020EHES0073 . tel-03168313

HAL Id: tel-03168313

<https://theses.hal.science/tel-03168313>

Submitted on 12 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



École des Hautes Études en Sciences Sociales

École doctorale n° 465 Économie Panthéon Sorbonne

Paris School of Economics

PhD. Thesis

Discipline: Analyse et politique économiques

Juni Singh

Essays on Social Networks and the Role of Information

Thesis supervised by: Philippe Jehiel

Date of defense: 29 June, 2020

Referees: Yann Bramoullé, Aix-Marseille University

Edoardo Gallo, University of Cambridge

Jury: Abhijit Banerjee, Massachusetts Institute of Technology

Margherita Comola, Université Paris Sud, PSE

Nicolas Jacquement, Université Paris 1, PSE

Supervisor: Philippe Jehiel, Ecole des Ponts - Paris Tech, PSE

Acknowledgment

This PhD experience has been a roller coaster ride and I thank everyone who were on board. Firstly, my supervisor, Philippe Jehiel to have believed in me and provided me with honest and prompt feedback even when times were hard. He has been a great mentor over the years and helped me become a better researcher. Secondly, I'd like to thank my committee members for their constant support: Nicolas Jacquement for the advice on the experimental design and analysis; Philippe Ketz for the help with econometrics and MATLAB; Margherita Comola for the guidance on network methods and analysis. I would also like to thank my Referees Yann Bramoulle and Edoardo Gallo for their critical reviews. I am grateful to Abhijit Banerjee for his guidance and support, without which the PhD experience would not have been the same. The two months at MIT as a visiting student were constructive and I would like to thank the Development group at MIT for making me feel at home.

I am also very thankful to everyone who made the lab and lab in the field experiments possible without which this PhD would not be possible. Firstly, my co-author Giulio Iacobelli who has been through the ups and downs of the field experiment and has provided immense technical and emotional support. Living in a Nepali village for 2 months and handling the team and didis was epic, Grazie Jules. I couldn't have thought of a better co-author. For financing the lab in the field experiment, I'd like to thank CEPREMAP, LABEX OSE, Sarah Andrieux Fund, and development group (Sylvie Lambert and Karen Macours), theory group (Olivier Tercieux, Olivier Compte), trade and institutions group (Thierry Verdier) at PSE. I am also grateful for support from the Belgian French-speaking community ARC project n°15/20-072 of Saint-Louis University - Brussels (Ana Mauleon and Vincent Vanntelbosch). Merci Gaelle Feneloux and Valerie Labadie in providing administrative support for the field experiment. I thank our partner organization Rooster logic and the enumerators for providing excellent field support. For the lab experiment, I would like to thank ERC grant no. 742816 and PSE Research grants. For the implementation of the experiment, I am grateful to Maxim Frolov for programming and J.C Vergnaud for administrative support.

I would like to thank the administrative staff at PSE for having helped me with the paperwork, surviving French administration would not have been possible without them. In particular, Beatrice Havet without whom PSE would never be the same, in addition to the reimbursements, the laughter shared brightened my day. I would also like to thank Marie Font for help with handling funds, dealing with money never seemed easier. Julien Thelot, for helping me figure out titre de Sejour issues when the prefecture remarked c'est pas mon probleme madame. Veronique Guilliton for always reminding me to be efficient and for the job market support. Marie Philipon and Sylvain Riffe Stern for random conversations and always being open to new ideas including the musical concert for 10 years of PSE. And last

but not the least, Stephane Brice and Radja Aroquiaradja for the technical support and for quick response whenever there was a problem with the server or computer.

Next, I would like to thank all the professors, at PSE and outside who have helped me progress with my research. They were always welcoming and felt like just an email away. Firstly, I'd like to thank Sylvie Lambert and Karen Macours who welcomed me to the development group. Whenever I had a research question to discuss, they always made time. In particular Sylvie for always being there to discuss the PhD roller coaster ride. I would like to thank François Libois for his constant support and feedback on the projects and otherwise. I would also like to thank the following people for providing me with feedback: David Margollis, Luc Behgel, Akikio Suwa-Eissenman, Jean François Laslier (also for the music), Andrew Clark (for the WIP seminar and Indian lunch after), Jean Marc Tallon, Béatrice Boulu-Reshef, Fabrice Etilé, Francis Bloch, Matt Leduc, Gabrielle Demange, Ariane Lambert-Mogiliansky and Kenneth Hounghedji. In Boston, I'd like to thank Esther Duflo, Emily Breza, Ben Golub and Cynthia Kinnan for their feedback on the project.

During this PhD journey, I've had the chance to meet some amazing people. Alexia, Shaden and Cristina for being pillars of emotional support and always being there to talk about n'importe quoi. This journey would not have been the same without you guys. Sarah and Juliette for being my tea buddies and Yasmine, Maiting, Rozenn, Lisa, Andrea, Paula, Emmanuela, Elias, Dalal, Monica and Caroline for being the 6th floor gang. A special thanks to Maiting for helping me polish the Introduction. I would also like to thank Luis GB for his support, random conversations and taking French lessons together. Lennart for the wonderful soul he is, helping with the model and Julien for being a mentor towards the beginning. I am also grateful to have come across all the other amazing people who have been a part of my journey in one way or the other: Laura, Clement, Simon, Paul, Johnathan, Yaz, Martin, Clara, Anthony, Marco, Fanny and Paul. In Boston, I would like to thank Clementine and Siri for supporting me.

I would like to thank all my friends outside PSE who have made PhD life so much fun with good music and food, in particular Chiara, Lama, Claudia, Lipsa, Leonardo and Andrea. I am also grateful to my family at La Houle for their constant support. Finally, I would like to thank my parents Mandira S. Shrestha and Bal Krishna Man Singh and siblings (Manzari and Ravi) for always believing in me and motivating me throughout this journey. I would also like to thank my partner Nicolas Gefflot who has seen all the ups and downs of this journey and who has been through thick and thin, motivating me to always be better.

Contents

Acknowledgement	i
1 General Introduction	1
2 Multi-state choices with Aggregate Feedback on Unfamiliar Alternatives	11
1 Introduction	12
2 Related Literature	15
3 Background and theory	17
3.1 Quantal valuation equilibrium	18
3.2 Learning Models	20
3.3 Similarity-based reinforcement learning	20
3.4 Generalized Bayesian Learning Model	22
4 Results	24
4.1 Further Description of the Experimental Design	24
4.2 Preliminary findings	25
4.3 Statistical estimations	26
4.4 Comparing the Reinforcement learning model to the data	28
4.5 Individual level heterogeneity	28
5 Robustness Checks	29
6 Conclusion	31
7 Appendix	38
3 Multi-state choices with Aggregate Feedback on Unfamiliar Alternatives: a Followup	43
1 Introduction: Why a follow up?	44
2 Experimental design: How is it different?	45
2.1 Context	45
2.2 Design	46
3 Results	47
3.1 Preliminary findings	47
3.2 Statistical estimations	48
3.3 Individual level heterogeneity	51
4 Appendix	55
4 Endogenous Institutions: a Network Experiment in Nepal	61
1 Introduction	62
2 Experiment	65
2.1 Overview: Networks and Data	65
2.2 Experimental context	67

2.3	Design	68
3	Results	69
3.1	Preliminary findings and limitations	70
3.2	Statistical Estimation	70
4	The model	75
5	Conclusion	79
6	Appendix	86
5	Summary	96
	List of Figures	104
	List of Tables	105

1 General Introduction

We make most decisions under uncertainty, that is without having a precise idea of the final outcome. When information is not complete, we have to rely on hearsay, rules of thumb, past experience or the experience of others. We form beliefs based on available information and these beliefs in turn shape our actions and choices. They affect the products we choose, our preferences and level of cooperation. This thesis focuses on two main research themes a) how do individuals form beliefs from the experience of others and b) why does cooperation vary across groups and how can it be improved.

To answer the first question, I use concepts from bounded rationality, which is the idea that the cognitive limitations of the mind limits the rationality of individuals (Simon, 1982). It is different from the usual assumption made in economics of individuals being rational and working towards maximizing their utility. Given the informational constraints and limitations in processing information, bounded rationality presents a more realistic picture of human behavior. We start with an initial belief about the available options. In the face of new incoming information, we learn and update our existing beliefs. Based on the new belief, we take actions leading to a payoff. How we learn is explained by two broad classes of theories: belief based learning and reinforcement learning. Belief based learning assumes that we make decisions based on beliefs about the behavior of others and calculate an expected payoff. With reinforcement learning, we assume learning solely from our past payoffs. We tend to look back at similar situations and extrapolate that experience to current decisions. In the first half of my thesis, I work with my supervisor Philippe Jehiel to combine bounded rationality with learning.

In Chapter 1, we are interested in understanding how individuals make their decisions in multi-state binary decision problems. In particular, we study cases when decision makers have precise state-specific information about the performance of one alternative and less precise information about the other. Less precise in the sense that decision makers only receive aggregate feedback about the performance of that alternative instead of state-specific information. We allow a set of agents to act repeatedly in such environments to understand the steady-state effects of providing agents with aggregate feedback about some alternatives. To better understand the feedback structure consider an example of adoption of a new technology by farmers. A farmer is familiar with the performance of the current technology but not about the new one. Before deciding to adopt the new technology, the farmer collects information about it. The feedback is from other farmers who have already adopted it. Yet, due to the heterogeneity of the soil or ability of farmers, what works well for one farmer need not perform in the same way for another. Thus, the feedback received about the new technology is coarse in the sense that it is aggregated over different situations (states in the decision theoretic terminology) as compared to the information for the current technology.

We conduct a lab experiment with the following experimental setting. There are two states, $s = 1, 2$. In each state, the decision maker has to choose between two urns identified with a color, Blue and Red in state $s = 1$, Green and Red in state $s = 2$ where the Red urns have different payoffs in each state $s = 1$ and 2. Each urn is composed of ten balls, black or white. When an urn is picked, one ball is drawn at random from this urn (and it is immediately replaced afterwards). If a black ball is drawn, this translates into a positive payment whereas a white ball offers no payment. We conduct three treatments varying the composition of the Red urns across states and keeping the composition of Blue and Green fixed. For the Blue and Green urns, we provide information on the number of black and white balls obtained after 100 random draws, giving the agents an idea of its composition in the spirit of familiar alternatives. The Red urns correspond to the unfamiliar choice in the above example (new technology for the farmer) and no initial information is provided.

There are twenty subjects and 70 rounds. In each round, ten subjects are randomly assigned state 1 and the other ten make a choice of urn in state 2. There are permutations of subjects between rounds so that every subject is in each state $s = 1$ or 2 the same proportion of time. Between rounds, subjects receive feedback about the number of times the Green, Blue and Red urns were picked by the various agents in the previous round, and for each color of urn, they are informed of the number of black balls that were drawn. In the case of the Red urns, this number is the total the number of black balls drawn from both the Red urns picked in state $s = 1$ and the Red urns picked in state $s = 2$ mimicking the kind of aggregate information suggested in the motivating examples. It should be highlighted that subjects were explicitly told that the compositions of the Red urn in state $s = 1$ and in state $s = 2$ need not be the same. We consider three treatments that differ in the composition of the Red urns but we fix the compositions of the Blue and Green urns in all treatments. The initial conditions in these various treatments are thus identical and any difference in behavior observed in later periods can be safely attributed to the difference in the feedback received by subjects across treatments.

Faced with this environment, we use reinforcement learning with a component of bounded rationality to understand the decision-making process. In particular, we use a solution concept proposed by Jehiel and Samet (2007) called the valuation equilibrium aimed at capturing the limiting outcomes of this type of similarity-based reinforcement learning models. Via reinforcement learning, subjects could assess the strength of the various urns by considering the proportions of black balls drawn in the corresponding urns (aggregating past feedback). However, for the Red urns feedback is aggregated over states $s = 1$ and 2, and so similarity-based reinforcement suggests assigning a common value to Red in both states. This value depends on feedback from both states and how frequently Red was picked in each state. An alternative way to learn about the urns could be by forming beliefs about the compositions of the two Red urns relying on some form of Bayesian updating to

adjust the beliefs after the additional feedback. In the Bayesian model, subjects construct an intermediate object which is a belief about the composition of the various urns, whereas the reinforcement model mediates payoffs more directly. We find that the similarity-based learning model fits the data better than the Bayesian model.

In Chapter 2, a natural followup to this question would be to see how learning differs when individuals can generate their own feedback instead of endogenous group feedback. In the context of our farmer example, instead of relying on information from others, the farmer would be allowed a trial window to use and learn about the new technology. It is possible that even under this scenario, the farmer fails to learn correctly given concerns of imperfect recall. We conduct additional experiments varying the feedback generation where each player plays for a total of 20 rounds with 10 repetitions. After each repetition, the player receives feedback about the number of times the Green, Blue and Red urns were picked by her in the previous round, and for each color, she is informed of the number of black balls that were drawn. The feedback for Red is still aggregated over states $s = 1$ and 2 . We find similar findings as in the original experiment where players follow similarity-based reinforcement learning. We find a slightly higher proportion of Bayesian learners compared to the original experiment.

Given the importance of the information we have access to, it would be important to understand what determines the information we get. We miss an important part by modeling individuals without taking into consideration how their beliefs are shaped by their peers. An immediate answer would be social networks. It is responsible in determining information flows, and thus economic outcomes and plays a central role in determining our beliefs, social norms, transactions and preferences (Grief, 1989). We could improve the first experiments by incorporating this factor, but for now we leave it for future research. The second theme of this dissertation is cooperation and how closeness and the endogenous choice of monitor affects cooperation in social networks. We study this in the context of a developing country where individuals rely on "*local connections*" and personalized relationships for risk sharing, public good provision and information delivery (Kranton, 1996). In such a setting, it becomes crucial to understand a) how social structures affect behavior and b) how these structures are formed in the first place. With my co-author Giulio Iacobelli, I study the first question in the third chapter of my thesis.

We conduct a lab in the field experiment in 19 villages in rural Nepal. Villagers play a cooperation game and can choose to elect a monitor through majority voting. This monitor can impose higher cooperation through reputational concerns. Reputation and social image are a strong driver for desired behavior in these villages. In the event of someone monitoring, the fear of being reported outside the experimental setup may induce higher cooperation (Andreoni et al., 2006). Further, the demand for monitoring itself may depend on the social structure. We estimate the demand for monitoring, relate it to the network

structure and study its impact on cooperation by tackling three questions. First, do individuals change their demand for monitoring as a function of the social composition of the group they interact with? Second, do monitors who are endogenously chosen spur cooperative behavior compared to those assigned exogenously? Third, is the monitor choice perceived as a signal of intra-group trust?

To answer these questions, we first conduct an intensive network survey a la Banerjee et al., 2015. We ask questions about advice, trust, friendship and financial relationships such as "who do you spend your free time with", "in case of an emergency, who would you rely on" and "who do you borrow money from". Based on data from the survey, we build an undirected social network where a connection between two people is established if any one of them names the other. Next, we identify the person in the network with the highest and lowest eigenvector centrality. In other words, we identify the most influential and least influential individuals in the village. Eigenvector centrality is a network measure that gives an idea of the influence of an individual by taking into consideration the friends of a friend. This is more robust than considering only the total number of connections each person has (referred to as degree in network economics). These two individuals are assigned the role of monitor candidates and the rest of the village is divided into groups of three with varying network distance. Groups of three are an optimal choice to study questions of mutual trust as it maximizes the behavioral contrast between groups which can sustain high levels of cooperation and those which cannot (Jackson et al., 2012).

The division into groups is made based on network distance. Distance refers to the shortest path to reach from one individual to the other. Players interact both in a closely-knit group and in a group with socially distant individuals. The closely-knit groups are often homogeneous in terms of caste. We refer to the closely-knit group as "dense" and the group with acquaintances as "sparse". In formal network terms "dense" implies each individual is at most at distance 2 (average path length < 1.6) and "sparse" implies each individual is at least at distance 4 (average path length > 4). We allow individuals to either vote for no monitor or to choose from two monitor candidates belonging to their village: a high central monitor (very prominent individual) or a low central monitor (less prominent). Voting for a monitor is costly. Players play a contribution game both with a monitor chosen by the group (endogenous treatment) and an externally assigned monitor (exogenous treatment). The monitor does not punish but only observes the contributions of each player, which would otherwise be private information. Players have an initial endowment and need to decide between how much to contribute to a common pot and how much to keep for themselves. Total contribution in the common pot is augmented by 50 percent and divided equally among the three players irrespective of initial contribution. Each individual, thus, plays both in a dense and in a sparse group. And in each group, both with exogenous and endogenous monitors. The order of all treatments is randomized. We also build a theoretical model supporting our experimental findings.

We find that socially distant groups are more likely to elect a high central monitor than closely-knit groups. Individuals are 40% less likely to elect a monitor with their close peers compared to when they are in groups with socially distant members. With respect to monitoring by an externally assigned monitor, we find that cooperation in the sparse group increases by 15.8%. Comparing endogenous versus exogenous monitoring, we find that letting individuals choose their monitor is important for sparse groups. When sparse groups can choose their monitors, choosing a socially prominent monitor increases contributions by 8.6% while choosing no monitor increases contribution by 21.9%. Lastly, since the chosen monitor is only revealed in the endogenous case, we find that this acts as a signal of intra-group trust. We average contribution across both endogenous and exogenous treatments. If sparse groups are in the endogenous treatment first, contribution increases significantly by 19.7% when no monitor is chosen. On the contrary, when dense groups play the endogenous treatment first and elect a high central monitor, contribution decreases by 13.3%. We model this using varying level of altruism across groups.

Introduction générale

Nous prenons la plupart des décisions dans l'incertitude, c'est-à-dire sans avoir une idée précise du résultat final. Lorsque l'information n'est pas complète, nous devons nous fier aux ouï-dire, aux règles empiriques, à l'expérience passée ou à celle des autres. Nous formons des croyances à partir des informations disponibles et ces croyances façonnent à leur tour nos actions et nos choix. Elles influencent les produits que nous choisissons, nos préférences et notre niveau de coopération. Cette thèse se concentre sur deux thèmes de recherche principaux : a) comment les individus forment-ils des croyances à partir de l'expérience des autres et b) pourquoi la coopération varie-t-elle d'un groupe à l'autre et comment peut-elle être améliorée.

Pour répondre à la première question, j'utilise les concepts de rationalité limitée, c'est-à-dire l'idée que les limitations cognitives de l'esprit limitent la rationalité des individus (Simon, 1982), ce qui est différent de l'hypothèse habituelle en économie selon laquelle les individus sont rationnels et travaillent à maximiser leur utilité. Étant donné les contraintes et les limites du traitement de l'information, la rationalité limitée présente une image plus réaliste du comportement humain. Nous partons d'une conviction initiale sur les options disponibles. Face aux nouvelles informations qui nous parviennent, nous apprenons et mettons à jour nos croyances existantes. Sur la base de cette nouvelle croyance, nous prenons des mesures qui se traduisent par un gain. La façon dont nous apprenons est expliquée par deux grandes classes de théories : l'apprentissage basé sur les croyances et l'apprentissage par renforcement. L'apprentissage basé sur les croyances suppose que nous prenons des

décisions basées sur des croyances concernant le comportement des autres et que nous calculons un gain attendu. Avec l'apprentissage par renforcement, nous supposons que nous apprenons uniquement à partir de nos gains passés. Nous avons tendance à revenir sur des situations similaires et à extrapoler cette expérience aux décisions existantes. Dans la première moitié de ma thèse, je travaille avec mon directeur de thèse Philippe Jehiel pour combiner la rationalité limitée et l'apprentissage.

Dans le chapitre 1, nous cherchons à comprendre comment les individus prendraient leurs décisions dans des problèmes de décision binaire multi-états. En particulier, nous étudions les cas où les décideurs disposent d'informations précises, spécifiques à un état, sur les performances d'une alternative et d'informations moins précises sur l'autre. Les informations moins précises prennent la forme que le décideur reçoit un retour d'information global (non spécifique à un état) sur les performances de cette alternative. Nous permettons à un ensemble d'agents d'agir de manière répétée dans de tels environnements afin de comprendre les effets en régime permanent de la transmission aux agents d'un retour d'information global sur certaines alternatives. Pour mieux comprendre la structure du retour d'information, prenons un exemple d'adoption d'une nouvelle technique par des agriculteurs. Un agriculteur connaît les performances de sa technique existante, mais pas tellement celles de la nouvelle technique. Avant de décider d'adopter la nouvelle technique, l'agriculteur recueille des informations à son sujet. Le retour d'information provient d'autres agriculteurs qui l'ont déjà adoptée. Cependant, en raison de l'hétérogénéité des sols ou des capacités des agriculteurs, ce qui fonctionne bien pour un agriculteur ne doit pas nécessairement fonctionner de la même manière pour un autre. Ainsi, les informations reçues sur la nouvelle technique sont grossières dans le sens où elles sont agrégées sur différentes situations (états dans la terminologie de la théorie de la décision) par rapport aux informations détenues pour l'ancienne technologie.

Nous effectuons une expérience en laboratoire dans le cadre suivant. Il y a deux états, $s = 1, 2$. Dans chaque État, le décideur doit choisir entre deux urnes identifiées par une couleur, Bleu et Rouge en état $s = 1$, Vert et Rouge dans l'état $s = 2$ où les urnes rouges ont des implications différentes dans les états $s = 1$ et 2 . Chaque urne est contenue de dix boules, noires ou blanches. Lorsqu'une urne est choisie, une boule est tirée au hasard (et elle est immédiatement remplacée par la suite). Si une boule noire est tirée, cela se traduit par un paiement positif alors qu'une boule blanche n'offre aucun paiement. Nous procédons à trois traitements en faisant varier la composition des urnes rouges selon les états et en maintenant fixe la composition des urnes bleues et vertes. Pour les urnes bleues et vertes, nous fournissons des informations sur le nombre de boules noires et blanches obtenues après 100 tirages au sort, ce qui leur donne une idée de leur composition dans l'esprit des alternatives habituelles. Les urnes rouges correspondent au choix inconnu dans l'exemple ci-dessus et aucune information initiale n'est fournie.

Il y a vingt sujets et 70 tours. Dans chaque tour, les sujets sont assignés à des états au hasard et font un choix d'urne dans l'état 1 et les dix autres font un choix d'urne dans l'état 2. Il y a des permutations de sujets entre les tours de sorte que chaque sujet se trouve dans chaque état $s = 1$ ou 2 la même proportion de temps. Entre les tours, les sujets reçoivent des informations sur le nombre de fois que les urnes verte, bleue et rouge ont été choisies par les différents agents au cours du tour précédent, et pour chaque couleur d'urne, ils sont informés du nombre de boules noires qui ont été tirées. Dans le cas des urnes rouges, ce nombre additionne le nombre de boules noires tirées des urnes rouges choisies dans l'état $s = 1$ et des urnes rouges choisies dans l'état $s = 2$, imitant ainsi le type d'information agrégée suggéré dans les exemples de motivation. Il convient de souligner que les sujets ont été explicitement informés que les compositions de l'urne rouge en état $s = 1$ et en état $s = 2$ n'ont pas à être identiques. Nous considérons trois traitements qui diffèrent dans la composition des urnes rouges mais notons que nous fixons les compositions des urnes bleues et vertes dans tous les traitements. Les conditions initiales dans ces différents traitements sont donc identiques et toute différence de comportement observée dans les périodes ultérieures peut être attribuée sans risque à la différence dans le retour d'information reçu par les sujets au cours des traitements.

Face à un tel environnement, nous utilisons l'apprentissage par renforcement avec une composante de rationalité limitée pour comprendre le processus de décision. En particulier, nous utilisons un concept de solution proposé par Jehiel et Samet (2007) appelé "valuation equilibrium" visant à saisir les résultats limites de tels modèles d'apprentissage par renforcement basé sur la similarité. Par le biais de cette méthode, les sujets pourraient évaluer la force des différentes urnes en considérant les proportions de boules noires tirées dans les urnes correspondantes (en agréant les réactions passées). Cependant, pour les urnes rouges, le feedback est agrégé sur les états $s = 1$ et 2 , et donc le renforcement basé sur la similarité suggère d'attribuer une valeur commune au rouge dans les deux états. Cette valeur dépend du feedback des deux états et de la fréquence à laquelle le rouge a été choisi dans chaque état. Une autre façon de connaître les urnes pourrait être de former des croyances sur la composition des deux urnes rouges en s'appuyant sur une forme de mise à jour Bayesian updating pour ajuster les croyances après le retour d'information supplémentaire. Dans le modèle bayésien, les sujets construisent un objet intermédiaire qui est une croyance sur la composition des différentes urnes, alors que le modèle de renforcement sert de médiateur plus directement. Nous constatons que le modèle d'apprentissage basé sur les similitudes correspond mieux aux données que le modèle bayésien.

Dans le chapitre 2, une suite naturelle à cette question serait de voir comment l'apprentissage diffère lorsque les individus peuvent générer leur propre feedback au lieu d'un feedback de groupe endogène. Dans le contexte de notre exemple d'agriculteur, au lieu de se fier aux informations des autres, l'agriculteur aurait une période d'essai pour utiliser la nouvelle technique et en apprendre davantage à son sujet. Il est possible que,

même dans ce scénario, l'agriculteur ne parvienne pas à apprendre correctement en raison d'une comparaison imparfaite à la technique précédente. Nous menons des expériences supplémentaires en faisant varier la génération de rétroaction, où chaque joueur joue un total de 20 tours avec 10 répétitions. Après chaque répétition, le joueur reçoit un retour d'information sur le nombre de fois qu'il a choisi les urnes verte, bleue et rouge au cours du tour précédent, et pour chaque couleur de l'urne, il est informé du nombre de boules noires qui ont été tirées. Le feedback pour le Rouge est toujours agrégé sur les états $s = 1$ et 2. Nous constatons des résultats similaires à ceux de l'expérience originale où les joueurs suivent un apprentissage de renforcement basé sur les similarités. Nous constatons une proportion légèrement plus élevée d'apprenants bayésiens par rapport à l'expérience originale.

Étant donné l'importance des informations auxquelles nous avons accès, il serait important de comprendre ce qui détermine le feedback que nous recevons. Nous passons à côté d'un élément important en modelant les individus sans prendre en considération la façon dont leurs croyances sont façonnées par leurs connexions. Une réponse immédiate serait les réseaux sociaux. Il est responsable de la détermination des flux d'information, donc des résultats économiques et joue un rôle central dans la détermination de nos croyances, des normes sociales, des transactions et de nos préférences (Grief, 1989). Nous pourrions améliorer les premières expérimentations en intégrant ce facteur, mais pour l'instant nous le laissons pour les recherches futures. Le deuxième thème de cette thèse est la coopération et comment la proximité et le choix endogène du moniteur affectent la coopération. Nous l'étudions dans le contexte d'un pays en développement où les individus comptent sur le "*connexions locales*" et les relations personnalisées pour le partage des risques, la fourniture de biens publics et la transmission d'informations (Kranton, 1996). Dans un tel contexte, il devient crucial de comprendre a) comment les structures sociales affectent le comportement et b) comment ces structures sont formées en premier lieu. Avec mon co-auteur Giulio Iacobelli, j'étudie la première question dans le troisième chapitre de ma thèse.

Nous menons une expérience en laboratoire sur le terrain dans 19 villages du Népal rural. Les villageois jouent à un jeu de coopération et peuvent choisir d'élire un moniteur par un vote à la majorité. Ce moniteur peut imposer une plus grande coopération par souci de réputation. La réputation et l'image sociale sont des drivers important des comportements et interaction sociale. Dans le cas d'une surveillance, la crainte d'être dénoncé en dehors du lab peut induire une plus grande coopération (Andreoni et al., 2006). En outre, la demande de surveillance elle-même peut dépendre de la structure sociale. Nous estimons la demande de surveillance, la mettons en relation avec la structure du réseau et étudions son impact sur la coopération en abordant trois questions ultérieures. Premièrement, les individus modifient-ils leur demande de surveillance en fonction de la composition sociale du groupe avec lequel ils interagissent ? Deuxièmement, les surveillants choisis de manière endogène encouragent-ils le comportement coopératif par rapport à ceux qui sont assignés de manière exogène ? Troisièmement, l'élection des surveillants est-elle perçue comme un

signal de confiance intra-groupe ?

Pour répondre à ces questions, nous réalisons tout d'abord une enquête intensive sur le réseau à la Banerjee et al., 2015. Nous posons des questions sur les conseils, la confiance, l'amitié et les relations financières, par exemple "avec qui passez-vous votre temps libre", "en cas d'urgence, à qui vous fieriez vous" et "à qui vous empruntez de l'argent". Sur la base des données de l'enquête, nous construisons un réseau social non dirigé où une connexion entre deux personnes est établie si l'une d'entre elles nomme l'autre. Ensuite, nous identifions la personne dans le réseau ayant la centralité de vecteur propre la plus élevée et la plus basse. En d'autres termes, nous identifions les personnes les plus influentes et les moins influentes du village. La centralité des vecteurs propres est une mesure de réseau qui donne une idée de l'influence d'un individu en prenant en considération les amis d'un ami. Cette mesure est plus robuste que si l'on considère uniquement le nombre total de connexions que chaque personne possède (on parle alors de degré en économie des réseaux). Ces deux personnes se voient attribuer le rôle de candidats moniteurs et le reste du village est ensuite divisé en groupes de trois avec une distance de réseau variable. Les groupes de trois sont un choix optimal pour étudier les questions de confiance mutuelle dans la mesure où il est possible de maximiser le contraste comportemental entre les groupes qui peuvent maintenir des niveaux élevés de coopération et ceux qui ne le peuvent pas (Jackson et al., 2012).

La division en groupes se fait sur la base de la distance du réseau sociale. La distance fait référence au chemin le plus court pour se rendre d'un individu à l'autre. Les joueurs interagissent à la fois dans un groupe étroitement lié et dans un groupe avec des individus socialement distants. Les groupes étroitement liés sont souvent homogènes en termes de caste. Nous appelons le groupe très soudé "dense" et le groupe composé de simples connaissances "éparses". En termes de réseau formel, "dense" implique que chaque individu se trouve au maximum à la distance 2 (longueur moyenne du chemin $< 1,6$) et "éparse" implique que chaque individu se trouve au minimum à la distance 4 (longueur moyenne du chemin > 4). Nous permettons aux individus soit de voter pour aucun moniteur, soit de choisir parmi deux candidats moniteurs appartenant à leur village : un moniteur central haut (individu très proéminent) ou un moniteur central bas (moins proéminent). Voter pour un moniteur est coûteux. Les joueurs jouent un jeu de contribution à la fois avec un moniteur choisi par le groupe (traitement endogène) et un moniteur attribué de l'extérieur (traitement exogène). Le moniteur ne sanctionne pas, mais se contente d'observer les contributions de chaque joueur, qui seraient autrement des informations privées. Les joueurs ont une provision initiale et doivent décider du montant de leur contribution à un pot commun et du montant à garder pour eux-mêmes. La contribution totale au pot commun est augmentée de 50 % et divisée en parts égales entre les trois joueurs, indépendamment de la contribution initiale. Ainsi, chaque individu joue à la fois dans un groupe dense et dans un groupe clairsemé. Et dans chaque groupe, à la fois avec des institutions de surveillance exogènes et endogènes. L'ordre de tous les traitements est randomisé. Nous construisons

également un modèle théorique à l'appui de nos résultats expérimentaux.

Nous constatons que les groupes socialement éloignés sont plus susceptibles que les groupes étroitement liés d'élire un moniteur central élevé. Les individus sont 40 % moins susceptibles d'élire un moniteur avec leurs amis proches que lorsqu'ils sont en groupe avec des membres socialement distants. En ce qui concerne la surveillance par un moniteur externe, nous constatons que la coopération dans le groupe épars augmente de 15,8 %. En comparant la surveillance endogène et exogène, nous constatons que laisser les individus choisir leur moniteur est important pour les groupes épars. Lorsque les groupes épars peuvent choisir leurs moniteurs, le choix d'un moniteur socialement important augmente les contributions de 8,6 %, tandis que le choix d'aucun moniteur augmente les contributions de 21,9 %. Enfin, comme le moniteur choisi n'est révélé que dans le cas endogène, nous constatons que cela agit comme un signal de confiance intra-groupe. On fait la moyenne des contributions pour les traitements endogène et exogène. Si les groupes épars sont dans le traitement endogène en premier, la contribution augmente de manière significative de 19,7% lorsqu'aucun moniteur n'est choisi. Au contraire, lorsque les groupes denses suivent d'abord le traitement endogène et choisissent un moniteur central élevé, la contribution diminue de 13,3 %. Nous modélisons cela en utilisant un niveau variable d'altruisme entre les groupes.

2 Multi-state choices with Aggregate Feedback on Unfamiliar Alternatives

Abstract

This paper studies a multi-state binary choice experiment in which in each state, one alternative has well understood consequences whereas the other alternative has unknown consequences. Subjects repeatedly receive feedback from past choices about the consequences of unfamiliar alternatives but this feedback is aggregated over states. Varying the payoffs attached to the various alternatives in various states allows us to test whether unfamiliar alternatives are discounted and whether subjects' use of feedback is better explained by similarity-based reinforcement learning models (in the spirit of the valuation equilibrium, Jehiel and Samet 2007) or by some variant of Bayesian learning model. Our experimental data suggest that there is no discount attached to the unfamiliar alternatives and that similarity-based reinforcement learning models have a better explanatory power than their Bayesian counterparts.

Key words: Ambiguity, Bounded Rationality, Experiment, Learning, Coarse feedback, Valuation equilibrium

JEL Classification: D81, D83, C12, C91

1 Introduction

In many situations, the decision maker faces a choice between two alternatives one of them being more familiar and thus easier to evaluate and another one being less familiar and thus harder to assess. There is generally some information about the less familiar alternative, but this information is typically coarse not being entirely relevant to the specific context of interest.

To give a concrete application, think of the adoption of a new technology by farmers. A farmer has a lot of information about the performance of the current technology but not so much about the new one. The farmer may collect information about the new technology by asking around other farmers who would have previously adopted it. But due to the heterogeneity of the soil and/or the heterogeneity in the ability of the farmers, what works well/poorly for one farmer need not perform in the same way for another. Thus, the feedback received about the new technology is coarse in the sense that it is aggregated over different situations (states in the decision theoretic terminology) as compared to the information held for the old technology.¹ Another example may concern hiring decisions.² Consider hiring for two different jobs, one requiring high skill going together with higher education level and the other requiring lower skills, and assume potential candidates either come from a majority group or a minority group (as determined by nationality, color, caste or religion, say). Presumably, there is a lot of familiarity with the majority group allowing in this group to distinguish the productivity as a function of education as well as past experiences. However, in the minority group information is more likely to be coarse and perceived productivity in that group may not be as easy to relate to education or past experiences.

We are interested in understanding how decision makers make their decisions in multi-state binary decision problems in which decision makers would have precise state-specific information about the performance of one alternative and less precise information about the other alternative. The less precise information takes the form that the decision maker receives aggregate (not state-specific) feedback about the performance of that alternative. Our interest lies in allowing a set of agents to act repeatedly in such environments so as to understand the steady state effects of having agents provided with coarse feedback about some alternatives.

To shed light on this, we consider the following experimental setting. There are two states, $s = 1, 2$. In each state, the decision maker has to choose between two urns identified

¹Ryan and Gross (1946) propose an early study of the diffusion of new technology adoption in the farming context. See also Young (2009) for a study focused on the diffusion dimension.

²This example is inspired by Fryer and Jackson (2008)'s discussion of discrimination and categorization.

with a color, Blue and Red in state $s = 1$, Green and Red in state $s = 2$ where the Red ³ urns have different payoff implications in states $s = 1$ and 2 . Each urn is composed of ten balls, black or white. When an urn is picked, one ball is drawn at random from this urn (and it is immediately replaced afterwards). If a black ball is drawn this translates into a positive payment. If a white ball is drawn there is no payment. We conduct three treatments varying the composition of the Red urns across states and keeping the composition of Blue and Green fixed. One hundred initial draws are made for the Blue and Green urns with no payoff implication for participants, and all subjects are informed of the corresponding compositions of black and white balls drawn from these urns. Thus, as seen in Table 1, subjects have a precise initial view about the compositions of the Blue and Green urns (these urns correspond to the familiar choices in the motivating examples provided above). In the experiment, the Blue urn has 3 black balls out of ten and the Green urn has 7 black balls out of ten.

Concerning the Red urns, there is no initial information. The Red urns correspond to the unfamiliar choices in the above examples. To guide their choices, subjects are provided with feedback about the compositions of the red urns as reflected by the colors of the balls that were previously drawn when a red urn either in state $s = 1$ or 2 was chosen. More precisely, there are twenty subjects and 70 rounds. In each round, ten subjects make a choice of urn in state 1 and the other ten make a choice of urn in state 2. There are permutations of subjects between rounds so that every subject is in each state $s = 1$ or 2 the same proportion of time. Between rounds, subjects receive feedback about the number of times the Green, Blue and Red urns were picked by the various agents in the previous round, and for each color of urn, they are informed of the number of black balls that were drawn. A typical feedback screen is shown in Figure 1. Note that in the case of the Red urns, this number aggregates the number of black balls drawn from both the Red urns picked in state $s = 1$ and the Red urns picked in state $s = 2$ mimicking the kind of coarse information suggested in the motivating examples. It should be highlighted that subjects were explicitly told that the compositions of the Red urn in state $s = 1$ and in state $s = 2$ need not be the same.

We consider three treatments T1, T2, T3 that differ in the composition of the Red urns as depicted in Figure 2, but note that we maintain the compositions of the Blue and Green urns in all treatments. The initial conditions in these various treatments are thus identical and any difference of behaviors observed in later periods can safely be attributed to the difference in the feedback received by the subjects across the treatments. In treatment 1, the best decision in both states $s = 1$ and 2 would require the choice of the Red urn, but averaging the composition of the Red urns across the two states leads to a less favorable

³In the instruction in Appendix A, Red in State 1 is referred to as Red1 and in State 2 is referred to as Red2. It is explicitly mentioned that these are two different urns with different payoff implications.

composition than the Green urn. In treatment 2, the best decision would require picking the Red urn in state 1 but not in state 2, but the average composition of the two Red urns dominates that of both the Blue and Green urns. Finally, in treatment 3, the best decision would require picking the Red urn in state 2 but not in state 1. Faced with such an environment, what could be the decision making process followed by subjects? We see the following possible approaches.

First, in the tradition of reinforcement learning (see Barto and Sutton 1998 or Fudenberg and Levine 1998 for textbooks), subjects could assess the strength of the various urns by considering the proportions of black balls drawn in the corresponding urns (aggregating past feedback in some way). One key difficulty in our context is that there is no urn specific feedback for the Red urns as the feedback is aggregated over states $s = 1$ and 2 , and so the standard reinforcement learning models which attach a different strength to every possible strategy do not apply. But, following Jehiel and Samet (2007), one could extend the approach by considering similarity-based reinforcement learning models in which a single valuation would be attached to the Red urns whether in state $s = 1$ or 2 (and reinforced accordingly) and the two Red urns would be considered alike in terms of strength by the learning subjects. Jehiel and Samet (2007) have proposed a solution concept called the valuation equilibrium aimed at capturing the limiting outcomes of such similarity-based reinforcement learning models. In our context, there would be a valuation for each color, Blue, Red and Green; a subject would pick the urn with a color attached to the highest valuation; the valuation attached to the Blue and the Green urns would be 0.3 and 0.7 respectively as reflected by the true compositions of these urns; the valuation of Red would be an average of the proportion of black balls in the Red urns in states $s = 1$ and $s = 2$ where the weight assigned to the Red urns in the various states should respect the proportion of times the Red urn is picked in states $s = 1$ and 2 .⁴ The valuation equilibrium would predict that in treatment 1 (T1), the Red urn is picked in state 1 but not in state 2; in treatment 2 (T2), the Red urns are picked both in states 1 and 2; in treatment 3 (T3), the Red urns are picked neither in state 1 nor 2. In our estimation, we will consider a noisy version of such a model in which subjects rely on noisy best-response in the vein of the logit model (as popularized by McKelvey and Palfrey (1995) in experimental economics).

Second, subjects could form beliefs about the compositions of the two Red urns relying on some form of Bayesian updating to adjust the beliefs after they get additional feedback. Note that being informed of the number of times the Blue and the Green urns were picked in the last round is also informative so as to determine if feedback about the Red urns concerns more state 1 or state 2 (as for example a strong imbalance in favor of the Green urns

⁴For the sake of illustration, if the Red urns are picked with the same proportion in states $s = 1$ and 2 , the valuation should be the unweighted average proportion of black balls in the two Red urns. If the Red urn is only picked in state 1 (resp 2), the valuation of Red should be the proportion of black balls of the Red urn in state 1 (resp 2).

as opposed to the Blue urns would be indicative that the previous red choices corresponded more to state 1). Of course, such a Bayesian approach heavily depends on the initial prior. When estimating such a model, we will assume that subjects consider a uniform prior over a support that we will estimate, and as for the reinforcement learning model we will assume that subjects employ a noisy best response of the logit type.

Another key consideration is that the feedback concerning the Red urns is ambiguous to the extent that it does not distinguish between states $s = 1$ and 2 . Following the tradition of Ellsberg (1961), one may suspect then that subjects would apply an ambiguity discount to the Red urns (see Gilboa and Schmeidler (1989) for an axiomatization of ambiguity aversion). In the terminology of Epstein and Schneider (2007) or Epstein and Halevy (2019), the coarse feedback about the composition of the Red urns can be viewed as an ambiguous signal. To cope with the ambiguous nature of the feedback in a simple way, we propose adding to the previous models (the similarity-based reinforcement learning and the Bayesian model) an ambiguity discount to the assessment of the Red urns. In the statistical exercise, the ambiguity discount is estimated for each learning model on the basis of the observed data, and a key question of interest is whether a non-null discount is applied to the Red urns in this context. Beyond the estimation exercise within each approach, another objective is to analyze which of the similarity-based reinforcement learning or the generalized Bayesian learning model explains best the observed data.

Our main findings are as follows. First, in our estimation of the similarity-based reinforcement learning model, we find that there is no ambiguity discount. That is, despite the inherent ambiguity of the feedback received about the Red urns, the Red urns are not discounted more than the familiar urns. This is similar to what is being assumed in the valuation equilibrium approach, even if to account for the steady state of the learning model that we propose, there is a need to extend the notion of valuation equilibrium to allow for noisy best-responses. Second, we find that the similarity-based reinforcement learning model explains the observed data much better than the generalized Bayesian learning model. In the last part, we discuss various robustness checks of the main findings. To the extent that the valuation equilibrium has properties very different from those arising with ordinary maximization (see Jehiel and Samet, 2007), we believe our experimental finding calls for pursuing further the implications of valuation equilibrium in economic contexts involving familiar and unfamiliar choices beyond the stylized lab examples considered in our experiment.

2 Related Literature

While the experimental literature on ambiguity is vast, there are only few experimental papers looking at ambiguous signals as we do (beyond Epstein and Halevy, we are only aware of Fryer et al (2019)). Note though that our experiment has a distinctive feature

not present in the previous experiments on ambiguous signals. In our setting, the nature of the ambiguity of the received signals (feedback) is endogenously shaped by the choice of subjects (if the Red urn is only chosen in state 1, there is no ambiguity as the feedback about Red urns is then clearly only informative about the composition of the Red urn in state 1; by contrast ambiguity seems somehow maximal if the Red urn is picked with the same frequency in the two states). This endogenous character of the ambiguity has no counterpart in the previous experiments on ambiguity, as far as we know. Our paper is related to other strands of literature beyond the references already mentioned. A first line of research related to our study is the framework of case-based decision theory as axiomatized by Gilboa and Schmeidler (1995). Compared to case-based decision theory, in the valuation equilibrium approach, the similarity weights given to the various actions in the various states happen to be endogenously shaped by the strategy used by the subjects, an equilibrium feature that is absent from the subjective perspective adopted in Gilboa and Schmeidler.

Another line of research related to our study includes the possibility that the strategy used by subjects would not distinguish behaviors across different states (Samuelson (2001), Mengel (2012) for theory papers and Grimm and Mengel (2012), Cason et al (2012) or Cownden et al. (2018) for experiments). Our study differs from that line of research in that subjects do adjust their behavior to the state but somehow mix the payoff consequences of some actions (the unfamiliar ones) obtained over different states, thereby revealing that our approach cannot be captured by a restriction on the strategy space. Another line related to our study is that of the analogy-based expectation equilibrium (Jehiel (2005) and Jehiel and Koessler (2008)) in which beliefs about other players' behaviors are aggregated over different states. Our study differs from that literature in that we are considering decision problems and not games. Yet, viewing nature as a player would allow to see closer connections between the two approaches. To the best of our knowledge, no experiment in the vein of the analogy-based expectation equilibrium has considered environments similar to the one considered here.

A related experimental literature includes a recent strand concerned with selection neglect. Experimental papers in this vein include Esponda and Vespa (2018), Enke (2019) or Barron et al. (2019). These papers conclude in various applications that subjects tend to ignore that data they see are selected. In our setting, the data related to Red are selected, and one can argue that subjects by behaving in agreement with the (generalized) valuation equilibrium do not seem to account for selection. Another related recent strand of experimental literature is concerned with the failure of contingent reasoning and/or some form of correlation neglect (see Enke and Zimmerman (2019), Martinez-Marquina et al (2019) or Esponda and Vespa (2019)). Some of these papers (see in particular Martinez-Marquina et al.) conclude that hypothetical thinking is more likely to fail in the presence of uncertainty, which somehow agrees with our finding that in the presence of aggregate feedback, subjects

find it hard to disentangle the value of choosing Red in the two states.

There is a number of contributions comparing reinforcement learning models to belief-based learning models in normal form games. While some of these contributions conclude that reinforcement learning models explain better the observed experimental data than belief-based learning models (Roth and Erev 1998, Camerer and Ho 1999), others suggest that it is not so easy to cleanly disentangle between these models (Salmon 2001, Hopkins 2002, Wilcox 2006). Our study is not much related to this debate to the extent that we consider decision problems and not games and that subjects do not immediately experience the payoff consequences of their choices (the feedback received concerns all subjects in the lab and subjects are only informed at the end how much they themselves earned). Relatedly the feedback received about some possible choices is aggregated over different states, which was not considered in the previous experimental literature. Despite these differences, relating Bayesian learning models to belief-based learning models, our results suggest that these perform less well than their reinforcement learning counterpart in our context, as in these other works.

Finally, one should mention the experimental work of Charness and Levin (2005) who consider decision problems in which, after seeing a realization of payoff in one urn, subjects have to decide whether or not to switch their choices of urns. In an environment in which subjects have a probabilistic knowledge about how payoffs are distributed across choices and states (but have to infer the state from initial information), Charness and Levin observe that when there is a conflict between Bayesian updating and Reinforcement learning, there are significant deviations from optimal choices. While the conclusion that subjects may rely on reinforcement learning more than on Bayesian reasoning is somehow common in their study and our experiment, the absence of ex ante statistical knowledge about the distribution of payoffs across states in our experiment makes it clearly distinct from Charness and Levin's experiment. In our view, the absence of ex ante statistical knowledge fits better the motivating economic examples mentioned above.

3 Background and theory

In the context of our experiment, this section defines a generalization of the valuation equilibrium allowing for noisy best-responses in the vein of the quantal response equilibrium (McKelvey and Palfrey, 1995). We next propose two families of learning models, a similarity-based reinforcement learning model (allowing for coarse feedback on some alternatives and an ambiguity discount attached to those)⁵ as well as a generalized Bayesian model (also allowing for noisy best-responses and a discount on alternatives associated to

⁵When there is no ambiguity discount, the long run properties of the similarity-based reinforcement learning model correspond to the generalized valuation equilibrium.

coarse feedback). The learning models will then be estimated and compared in terms of fit in light of our experimental data.

3.1 Quantal valuation equilibrium

In the context of our experiment, there are two states $s = 1$ and 2 that are equally likely. In state $s = 1$, the choice is between *Blue* and *Red*₁. In state $s = 2$, the choice is between *Green* and *Red*₂. The payoffs attached to these four alternatives are denoted by $v_{Blue} = 0.3$, v_{Red_1} , v_{Red_2} and $v_{Green} = 0.7$ where v_{Red_1} and v_{Red_2} are left as free variables to accommodate the payoff specifications of the various treatments.

A strategy for the decision maker can be described as $\sigma = (p_1, p_2)$ where p_i denotes the probability that *Red* _{i} is picked in state $s = i$ for $i = 1, 2$. Following the spirit of the valuation equilibrium (Jehiel and Samet, 2007), a single valuation is attached to *Red*₁, *Red*₂ so as to reflect that subjects in the experiment only receive aggregate feedback about the payoff obtained when a Red urn is picked either in state $s = 1$ or 2 . Accordingly, let $v(Red)$ be the valuation attached to *Red*. Similarly, we denote by $v(Blue)$ and $v(Green)$ the valuations attached to the Blue and Green urns, respectively.

In equilibrium, we require that the valuations are consistent with the empirical observations as dictated by the equilibrium strategy $\sigma = (p_1, p_2)$. This implies that $v(Blue) = v_{Blue}$, $v(Green) = v_{Green}$ and more interestingly that

$$v(Red) = \frac{p_1 \times v_{Red_1} + p_2 \times v_{Red_2}}{p_1 + p_2} \quad (1)$$

whenever $p_1 + p_2 > 0$. That is, $v(Red)$ is a weighted average of v_{Red_1} and v_{Red_2} where the relative weight given to v_{Red_1} is $p_1/(p_1 + p_2)$ given that the two states $s = 1$ and 2 are equally likely and *Red* _{i} is picked with probability p_i for $i = 1, 2$. Based on the valuations $v(Red)$, $v(Blue)$ and $v(Green)$, the decision maker is viewed as picking a noisy best-response where we consider the familiar logit parameterization (with coefficient λ). Formally,

Definition: A strategy $\sigma = (p_1, p_2)$ is a quantal valuation equilibrium if there exists a valuation system $(v(Blue), v(Green), v(Red))$ where $v(Blue) = 0.3$, $v(Green) = 0.7$, $v(Red)$ satisfies (1), and

$$\begin{aligned} p_1 &= \frac{e^{\lambda v(Red)}}{e^{\lambda v(Red)} + e^{\lambda v(Blue)}} \\ p_2 &= \frac{e^{\lambda v(Red)}}{e^{\lambda v(Red)} + e^{\lambda v(Green)}} \end{aligned}$$

It should be stressed that the determination of $v(Red)$, p_1 and p_2 are the results of a fixed point as the strategy $\sigma = (p_1, p_2)$ affects $v(Red)$ through (1) and $v(Red)$ determines the strategy $\sigma = (p_1, p_2)$ through the two equations just written.

We now briefly review how the quantal valuation equilibria look like in the payoff specifications corresponding to the various treatments. In this review, we consider the limiting case in which λ goes to ∞ (thereby corresponding to the valuation equilibria as defined in Jehiel and Samet, 2007).

Treatment 1: $v_{Red_1} = 0.4$ and $v_{Red_2} = 0.8$

In this case, clearly $v(Red) > v(Blue) = 0.3$ (because $v(Red)$ is some convex combination between 0.4 and 0.8). Hence, the optimality of the strategy in state $s = 1$ requires that the Red urn is always picked in state $s = 1$ ($p_1 = 1$). Regarding state $s = 2$, even if Red_2 were picked with probability 1, the resulting $v(Red)$ that would satisfy (1) would only be $\frac{0.4+0.8}{2} = 0.6$, which would lead the decision maker to pick the Green urn in state $s = 2$ given that $v(Green) = 0.7$. It follows that the only valuation equilibrium in this case requires that $p_2 = 0$ so that the Red urn is only picked in state $s = 1$ (despite the Red urns being payoff superior in both states $s = 1$ and 2). In this equilibrium, consistency (i.e., equation (1)) implies that $v(Blue) < v(Red) = 0.4 < v(Green)$.

Treatment 2: $v_{Red_1} = 1$, $v_{Red_2} = 0.6$

In this case too, $v(Red) > v(Blue) = 0.3$ (because any convex combination of 0.6 and 1 is larger than 0.3) and thus $p_1 = 1$. Given that $v_{Red_2} < v_{Red_1}$, this implies that the lowest possible valuation of *Red* corresponds to $\frac{1+0.6}{2} = 0.8$ (obtained when $p_2 = 1$). Given that this value is strictly larger than $v(Green) = 0.7$, we obtain that it must that $p_2 = 1$, thereby implying that the Red urns are picked in both states. Valuation equilibrium requires that $p_1 = p_2 = 1$ and consistency implies that $v(Blue) < v(Green) < v(Red) = 0.8$.

Treatment 3: $v_{Red_1} = 0.1$, $v_{Red_2} = 0.9$

In this case, we will show that the Red urns are not picked neither in state 1 nor in state 2. To see this, assume by contradiction that the Red urn would (sometimes) be picked in at least one state. This should imply that $v(Red) \geq v(Blue)$ (as otherwise, the Red urns would never be picked neither in state $s = 1$ nor 2). If $v(Red) < v(Green)$, one should have that $p_2 = 0$, thereby implying by consistency that $v(Red) = v_{Red_1} = 0.1$. But, this would contradict $v(Red) \geq v(Blue) = 0.3$. If $v(Red) \geq v(Green)$, then $p_1 = 1$ (given that $v(Red) > v(Blue)$), and thus by consistency $v(Red)$ would be at most equal to $\frac{0.1+0.9}{2} = 0.5$ (obtained when $p_2 = 1$). Given that $v(Green) = 0.7 > 0.5$, we get a contradiction, thereby implying that no Red urn can be picked in a valuation equilibrium.

As explained above the value of $v(Red)$ in the valuation equilibrium varies from being below $v(Blue)$ in treatment 3 to being in between $v(Blue)$ and $v(Green)$ in treatment 1 to being above $v(Green)$ in treatment 2, thereby offering markedly different predictions according to the treatment in terms of long run choices. Allowing for noisy as opposed to exact best-responses would still allow to differentiate the behaviors across the treatments but in a less extreme form (clearly, if $\lambda = 0$ behaviors are random and follow the lottery 50 : 50 in every state and every treatment, but for any $\lambda > 0$, behaviors are different across

treatments).

3.2 Learning Models

We will consider two families of learning models to explain the choice data observed in the various treatments of the experiment: A similarity-based version of reinforcement learning model in which choices are made on the basis of the valuations attached to the various colors of urns and valuations are updated based on the observed feedback, and a Bayesian learning model in which subjects update their prior belief about the composition of the Red urns based on the feedback they receive. In each case, we will assume that subjects care only about their immediate payoff and do not integrate the possible information content that explorations outside what maximizes their current payoff could bring. This is -we believe- justified to the extent that in the experiment there are twenty subjects making choices in parallel and that the feedback is anonymous making the informational value of the experimentation by a single subject rather small (it would be exactly 0 if we were to consider infinitely large populations of subjects and we are confident it is negligible when there are twenty subjects).

3.3 Similarity-based reinforcement learning

Standard reinforcement learning models assume that strategies are reinforced as a function of the payoff obtained from them. In the context of our experiment, subjects receive feedback about how the choices made by all subjects in the previous period translated into black (positive payoff) or white (null payoff) draws. More precisely, the feedback concerns the number⁶ of Black balls drawn when a *Blue*, *Green* or *Red* urn was picked in the previous period as well as the number of times an urn with that color was then picked. Unlike standard reinforcement learning, payoff obtained from some actions are coarse in our setting and hence similarity- based reinforcement. Accordingly, at each time $t = 2, \dots, 70$, one can define for each possible color $C = B, R, G$ (for Blue, Red, Green) of urn(s) that was picked at least once at $t - 1$:

$$UC_t = \frac{\#(\text{Black balls drawn in urns with color } C \text{ at } t - 1)}{\#(\text{an urn with color } C \text{ picked at } t - 1)}. \quad (2)$$

UC_t represents the strength of urn(s) with color C as reflected by the feedback received at t about urns with such a color. Note the normalization by $\#(\text{an urn with color } C \text{ picked at } t - 1)$ so that UC_t is comparable to a single payoff attached to choosing an urn with color C .

We will let BC_t denote the value attached to an urn with color C at time t and BC_{init} denote the initial value attached to an urn with that color. For *Green* and *Blue* there is

⁶The symbol $\#$ is used to refer to number.

initial information and it is natural to assume that

$$\begin{aligned} BB_{init} &= \frac{30}{100} = 0.3 \\ BG_{init} &= \frac{68}{100} = 0.68 \end{aligned}$$

whereas for *Red*, the initial value BR_{init} is a priori unknown and it will be estimated in light of the observed choice data.

Dynamics of BC_t :

Concerning the evolution of BC_t , we assume that for some (ρ_U, ρ_F) , we have:⁷

$$\begin{aligned} BR_t &= \rho_U \times BR_{t-1} + (1 - \rho_U) \times UR_t \\ BB_t &= \rho_F \times BB_{t-1} + (1 - \rho_F) \times UB_t \\ BG_t &= \rho_F \times BG_{t-1} + (1 - \rho_F) \times UG_t \end{aligned}$$

In other words, the value attached to color C at t is a convex combination between the value attached at $t-1$ and the strength of C as observed in the feedback at t . Observe that we allow the weight to be assigned to the feedback to be different for the Red urns on the one hand and the Blue and Green urns on the other to reflect the idea that when a choice is better known as is the case for more familiar alternatives (here identified with urns *Blue* and *Green*) the new feedback may be considered as less important to determine the value of it. Accordingly, we would expect that ρ_F is larger than ρ_U , and we will be concerned whether this is the case in our estimations.⁸

Choice Rule:

Given that the feedback concerning the Red urns is aggregated over states $s = 1$ and 2 , there is extra ambiguity as to how well BR_t represents the valuation of *Red*₁ or *Red*₂ as compared to how well BG_t or BB_t represent the valuations of *Blue* and *Green*. The valuation equilibrium (or its quantal extension as presented above) assumes that BR_t is used to assess the strength of *Red*_s whatever the state $s = 1, 2$. In line with the literature on ambiguity aversion as experimentally initiated by Ellsberg (1961), it is reasonable to assume that when assessing the urn *Red*_s, $s = 1, 2$, subjects apply a discount $\delta \geq 0$ to

⁷In case no urn of color C was picked at $t-1$, then $UC_t = BC_{t-1}$ so that $BC_t = BC_{t-1}$.

⁸There are many variants that could be considered. For example, one could have made the weight of the new feedback increase linearly or otherwise with the number of times an urn with that color was observed. One could also have considered that the weight on the feedback is a (decreasing) function of t so as to reflect that as more experience accumulates, new feedback becomes less important. These extensions did not seem to improve how well we could explain the data and therefore, we have chosen to adopt the simpler approach described in the main text.

BR_t .⁹ Allowing for noisy best-responses in the vein of the logit specification, this would lead to probabilities p_{1t} and p_{2t} of choosing Red_1 and Red_2 as given by

$$p_{1t} = \frac{e^{\lambda(BR_t - \delta)}}{e^{\lambda(BR_t - \delta)} + e^{\lambda BB_t}}$$

$$p_{2t} = \frac{e^{\lambda(BR_t - \delta)}}{e^{\lambda(BR_t - \delta)} + e^{\lambda BG_t}}$$

The learning model just described is parameterized by $(\rho_U, \rho_F, \delta, \lambda, BR_{init})$. In the next Section, these parameters will be estimated pooling the data across all three treatments using the maximum likelihood method. Particular attention will be devoted to whether $\delta > 0$ is needed to explain better the data, whether $\rho_F > \rho_U$ as common sense suggests, as well as to the estimated value of λ and the obtained likelihood for comparison with the Bayesian model to be described next.¹⁰

3.4 Generalized Bayesian Learning Model

As an alternative learning model, subjects could form some initial prior belief regarding the compositions of Red_1 and Red_2 , say about the chance that there are k_i black balls out of 10 in Red_i , and update these beliefs after seeing the feedback using Bayes' law.

Let us call $\beta_{init}(k_1, k_2)$ the initial prior belief of subjects that there are k_i black balls out of 10 in Red_i . In the estimations, we will allow the subjects to consider that the number of black balls in either of the two Red urns can vary between k_{inf} and k_{sup} with $0 \leq k_{inf} \leq k_{sup} \leq 10$ and we will consider the uniform distribution¹¹ over the various possibilities. That is, for any $(k_1, k_2) \in [k_{inf}, k_{sup}]^2$

$$\beta_{init}(k_1, k_2) = \frac{1}{(k_{sup} - k_{inf} + 1)^2},$$

and $\beta_{init}(k_1, k_2) = 0$ otherwise. The values of k_{inf} and k_{sup} will be estimated.

Dynamics of the beliefs:

⁹One possible rationale following the theoretical construction of Gilboa and Schmeidler (1989) is that the proportion of Black balls in Red_1 and Red_2 is viewed as being in the range $[BR - \delta, BR + \delta]$ and that subjects adopt a maxmin criterion, leading them to consistently use $BR - \delta$ to assess both Red_1 and Red_2 . More elaborate specifications of ambiguity would be hard to estimate given the nature of our data.

¹⁰If $\delta = 0$, it can be shown that (p_{1t}, p_{2t}) converges to the quantal valuation equilibrium as defined in subsection 3.1.

¹¹This choice of prior is somewhat arbitrary. For robustness check, we have also considered the case of triangular priors centered around 5, which together with the uniform prior considered in the main analysis allows us to span a wider range of priors. As it turns out the uniform prior leads to a better fit so that our main finding that the generalized reinforcement learning model outperforms the Bayesian learning model would a fortiori be true had we considered the triangular prior.

To simplify the presentation a bit, we assume there is no learning on the urns *Blue* and *Green* for which there is substantial initial information. At time $t + 1$, the feedback received by a subject can then be formulated as (b, g, n) where b, g are the number of blue and green urns respectively that were picked at t , and n is the number of black balls drawn from the *Red* urns. In the robustness checks, we allow for Bayesian updating also on the compositions of the Blue and Green urns, and obtain that adding learning on those urns does not change our conclusion.

To further simplify the presentation, we assume that in the feedback subjects are exposed to, there is an equal number of states $s = 1$ and $s = 2$ decisions assumed by the subjects (allowing the subjects to treat these numbers as resulting from a Bernoulli distribution would not alter our conclusions, see the robustness check section for elaborations). In this case, the feedback can be presented in a simpler way, because knowing (b, g, n) now allows subjects to infer that $m_1 = 10 - b$ choices of *Red* urns come from state $s = 1$ and $m_2 = 10 - g$ choices of *Red* urns come from state $s = 2$. Accordingly, we represent the feedback as (m_1, m_2, n) where m_i represents the number of *Red_i* that were picked. Clearly, the probability of observing m_1, m_2, n when there are k_1 and k_2 black balls in *Red₁* and *Red₂* respectively is given by:

$$Pr(m_1, m_2, n | k_1, k_2) = \sum_{\substack{n_1 \leq m_1 \\ n_2 \leq m_2 \\ n_1 + n_2 = n}} \binom{m_1}{n_1} \binom{m_2}{n_2} (k_1/10)^{n_1} (1 - k_1/10)^{m_1 - n_1} (k_2/10)^{n_2} (1 - k_2/10)^{m_2 - n_2}$$

$$\text{where } \binom{a}{b} = \frac{a!}{(a-b)!b!} \text{ for integers } a, b \text{ with } a \geq b.$$

The posterior at $t + 1$ about the probability that there are k_1 and k_2 black balls out of ten in *Red₁* and *Red₂* after observing (m_1, m_2, n) at t is then derived from Bayes' law by

$$\beta_{t+1}(k_1, k_2) = \frac{\beta_t(k_1, k_2) \cdot \Pr(m_1, m_2, n | k_1, k_2)}{\sum_{r_1, r_2} \beta_t(r_1, r_2) \cdot \Pr(m_1, m_2, n | r_1, r_2)}.$$

with $\beta_1(k_1, k_2) = \beta_{init}(k_1, k_2)$.

Define $v_t^{Bayes}(Red_i) = \sum_{k_i} \frac{k_i}{10} \beta_t(k_i)$ where $\beta_t(k_i) = \sum_{k_{-i}} \beta_t(k_i, k_{-i})$ as the time t expected proportion of black balls in *Red_i* given the distribution β_t .

Choice Rule:

As for the similarity-based reinforcement learning model, we allow for noisy best responses and we introduce an ambiguity discount δ for the evaluation of the Red urns.¹²

¹²Some might dispute that the ambiguity discount is not so much in the spirit of the Bayesian model in which case one should freeze this parameter to be 0.

Accordingly, the probabilities p_{1t} and p_{2t} of choosing Red_1 and Red_2 at time t in the generalized Bayesian learning model are given by:

$$p_{1t} = \frac{e^{\lambda(v_t^{Bayes}(Red_1) - \delta)}}{e^{\lambda(v_t^{Bayes}(Red_1) - \delta)} + e^{\lambda v(Blue)}}$$

$$p_{2t} = \frac{e^{\lambda(v_t^{Bayes}(Red_2) - \delta)}}{e^{\lambda(v_t^{Bayes}(Red_2) - \delta)} + e^{\lambda v(Green)}}$$

where as our simplification implies we assume that $v(Blue) = 0.3$ and $v(Green) = 0.7$.¹³

Studying the dynamics of the above Bayesian learning model is a bit cumbersome for general specifications of $(k_{inf}, k_{sup}, \lambda, \delta)$. But to illustrate how it leads to predictions markedly different from those of the valuation equilibrium, consider the case in which $\delta = 0$, $k_{inf} = 0$, $k_{sup} = 10$ and $\lambda = \infty$. Then in all treatments, Red_2 is not chosen to start with given that it is perceived to deliver 0.5 in expectation, which is less than 0.7. As a result, subjects can safely attribute the feedback they receive about Red to be coming from Red_1 . This in turn implies that (considering the limiting case with large population of subjects) subjects eventually learn the value of Red_1 and never play Red_2 . Thus, subjects play Red_1 in treatments 1 and 2 but give up playing Red_1 in treatment 3, and they never play Red_2 in any of the treatments (by contrast, Red_2 was played in treatment 2 in the valuation equilibrium).

More generally, the proposed (generalized) Bayesian learning model is parameterized by $(k_{inf}, k_{sup}, \lambda, \delta)$. In the next section, these parameters will be estimated by the maximum likelihood method in light of the collected data.

4 Results

4.1 Further Description of the Experimental Design

The computerized experiments were conducted in the Laboratory at Maison de Sciences Economiques (MSE) between March 2015 and November 2016, with some additional sessions running in March 2017. Upon arrival at the lab, subjects sat down at a computer terminal to start the experiment. Instructions were handed out and read aloud before the start of each session. The experiment consisted of three main treatments which varied in the payoffs of the Red urns as explained above. In addition we had two other treatments referred to as controls in which subjects received state-specific feedback about the Red urns, i.e the feedbacks for Red_1 and Red_2 appeared now in two different columns, for the two payoff specifications of treatments 1 and 2. The purpose of these control treatments was to check whether convergence to optimal choices was observed in such more standard feedback

¹³As previously mentioned, we present a model that allows subjects to update $v(Blue)$ and $v(Green)$ according to Bayes rule in the robustness checks.

scenarios.

Each session involved 18-20 subjects¹⁴ and four sessions were run for each treatment and control. Overall, 235 subjects drawn from the participant pool at the MSE -who were mostly students- participated in the experiment. Each session had seventy rounds. In all treatments, all sessions, and all rounds, subjects were split up equally into two states, State 1 and State 2. Subjects were randomly assigned to a new state at the start of each round. The subjects knew the state they were assigned to, but did not know the payoff attached to the available actions in each state.¹⁵ In each state, players were asked to choose between two actions as detailed in Figure 1. The feedback structure for the main treatments was as explained above. For the control group, the information structure was disaggregated. We use this as a baseline to show that under simpler feedback structure, individuals learn optimally the best available option.

Subjects were paid a show-up fee of 5 €. In addition to this, they were given the opportunity to earn 10 € depending on their choice in the experiment. Specifically, for each subject, two rounds were drawn at random and a subject earned an extra 5 € for each black ball that was drawn from their chosen urn in these two rounds. The average payment was around 11 € per subject, including the turn-up fee. All of the sessions lasted between 1 hour and 1.5 hour, and subjects took longer to consider their choices at the start of the experiment.

4.2 Preliminary findings

We first present descriptive statistics and next present the structural analysis. In Figure 3, we report how the choices of urns vary with time and across treatments. Across all these sessions, initially, subjects are more likely to choose the Red urn than the Blue urn in state 1 and they are more likely to choose the Green urn than the Red urn in state 2. This is, of course, consistent with most theoretical approaches including the ones discussed above given that the Green urn is more rewarding than the Blue urn and the Red urns look (at least initially) alike in states 1 and 2.

The more interesting question concerns the evolution of choices. Roughly, in state 1, we see toward the final rounds, a largely dominant choice of the Red urns in treatments 1 and 2 whereas Red in state 1 is chosen less than half the time in treatment 3.

Concerning state 2, we see that in the final rounds, the Red urns are rarely chosen in treatments 1 and 3 and chosen with high frequency in treatment 2. The qualitative differences of the choices in the final rounds among the three treatments and the two states are in line with the prediction of the valuation equilibrium even if some noise in the best-response

¹⁴Note that when 18 subjects participated in the session, Bayes updating was modified accordingly.

¹⁵For urns *Blue* and *Green*, they had initial information, as explained in the Introduction.

is obviously needed especially for treatment 3 in state 1 to explain why about 40% of choices correspond to Red.¹⁶

In Figure 4, with state-specific feedback for the Red urns, we see a clear trend toward the optimal choices even if some noise would be needed to explain why only 49% of choices correspond to Red in state 2 in Control 1. In contrast to the feedback structure in the treatment group, we see that disaggregating feedback on the Red urns across states, players learn the optimal choice. In line with section 3.1, the fine feedback helps the agent attach a valuation $v(Red1)$ and $v(Red2)$ separately for the *Red* urns in the two states instead of a joint valuation $v(Red)$. Due to this finer feedback structure, the simple heuristic of reinforcement learning leads to an optimal choice, unlike in the control treatments in which an analogous reinforcement learning heuristic leads to valuation equilibrium.

4.3 Statistical estimations

Similarity-based reinforcement learning

The estimations of the parameters of the similarity-based reinforcement learning model together with the corresponding log likelihood¹⁷ is given in the following Table 2. Concerning the likelihood, by way of comparison, a complete random choice model where in every state, subjects would randomize 50:50 between the two choices would result in a negative log likelihood of $L=11402$, which is much higher than 7626.6. More generally, the similarity-based reinforcement learning model explains data much better than any model in which behavior would not be responsive to feedback.¹⁸ We now discuss the most salient aspects of the estimations.

The finding that $\rho_F > \rho_U$ seems natural as mentioned above, to the extent that for the familiar urns, the feedback should affect less how the valuations are updated.

The finding that BR_{init} is slightly below 0.5 may be interpreted along the following lines. In the absence of any information, an initial value of 0.5 would be the one dictated by the principle of insufficient reason, but the uncertainty attached to the unfamiliar urns may lead to some extra discount in agreement with some form of ambiguity aversion as reported in Ellsberg.¹⁹

¹⁶We note that the large share of Red chosen in state 2 of treatment 2 is not in line with the noisy version of the Bayesian learning model as explained above.

¹⁷Likelihood throughout the paper refers to the negative of the log likelihood. Thus, the lower the likelihood, the better the model (See textbook Train 2003 for further details). Standard errors are reported in brackets.

¹⁸Optimizing on the probability of Red_1 vs Red_2 in such a model, would lead to assume that Red_1 is chosen with probability $p_1=0.7$ and Red_2 is chosen with probability $p_2=0.3$ with a negative log likelihood of $L=9899.6$ which is much higher than 7626.6

¹⁹The difference $0.5 - 0.44 = 0.06$ can be interpreted as measuring the ambiguity aversion of choosing an unfamiliar urn when no feedback is available.

The most interesting observation concerns δ which is estimated to be 0. Even though the feedback for the Red urns is ambiguous (because it is aggregated over the two states), the valuations for Red are not discounted as if subjects were ambiguous neutral from that perspective. Thus, what our estimation suggests is that while there may be some (mild) initial ambiguity aversion relative to the unfamiliar choices (as reflected by BR_{init} being smaller than 0.5), no ambiguity discount seems to be applied to the valuation of Red despite the ambiguity attached to the feedback received about the Red urns.

Generalized Bayesian learning model:

The estimated parameters for the generalized Bayesian learning model is given in Table 3. The value of $\delta = 0.003$ implies that with the Bayesian model, the subjects show some mild form of ambiguity aversion. However we cannot statistically reject the hypothesis that $\delta = 0$, which implies that with the Bayesian model too, there is no significant ambiguity discount similarly to what we found in the similarity-based reinforcement learning estimations. For the support of initial prior, we found that $k_{inf} = 3$ and $k_{sup} = 7$.²⁰ We also note that the value of λ is slightly higher than that for the reinforcement model.

Comparing the two models:

Maybe the most important question is which of the Bayesian learning model or the reinforcement learning model explains the experimental data best. We use three methods of comparisons, all establishing that the reinforcement learning model outperforms the Bayesian learning model. First, looking at the likelihood of the two models, we see that the Bayesian learning Model explains less well the data than the similarity-based reinforcement learning model. Second, to account for the difference in the number of parameters between the two models, we use the Bayesian Information Criterion (BIC) or Schwarz criterion (also SBC, SBIC). BIC is a criterion for model selection among a finite set of models where the model with the lower BIC is closer to the data generating process. It is based, in part, on the likelihood function to determine the goodness of fit in the two models accounting for the no. of parameters, formally defined as

$$BIC = \ln(n)k - 2 \ln(L^*)$$

where L^* is value of maximized likelihood of model M, n is the number of observations, k is the number of parameters estimated by the model. As seen from Table 4, we can conclude that the reinforcement model performs better than the Bayesian one in explaining the data.

Finally, we perform a Vuong test²¹ to compare the performance of the two models statistically. Under the null hypothesis H_0 , that both models perform equally well, we

²⁰The value of the bounds correspond to $v_{Blue}=0.3$ and $v_{Green}=0.7$ respectively and so one may speculate that maybe the compositions of the familiar urns serve as anchoring the support of the priors. Observe that because best-responses are noisy, the derived support does not imply that the Red urn is always picked in state 1 and never picked in state 2.

²¹See Merkel et al. 2016 for more details.

conclude that the null can be rejected in favor of the reinforcement model. Specifically,

$$H_0 = E(L(\theta_R; x_d)) = E(L(\theta_B; x_d))$$

$$H_a = E(L(\theta_R; x_d)) \neq E(L(\theta_B; x_d))$$

where x_d is the collection of observed individual data points, θ_R is the set of parameters estimated via reinforcement learning, θ_B is the set of parameters estimated via Bayesian learning, $L(\theta_R; x_d)$ is the log likelihood under reinforcement model and $L(\theta_B; x_d)$ is the log likelihood under Bayesian learning model for each data point d . The Vuong statistics is then defined by

$$V_{stat} = \sqrt{N} \frac{\bar{m}}{S_m}$$

where $\bar{m} = E(L(\theta_R; x_d)) - E(L(\theta_B; x_d))$ for each individual d , N is the total number of observations and S_m is the sample standard deviation.

V_{stat} tests the null hypothesis (H_0) that the two models are equally close to the true data generating process, against the alternative H_1 that one of the model is closer.²² The obtained $V_{stat} = 25.01$ being large and positive implies that the reinforcement model is a better fit to our experimental data than the Bayesian model. This is in line with the findings derived with the BIC.

4.4 Comparing the Reinforcement learning model to the data

It is of interest to see how the obtained frequencies of choices as generated by the similarity-based reinforcement learning model with estimations as reported in Table 2 compare to the observed frequencies from our experimental data. In Figure 5, we report the simulated frequencies of urn choices using the reinforcement model across all time periods and treatments. Across all these sessions, our simulated frequencies remain close to the actual frequencies with a slightly less good fit in Treatment 1. Allowing for a different λ in Treatment 1, we observe that a larger lambda significantly improves the fit in this treatment as shown in Appendix C.

4.5 Individual level heterogeneity

While we have established that the similarity-based reinforcement learning model explains the data better than its Bayesian counterpart, it is of interest to see how the fit as generated by such a model with estimations as reported in Section 4.3 vary across individuals. The

²²Vuong test compares the predicted probabilities of two non nested models. It computes the difference in likelihood for each observation i in the data. A high positive V_{stat} implies Model 1 is better than Model 2 where $\bar{m} = \log(Pr(x_i|Model1)) - \log(Pr(x_i|Model2))$

pool of 235 subjects had very similar backgrounds across treatments. We fit the parameters obtained from the two learning models to each individual across 70 rounds. That is, for each individual, we ask for each of the two learning models, the generalized reinforcement learning model with parameters as in Table 2 or the generalized Bayesian model with parameters as in Table 3 allows to better explain the observed choices of the individual over the 70 rounds given the feedback he has been exposed to. And for the model comparison, we use the Vuong test as explained above. Overall, we find that 82.97% of players can be categorized as reinforcement learners and the rest as Bayesian learners.

Comparing across treatments, we find that T1 has the highest proportion of Reinforcement learners (93.5%). In T2, the proportion of similarity-based reinforcement learners is 79.22% followed by 76.25% in T3. The higher proportion of Reinforcement learners in T1 can be related to the previous observation that in aggregate in the Reinforcement learning model with lower level of noise allows for a better fit in T1.

We have also compared the overall payoff as measured by the sum of black balls obtained over the 70 rounds across the population of subjects categorized as reinforcement learners or Bayesian learners. For reinforcement learners, the average was 42.09 black balls whereas, for Bayesian learners, an average of 40.55 black balls was observed. This difference is not statistically significant.

5 Robustness Checks

As many variants of reinforcement learning models and Bayesian models could be considered, we review a few of these here and suggest that our basic conclusions remain the same in these variants. In each case, the reported estimation relies on the same methodology as above.

Similarity-based reinforcement learning model

Regarding reinforcement models, we consider the following variants. First, we allow the speed of adjustment of the valuation of the Red urns to differ across the two states as a large imbalance in the number of green urns as opposed to blue urns in the feedback is indicative that the feedback concerned more Red urns in state 1 than in state 2.²³ Specifically, we now introduce a new parameter μ and specify the weight on the previous valuation to satisfy

$$\begin{aligned}\rho_{U1} &= \rho_U \times [1 - \mu \cdot (\frac{NB}{NG + NB} - 0.5)] \\ \rho_{U2} &= \rho_U \times [1 - \mu \cdot (\frac{NG}{NG + NB} - 0.5)]\end{aligned}$$

²³This is in some sense making use of some qualitative features of the Bayesian model to improve the reinforcement learning model.

where NB and NG are the respective numbers of Blue and Green urns appearing in the feedback. One would expect μ to be negative so that when NB is observed to be smaller than NG , subjects infer the new feedback on the Red urns is more informative on the composition of Red in state 1 than Red in state 2. The estimations of this extended model are reported in Table 5. Our estimation yields $\mu > 0$ but note that it is not significant and that $\mu=0$ cannot be rejected. Thus, this extended model does not explain the data better than our previously proposed version.

A different idea somewhat related to the one just discussed is that subjects would apply a different discount to the Red urn in state 1 and 2 maybe because they would consider the feedback for the Red urns to be more indicative of Red in state 1 than in state 2 (again maybe because of the imbalance of the number of Blue and Green urns in the feedback). This leads us to consider an extended version with two different discounts δ_1 and δ_2 for Red in state 1 and 2 while keeping the other aspects of the dynamics unchanged as compared to the main reinforcement learning model. That is, the only change in this variant is in the choice rule.

Choice Rule:

$$p_{1t} = \frac{\exp^{\lambda(BR_t - \delta_1)}}{\exp^{\lambda(BR_t - \delta_1)} + \exp^{\lambda BB_t}}$$

$$p_{2t} = \frac{\exp^{\lambda(BR_t - \delta_2)}}{\exp^{\lambda(BR_t - \delta_2)} + \exp^{\lambda BG_t}}$$

The estimated parameters for this variant are reported in Table 6.

In this variant, we see a slight discount for Red_2 but not for Red_1 . The likelihood for this model is better than for the original model and the hypothesis $\delta_1 = \delta_2 = 0$ is rejected under significance level 0.01. While this extension has a slightly better explanatory power, we find only a modest level of ambiguity aversion applied to the urn Red in state 2 when allowed to differ from the ambiguity aversion to the urn Red in state 1.²⁴

Generalized Bayesian learning model

For the Bayesian model, one could argue that instead of fixing $v(Blue) = 0.3$ and $v(Green) = 0.7$, the values of the Blue and Green urns could be updated similarly to the Red urns.²⁵ We have estimated such an extended model taking the same prior parameterized

²⁴We also considered the possibility that subjects would use a different slope to appreciate payoffs above 0.5 and payoffs below 0.5 in the spirit of prospect theory (with a reference payoff fixed at 0.5), but such a variant did not result in an improvement of the likelihood, hence we do not report it here (see Tversky and Kahneman (1974), (1979) for the introduction of prospect theory).

²⁵The initial information provided about those urns would of course be used.

by the support $[k_{\inf}, k_{\sup}]$ for all the urns in Table 7. This model performs better than the generalized Bayesian one in terms of likelihood. However, this extended model is still statistically dominated by the similarity-based reinforcement learning model.²⁶

A more elaborate version of the Bayesian approach would be to take into account the probability of having r_i state i in the 20 observation of the feedback (instead of assuming that in each round, there are exactly 10 subjects assigned to each state). Accordingly, we now represent the feedback as (b, g, n) where b, g are the number of draws from blue and green urns and n is the number of black balls in *Red*. We modify the generalized Bayesian model by taking into account the probability of having x states $s = 1$ out of 20. Formally,

$$Pr(b, g, n|k_1, k_2) = \sum_x \binom{20}{x} \left(\frac{1}{2^{20}} \right) Pr(m_1 = x - b, m_2 = 20 - x - g, n|k_1, k_2)$$

where x is the number of times state $s = 1$ was observed in one round, $Pr(m_1 = x - b, m_2 = 20 - x - g, n|k_1, k_2)$ is defined as in section 3.2.2 where the total number of players in each session is 20.

The dynamics of beliefs is now given by

$$\beta_{t+1}(k_1, k_2) = \frac{\beta_t(k_1, k_2) \cdot Pr(b, g, n|k_1, k_2)}{\sum_{r_1, r_2} \beta_t(r_1, r_2) \cdot Pr(b, g, n|r_1, r_2)}.$$

with $\beta_1(k_1, k_2) = \beta_{init}(k_1, k_2)$. The other ingredients of the Bayesian learning model are identical to those considered in section 3.2.2.

After running the estimation of this model in Table 8, we note that the corresponding likelihood further improved compared to the other two Bayesian models. However, even with the improved likelihood, the model still under performs when compared to the reinforcement model, and the Vuong test still statistically favors the similarity-based reinforcement learning model.²⁷

6 Conclusion

In this paper, we have considered the choices to be made between familiar alternatives and unfamiliar alternatives for which the obtained feedback is aggregated over different states of the economy. The literature on ambiguity aversion would suggest that the unfamiliar alternatives would be discounted as compared to the familiar ones, but that literature has

²⁶The Vuong test was conducted with the null hypothesis that both models explain the data equally well. The null was rejected in favor of the similarity-based reinforcement learning model with $V_{stat} = 24.72$.

²⁷The Vuong test was conducted with the null hypothesis that both models explain the data equally well. The null was rejected in favor of the similarity-based reinforcement learning model with $V_{stat} = 20.18$.

largely ignored how behaviors would change in the face of continuously coming new feedback that would remain aggregated over different states.

Several competing learning models could be considered to tackle the choices in the face of new feedback: either extensions of reinforcement models in the spirit of the valuation equilibrium (Jehiel and Samet, 2007) or Bayesian models in which subjects would start with some diffuse priors and update as well as they can, based on the coarse feedback they receive. Clearly, ideas of ambiguity aversion can be combined with such learning models along with the idea that subjects make noisy best-responses to their representations of the alternatives, as routinely done in the empirical literature (discrete choice models as considered by McFadden) or in the experimental literature (quantal response equilibrium as defined by McKelvey and Palfrey, 1995).










Our results indicate that the similarity-based reinforcement learning models outperform their Bayesian counterparts and that little discount seems to be applied to unfamiliar choices even when the feedback relative to them is aggregated over different states. As in other experimental findings, our results also indicate that subjects' choices are noisy, which we have tackled by assuming that subjects employ noisy best-responses. We believe such a work could be viewed as a starting point for an ambitious research agenda that aims at understanding how subjects make choices in the face of a mix of coarse and precise (state-specific) feedback. It seems well suited to cope with a number of choice problems in which one alternative is familiar and another one is not. Questions of whether subjects seek to generate state-specific feedback (and when) should also be part of this broader agenda.

Figures

Période n°9. État 2

In this period, you will make a choice in State 2.
 A cette période, vous faites un choix dans l'état 2.

Balls drawn from the urn of each colour in the previous period:
 Balles tirées dans l'urne de chaque couleur lors de la période précédente :

Urne	 (Blue)	 (Red)	 (Green)
Number of balls of each colour Nombre de balles de chaque couleur is the ball with positive payoff (● est une balle gagnante)	$\begin{array}{r} 2 \\ - \\ 1 \end{array}$  	$\begin{array}{r} 7 \\ - \\ 3 \end{array}$  	$\begin{array}{r} 5 \\ - \\ 2 \end{array}$  

Which urn do you want to choose?
 Quelle urne choisissez-vous ?



☒ 
☐ 

Figure 1 Feedback structure for treatment sessions

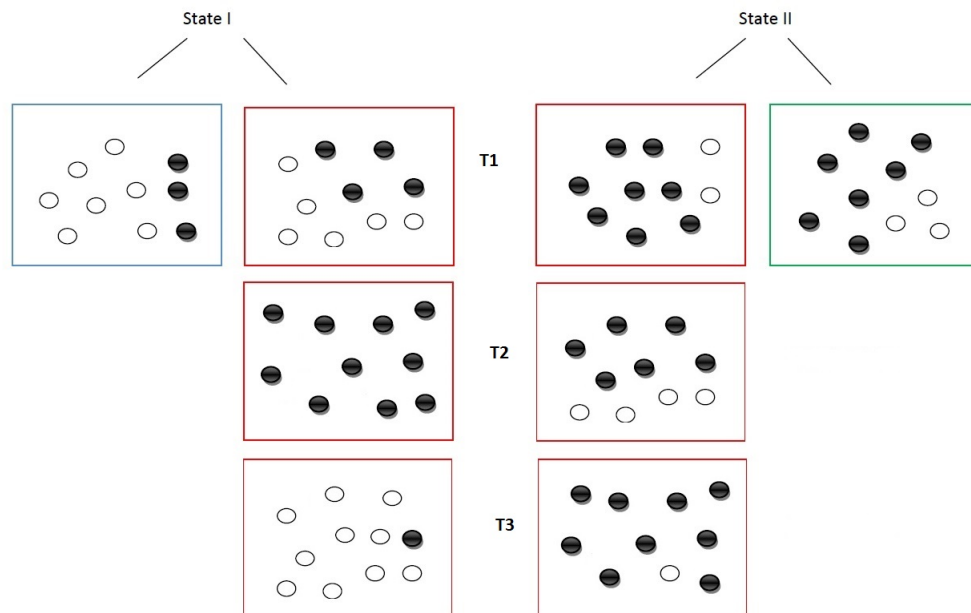
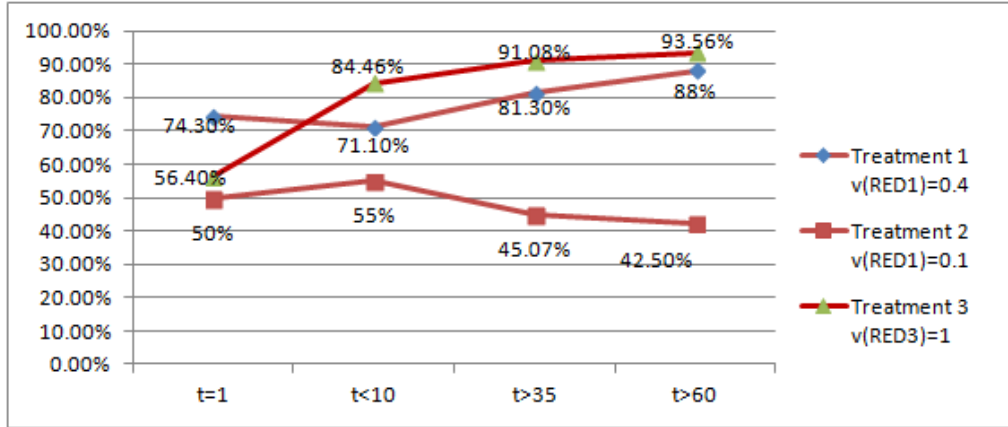
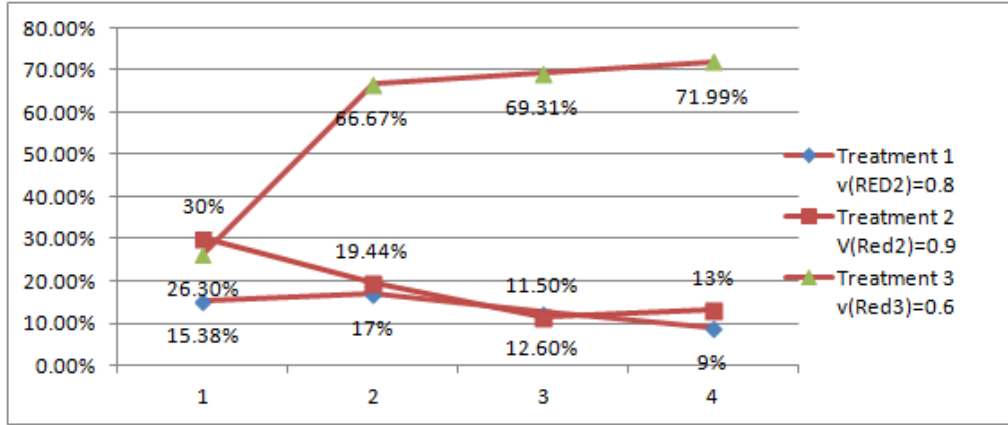


Figure 2 Set up of the different treatment sessions

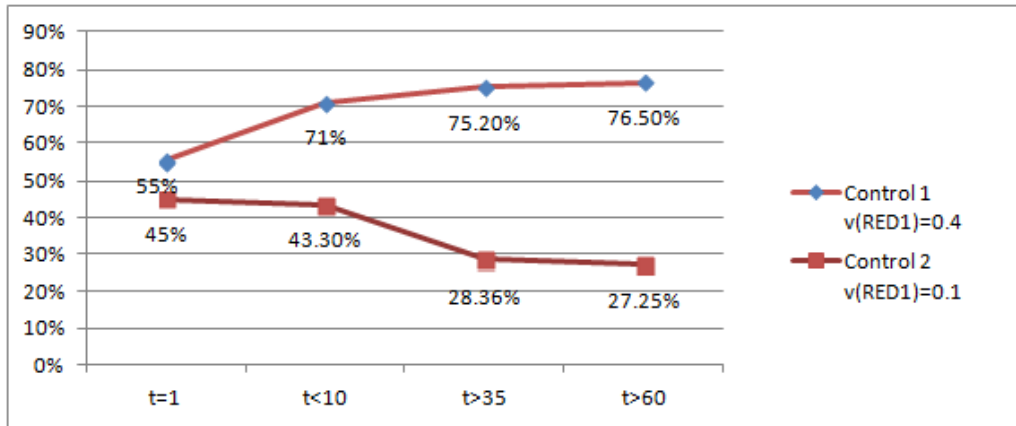


(a) Red1 across treatments

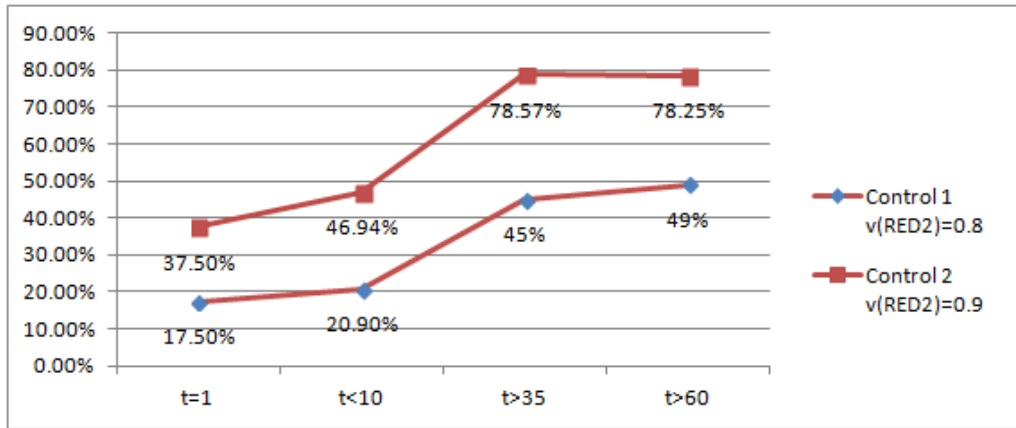


(b) Red2 across treatments

Figure 3 Evolution of choice across treatments with aggregated feedback

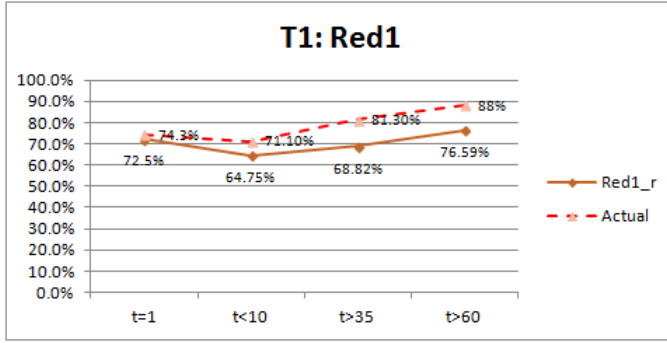


(a) Red1 in control

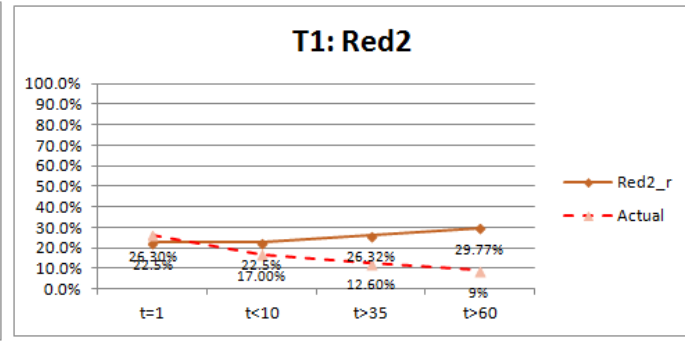


(b) Red2 in control

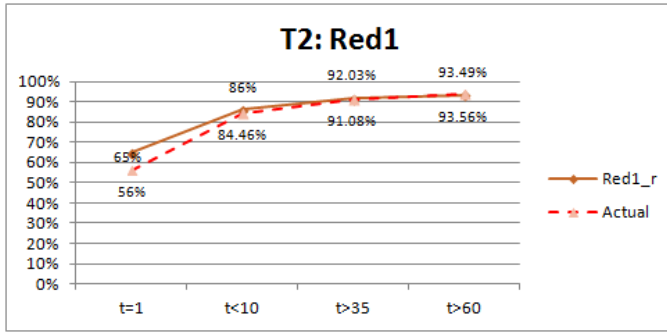
Figure 4 Evolution of choice across treatments with dis-aggregated feedback



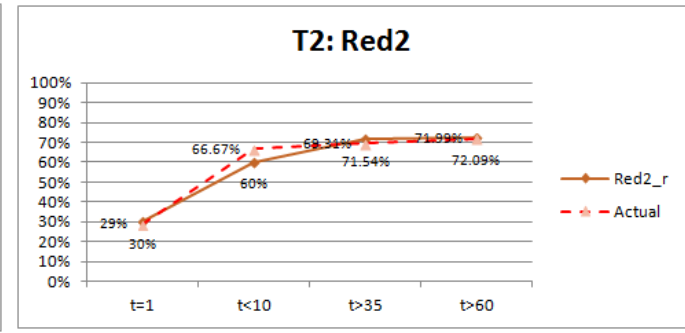
(a) Treatment 1: State 1



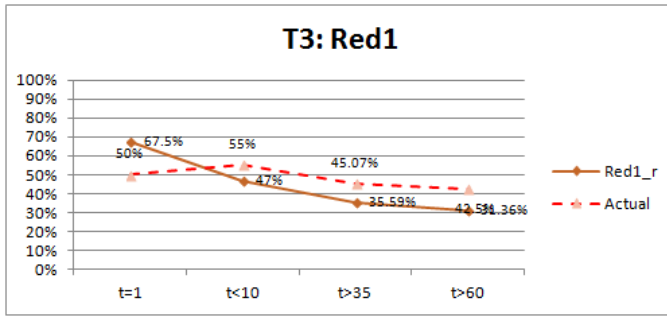
(b) Treatment 1: State 2



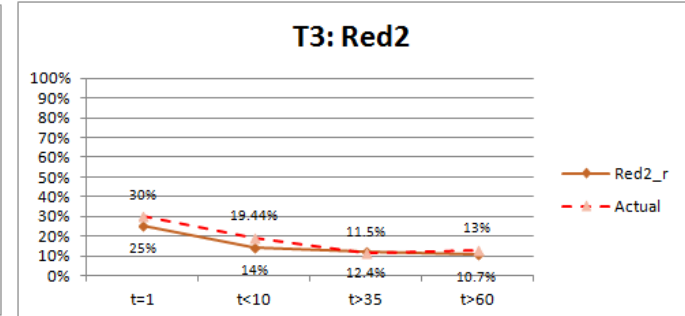
(c) Treatment 2: State 1



(d) Treatment 2: State 2



(e) Treatment 3: State 1



(f) Treatment 3: State 2

Figure 5 Simulated choices using the reinforcement model

Tables

Blue	30 black (B)	70 white (W)
Green	68 black (B)	32 white (W)

Table 1 Information about the relative payoff of urns Blue and Green after 100 random draws is reported at the start of each session.

ρ_U	ρ_F	δ	BR_{ini}	λ	L
0.43	0.599	0.00	0.42	5.24	7626.6
[0.4, 0.49]	[0.55, 0.64]	[0, 0.0009]	[0.38, 0.49]	[5.04, 5.39]	-

Note: Confidence interval at 95% are reported in brackets for the restricted estimators. (See Ketz 2018 for details).

Table 2 Parameters for similarity-based reinforcement learning model

λ	k_{inf}	k_{sup}	δ	L
7.488	3	7	0.003	8816.2
[7.43, 7.52]	-	-	[0.001, 0.005]	-

Table 3 Parameter for Bayesian model with bounds

Valuation model	Bayesian model
1.52×10^4	1.763×10^4

Note: BIC is a model selection criterion partly based on the likelihood function to determine the goodness of fit in the two models accounting for the no. of parameters. Lower the BIC value, better the fit. Valuation model therefore fits our data better.

Table 4 BIC values for the two competing models

ρ_U	ρ_F	δ	BR_{ini}	λ	μ	L
0.45	0.6	0.00	0.45	5.2	0.108	7626.3
[0.40, 0.49]	[0.55, 0.64]	[0, 0.001]	[0.39, 0.5]	[5.02, 5.37]	[-0.26, 0.48]	-

Table 5 Parameters for variant1 of similarity-based reinforcement learning model

ρ_U	ρ_F	δ_1	δ_2	BR_{ini}	λ	L
0.39 [0.34, 0.44]	0.54 [0.49, 0.58]	0.00 [0, 0.0007]	0.04 [0.03, 0.05]	0.46 [0.39, 0.52]	4.99 [4.8, 5.17]	7610.6 -

Table 6 Parameters for variant2 of similarity-based reinforcement learning model

λ	k_{inf}	k_{sup}	δ	Likelihood
8.69 [8.4, 8.9]	3 (-)	7 (-)	0.008 [0.003, 0.013]	8583.8 (-)

Table 7 Parameter for Bayesian model with Blue and Green updating

λ	k_{inf}	k_{sup}	δ	Likelihood
6.56 [6.57, 6.64]	3 (-)	7 (-)	0 [0, 0.008]	8584.4 (-)

Table 8 Parameter for elaborate Bayesian model

7 Appendix

Appendix A

Instruction sheet for the players (In the lab the instructions were in French):

Control Group:

Welcome to the experiment and I thank you for your participation. Please listen to these instructions carefully. If you have any questions kindly raise your hand and it shall be addressed. You receive 5 euros for participating and then your payoff depends on your performance in the experiment.

The Experiment:

The experiment consists of 70 rounds. It is a simple decision task. There are two situations you may face referred to as states 1 and 2. In each state, you have to choose one of two urns. Each urn is composed of ten balls either black or white in color. When you choose an urn, one of the balls in the urn is drawn at random (by the computer) and it is immediately replaced after the computer has noted the color of the ball. If the ball drawn is Black, you can receive extra payment (see below for details) whereas if the ball drawn is White you receive no payment.

The two urns available in state 1 are Blue and Red, respectively. The two urns available in state 2 are Green and Red, respectively. While the compositions of the various urns

remain the same throughout the experiment, note that the compositions of the Red urn in state 1 need not be the same as the composition of the Red urn in state 2. These are two different urns.

As the experiment goes, on your computer screen, you will be informed whether you have to make a choice of urns in state 1 (Blue or Red) or in state 2 (Green or Red). The sequence of choices from states 1 or 2 is decided randomly by the computer. Your task is to choose one urn out of the two in each state.

Note: We drew 100 times a ball (replacing the ball in the urn after each draw) out of the Blue and Green urn. We obtained the following composition

Blue	30 Black	70 White
Green	68 Black	32 White

At the beginning of the experiment:

- Your terminal is randomly assigned a State of the world. If in State 1, you choose between a Red and Blue urn. If in State 2, you choose between a Red and Green urn.
- After you choose the color of the urn that you want to pick, you click on the screen. A ball (the color of which could be either Black or White) will be drawn from that urn by the computer. You will not know the color of the ball drawn. This implies you will not have the information for your choice.
- Once all participants have made their choices, we provide you with some feedback. The total number of black and white balls drawn in previous rounds by all subjects according to the color of the urn (Blue, Red1, Red2, Green).
- Following the feedback, your terminal is randomly assigned a state of the world again. The state may vary from the previous round or remain the same.
- We then repeat the same experiment again until the completion of the 70 rounds.

For determining your payoff, two of the rounds will be randomly chosen at the end of the experiment. If one of your balls in these two rounds is Black, you will get an extra 5 euros. If both of your balls in these two rounds are Black, you will have an extra 10 euros. Otherwise (if both balls are White), you will have no extra return. So if you have no questions let us begin!

Treatment Group:

Welcome to the experiment and I thank you for your participation. Please listen to these instructions carefully. If you have any questions kindly raise your hand and it shall be addressed. You receive 5 euros for participating and then your payoff depends on your performance in the experiment.

The Experiment:

The experiment consists of 70 rounds. It is a simple decision task. There are two situations you may face referred to as states 1 and 2. In each state, you have to choose one of two urns. Each urn is composed of ten balls either black or white in color. When you choose an urn, one of the balls in the urn is drawn at random (by the computer) and it is immediately replaced after the computer has noted the color of the ball. If the ball drawn is Black, you can receive extra payment (see below for details) whereas if the ball drawn is White you receive no payment.

The two urns available in state 1 are Blue and Red, respectively. The two urns available in state 2 are Green and Red, respectively. While the compositions of the various urns remain the same throughout the experiment, note that the compositions of the Red urn in state 1 need not be the same as the composition of the Red urn in state 2. These are two different urns.

As the experiment goes, on your computer screen, you will be informed whether you have to make a choice of urns in state 1 (Blue or Red) or in state 2 (Green or Red). The sequence of choices from states 1 or 2 is decided randomly by the computer. Your task is to choose one urn out of the two in each state.

Note: We drew 100 balls randomly out of the Blue and Green urn which gave us the composition

Blue	30 Black	70 White
Green	68 Black	32 White

At the beginning of the experiment:

- Your terminal is randomly assigned a State of the world. If in State 1, you choose between a Red and Blue urn. If in State 2, you choose between a Red and Green urn.
- After you choose the color of the urn that you want to pick, you click on the screen. A ball (color of which could be either Black or White) will be picked up from that urn. You will not know the color of the ball drawn. This implies you will not have the information of your choice
- Once every participant has made their choice, we provide you with the feedback. The no. of black and white balls drawn from each colored urn (Blue, Red, Green) across states based on only the previous round draw is reported. *Note that the information for Red corresponds to the no. of balls picked in State 1 and 2.*
- Following the feedback, your terminal is randomly assigned a state of the world again. The state may vary from the previous round or remain the same. Note that the composition of the urn is however fixed throughout the experiment.
- We then repeat the same experiment again till we complete 70 rounds.

For determining your payoff, two of the rounds will be randomly chosen at the end of the experiment. If you have picked up B in that particular round, you end up with 5 euros more for each B otherwise no returns. So if you have no questions let us begin!

Appendix B

The learning model we described is parameterized by $(\rho_R, \rho_B, \delta, \lambda, BR_{init})$. The diagrams below show that these parameters are normally distributed via Monte Carlo simulations with 1000 iterations and $n=240$.²⁸

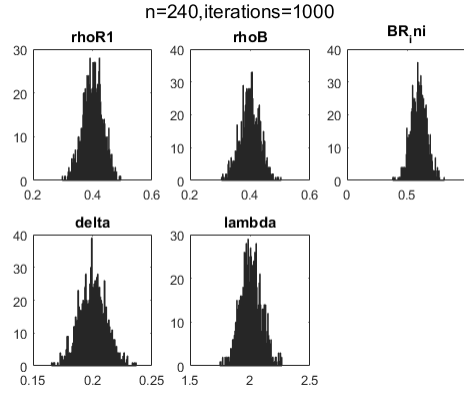


Figure 6 Results for Montecarlo simulations for Model 1.

²⁸This is in line with number of players in our actual experiment.

Appendix C

The figure shows simulated proportions of choices for the reinforcement model over 70 rounds with the estimated parameters for Treatment 1. Instead of using the noise parameter, $\lambda = 5.23$, we use $\lambda = 7$ to introduce less noise. This improves the fit of the simulated data with the actual one.

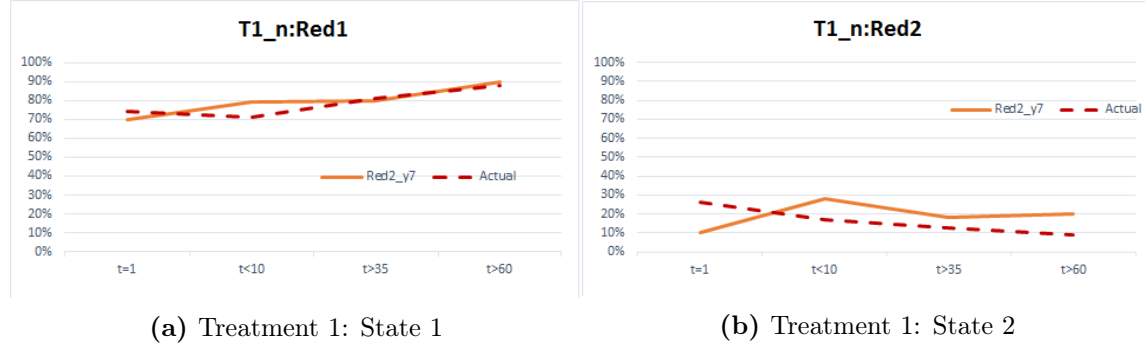


Figure 7 Improving the fit of simulated data in T1 with a lower noise parameter λ

3 Multi-state choices with Aggregate Feedback on Unfamiliar Alternatives: a Followup

Abstract

This paper follows up on a multi-state binary choice experiment in which in each state, one alternative has well understood consequences whereas the other alternative has unknown consequences. Subjects repeatedly receive feedback from past choices about the consequences of unfamiliar alternatives but this feedback is aggregated over states. The follow up experiment allows subject to experiment and generate individual level feedback unlike the group feedback in the original one. Varying the payoffs attached to the various alternatives in various states allows us to test whether unfamiliar alternatives are discounted via three different treatments. Our experimental data suggest similar findings as in the original experiment where players follow similarity-based reinforcement learning. We find a slightly higher proportion of Bayesian learners compared to the original experiment. is no difference in choice in the followup.

Key words: Ambiguity, Bounded Rationality, Experiment, Learning, Coarse feedback, Valuation equilibrium

JEL Classification: D81, D83, C12, C91

1 Introduction: Why a follow up?

Our previous paper studies a multi-state binary choice experiment in which in each state, one alternative has well understood consequences whereas the other alternative has unknown consequences. Subjects repeatedly receive feedback from past choices about the consequences of unfamiliar alternatives but this feedback is aggregated over states and across choices of everyone. We vary the payoffs attached to the various alternatives in various states to test whether unfamiliar alternatives are discounted and whether subjects' use of feedback is better explained by similarity-based reinforcement learning models (in the spirit of the valuation equilibrium, Jehiel and Samet 2007) or by some variant of Bayesian learning model. Our experimental data suggest that there is no discount attached to the unfamiliar alternatives and that similarity-based reinforcement learning models have a better explanatory power than their Bayesian counterparts.

A natural follow up to the question would be to understand how these findings would change in the face of the feedback being individual specific in the spirit of learning by doing. Going back to one of the concrete example mentioned in our previous paper, think of the adoption of a new technology by farmers. A farmer has a lot of information about the performance of the current technology but not so much about the new one. The farmer may collect information about the new technology by asking around other farmers who would have previously adopted it. But due to the heterogeneity of the soil and/or the heterogeneity in the ability of the farmers, what works well/poorly for one farmer need not perform in the same way for another. Thus, the feedback received about the new technology is coarse in the sense that it is aggregated over different situations (states in the decision theoretic terminology) as compared to the information held for the old technology.²⁹ Now consider an alternative scenario where the farmer instead of relying on others experience, himself uses the new technology over different situations. In principal, the farmer should be able to experiment and observe performance of the unfamiliar technology separately in the two states. However, an external factor like weather condition at the time the new technology is adopted might affect the state. There could also be concerns of imperfect recall regarding which state the performance data is coming from. Hence, the information still remains coarse in the sense that it is aggregated over different situations.

We are interested in understanding how decision makers would vary their decisions in multi-state binary decision problems in which decision makers would have precise state-specific information about the performance of one alternative and less precise information about the other alternative. The less precise information takes the form that the decision maker is responsible to experiment and generate aggregate (not state-specific) feedback

²⁹Ryan and Gross (1946) propose an early study of the diffusion of new technology adoption in the farming context. See also Young (2009) for a study focused on the diffusion dimension.

about the performance of that alternative. In course of presentations³⁰, we were often asked if allowing experimentation rather than aggregate group feedback could change the behavior of subjects. Feedback can be generated in many different ways and via this followup, we try to specifically understand if the aggregated valuation equilibrium still holds true when subjects are given the opportunity to be responsible for their own feedback. Similar to the earlier paper, our interest lies in allowing a set of agents to act repeatedly in such environments so as to understand the steady state effects.

We have established that subjects in the face of repeated aggregated information learn with a similarity-based reinforcement learning when the feedback is generated by group choices rather than individual. The feedback on the unknown alternative when allowed to be generated via individual experimentation, subjects continue to learn with similarity-based reinforcement learning. In the follow up, the performance of the unknown alternative still remains aggregated (not state-specific) but every individual is now able to experiment and generate their own set of feedback. In the next section, we detail the experiment in greater details and point out the differences from the previous case. This is followed by comparing results of the old and the follow up design.

2 Experimental design: How is it different?

2.1 Context

The computerized experiments were conducted in the Laboratory at Maison de Sciences Economiques (MSE) in September 2019. Upon arrival at the lab, subjects sat down at a computer terminal to start the experiment. Instructions were handed out and read aloud before the start of each session. The experiment consisted of three main treatments which varied in the payoffs of the Red urns as will be explained below. In addition we had two other treatments referred to as controls in which subjects received state-specific feedback about the Red urns, i.e the feedbacks for Red_1 and Red_2 appeared now in two different columns, for the two payoff specifications of treatments 1 and 2. The purpose of these control treatments was to check whether convergence to optimal choices was observed in such more standard feedback scenarios.

Each treatment involved 76 participants. Overall, 328 subjects drawn from the participant pool at the MSE -who were mostly students- participated in the experiment. Each session had ten repetition of twenty rounds. In all treatments, all repetitions, and all rounds, subjects were randomly assigned to a new state at the start of each round. The subjects knew the state they were assigned to, but did not know the payoff attached to the available

³⁰We would sincerely like to thank our anonymous referee for suggesting an alternative way the feedback could be generated,

actions in each state.³¹ In each state, players were asked to choose between two actions. The feedback structure for the main treatments is explained in the design section. For the control group, the information structure was disaggregated. We use this as a baseline to show that under simpler feedback structure, individuals learn optimally the best available option.

Subjects were paid a show-up fee of 5 €. In addition to this, they were given the opportunity to earn 12 € depending on their choice in the experiment. Specifically, for each subject, six rounds were drawn at random and a subject earned an extra 2 € for each black ball that was drawn from their chosen urn in these six rounds. The average payment was around 11 € per subject, including the show-up fee. All of the sessions lasted around an hour.

2.2 Design

We consider the following experimental setting. There are two states, $s = 1, 2$. In each state, the decision maker has to choose between two urns identified with a color, Blue and Red in state $s = 1$, Green and Red in state $s = 2$ where the Red urns have different payoff implications in states $s = 1$ and 2 . Each urn is composed of ten balls, black or white. When an urn is picked, one ball is drawn at random from this urn (and it is immediately replaced afterwards). If a black ball is drawn this translates into a positive payment. If a white ball is drawn there is no payment. One hundred initial draws are made for the Blue and Green urns with no payoff implication for participants, and all subjects are informed of the corresponding compositions of black and white balls drawn from these urns. Thus, as seen in Table 9, subjects have a precise initial view about the compositions of the Blue and Green urns (these urns correspond to the familiar choices in the motivating example provided above). In the experiment, the Blue urn has 3 black balls out of ten and the Green urn has 7 black balls out of ten.

Concerning the Red urns, there is no initial information. The Red urns correspond to the unfamiliar choices in the above examples. To guide their choices, subjects are provided with feedback about the compositions of the red urns as reflected by the colors of the balls that were previously drawn when a red urn either in state $s = 1$ or 2 was chosen. More precisely, each subject plays 10 repetitions with 20 rounds in each repetition. In each round, subjects make a choice of urn in either state 1 or 2. There are permutations of subjects between rounds so that every subject is in each state $s = 1$ or 2 the same proportion of time. Between repetitions, subjects receive individual feedback about the number of times the Green, Blue and Red urns were picked by them individually in the previous repetition (previous 20 rounds), and for each color of urn, they are informed of the number of black balls that were drawn. A typical feedback screen is shown in Figure 8. In the case of the

³¹For urns *Blue* and *Green*, they had initial information, as explained the design section.

Red urns, this number aggregates the number of black balls drawn from both the Red urns picked in state $s = 1$ and the Red urns picked in state $s = 2$ mimicking the kind of coarse information suggested in the motivating examples. It should be highlighted that subjects were explicitly told that the compositions of the Red urn in state $s = 1$ and in state $s = 2$ need not be the same.

Note that in the original experiment, subjects faced similar situation with the only difference being how the feedback is generated. In the original experiment, unlike the follow up where aggregated feedback is generated individually after 20 rounds of choice, subjects at the end of each round were informed of the number of black balls drawn from each urn by everyone in the session. There were around 20 individuals per session. In the case of the follow up, despite the feedback being aggregated across states, players have an opportunity to experiment and learn more about the composition of the Red urns.

We consider three treatments T1, T2, T3 that differ in the composition of the Red urns as depicted in Figure 9, but note that we maintain the compositions of the Blue and Green urns in all treatments. The initial conditions in these various treatments are thus identical and any difference of behaviors observed in later periods can safely be attributed to the difference in the feedback received by the subjects across the treatments. In treatment 1, the best decision in both states $s = 1$ and 2 would require the choice of the Red urn, but averaging the composition of the Red urns across the two states leads to a less favorable composition than the Green urn. In treatment 2, the best decision would require picking the Red urn in state 1 but not in state 2, but the average composition of the two Red urns dominates that of both the Blue and Green urns. Finally, in treatment 3, the best decision would require picking the Red urn in state 2 but not in state 1.

3 Results

We compare the findings in the follow up experiments to the original ones. In order to review the optimal strategies under aggregated and unaggregated feedback, we present Table 10. We want to understand if the valuation reinforcement model can still explain the behavior of subjects given that the feedback is more informative yet still aggregated. In other words, we want to test if subjects successfully learn the composition of Red1 and Red2 separately due to experimentation despite the feedback being aggregated.

3.1 Preliminary findings

In Figure 10, we report how the choices of urns vary with repetitions and across treatments. Across all the sessions, initially, subjects are more likely to choose the Red urn than the Blue urn in state 1 and they are more likely to choose the Green urn than the Red urn in

state 2. This is, of course, consistent with most theoretical approaches including the ones discussed above given that the Green urn is more rewarding than the Blue urn and the Red urns look (at least initially) alike in states 1 and 2.

The more interesting question concerns the evolution of choices. Roughly, in state 1, we see toward the final rounds, a largely dominant choice of the Red urns in treatments 1 and 2 whereas Red in state 1 is chosen less than half the time in treatment 3. Concerning state 2, we see that in the final rounds, the Red urns are rarely chosen in all treatments. The qualitative differences of the choices in the final rounds among the three treatments and the two states are in line with the prediction of the valuation equilibrium except in treatment 2.

This graph is also interesting to understand information generation under ambiguity. In State 2, where the outside option, Green urn has a higher value than that of the Blue urn, subjects experiment less with the ambiguous option. The inverse is true for State 1. Comparing with previous graphs in Figure 11, we see that the trend is similar across all treatments except in Treatment 2. The final convergence is however starker in the original experiment than in the followups. This could be explained by the fact that individual level feedback generation and experimentation leads to more noise hence the convergence is less starker.

Explaining the difference in Treatment 2:

In the followup, feedback is individual specific and subjects can in principle experiment to learn the exact composition of the Red urns across states separately. This learning state by state should show a strong imbalance in the number of times Red is picked in State1 and State2. In Figure 12 we plot the imbalance between the no. of times Red was chosen in State 1 and 2 in the first repetition where no feedback on the Red urns were provided. The red lines are the imbalance in the original treatments across sessions.

We see that in T2, compared to the original treatment, there is more imbalance where most people either choose Red always in State 1 but not in State 2. The distribution of the difference is skewed to the right unlike the difference in original treatment marked by red lines on 2-4. The imbalance in the other two treatments are closer to the original treatment. In T1, the original treatment had a difference between 5-7 which is closer to the followup. In T3, the difference is fairly spread with a peak at 0. Note that 0 could mean none of the Reds are chosen.

3.2 Statistical estimations

Similarity-based reinforcement learning

The estimations of the parameters of the similarity-based reinforcement learning model together with the corresponding log likelihood³² are given in the Table 11. The model is

³²Likelihood throughout the paper refers to the negative of the log likelihood. Thus, the lower the

explained in great details in the previous paper. Concerning the likelihood, by way of comparison, a complete random choice model where in every state, subjects would randomize 50:50 between the two choices would result in a negative log likelihood of $L=1580.4$, which is much higher than 1340.3 obtained in Table 11. More generally, the similarity-based reinforcement learning model explains data much better than any model in which behavior would not be responsive to feedback. We now discuss the most salient differences in the estimations.

The finding that $\rho_F > \rho_U$ seems natural as mentioned above, to the extent that for the familiar urns, the feedback should affect less how the valuations are updated. The values in followup are similar to the original estimation. ρ_U seems to be slightly lower which seems natural given the fact that the followup allows greater experimentation hence subjects tend to rely more on current period feedback rather than previous belief. The finding that BR_{init} is slightly below 0.5 is similar to the original experiment and can be interpreted along the following lines. In the absence of any information, an initial value of 0.5 would be the one dictated by the principle of insufficient reason, but the uncertainty attached to the unfamiliar urns may lead to some extra discount in agreement with some form of ambiguity aversion as reported in Ellsberg.³³

The most interesting observation concerns δ which is estimated to be 0.08 in contrast to $\delta = 0$ in the original experiment. This difference could be due to the fact that Red is chosen less frequently in State 2 therefore leading to a different trend in T2. Even though the feedback for the Red urns is ambiguous (because it is aggregated over the two states), the valuations for Red are not much discounted as if subjects were almost ambiguous neutral from that perspective. Thus, our estimation suggests that even with more informative feedback and individual experimentation, very little ambiguity discount seems to be applied to the valuation of Red.

Given Figure 12, we next estimate a model that takes the imbalance in the choice of Red across states into account. We consider the following variant where we allow the speed of adjustment of the valuation of the Red urns to differ across the two states. A large imbalance in the number of green urns as opposed to blue urns in the feedback is indicative that the feedback concerned Red urns in state 1 more than in state 2. Specifically, we now introduce a new parameter μ and specify the weight on the previous valuation to satisfy

likelihood, the better the model (See textbook Train 2003 for further details). Standard errors are reported in brackets.

³³The difference $0.5 - 0.44 = 0.06$ can be interpreted as measuring the ambiguity aversion of choosing an unfamiliar urn when no feedback is available.

$$\rho_{U1} = \rho_U \times [1 - \mu \cdot (\frac{NB}{NG + NB} - 0.5)]$$

$$\rho_{U2} = \rho_U \times [1 - \mu \cdot (\frac{NG}{NG + NB} - 0.5)]$$

where NB and NG are the respective numbers of Blue and Green urns appearing in the feedback. Given that one could infer from NB being smaller than NG that there were more draws from Red1 than from Red2, one would expect more weight to be assigned to the feedback in the reassessment of Red in state 1 i.e $\mu < 0$. The estimations of this extended model are reported in Table 12. As expected, we find that $\mu < 0$. There is a significant gain in likelihood and $\delta = 0$ cannot be rejected. Thus, this extended model explains the data better than our previously proposed version.

Generalized Bayesian learning model:

Subjects could form some initial prior belief regarding the compositions of Red1 and Red2, say about the chance that there are k_i black balls out of 10 in Red_i , and update these beliefs after seeing the feedback using Bayes' law. The estimated parameters for the generalized Bayesian learning model are given in Table 13.

The value of $\delta = 0.0116$ implies that with the Bayesian model, the subjects show some mild form of ambiguity aversion. However we cannot statistically reject the hypothesis that $\delta = 0$, which implies that with the Bayesian model, there is some ambiguity discount. For the support of initial prior, we found that $k_{inf} = 3$ and $k_{sup} = 7$.³⁴

Comparing the two models:

Maybe the most important question is with the new feedback structure which is more informative than the original design, does Bayesian learning model explains the experimental data better than the reinforcement one. We consider two methods of comparisons, all establishing that the reinforcement learning model outperforms the Bayesian learning model. First, looking at the likelihood of the two models, we see that the Bayesian learning model explains less well the data than the similarity-based reinforcement learning model. Second, we perform a Vuong test³⁵ to compare the performance of the two models statistically. Under the null hypothesis H_0 , that both models perform equally well, we conclude that the null can be rejected in favor of the reinforcement model. Specifically,

$$H_0 = E(L(\theta_R; x_d)) = E(L(\theta_B; x_d))$$

³⁴The value of the bounds correspond to $v_{Blue}=0.3$ and $v_{Green}=0.7$ respectively and so one may speculate that maybe the compositions of the familiar urns serve as anchoring the support of the priors. Observe that because best-responses are noisy, the derived support does not imply that the Red urn is always picked in state 1 and never picked in state 2.

³⁵See Merkel et al. 2016 for more details.

$$H_a = E(L(\theta_R; x_d)) \neq E(L(\theta_B; x_d))$$

where x_d is the collection of observed individual data points, θ_R is the set of parameters estimated via reinforcement learning, θ_B is the set of parameters estimated via Bayesian learning, $L(\theta_R; x_d)$ is the log likelihood under reinforcement model and $L(\theta_B; x_d)$ is the log likelihood under Bayesian learning model for each data point d . The Vuong statistics is 5.87 which is positive and hence we can conclude that the reinforcement model explains the data better than the Bayesian one.

3.3 Individual level heterogeneity

While we have established that the similarity-based reinforcement learning model explains the data better than its Bayesian counterpart, it is of interest to see how the fit as generated by such a model with estimations as reported in Section 3.2 vary across individuals. The pool of 228 subjects had very similar backgrounds across treatments. We fit the parameters obtained from the two learning models to each individual across 10 repetitions. That is, for each individual, we ask for each of the two learning models, the generalized reinforcement learning model with parameters as in Table 11 or the generalized Bayesian model with parameters as in Table 13 allows to better explain the observed choices of the individual over the 200 rounds given the feedback he has been exposed to. And for the model comparison, we use the Vuong test as explained above. Overall, we find that 64.03% of players can be categorized as reinforcement learners and the rest as Bayesian learners. Comparing across treatments, we find that T3 has the highest proportion of Reinforcement learners (72.3%). In T2, the proportion of similarity-based reinforcement learners is 53.9% followed by 65.7% in T3.




We have also compared the overall payoff as measured by the sum of black balls obtained over the 200 rounds across the population of subjects categorized as reinforcement learners or Bayesian learners. For reinforcement learners, the average was 105.3 black balls whereas, for Bayesian learners, an average of 112.7 black balls was observed. This difference is significant (p-value 0.019) with Bayesian learners performing better.

Figures

Période n°9 : Etat 2

In this period, you will make a choice in State 2.
 A cette période, vous faites un choix dans l'état 2.

Balls drawn from the urn of each colour in the previous period:
 Balles tirées dans l'urne de chaque couleur lors de la période précédente :

Urn	 (Blue)	 (Red)	 (Green)
Number of balls of each colour Nombre de balles de chaque couleur	2	7	5
is the ball with positive payoff (est une balle gagnante)	1	3	2

Which urn do you want to choose?
 Quelle urne choisissez-vous ?



☒ 
☐ 

Figure 8 Feedback structure for treatment sessions

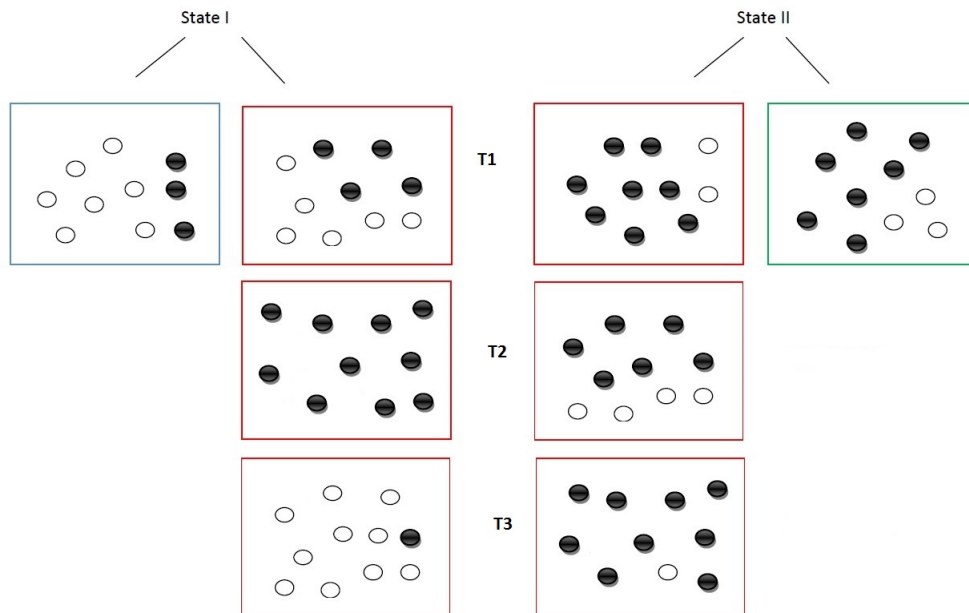
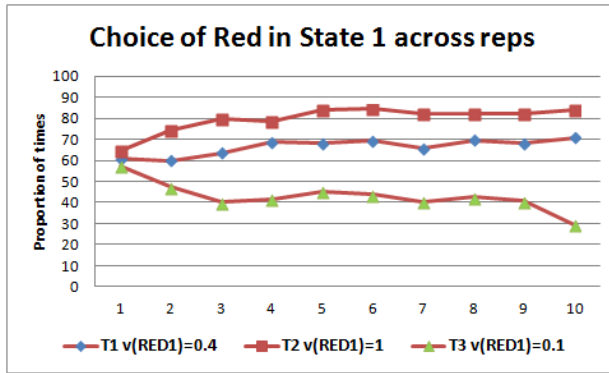
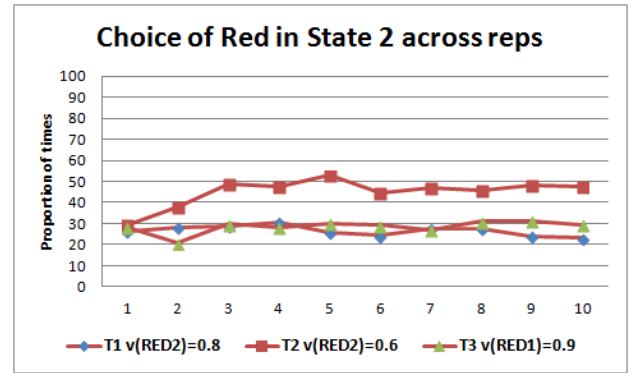


Figure 9 Set up of the different treatment sessions

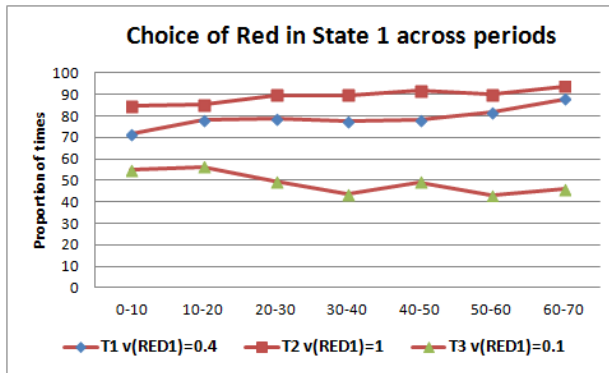


(a) Red1 across treatments

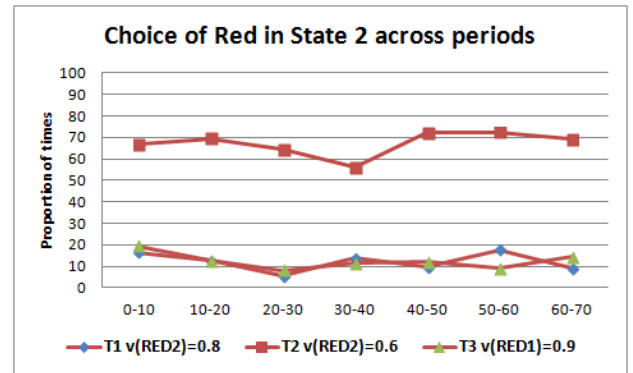


(b) Red2 across treatments

Figure 10 Evolution of choice across treatments with aggregated feedback: followup design



(a) Red1 across treatments



(b) Red2 across treatments

Figure 11 Evolution of choice across treatments with aggregated feedback: original design

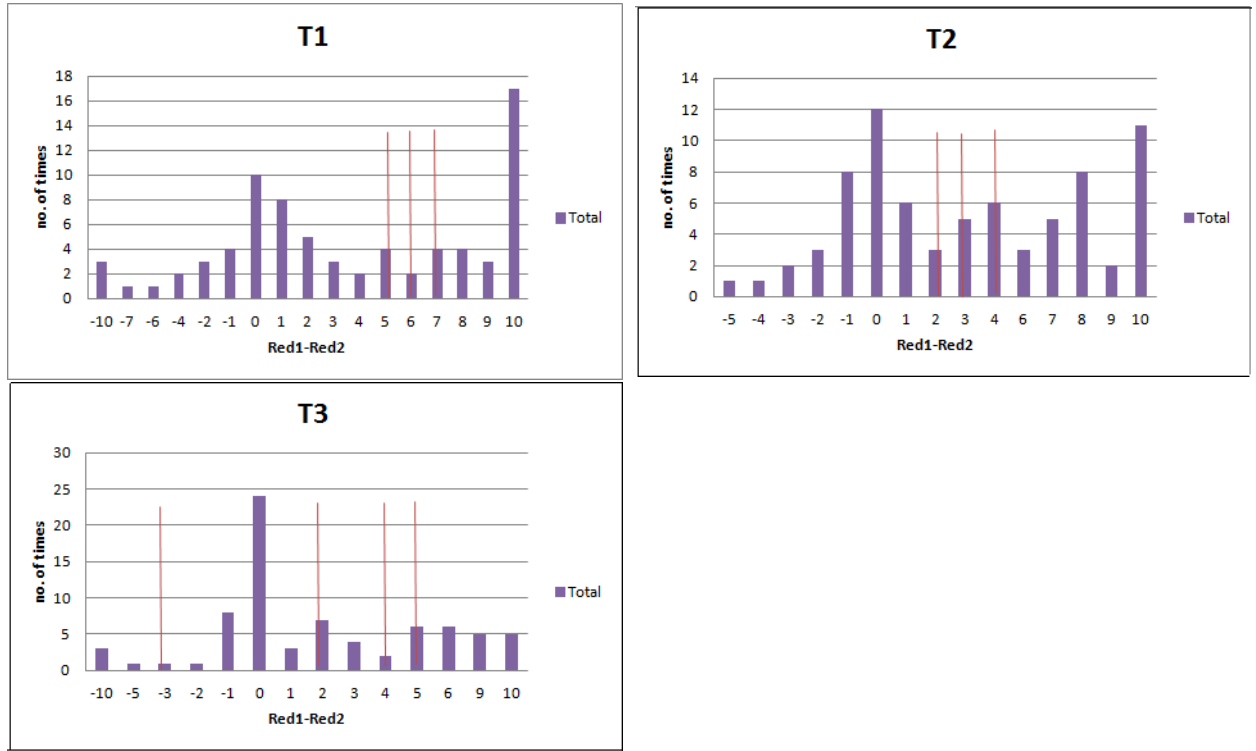


Figure 12 Imbalance in the choice of Red across states

Tables

Blue	30 black (B)	70 white (W)
Green	68 black (B)	32 white (W)

Table 9 Information about the relative payoff of urns Blue and Green after 100 random draws is reported at the start of each session.

	T1	T2	T3
Nash	(R, R)	(R, G)	(B, R)
Valuation	(R, G)	(R, R)	(B, G)

Table 10 Optimal choice under the two solution concept.

ρ_U	ρ_F	δ	BR_{ini}	λ	L
0.204	0.64	0.08	0.46	3.032	1340.3
[0.17, 0.23]	[0.54, 0.730]	[0.07, 0.09]	[0.41, 0.51]	[2.82, 3.25]	-

Note: Confidence interval at 95% are reported in brackets for the restricted estimators. (See Ketz 2018 for details).

Table 11 Parameters for similarity-based reinforcement learning model

ρ_U	ρ_F	δ	BR_{ini}	λ	μ	L
0.45	0.73	0.002	0.28	3.55	-2.53	1278.9
[0.33, 0.56]	[0.63, 0.824]	[0, 0.051]	[0.148, 0.41]	[3.18, 3.93]	[-2.82, -2.19]	-

Note: Confidence interval at 95% are reported in brackets for the restricted estimators. (See Ketz 2018 for details).

Table 12 Parameters for variant1 of similarity-based reinforcement learning model

λ	k_{inf}	k_{sup}	δ	L
3.08	3	7	0.0116	1429.3
[3.04, 3.12]	-	-	[0, 0.01]	-

Table 13 Parameter for Bayesian model with bounds

4 Appendix

Appendix A: Instructions

Instruction sheet for the players (In the lab the instructions were in French):

Control Group:

Welcome to the experiment and I thank you for your participation. Please listen to these instructions carefully. If you have any questions kindly raise your hand and it shall be addressed. You receive 5 euros for participating and then your payoff depends on your performance in the experiment.

The Experiment:

The experiment consists of 10 repetitions. In each round, there is a simple decision task. Every repetition has 20 rounds. There are two situations you may face referred to as states 1 and 2. In each state, you have to choose one of two urns. Each urn is composed of ten balls either black or white in color. When you choose an urn, one of the balls in the urn is drawn at random (by the computer) and it is immediately replaced after the computer has

noted the color of the ball. If the ball drawn is Black, you can receive extra payment (see below for details) whereas if the ball drawn is White you receive no payment.

The two urns available in state 1 are Blue and Red, respectively. The two urns available in state 2 are Green and Red, respectively. While the compositions of the various urns remain the same throughout the experiment, note that the compositions of the Red urn in state 1 need not be the same as the composition of the Red urn in state 2. These are two different urns.

As the experiment goes, on your computer screen, you will be informed whether you have to make a choice of urns in state 1 (Blue or Red) or in state 2 (Green or Red). The sequence of choices from states 1 or 2 is decided randomly by the computer. Your task is to choose one urn out of the two in each state.

Note: We drew 100 times a ball (replacing the ball in the urn after each draw) out of the Blue and Green urn. We obtained the following composition

Blue	30 Black	70 White
Green	68 Black	32 White

At the beginning of the experiment:

- Your terminal is randomly assigned a State of the world. If in State 1, you choose between a Red and Blue urn. If in State 2, you choose between a Red and Green urn.
- After you choose the color of the urn that you want to pick, you click on the screen. A ball (color of which could be either Black or White) will be picked up from that urn. You will not immediately know the color of the ball drawn.
- Once you have made your choice, your terminal is randomly assigned a state of the world again. The state may vary from the previous round or remain the same.
- After having played 20 rounds, a supplementary information appears on your screen: For each colored urn (Blue, Red1, Red2, Green), you are informed the no. of black and white balls drawn by you in the 20 previous rounds corresponding to the color of the urn.
- We then repeat the same experiment 10 times.

For determining your payoff, of the rounds will be randomly chosen at the end of the experiment. If you have picked up a black ball in that particular round, you end up with 2 euros more for each black ball otherwise no returns. So if you have no questions let us begin!

Treatment Group:

Welcome to the experiment and I thank you for your participation. Please listen to these instructions carefully. If you have any questions kindly raise your hand and it shall

be addressed. You receive 5 euros for participating and then your payoff depends on your performance in the experiment.

The Experiment:

The experiment consists of 10 repetitions. In each round, there is a simple decision task. Every repetition has 20 rounds. There are two situations you may face referred to as states 1 and 2. In each state, you have to choose one of two urns. Each urn is composed of ten balls either black or white in color. When you choose an urn, one of the balls in the urn is drawn at random (by the computer) and it is immediately replaced after the computer has noted the color of the ball. If the ball drawn is Black, you can receive extra payment (see below for details) whereas if the ball drawn is White you receive no payment.

The two urns available in state 1 are Blue and Red, respectively. The two urns available in state 2 are Green and Red, respectively. While the compositions of the various urns remain the same throughout the experiment, note that the compositions of the Red urn in state 1 need not be the same as the composition of the Red urn in state 2. These are two different urns.

As the experiment goes, on your computer screen, you will be informed whether you have to make a choice of urns in state 1 (Blue or Red) or in state 2 (Green or Red). The sequence of choices from states 1 or 2 is decided randomly by the computer. Your task is to choose one urn out of the two in each state.

Note: We drew 100 balls randomly out of the Blue and Green urn which gave us the composition

Blue	30 Black	70 White
Green	68 Black	32 White

At the beginning of the experiment:

- Your terminal is randomly assigned a State of the world. If in State 1, you choose between a Red and Blue urn. If in State 2, you choose between a Red and Green urn.
- After you choose the color of the urn that you want to pick, you click on the screen. A ball (color of which could be either Black or White) will be picked up from that urn. You will not immediately know the color of the ball drawn.
- Once you have made your choice, your terminal is randomly assigned a state of the world again. The state may vary from the previous round or remain the same.
- After having played 20 rounds, a supplementary information appears on your screen: For each colored urn (Blue, Red, Green), you are informed the no. of black and white balls drawn by you in the 20 previous rounds corresponding to the color of the urn. Note that information for Red represents a total of balls from both State 1 and State 2.
- We then repeat the same experiment 10 times.

For determining your payoff, six of the rounds will be randomly chosen at the end of the experiment. If you have picked up a black ball in that particular round, you end up with 2 euros more for each black ball otherwise no returns. So if you have no questions let us begin!

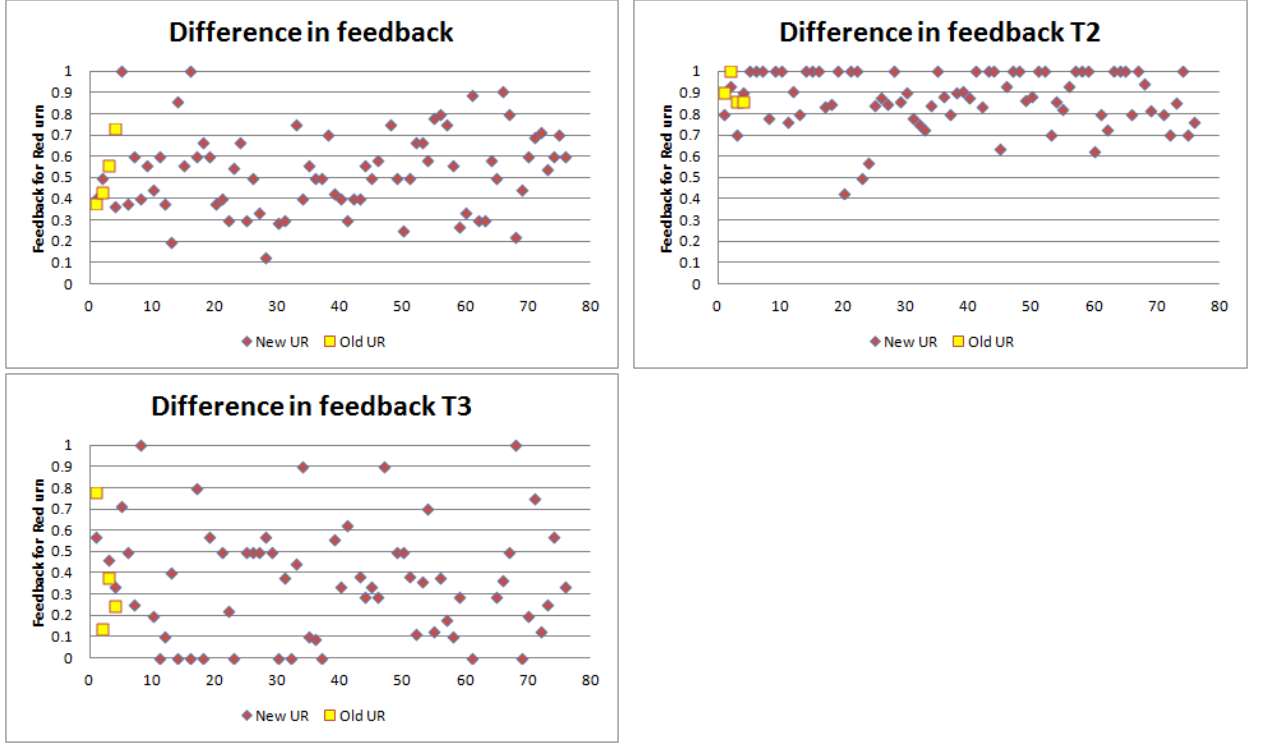
Appendix B: Explaining the difference in Treatment 2:

In order to understand the difference in trend for T2, we try to look at a different measure, the spread of feedback for the Red urns across treatments in the first round. Note that these choices were made initially without feedback. In Fig 6, the yellow squares represent feedback in the original experiment across sessions and the red dots represent feedback in the followup for each person. The graph plots UR_1 where

$$UC_t = \frac{\#(\text{Black balls drawn in urns with color } C \text{ at } t-1)}{\#(\text{an urn with color } C \text{ picked at } t-1)}. \quad (3)$$

UC_t represents the strength of urn(s) with color C as reflected by the feedback received at t about urns with such a color. Note the normalization by $\#(\text{an urn with color } C \text{ picked at } t-1)$ so that UC_t is comparable to a single payoff attached to choosing an urn with color C .

For T1, the value of Red is fairly concentrated between 0.3 and 0.7. For T2, it is concentrated between 0.7 and 1. Our hunch is that most of the information in this treatment is coming from Red in State 1 hence imbalance in learning across the two states. In T3, it is concentrated between 0.1 and 0.4.



(a) Red2 across treatments

Figure 13 Evolution of choice across treatments with aggregated feedback: followup design

Appendix C: Control

In Figure 14, with state-specific feedback for the Red urns, we see a clear trend toward the optimal choices even if some noise would be needed to explain why only 30% of choices correspond to Red in state 2 in Control 1. In contrast to the feedback structure in the treatment group, we see that disaggregating feedback on the Red urns across states, players learn the optimal choice. The fine feedback helps the agent attach a valuation $v(\text{Red1})$ and $v(\text{Red2})$ separately for the Red urns in the two states instead of a joint valuation $v(\text{Red})$. Due to this finer feedback structure, the simple heuristic of reinforcement learning leads to an optimal choice, unlike in the aggregated treatments in which an analogous reinforcement learning heuristic leads to valuation equilibrium.

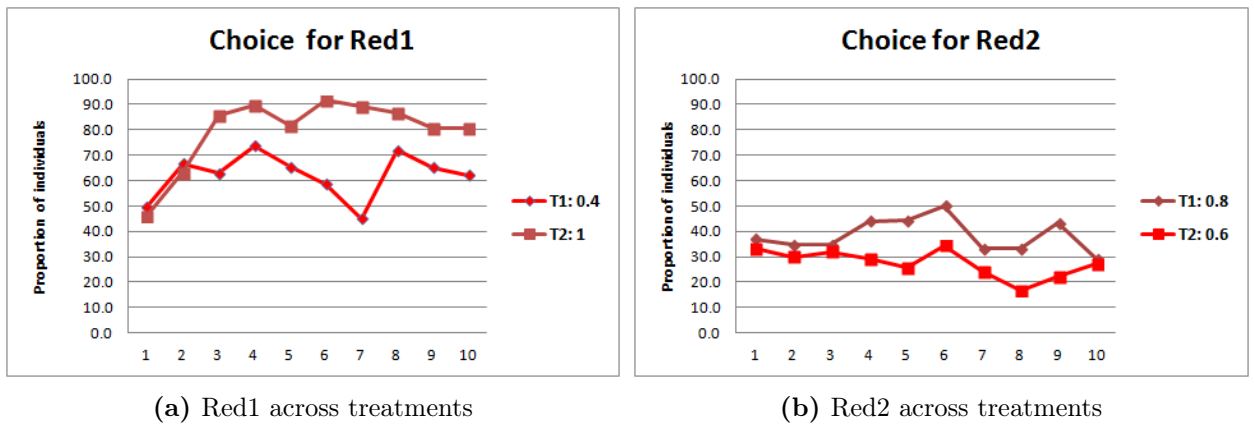


Figure 14 Evolution of choice across treatments with dis-aggregated feedback: Control

4 Endogenous Institutions: a Network Experiment in Nepal

Abstract

In developing countries where formal institutions are often weak, the community is responsible to enforce local agreements. In such settings, peer monitoring represents a natural mechanism for the enforcement of agreements. This paper studies the demand for monitoring and its effectiveness in sustaining cooperation across social groups. We map social networks of 19 villages in rural Nepal and conduct an experiment to explore the role of the endogenous choice of monitors. Individuals play in groups of three, both with their close friends and with people socially distant in their network. They receive the opportunity to anonymously choose their preferred "institution". We combine a theoretical model and a unique lab-in-the-field experiment to show that closely knit groups are significantly more likely not to choose a monitor, while sparse groups tend to prefer a monitor who is highly central in the network. The endogenous selection of monitoring ensures higher cooperation compared to an exogenous assignment, but only in sparse groups. Further, we observe that the outcome of the vote acts as a signal of intra-group trust.

Key words: Network, Peer Monitoring, Experiment, Public good game, Sparseness

JEL Classification: C93, L14, P48

1 Introduction

The engagement of local communities is a standard mechanism for channeling development programs. The benefit of this approach lies in increased participation of communities and in deeper adaptation to local needs. Yet, the efficiency of such initiatives crucially depend on the institutional environment and on the level of cooperation. Development practitioners recognize the increased risk of free riding when institutions are imposed without involving the community (Mansuri and Rao, 2004; Narayan, 1995). In this context, it is important to understand which informal institutions sustain and promote cooperation. Peer monitoring represents a natural mechanism for the enforcement of social norms in such a setting. Previous research has focused on the effectiveness of peer monitors but the demand for such a mechanism is understudied. Our paper fills this gap in two main directions. First, it studies the demand for peer monitors across social groups. Second, it explores how endogenous monitors affect cooperation.³⁶ We focus on the role of social networks in the choice of these monitors and its impact on cooperation.

In developing countries, given the weakness of formal institutions, social networks represent a relevant tool to regulate transactions between individuals. Communities mostly rely on interactions with socially proximate peers and third party institutions are often used to enforce contracts among socially distant individuals. Ostrom, 1991 suggests how mechanisms based on social ties have sustained cooperation³⁷ rather than mere punishment.³⁸ We conduct a lab in the field experiment in 19 villages in rural Nepal to understand third party monitoring and cooperation. Villagers play a cooperation game and can choose to elect a monitor through majority voting. This monitor can impose higher cooperation through reputational concerns. Concerns of social image are important drivers of cooperative behavior. Third party monitoring can substitute for social density by emphasizing on these very concerns (Andreoni et al., 2006; Greif, 1989; Bowles, 2008). Further, the impact of the monitor may depend on his position in the social network³⁹ and on the political process whereby it is assigned to groups. We explore the effects of monitoring induced by reputation concerns rather than by material punishment. Two main pieces of evidence emerge from previous research. First, dense groups are able to sustain more cooperation than socially distant groups. Second, monitors have the power to relax the inefficiencies

³⁶See Olken, 2007; Bjorkman et al., 2009 and Gelade, 2018. Unlike direct reporting in these papers, our experiment relies on reputational concerns affected via gossip by the chosen monitor.

³⁷In rural Nepal, community-based organizations (Forest User Associations, Water User Associations and Cooperatives) presents a good example of the success of such reputational mechanisms.

³⁸There is a huge literature that shows the effect of punishment on public good games. For more details see Fehr and Gächter, 2000; Charness et al., 2008; Beersma and Kleef, 2011; Wu et al., 2015, 2016; Sommerfeld et al., 2007; Glockner et al., 2007; Galbiati and Vertova, 2008; Fiedler Harvey, 2016; Fonseca, 2018 and Fehr and Sutter, 2019.

³⁹Central individuals in the network are shown to be particularly effective in monitoring due to their higher ability of spreading information in the form of gossip (Ballester et al., 2006 and Banerjee et al., 2013)

arising from contractual incompleteness in the context of socially distant groups (Breza and Chandrashekhar, 2018; Breza et al., 2016; Chandrashekhar et al., 2018). These studies focus on the impact of exogenously assigned monitors on cooperation in groups.

The aim of this paper is to bring the literature forward by allowing individuals to endogenously elect their preferred monitor and by studying the induced cooperative behavior. We estimate the demand for monitoring, relate it to the network structure and study its impact on cooperation by tackling three questions. First, do individuals change their demand for monitoring as a function of the social composition of the group they interact with? Second, do monitors who are endogenously chosen spur cooperative behavior compared to those assigned exogenously? Third, is the election of monitors perceived as a signal of intra-group trust? To answer these questions, we conduct a lab in the field experiment in rural Nepal and build a theoretical model supporting our experimental findings.

First, we ask whether groups with different social proximity elect different third party institutions. In line with the literature, we offer three monitoring options, according to a measure of social prominence. We present strong evidence that socially distant groups are more likely than closely-knit groups to elect a high central monitor. We find that individuals are 40% less likely to elect a monitor with their close peers compared to when they are in groups with socially distant members. This supports the idea that contractual incompleteness can be mitigated by social density, but socially distant individuals need third party monitors to enforce social norms and increase efficiency.⁴⁰

Secondly, we investigate whether the political process by which the institution is chosen matters for cooperative behavior. Interestingly, we find that a monitor that is democratically elected has strong positive effects on cooperation compared to a randomly assigned monitor. Previous experimental evidence in economics (Sutter et al., 2010; Tyran and Feld, 2006; Dal Bo et al., 2010) and sociology (Grossman and Baldassarri, 2012) show that cooperation is higher when players are given the opportunity to choose the institution rather than having an externally imposed one. We dig deeper in this dimension, and we offer evidence that the positive impact of endogenous institutions is limited to socially distant groups. More precisely, the magnitude of the increase in contribution ranges from 8.6 % when a socially prominent monitor is chosen to 21.9 % when no monitor is chosen. In socially close groups, the point estimate is negative and not significant. The possible mechanism underlying this “democracy premium” can be explained by an increased sense of agency and control, increased sense of authority (Greif, 2006) and stronger worthiness of authority (Zelditch, 2001).

⁴⁰In line with this reasoning Glaeser et al., 2000 establish that groups with shorter social distance. In the same spirit, experiments in the lab (Hoffman et al., 1996; Leider et al., 2009; Goeree et al., 2010) and in the field (Etang et al., 2011) show that cooperation increases with decreasing social distance.

Third, we present the first evidence of heterogeneous impact of asymmetric information in different social groups. The theoretical literature of principal-agent models (Herold, 2010) shows that the proposal of a complete contract can signal distrust. The principal may thus prefer to leave the contract incomplete rather than to signal distrust by proposing a complete contract. On this basis, we explore if group members perceive the election outcome as a signal of intra-group trust. Individuals in closely-knit groups are particularly affected by the outcome of the vote. We find that the group interprets the outcome of the vote as a strong signal of trust when no monitor is chosen. This entails a significant increase in contribution by 19.7% in sparse groups. On the contrary, dense groups perceive the election of a high central monitor as a signal of mistrust leading to a decrease in contribution by 13.3%. This finding sheds light on how the effects of signaling are heterogeneous when we take into account the social structure of the agents.

To answer these questions, we first conduct an intensive network survey a la Banerjee et al., 2015. We ask questions about advice, trust, friendship and financial relationships such as "who do you spend your free time with", "in case of an emergency, who would you rely on" and "who do you borrow money from". Based on data from the survey, we build an undirected social network where a connection between two people is established if any one of them names the other. Next, we identify the person in the network with the highest and lowest eigenvector centrality. In other words, we identify the most influential and least influential individuals in the village. The rest of the village is divided into groups of three with varying network distance. Groups of three are an optimal choice to study questions of mutual trust inasmuch it is possible to maximize the behavioral contrast between groups which can sustain high levels of cooperation and those which cannot (Jackson et al., 2012). Players interact both in a closely knit group and in a group with socially distant individuals. The closely knit groups are often homogeneous in terms of caste.

We refer to the closely knit group as "dense" and to the group with acquaintances as "sparse". In formal network terms "dense" implies each individual is at most at distance 2 (average path length < 1.6) and "sparse" implies each individual is at least at distance 4 (average path length > 4). We allow individuals to either vote for no monitor, or to choose one from two monitor candidates belonging to their village: a high central monitor (very prominent individual) or a low central monitor (less prominent). Players play a contribution game both with a monitor chosen by the group (endogenous treatment) and an externally assigned monitor (exogenous treatment). The monitor does not materially punish but only observes the contributions of each player, which would otherwise be private information. Players have an initial endowment and need to decide between how much to contribute to a common pot and how much to keep for themselves. Total contribution in the common pot is augmented by 50 percent and divided equally among the three players irrespective of initial contribution. Each individual therefore plays both in a dense and in a sparse group, and in each group both with exogenous and endogenous monitoring institution. The order

of all treatments is randomized.

We vary both the social composition of groups and how monitors are assigned to study the demand for monitoring and cooperation. The rest of this paper is organized as follows: Section 2 presents the experimental protocol and the data collection process. Section 3 describes the experiment’s results and the econometric specifications we use. In Section 4, we build a theoretical model supporting our empirical results. We discuss the results and conclude in Section 5.

2 Experiment

2.1 Overview: Networks and Data

We start by mapping the social network of villages, with a special focus on relations of trust. Given the location of these villages ⁴¹, mutual trust fundamentally shapes social interactions and the contribution to local public goods. As a first step, we assigned a unique identification code to each woman in the census. We started interviewing very few individuals, who would give us names of their closest friends and we administered the network questionnaire to those women who were nominated in the first round. This process was repeated iteratively until either all women were covered or no new individual was nominated - the elicited network is "closed". This technique has the advantage to be faster than the standard network elicitation method and simplifies considerably the issue of homonyms. Each woman was asked at least three connections for each question. The questionnaire consisted of a set of questions designed to elicit social networks, inspired by Banerjee et al., 2015. These questions are meant to elicit ties of friendship and trust and span along various dimensions of social interactions. A link between two individuals i and j is established when either i nominates j or vice versa in any of the questions. We then aggregate and collapse the networks obtained from different questions into one, undirected network. Once a network is fully mapped, it is possible to visualize it and extract important statistics that are central in our experimental design. Figure 15 in the Appendix is a snapshot of the network of a village where we conducted the experiment.

The network we obtain is thus a good representation of the social structure of the community and it is an essential variable of our study. More precisely, we use the network to create groups of contrasting social density for every participant and interact it with variations along two dimensions: monitoring centrality and the political process by which monitors are assigned, either democratically elected or exogenously given. We focus on networks of only women due to the high emigration rate of men either to Kathmandu or abroad, as shown by our pilot experiment conducted in the spring 2018. In the districts

⁴¹The villages are situated at 1200m above sea level in the mid hills of Nepal. They are a four hours drive away from Kathmandu

we worked in, social networks are often gender specific and women play a preponderant role: they are responsible for households' finances, for agricultural production and for their children.

We look at how social density influences the demand for monitors and how it ultimately affects individual contribution to public goods. Groups are formed in order to maximize the number of participants who play in both dense groups, i.e. groups of average path length less than 1.6,⁴² and in sparse groups, i.e. with average path length higher than 4. In other words, being in a dense group implies that the members of the group are at no more than 2 steps away from each other whereas in the sparse group they are at least 4 steps away. The cutoffs defining dense and sparse have been carefully chosen in order to amplify the respective contrast in trust and reputation while maximizing the number of observations. The starker is the difference between dense and sparse groups, the more different will be the behavioral response in the different treatments. Figure 21 in Appendix B shows the distribution of average path length of all groups we formed. We over sampled dense groups to make a reliable comparison with sparse ones. Players in the dense groups often belong to the same caste and have similar characteristics. We end up with 503 women who played in both sparse and dense groups, as defined by our thresholds. The summary statistics are presented in Table 20. In total, we have four observations for each participant, for a total of 2012 observations.

We choose monitors candidate with respect to their Eigenvector centrality and their assignment to groups can be determined by either democratic election or random exogenous assignment. In order to neatly disentangle the different possible channels that might drive behavior, we set up an experiment where groups of three individuals are asked to privately vote for their preferred monitor and then play twice a standard public good game. The experimental session is sequenced as follows: first, players are assigned to a group formed either by their closest friends or by socially distant peers. The order of assignment to these two group compositions is randomized. Secondly, after being assigned their groups, players privately vote for their preferred monitor. Third, the choice of monitor is immediately followed by a contribution game. Each individual plays 2 rounds of a public good game within each group, once played with the elected monitor and once with a randomly picked monitor option, where we randomize the order of the two treatments. Groups are then reshuffled so that the same player is then placed in a different group composition (dense or sparse) and the game unfolds again as explained above. In total, each individual plays 4 rounds in two different groups (dense and sparse).

After participants play in the experimental sessions and receive payment for their perfor-

⁴²The dense groups would correspond to topography that is a triangle (average path=1), line (average path=1.3)

mance in the games, we administer a second questionnaire meant to capture caste, wealth, religion, membership to community based organizations and a set of other individual level characteristics. Participants are quite homogeneous in terms of wealth and networks display high level of clustering with respect to caste.

2.2 Experimental context

Nepali villages are often too remote to be reached easily or too sparse to ask their members to participate to group sessions in a fixed location. We decided to conduct our experiments in the mid-hills of Nepal in the district of Makwanpur, which is around four hours drive from Kathmandu. The municipalities we chose - Palung, Bajrabarahi, and Chitlang - present an economy almost uniquely focused on agriculture and the exploitation of natural resources. Dozens of community based organizations are active in the region and people are generally involved in at least one, they are familiar with issues of coordination and with the risks of free-riding. Villages are on average composed of 70 households, for an average of 120 women per village. We covered 19 villages with more than 2000 women between 18-60 years answering our network survey. We have a census of all inhabitants living in each village and administered the network questionnaire to every woman.

In partnership with a local research company based in Kathmandu, we hired a team of local enumerators. All enumerators were women, in order not to add any confounding factor in the network elicitation and in the experimental sessions. In each village, women who answered our network survey were invited via a phone call to take part in the experiment. We invited around 75% of interviewed women in each village and, based on a measure of (eigenvector) network centrality, we divided people into either players or monitor candidates. The individuals belonging to the top and bottom 5% of the centrality distribution were assigned the role of monitor candidates, while the others were assigned the role of players. Among those who were assigned the role of players, we oversampled groups in the periphery so as to avoid picking high centrality individuals in order to maximize the contrast between dense and sparse groups. This gives more power to the information transmission role of the monitor. As an incentive to participate, every player was given 100 Nrs (1€) along with the possibility to obtain additional money up to 200 Rs, as a function of their performance in the games. On average, the total gain was around 220 Rs. per individual which is half day's wage. Monitors were given a fixed sum of 250 Rs. for their participation.

The experiments were typically conducted early in the morning in schools close to each village. Women, as they arrived to schools, were assigned to either sparse or dense groups for the experiment. They were progressively sent to one of the classrooms to play the games. Once played, they got reassigned to play again with another group with a different group composition. The order of the dense and sparse groups was randomized. Typically, three sessions were run in parallel in separate classrooms with one session lasting for around 15

min. Two enumerators were in charge of each session: they read the instructions, conducted the game and noted down the choices of participants.

2.3 Design

In our experiment, we have three treatments variables. First, group composition. Groups can be composed of either close friends or people socially distant in the network. Second, centrality of monitors. In our experiment, we offer three monitoring options: high central monitors, low central monitors and no monitors. Third, the process whereby monitoring institutions are assigned: either democratically elected by the group or exogenously imposed. After assigning the role of high central and low central monitors, which remains fixed throughout the experiment, we divide the rest of the individuals into groups of three with varying group composition, either dense or sparse. Individuals play in groups of three in both dense and sparse treatment in a randomized order. In Figure 15, we show two possible groups for the player circled in green. She plays both with her closest friends - circled in red - and with individuals far in the social network circled in blue. By always reshuffling groups in such a way that every individual plays exactly in two different groups, we are able to extract individual fixed effects. This part of the design is of paramount importance because of the intrinsically endogenous nature of networks: the network position of player i is endogenous to her observable characteristics which are in turn affecting her contribution. This design allows a neat disentanglement of the endogenous position in the network from the contribution, through the extraction of fixed effects at the individual level.

At the start of each session, group players are gathered in a room where they can see each other, but no communication is allowed. Each member of the group receives 10 tokens of a different color, where the value of 1 token is marked at Rs 10. Each session is divided in two stages. In the first stage, each player privately casts a vote on her preferred monitoring option.⁴³ In Figure 16, step 1 represents the setting of the game. Players are given the option to choose between high central monitor (H), a low central monitor (L) or no monitor at all (NM). Note that this monitor is a fourth "player" that remains the same for all groups within a village. The cost of choosing the monitor is 20 Rs.⁴⁴ This cost makes always choosing a monitor a non - dominated strategy. The cost is paid by participants who vote to have a monitor (either high central or low central), irrespective of the voting outcome of the group.⁴⁵ The monitor is elected by a majority rule and the result of the vote is not immediately revealed. As seen in Step 2 of Figure 16, the group is then randomly assigned to either the endogenous treatment or the exogenous one. The

⁴³In case of a tie, the monitor choice was determined by a random draw. Ties represent around 6% of cases.

⁴⁴In line with the public good literature, the cost of the option was around 7% of the average earnings across all games.

⁴⁵If x votes for a monitor but no monitor is elected by the group, x stills pays the cost of voting for a monitor.

randomization is implemented by picking one out of two balls: if the ball drawn is green, the endogenous treatment is played first and the exogenous follows. If the ball drawn is pink then exogenous is played first followed by endogenous. The result of voting is only revealed just before playing the endogenous treatment. In the exogenous treatment, the group is randomly assigned either to a high central, low central or to no monitor treatment.

In the second stage of the experiment, the group plays a public good game where each player decides how many tokens to contribute to the public pot out of the 10. They are informed that the money in the public pot would be increased by 50% and then divided equally among them. As seen from Step 3 of Figure 16, once the contributions are made, the monitor -either elected or assigned- is called into the room to see how much each player contributed in the public pot. The monitor can distinguish the contributions belonging to each player by the different colors of the tokens they were endowed with. Moreover, the monitor does not have the power to impose fines and simply observes how much each player contributed. We exploit only the informational channel whereby the players' reputation can be affected (e.g. gossips, reporting etc.), following the assumption that it would drive a lot of real-life interaction in the village. We study how the fear of being reported on by the monitor outside the lab drives the behavior of people and how it consequently affects the demand for third party monitoring.⁴⁶ To sum up, the contribution game is played twice in the same group without receiving any feedback, once with the monitor option chosen by the group (endogenous) and once randomly assigned monitoring option (exogenous).

3 Results

The hypothesis is that the individual demand for peer monitoring varies depending on the identity of group members. In particular, we expect individuals in dense groups to not choose a monitor and to enforce co-operation on their own in the second stage of the game. This result would not hold for socially sparse groups where the temptation to free ride is higher, given the lower level of reciprocal trust. Thus, socially sparse groups might have a stronger incentive to pay the fixed cost of electing a monitor that is able to strengthen the reputation channel and the level of intra-group trust. The presence of a monitor - even more so for a high central one - increases the possibility of being reported on outside the lab in case of "defection". On the other hand, we expect socially close groups to be co-operative irrespective of the treatments.⁴⁷ We also expect to find a different impact between the endogenous and the exogenous monitors. The choice of the group is revealed only in the endogenous treatment.

⁴⁶Breza et al (2016 do not find a significant difference between information and punishment treatments)

⁴⁷Socially close or distant is characterized by the average social distance (path length) in a group

3.1 Preliminary findings and limitations

We start the analysis by looking at the individual level variation in the choice of the monitor. In Table 14, the numbers along the diagonal represent the percentage of individuals that always choose the same voting strategy irrespective of group composition. The largest proportion being 34.95% that always chooses to have no monitor, followed by 19.68% that always vote to have a high central monitor. The voting result shows substantial variation in voting strategy. Looking at the aggregate demand for peer monitoring, both dense and sparse groups vote more often to not have a monitor. Figure 17 shows that in dense groups, around 32% of players vote for a high central monitor, while in sparse groups more than 39% of players do so. Low central monitor is seldom chosen accounting for around 13% in both dense and sparse groups. For contribution, exogenous monitoring increases contribution only in sparse groups as seen from Table 15. We want to study how this differs when individuals have the power to demand their own monitoring institution. To begin with, we compare the outcomes under endogenous and exogenous institutions, clubbing all three monitor treatments together for the later in Figure 18. 'How' the monitoring institution is obtained matters only for sparse groups where endogenous monitoring in blue increases contribution compared to the exogenous one.

Before presenting the results, we want to highlight a possible limitation of the result given the caveats in the process of network elicitation. We ask for at least three "nominations" of friends. In most interviews, women named an average of 3-4 women which may not be fully exhaustive and may lead to networks that are much sparser than they actually are. This could imply an overestimation of social distance, i.e. individuals are actually socially closer than what they appear to be. It does not however represent a threat to the validity of our results. On the contrary, it implies that the estimated effects of our treatments represent a lower bound of the real effect.

3.2 Statistical Estimation

3.2.1 Impact of group on monitor voting: Sparse groups elect a high central monitor more often than dense

As suggested by the preliminary results shown in Figure 17, we conduct a Mann-Whitney test to understand whether the proportion of participants choosing a given monitor is significantly different across group compositions. We find that no monitor is chosen significantly more often in dense groups rather than in sparse groups (p-value 0.07) and that high central monitors are chosen more often in sparse rather than in dense groups (p-value 0.002).

In order to estimate how the demand for monitor varies depending on the group composition, we use a multi logit regression with individual and round fixed effects. Since players vote once in a dense and once in a sparse group in a random order, we can include both individual and round fixed effects, therefore exploiting a "within" design. The fixed effect

multi logit model is therefore defined by the logistic probability of choice of monitor y_{jt} , where $y_{jt}=0$: No monitor, $y_{jt}=1$: Low central monitor and $y_{jt}=2$: High central monitor. We take $y_{jt}=0$ as base category and can write the fixed effect logit as

$$Pr(y_{jt} = 1) = \frac{1}{1 + e^{-(\alpha + \beta_1 G_{jt} + \beta_2 X_g + \rho_j + \nu_t + \epsilon_{jt})}}$$

$$Pr(y_{jt} = 2) = \frac{1}{1 + e^{-(\alpha + \beta_2 G_{jt} + \beta_3 X_g + \rho_j + \nu_t + \epsilon_{jt})}}$$

where y_{jt} is the chosen level of monitoring, G_{jt} : dummy for group composition equal to 1 if the treatment is for dense groups, X_g : group characteristics ν_t : round fixed effect and ρ_j : individual fixed effects.

We present in Table 16 the results of the multinomial fixed effect regression of individual monitor choice (voting) on the social composition of the group (dense / sparse). In the first column, we find that dense groups are less likely to elect a high central monitor by 40% points compared to sparse groups. In the second column, we see that this is also true when we control for group characteristics (if the members have same caste and same education level). More details on voting as a function of individual characteristics is presented in Appendix C. This result is in line with our hypothesis that closer groups would prefer not to have monitoring whereas individuals in sparse groups would want a high level of monitoring.

3.2.2 Impact of different exogenous monitoring: High central monitor increases cooperation relative to no monitor in sparse groups

For contribution, we start with the baseline case where monitors are assigned exogenously and study the difference in contribution between sparse and dense groups. As seen from Table 15, in sparse groups, average contributions increase significantly (p-value 0.014) by Rs 7.4⁴⁸ (15.8% of the mean) in the presence of a high central monitor (H) as compared to no monitor (NM). In dense groups, there is a Rs 4.5 increase (8.3% of the mean) but the difference is not significant. This result is in line with the literature that suggests presence of a central monitor increases cooperation only in sparse groups (Breza et al, 2016). Further the cost of the monitor being 8% of the average payoff, it is optimal for sparse groups to vote for a monitor but not dense. Taking only the exogenous monitor treatment, we run a linear regression with fixed effects on the contribution with respect to the type of monitor that was assigned and the group composition. It takes the following form:

$$c_{jt} = \alpha + \beta_1 \cdot Dense + \beta_2 \cdot H + \beta_3 \cdot L + \beta_4 \cdot H * Dense + \beta_5 \cdot L * Dense + \rho_j + \nu_t + \epsilon_{jt}$$

⁴⁸Note that the value of 1 token is Rs 10. The regression is in terms of tokens but all the results are expressed in terms of Rs.

where c_{jt} : contribution of individual j in round t , $Dense$: dummy equal to 1 if the group is dense, H : dummy equal to 1 if a high central monitor is assigned, L : dummy equal to 1 if a low central monitor is assigned, ρ_j : individual fixed effect and ν_t : round fixed effects.

We are particularly interested in the coefficient β_2 that shows the effect of being assigned a high central monitor and β_4 that shows the difference in the effect across dense and sparse groups. In the first column of Table 17, the dependant variable is the monitor choice. We see that dense groups in general contribute Rs 13.7 higher (23% of the mean) than sparse ones. Next, contribution increases by Rs 7.25 (11% of the mean) in the presence of a High central monitor (H). As seen from the interaction term⁴⁹, the effect is starker in sparse groups. The second column shows similar effects when controlled for group characteristics. In both columns we control for the monitor choice of the individual.

3.2.3 Impact of endogenous v/s exogenous monitoring: Holding the monitoring institution fixed, for a given population, choosing a monitor increases cooperation only in sparse groups

To answer the impact of "how" the monitoring institution is chosen impacts contribution, we estimate a linear fixed effect regression. In the endogenous treatment, individuals 'select' into an institution that drives their contribution in addition to "how" the monitor is chosen. In order to overcome this selection problem, we keep monitoring fixed and compare groups which play both exogenous and endogenous treatment under the same monitor. Our identification strategy is to overcome selection by comparing the same group, with the same monitor treatment, differing only on how this monitor was obtained. Inspired by Dal bo et al., 2010, an individual i 's action in the game may depend on the group density $G \in \{\text{dense, sparse}\}$, elected monitor $M \in \{NM, H, L\}$, mechanism that selected the monitor $I \in \{\text{Endo, Exo}\}$ and her type α_i . The probability of cooperation is therefore determined by

$$P_i = f(M, G, I, \alpha_i)$$

We fix the group G and monitor M to determine the effect of the mechanism by which the monitor is elected. More formally,

$$E(P_i | G = \text{dense}, M = NM, \alpha_i, \text{Endo}) - E(P_i | G = \text{dense}, M = NM, \alpha_i, \text{Exo})$$

By doing so, we eliminate the threat of self-selection and we are able to disentangle the effect of the exogenous vs endogenous treatments. In terms of regression, it translates into

⁴⁹H x Dense being an interaction term represents [(H=1)-(H=0)|Dense]- [(H=1)-(H=0)|Sparse]

the following fixed effect equations.

$$\begin{aligned}
c_{jt} &= \alpha + \beta_1 \cdot (Endo \mid G = S, M = H) + \rho_j + \nu_t + \epsilon_{jt} \\
c_{jt} &= \alpha + \beta_1 \cdot (Endo \mid G = D, M = H) + \rho_j + \nu_t + \epsilon_{jt} \\
c_{jt} &= \alpha + \beta_1 \cdot (Endo \mid G = S, M = NM) + \rho_j + \nu_t + \epsilon_{jt} \\
c_{jt} &= \alpha + \beta_1 \cdot (Endo \mid G = D, M = NM) + \rho_j + \nu_t + \epsilon_{jt}
\end{aligned}$$

where c_{jt} : contribution of individual j in round t , $Endo$: is a dummy variable that takes value 1 if monitor is endogenously chosen, given group $G : \{D = dense, S = sparse\}$ and monitor choice $M : \{NM = No\ monitor, H = High\ central\ monitor\}$,⁵⁰ ρ_j : individual fixed effect and ν_t : round fixed effect. We are primarily interested in the coefficient β_1 that captures the effect of having an endogenous monitor as compared to being assigned exogenously.

Figure 19 shows the average contributions for sub samples that are free from selection effect. We see that for a sparse group, contribution increases under an endogenous monitoring setting as seen from the red bars. In particular, with endogenous no monitor, contribution increases significantly (p-value 0.009) by Rs 9.1 while with a high central monitor it increases by Rs 5 but not significantly. The change in dense groups across endogenous and exogenous monitoring institutions is not significantly different. We find in sparse groups that giving individuals opportunity to chose their own monitoring institution leads to better outcomes than externally imposing a third party monitor.

The first column in Table 18 reports results for individuals who self-selected into high monitoring institution in sparse groups followed by dense groups in the second column. The next two columns report the same but for the case where groups self selected into no monitoring. We see that a sparse group electing a high central monitor (H) endogenously increases contribution by Rs 5 (8.6% of the mean) whereas there is no effect for the dense group. Similarly, sparse group electing no monitor (NM) endogenously increases contribution by Rs. 9.13 (21.9% of mean) with no effect in the dense groups. This result presents evidence to believe that there exists a sort of "endogeneity premium": individuals facing the same monitoring institution behave differently depending on whether the institution is chosen by the group itself or imposed.

3.2.4 Impact of order endogenous/exogenous on contributions: Contributions are higher under EndoExo for dense NM and lower for dense H

Next, we look at the effect of order of endogenous and exogenous treatment on contribution. The result of the vote is only revealed in the Endogenous case, hence if Endogenous is

⁵⁰We also tried to do individual level analysis by looking at variation in monitor choice within groups. We find that 50% of the groups vote unanimously for the same monitor option hence not much power to study this effect.

played first, there is a possible extra information⁵¹ that could affect contribution in both the endogenous and exogenous treatment. We focus on No Monitor and High central monitors because of the very few observations we have for Low central monitor. In presenting this comparison, we plot the average contribution in treatments across monitoring condition and group composition that vary by the order in which endogenous and exogenous treatments were implemented. We can see in Figure 20 an evidence of a possible significant effect of the order, especially in dense groups. The election of a high central monitor in dense groups decreases contribution by Rs 9.7 significantly (p-value 0.09). When no monitor is elected, contribution increases by Rs 7.1 in sparse groups and Rs 9.4 in dense groups (p value 0.07). We run OLS regressions controlling for individual level characteristics. The dependant variable takes value 1 if the endogenous round is played first and 0 otherwise. We look at average contribution across the endogenous and exogenous treatments conditional on endogenous being played first versus if it is played second.

$$\begin{aligned}
c_{jt} &= \alpha + \beta_1 \cdot (Order \mid G = S, M = H) + \beta_i \cdot X + \epsilon_{jt} \\
c_{jt} &= \alpha + \beta_1 \cdot (Order \mid G = D, M = H) + \beta_i \cdot X + \epsilon_{jt} \\
c_{jt} &= \alpha + \beta_1 \cdot (Order \mid G = S, M = NM) + \beta_i \cdot X + \epsilon_{jt} \\
c_{jt} &= \alpha + \beta_1 \cdot (Order \mid G = D, M = NM) + \beta_i \cdot X + \epsilon_{jt}
\end{aligned}$$

where c_{jt} : contribution of individual j in round t , $Order$: is a dummy variable that takes value 1 if endogenous treatment is played first, given group $G : \{D = dense, S = sparse\}$ and monitor choice $M : \{NM = No\ monitor, H = High\ central\ monitor\}$, X : individual characteristics (caste, wealth, age and education). We are primarily interested in the coefficient β_1 that captures the effect of having played an endogenous monitor round first followed by the exogenous one.

In Table 19, we see that the effect of revealing the group's choice is stronger in dense groups as compared to the sparse. The average contribution decreases significantly by Rs 13.68 (13.3 % of mean) in dense groups that played endogenous first and elected a high central monitor (H). On the other hand, when groups play endogenous first and elect no monitor (NM), contribution increases by Rs 12.73 (19.7 % of mean) in dense groups and Rs 7.7 (7% of mean) in sparse. We hypothesize that signalling is stronger in dense groups, since they have a stronger prior about the altruism level in their group and are more likely to have future interactions. We model this in the section below.

⁵¹We hypothesize that this information could act as a signal of the level of trust in the group vis à vis each other.

4 The model

Through our data, we observe that the density of groups; either close in the network (dense) or far away (sparse) matters in demand for peer monitoring. Particularly, sparse groups vote more often to have a high central monitor. Through the model, we want to show this difference in demand for peer monitoring. Further, how the monitor is chosen matters for cooperation. The group having the choice to choose it's own monitor increases cooperation when no monitor is elected and decreases cooperation otherwise. The group's vote outcome may act as a signal about the level of altruism and trust in the group. The outcome of the vote being revealed only in the endogenous setting carries extra information that can change the belief of player i about the type of player j . We want to study the possible differences in contribution when endogenous is played first versus played second via a simple two player model where players vary in the level of altruism.

Types

We model the contribution behavior of individuals with an *altruism* parameter α . We can think of this parameter as determining the propensity of higher contribution. As people become more altruistic, i.e, the value of α increases, individuals are more likely to contribute a higher amount. Each player depending on the group can have a level of altruism $\alpha_i \in \{\alpha_l, \alpha_h\}$, where $\alpha_h > \alpha_l$. We further think of the parameter α as determining the choice of the monitor. Player i knows her own level of altruism α_i and forms a prior $\mu_{0i}(\alpha_j)$ on the level of altruism of the other player j . Let us assume that the subjective probability of individuals being type α_h is p in dense groups and q in sparse, where $p > q$ ⁵². This mimics the fact that in dense groups, people perceive their neighbors to be more altruistic compared to in a sparse group. Let d_i represent group composition where $d_i \in \{dense, sparse\}$.

The initial prior of Player i about j is then as follows:

$$\mu_{0i}(\alpha_j = \alpha_h) = \begin{cases} p & \text{if } d_i = \{dense\} \\ q & \text{if } d_i = \{sparse\} \end{cases} \quad (4)$$

Timing, Actions and Payoff

First, agents simultaneously vote for their preferred monitor $m_i \in \{0, 1\}$, where $m_i=0$ implies no monitor is chosen by individual i and $m_i=1$ means i votes for having the monitor. Once the participants cast their votes, a monitoring technology is assigned to the group according to the following rule

⁵²This assumption is in line with the literature on altruism where individuals contribute and cooperate more in closely knit dense groups. See Liedler et al. (2009) and Goeree et al. (2010).

$$m^* = \begin{cases} 1 & \text{if } m_i = m_j = 1 \\ 0 & \text{if otherwise} \end{cases} \quad (5)$$

where m^* denotes the outcome of the vote. Second, agents make their contribution decision $c_i \in \mathbb{R}$. The action profile of agent i is then (m_i, c_i) . The total contribution is increased by 50 % and divided equally among the group meaning the rate of return for the contribution game with two players is $\frac{3}{4}$.

The utility of Player i is a function of both c_i and c_j , of the level of altruism α_i and the rate of return in the contribution game $\frac{3}{4}$. We assume a convex cost c^2 in order to ensure there is an interior solution. We also believe this functional form characterized by increasing marginal cost well represents the behavioral burden of contributing. Further, in the spirit of psychological games⁵³, how much player i values the utility of player j depends on i 's belief about the altruism of player j , $\mu_{0i}(\alpha_j)$, since we believe that in this context belief dependant motivations deeply affect player i 's actions. In this regard, we take inspiration by Rabin (1993), who models reciprocity where players wish to act kindly (unkindly) as a function of belief about the other's type. The payoff of player i in the contribution game without a monitor is

$$U(\alpha_i | m^* = 0) = W - c_i - c_i^2 + \frac{3}{4}(c_i + c_j) + \alpha_i \cdot \mu_{0i}[W - c_j - c_j^2 + \frac{3}{4}(c_i + c_j)] \quad (6)$$

In the case where a monitor is elected, we add two terms to the above utility function: a monitor cost mc and a reputation cost $-\delta P(c_i < \theta)$. Voting for the monitor is costly and i pays mc if she votes for the monitor irrespective if the monitor gets elected or not. If elected, the monitor can impose a reputation cost on the players. We introduce a parameter $\delta > 1$ which represents the penalty from a contribution lower than the social norm θ in the presence of the monitor. For the sake of exposition, we use a fixed value of δ . However, we could incorporate varying power of the monitor depending on their centrality by allowing $\delta \in \{\delta_H, \delta_L\}$, where $\delta_H > \delta_L$ i.e. high central monitors are more effective in spreading information. The social norm is a stochastic parameter given that different monitors would have a different view of what a socially acceptable norm is. It is assumed to be uniformly distributed between $[0, \bar{\theta}]$ where $\bar{\theta}$ is the highest possible contribution. It can also be interpreted as a reference point (Kahneman and Tversky 1991) that varies with each monitor, i.e. it hinges on the distribution of θ . The probability of one's contribution being higher than the norm is then simply the cumulative distribution of all contributions which is $\frac{c_i}{\bar{\theta}}$ and the probability of contributing below the acceptable social - as perceived by the monitor - norm can thus be represented as

⁵³For a review on psychological game refer to Dufwenberg (2008) and Attanasiy and Nage (2008).

$$P(c_i < \theta) = \begin{cases} 1 - \frac{c_i}{\theta} & \text{if } c_i < \theta \\ 0 & \text{if otherwise} \end{cases} \quad (7)$$

The utility if a monitor is elected ($m^* = 1$) is then $U_i(\alpha_i | m^* = 1) =$

$$W - \hat{c}_i - \lambda \hat{c}_i^2 + \frac{3}{4}(\hat{c}_i + \hat{c}_j) - mc - \underbrace{\delta P(\hat{c}_i < \theta)} + \underbrace{\alpha_i \cdot \mu_{0i} [W - \hat{c}_j - \lambda \hat{c}_j^2 + \frac{3}{4}(\hat{c}_i + \hat{c}_j) - mc - \delta P(\hat{c}_j < \theta)]}_{(8)}$$

Moreover, players are Bayesian and i updates her prior about j 's type $\mu_{0i}(\alpha_j)$ via Bayesian rule to $\mu_{1i}(\alpha_j)$ depending on the outcome of the voting, m^* .

Equilibrium

We assume that the altruism parameter α_i of individual i fully determines her demand for peer monitoring. More formally, we consider an equilibrium of the form below. An (altruistic) player i of type α_h cares very much about the utility of the other player irrespective of j 's type. She would therefore prefer not to elect a monitor⁵⁴ in order to avoid the other player being punished via bad reputation in case of low contribution. For a player i of type α_l , however, the cost of electing a monitor and the negative reputation effects for both herself and j is outweighed by the increase in group contribution driven by the presence of the monitor. Agents then contribute differently depending on their type α_i , outcome of the vote m^* and updated belief μ_{1i} about player j , once the outcome of the vote is revealed. The separating equilibrium would then be

$$\sigma_i(\alpha_i) = \begin{cases} m_i = 0 & \text{if } \alpha_i = \alpha_h \\ m_i = 1 & \text{if } \alpha_i = \alpha_l \end{cases} \quad (9)$$

Given the equilibrium above, when $\alpha_i = \alpha_l$, i would always vote for a monitor and, given the voting rule m^* defined above, she is able to perfectly infer the voting choice of player j . In this case, i updates her prior to $\mu_{1i}(\alpha_j) = 1$ if $m^* = 0$ and $\mu_{1i}(\alpha_j) = 0$ if $m^* = 1$. On the other hand, type α_h always votes for $m_i = 0$ and no monitor is elected - $m^* = 0$ - irrespective of the vote of the other player. In this case player i cannot infer anything about j 's type and she sticks to the prior $\mu_{1i}(\alpha_j) = \mu_{0i}(\alpha_j)$. First, we solve the above set of equations and calculate the value of optimal contributions across the different scenarios. Secondly, given c_i , we study when the above separating equilibrium holds true.

⁵⁴Given that dense groups have higher subjective probability of being altruists, the demand of peer monitoring should be lower than that in sparse as seen in Fig 17.

We find that for type α_l , voting for the monitor is an optimal strategy for certain values of initial prior $\mu_{0i}(\alpha_j) < p_{\alpha_l}^*$. On the other hand, for type α_h it is always a dominant strategy to vote for no monitor. Therefore in order for there to be a separating equilibrium, it should be the case that people have a low prior on the proportion of altruists.

Proposition 1. *In the game with exogenous monitors, high-type players α_h contribute always more than low-type ones α_l . Moreover, at equilibrium the contributions of both players' types are higher in the presence of the monitor than without,*

$$\hat{c}_i^{exo} > c_i^{exo}$$

for $i = h, l$

where \hat{c}_i^{exo} indicates the optimal contribution when the contribution game is played in presence of a monitor and c_i^{exo} . The result is simply driven by the reputation effect of the monitor, which can entail the penalty δ in case of contributions lower than the social norm θ .

We now study the optimal contributions in the setting with the endogenous election of the monitor. In this case, the elected monitor serves as a signal of each other's types. Players are Bayesian and update their prior beliefs about the opponent's type, knowing their own vote in the election stage. The subjects play the game twice in random order, once with the endogenous monitor and once with the exogenous. When they play with the endogenously elected monitor first, they update their beliefs before the first contribution and we assume they will contribute identical amounts in the rounds, given their constant beliefs across rounds. On the other hand, when they play first with the exogenously assigned monitor, the updating happens only at the second round.

Proposition 2. *When the endogenous selection of monitors is allowed, the average contribution across the exogenous and the endogenous selection of monitors is higher if endogenous is played first and no monitor ($m^* = 0$) is elected by the group. If the group elects a monitor ($m^* = 1$), contribution decreases if endogenous is played first only in dense group.*

The election of no monitor signals that the group peers are high types α_h , thus pushing the posterior belief up to $\mu_{1i} = 1$. When this happens in the first round, contribution in both rounds increases since updating of the prior happens before players choose their optimal contribution. Surely, in this case it is higher then when the updating occurs only in the second round. On the other hand, when a monitor is elected in the first round, contribution decreases in both rounds. The decrease is however starker for dense groups due to higher initial prior. Hence, this effect holds only in dense groups.

Proposition 3. *Let us assume that δ is large enough and θ is small enough. Then, there exists a value of the initial prior $p_{\alpha_l}^*$ such that for $p < p_{\alpha_l}^*$, low types players α_l are better off voting for no monitor. Moreover, for $0 < p < p_{\alpha_l}^*$, there always exists the separating equilibrium σ , where low types α_l vote for the monitor and high types α_h vote for no monitor.*

We believe that the assumption of large δ is quite natural, given that in our context formal institutions are weak, and reputation concerns drive most of the social interactions. Similarly, the ex-ante level of cooperative behavior of these villages is modest, hence justifying the assumption of low values of θ . The mechanism underlying this proposition lies in the fact that high type players α_h always vote for no monitor, irrespective of the group they are in. Moreover, this proposition gives us reason to believe that a story of reciprocal altruism where players strategically vote for peer monitoring and contribute to a local public good, well describes the voting behavior we see in the experimental data, i.e. players vote more often for having a monitor in sparse groups (low p) rather than in dense ones.

5 Conclusion

By using original network data and a novel design, we try to understand how the varying demand of peer monitoring depends on group density and how this in turn affects cooperation. We divide the network into groups of three individuals with varying network distance, where dense implies each individual is at most at distance 2 (average path length < 1.6) and sparse implies each individual is at least at distance 4 (average path length > 4). To begin with, we show that dense groups prefer to not have a monitor whereas sparse groups choose to have a central one, reflecting variation in trust. Low central monitors are seldom chosen. In line with previous literature, when individuals are socially close (dense), they can sustain a higher level of cooperation without outside intervention. Dense groups contribute higher than the sparse group in the contribution game.

Next, we show "how" an institution is assigned matters for cooperation. The endogenous choice of monitoring increases cooperation in only sparse groups. Looking at the order of the monitor treatment, the outcome of the vote being revealed in endogenous treatment carries an additional information regarding individual preferences and hence, when revealed, acts as a signal to the group. When endogenous treatment is played first and no monitor is chosen by the group, individuals tend to contribute higher in both groups. However, when endogenous treatment is played first and a monitor is chosen, contribution decreases only in dense groups due to a stronger prior about the level of altruism. This is an interesting finding that suggests monitoring should be catered to the needs of the community. It is also in line with the argument that repeated interactions in dense groups imply higher concern for reputation.

Given the increased popularity of community-based interventions and focus on peer monitoring, it is important to understand the role social networks play in small scale societies. We propose here a theoretical framework followed by a simple experiment that show that the effect of a monitor can be very different depending on the density of the network. Our work opens avenues for further research. We would like to understand the choice of the monitors further by presenting individuals with a panel of monitor options rather than

just the high and low central ones.

Figures

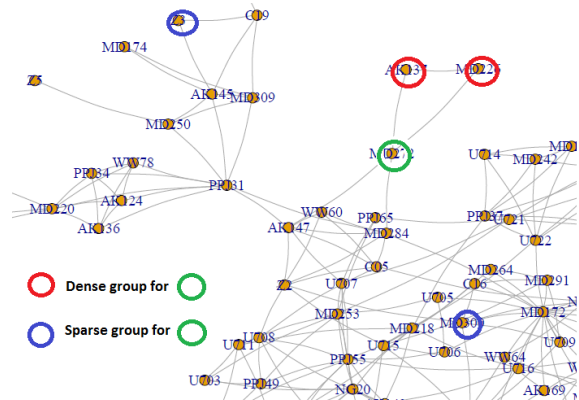


Figure 15 Example of formation of groups

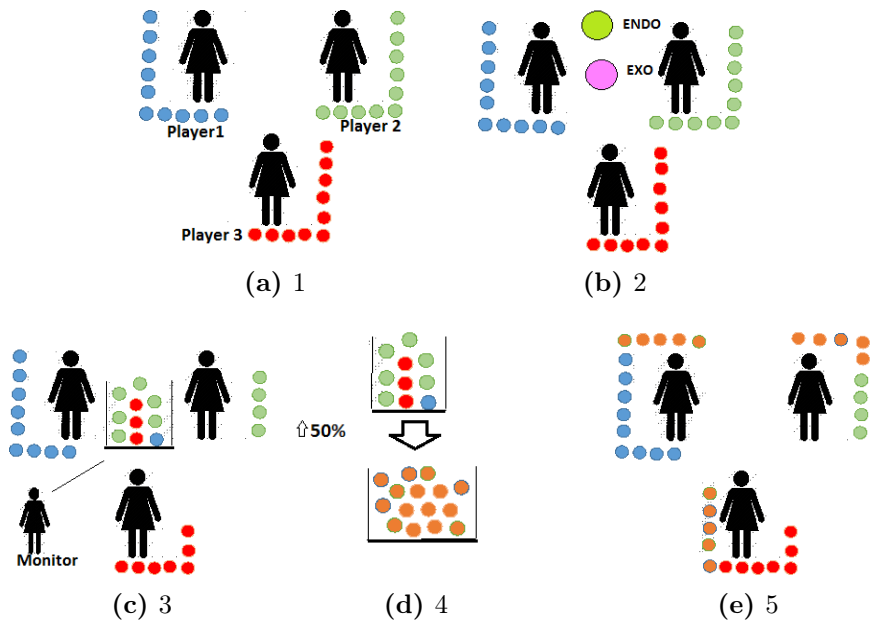


Figure 16 Experimental Design

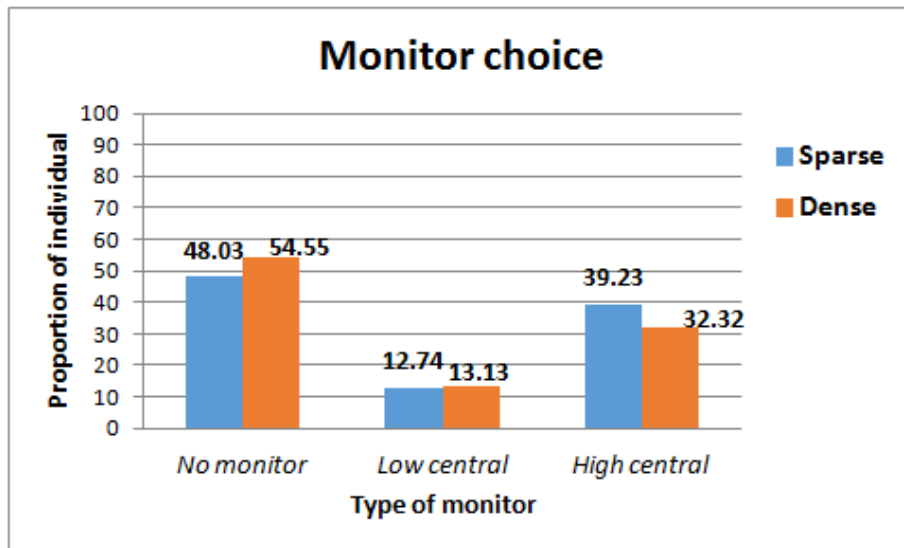


Figure 17 Percentage of individuals voting in Sparse and Dense groups

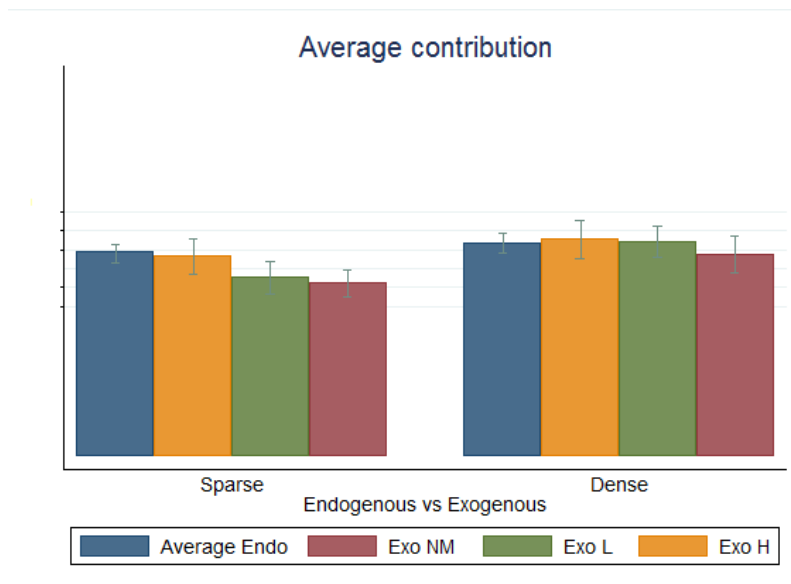


Figure 18 Average contribution endogenous v/s exogenous monitors with selection

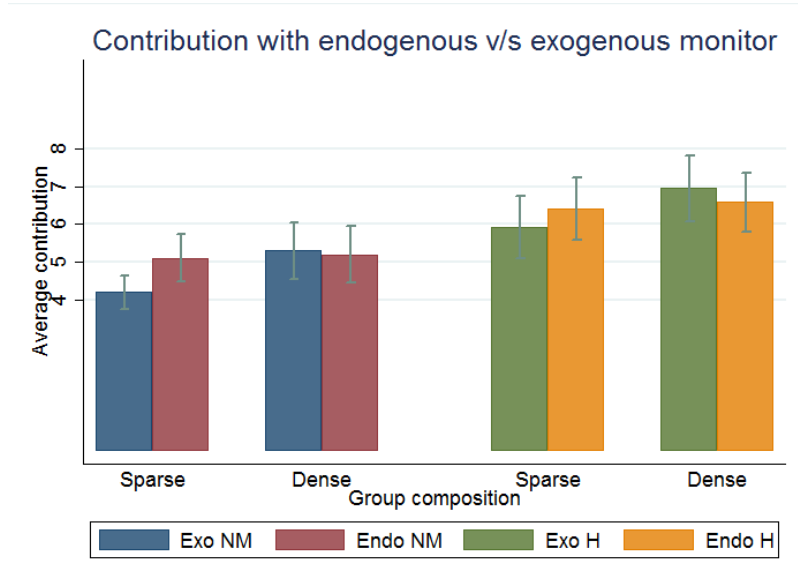


Figure 19 Contribution with endogenous v/s exogenous monitors without selection

Note: In the bar graph, y-axis represents group composition and x-axis represents average contribution. We focus on a sub sample where the same group plays under the same monitoring condition both exogenously and endogenously. The blue bar represents no monitor being assigned exogenously as compared to the red bar where monitor is assigned endogenously. The green bar represents exogenously assigned high central monitor as compared to the yellow bar.

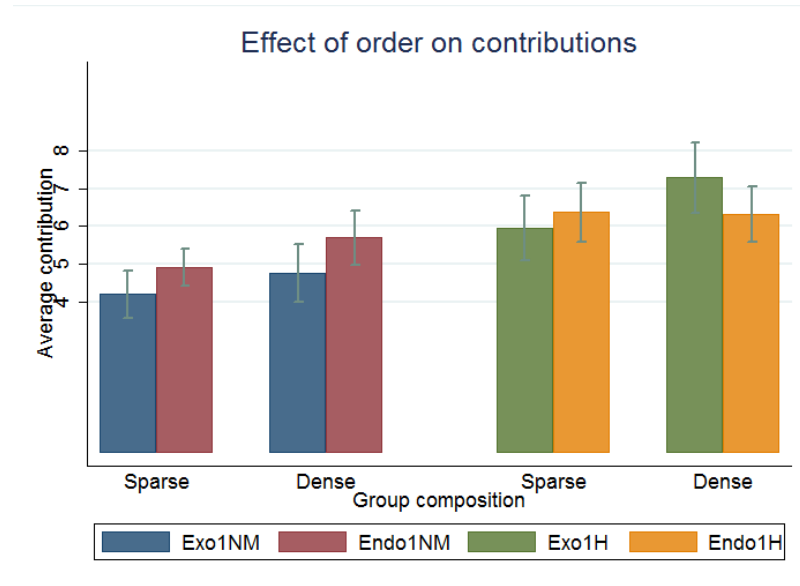


Figure 20 Order of endogenous and exogenous monitors

Note: In the bar graph, y-axis represents group composition and x-axis represents average contribution. We focus on a sub sample where the same group plays under the same monitoring condition both exogenously and endogenously. The blue bar represents no monitor being assigned exogenously as compared to the red bar where monitor is assigned endogenously. The green bar represents exogenously assigned high central monitor as compared to the yellow bar.

Tables

		Dense group		
		<i>No monitor</i>	<i>Low central</i>	<i>High central</i>
Sparse group	<i>No monitor</i>	34.95%	4.57%	9.34%
	<i>Low central</i>	5.17%	4.37%	2.98%
	<i>High central</i>	14.71%	4.17%	19.68%

Table 14 Variation in voting within individual across different groups

	NM	L	H
DENSE	5.39	5.71	5.84
SPARSE	4.67	4.76	5.41

Notes: Dense group contribute more than the sparse ones. In the presence of a high central monitor, contribution increases significantly in sparse groups.

Table 15 Average contribution in the exogenous treatment

	Monitor choice	Monitor choice
Low central		
Dense	-0.062 (0.24)	0.062 (0.20)
High central		
Dense	-0.407*** (0.18)	-0.466** (0.20)
N	503	459
Group characteristics	No	Yes

* p<0.10, ** p<0.05, *** p<0.01

Notes: No monitor is the base outcome. Monitor choice refers to the individual choice out of: No monitor, High central monitor and Low central monitor. Elected monitor is choice at the level of the group. Dense is a dummy variable that takes value 1 if the group is dense (average path length <2) and 0 otherwise. We control for individual and round fixed effects.

Table 16 Multilogit regression on monitor choice

	contribution	contribution
Dense	1.407*** (0.31)	1.614*** (0.33)
H	0.760** (0.36)	0.939** (0.38)
H \times Dense	-0.933** (0.43)	-1.088** (0.44)
L	0.646* (0.38)	0.748* (0.39)
L \times Dense	-0.972** (0.42)	-1.170*** (0.45)
N	503	459
Group characteristics	No	Yes

* p<0.10, ** p<0.05, *** p<0.01

Notes: Dense is a dummy variable that takes value 1 if the group is dense (average path length <2) and 0 otherwise. H is a dummy variable which is 1 if a High central monitor is elected and L is a dummy variable which is 1 if a Low central monitor is elected. We control for individual and round fixed effects along with individual choice.

Table 17 Fixed effect regression on average contribution under exogenous monitors

Contribution				
	Sparse(H)	Dense(H)	Sparse(NM)	Dense(NM)
Endogenous	0.50** (0.24)	-0.377 (0.25)	0.913*** (0.22)	-0.111 (0.20)
N	112	106	184	144

* p<0.10, ** p<0.05, *** p<0.01

Notes: Contribution is the amount given by individuals under each sub group. Sparse (H) refers to sparse groups (average path length > 4) who played both endogenous and exogenous treatment under a High central monitor (H). Dense NM refers to dense groups (average path length < 2) who played both endogenous and exogenous treatment under no monitor. Endogenous is a dummy that takes value 1 if contribution was made with choice of the group. We control for individual and round fixed effects.

Table 18 Endogenous v/s Exogenous Contribution Regression FE

	Contribution			
	Sparse(H)	Dense(H)	Sparse(NM)	Dense(NM)
Order	0.281 (0.58)	-1.368** (0.59)	0.770* (0.42)	1.273** (0.64)
N	92	94	158	120

* p<0.10, ** p<0.05, *** p<0.01

Notes: Contribution is the amount given by individuals under each sub group. Sparse (H) refers to sparse groups (average path length > 4) who played both endogenous and exogenous treatment under a High central monitor (H). Dense NM refers to dense groups (average path length < 2) who played both endogenous and exogenous treatment under no monitor. Order is a dummy that takes value 1 if endogenous treatment was played first. We control for individual characteristics and have round fixed effects.

Table 19 Effect of order on Contribution

6 Appendix

Appendix A Important clarification

The text in italic is not meant to be read aloud to experiment participants. It has the explanation of what experimenters should do. The remaining text that is not in italics is meant to be read aloud to experiment participants.

Experiment

Divide the research team into two groups: team A and team B. As participants enter the venue, team A must welcome them and locate their ID number based on their name from the individual identification list. The research team must then provide the participants with the consent forms, read the forms aloud, explain to them the contents of the forms and that the participants are free to leave at their discretion, answer any questions participants may have, and obtain their consent. [Go to Consent Form]

Then, team B should be ready to enter data on contributions.

Experiment begins

Thanks for coming today! We are researchers from Rooster Logic. You are participating in a study on daily decision-making. Today you will play a series of short games. The information gathered here will be confidential and used for research purposes only.

Overview

Today, we will ask you to play a game with two different groups of people for two rounds

each. You will randomly be placed in groups of three for the game, whose identity will be known. In each game, you and your group members will make some decisions. The result of these decisions will determine how much money you will earn today.

The games will represent situations and decisions you make every day in your life. You earn some money, you keep some money for yourself, you might give some money to your neighbors or friend, use the money to fund a common project etc.

Explanation of payment

Let us now discuss how you will make money today. First, you will receive 100 Rs. for simply participating in our games. Second, you will make money from the decisions made during the game.

You will play the same game with two different groups. In the beginning of each game, you will get some income in the form of tokens in a bag we call an “INCOME POT”. The game is easy and all that you need to do is decide how many tokens you want to keep for yourself and how many tokens you want to contribute to the “PUBLIC POT”. The total amount collected in the “PUBLIC POT” will be increased in value by 50%. In both games, the experimenter will collect the tokens that you want to contribute in two different “PUBLIC POT”.

At the end of the experiment, we will pick one “PUBLIC POT” out of the 4 and the total amount with the additional 50% increase will be equally divided among the four players in your corresponding group. You will receive equal share, irrespective of how much you put in the ‘PUBLIC POT’, Respectively, the tokens you decided to keep for yourself in the “INCOME POT” corresponding to that game will be yours.

Demonstrate: The “experimenter” should explain that they will be playing four rounds during the day with two different groups of people. Please show them the graphical image and explain how the contribution game works and how they would earn.

See then that the decisions you make in all rounds count but you will only be paid the amount in one randomly chosen game. Before I explain the game you will play today in detail, are there any questions?

Answer any questions that they may have.

Explanation of the game

The game I will explain to you is a very simple one. In this game, you will be matched randomly with 3 more people who you will interact with. You are not allowed to talk to each other throughout this game. At the beginning of the game, you and your partners will get some money that you can either keep for yourself or contribute to a common pot.

There are two stages in this game: First you will be given the choice to elect a monitor to oversee the contribution game that we just briefly explained. The monitor vote will be

followed by the contribution task. Let me explain in detail what the contribution task is.

At the beginning of each game, each of you will be given an initial income of Rs 100. All earnings during the games will be represented by tokens, each with a value of Rs 10. Then, each of you will be provided with 10 tokens that are worth Rs 100 in total. This cup will be known as “INCOME POT”.

Demonstrate procedure, the objective you should have in mind is that individuals acquire a sense of the physicality of the game.

Now, we will explain how you can use your income in the game. You can either keep the tokens for yourself in the ‘INCOME POT’ or you can contribute to the ‘PUBLIC POT’. The money that you decided to keep in the ‘INCOME POT’ will be yours. The tokens that you will put in the ‘PUBLIC POT’ will be added to the tokens that rest of your group put in the ‘PUBLIC POT’. The total amount contributed by the group will then increase in value by 50%.

The amount you contribute to the ‘PUBLIC POT’ will not be revealed to the rest of the members of your group. To contribute to the ‘PUBLIC POT’, you will give the number of tokens you want to contribute to the experimenter in the ‘PUBLIC POT’. Remember that 1 token is worth 10 Rs.

Demonstrate the procedure via the chart again. Explain to them that 2 tokens= 20 Rs

In the first stage, you will be given a chance to elect a monitor to oversee this contribution task. The monitor will observe the amount contributed by each individual to the ‘PUBLIC POT’ which is otherwise not known. In order to choose a monitor, you will put a tick next to one of the two choices: either having a monitor or not having a monitor. If you decided to have a monitor by putting a tick on the square, you will choose the name of the person you want to elect in the same sheet. If you decide to vote for having a monitor, you will be charged 10 Rs from the money you have been given for participation in the game.

Demonstrate the voting sheet to participants.

We will consider the choices of everyone in your group. The option that gets the highest number of votes will be chosen. Now, to see whether the majority choice will be implemented or an external option will be randomly assigned, we will pick a ball from this box without looking. In the box which we will call the ‘CHOICE BOX’.

We have two balls, 1 Pink and the other Green. We will pick a ball from the box, if a green ball is chosen, then the option chosen by the group will be implemented. If a pink ball is chosen instead, we will randomly assign one of the 3 options to your group.

Demonstrate the voting procedure to the participants with four enumerators. Make sure

they understand the use of the ‘CHOICE BOX’

Do we have any questions at this point? Have you understood the two stages of the game? Now, we will demonstrate the complete game.

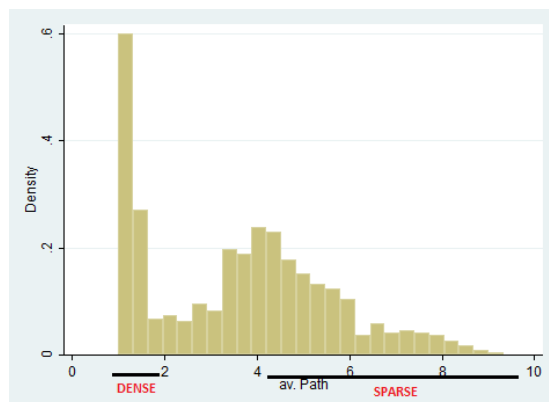
Five members of the team of experimenters should do the demonstration. Four of them should take the role contributors. The fifth person should represent himself and we will refer to him/her as the “experimenter.”

Do you have any questions?

Now, we will practice the game. Note that this will only be practice rounds and that you will not actually play with your actual partner. You will play the actual games with your actual partners after we explain the contribution game, practice them and we answer any question you might have about the games.

Participants play three rounds of the game and information is recorded exactly as if the game was actually being played.

Appendix B



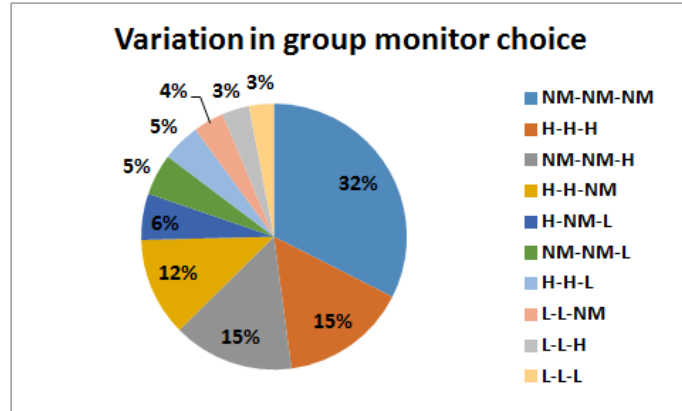
This is the distribution of average path length in the 1006 groups we formed. Average path length is defined as the average number of steps along the shortest paths for all possible pairs of the group. We over sampled closely knit groups with average path length < 2 (dense). Sparse group is defined as groups with average path length > 4 .

Figure 21 Distribution of average path length across groups

	Mean	Std.dev	N
<i>Individual Characteristics</i>			
age	35.8	11.43	503
education	3.06	3.85	503
no, of links	11.38	4.46	503
centrality	0.052	0.071	503
wealth index	-0.253	1.503	503
<i>Group Characteristics</i>			
Same caste	0.74	0.438	503
Same education	0.3801	0.485	503

Table 20 Summary statistics of the participants

Appendix C: Monitor choice



It shows the variation in individual choice within a group. NM: no monitor being chosen, L: low central monitor and H: high central monitor is chosen. In most groups, all three members vote for NM followed by all three group members voting for H .

Figure 22 Variation in individual choice within a group

	Monitor	
	Dense	Sparse
Age	-0.0004 (0.002)	-0.003* (0.001)
Caste	0.021 (0.02)	0.090*** (0.02)
Education	-0.030** (0.01)	-0.052*** (0.01)
Wealth	0.014 (0.01)	0.050*** (0.01)
Favor return strangers	0.002 (0.01)	-0.048*** (0.01)
Help friends	-0.009 (0.01)	0.038*** (0.01)
Centrality	0.181 (0.30)	0.566** (0.29)
Distance to H	-0.001 (0.01)	0.032*** (0.01)
Distance to L	0.005 (0.01)	0.010* (0.01)
N	842	842

* p<0.10, ** p<0.05, *** p<0.01

Notes: Monitor is a dummy that takes value 0 if no monitor is elected and 1 if either a high or low monitor is elected. The first column (Dense) regresses individual characteristics with outcome of the vote and the second column does the same but for sparse groups.

Table 21 OLS regression for monitor choice behavior

Appendix D: Model

Let us first consider the exogenous case with no signalling. Since the monitoring technology is randomly assigned and not chosen by the group, there is no update of the prior μ_{0i} .

For type α_h

$$\begin{aligned}
 U(\alpha_i = \alpha_h, \cdot, m^* = 1) &= p \cdot U(\hat{c}_h^{exo}, \hat{c}_h^{exo}, \alpha_h, \alpha_h) + (1 - p) \cdot U(\hat{c}_h^{exo}, \hat{c}_l^{exo}, \alpha_h, \alpha_l) \\
 U(\alpha_i = \alpha_h, \cdot, m^* = 0) &= p \cdot U(c_h^{exo}, c_h^{exo}, \alpha_h, \alpha_h) + (1 - p) \cdot U(c_h^{exo}, c_l^{exo}, \alpha_h, \alpha_l)
 \end{aligned}$$

For type α_l

$$\begin{aligned}
 U(\alpha_i = \alpha_l, \cdot, m^* = 0) &= p \cdot U(c_l^{exo}, c_h^{exo}, \alpha_l, \alpha_h) + (1 - p) \cdot U(\hat{c}_l^{exo}, \hat{c}_l^{exo}, \alpha_l, \alpha_l) \\
 U(\alpha_i = \alpha_l, \cdot, m^* = 1) &= p \cdot U(c_l^{exo}, c_h^{exo}, \alpha_l, \alpha_h) + (1 - p) \cdot U(\hat{c}_l^{exo}, \hat{c}_l^{exo}, \alpha_l, \alpha_l)
 \end{aligned}$$

where \hat{c}_i^{exo} denotes the contribution of player i when there is the monitor and c_i^{exo} when there is no monitor.

Solving for each contribution level c_l^{exo} , c_h^{exo} , \hat{c}_l^{exo} , \hat{c}_h^{exo} we get for the exogenous assignment of monitoring technology,

$$\begin{aligned}\hat{c}_l^{exo} &= \frac{3\alpha_l p - 1}{8} + \frac{\delta}{2\theta} & c_l^{exo} &= \frac{3\alpha_l p - 1}{8} \\ \hat{c}_h^{exo} &= \frac{3\alpha_h p - 1}{8} + \frac{\delta}{2\theta} & c_h^{exo} &= \frac{3\alpha_h p - 1}{8}\end{aligned}$$

In the endogenous case we have to take into account the election rule and now the monitor outcome (m^*) becomes a signal according to which players update their belief about player j 's type. We show that in both dense and sparse groups, average contribution in two rounds is higher if endogenous is played first and no monitor ($m^* = 0$) is elected by the group. In the event of a monitor ($m^* = 1$) being elected by the group, contribution decreases significantly only in dense groups. Given the updated priors, we can write the utility function for type α_h

$$\begin{aligned}U(\alpha_i = \alpha_h, \cdot, m^* = 1) &= U(\hat{c}_h^{end}, \hat{c}_l^{end}, \alpha_h, \alpha_l) \\ U(\alpha_i = \alpha_h, \cdot, m^* = 0) &= p \cdot U(c_h^{end}, c_l^{end}, \alpha_h, \alpha_h) + (1 - p) \cdot U(c_h^{end}, c_l^{end}, \alpha_h, \alpha_l)\end{aligned}$$

For type α_l

$$\begin{aligned}U(\alpha_i = \alpha_l, \cdot, m^* = 0) &= U(c_l^{end}, c_h^{end}, \alpha_l, \alpha_h) \\ U(\alpha_i = \alpha_l, \cdot, m^* = 1) &= U(\hat{c}_l^{end}, \hat{c}_h^{end}, \alpha_l, \alpha_l)\end{aligned}$$

Solving for each contribution level c_l^{end} , c_h^{end} , \hat{c}_l^{end} , \hat{c}_h^{end} we get,

$$\begin{aligned}\hat{c}_l^{end} &= -\frac{1}{8} + \frac{\delta}{2\theta} & c_l^{end} &= \frac{3\alpha_l - 1}{8} \\ \hat{c}_h^{end} &= -\frac{1}{8} + \frac{\delta}{2\theta} & c_h^{end} &= \frac{3\alpha_h p - 1}{8}\end{aligned}$$

To establish the separating equilibrium,

For type: $\alpha_i = \alpha_l$

To show that voting for a monitor is optimal we calculate ex-ante utility of player i using prior beliefs via ,

$$\begin{aligned}U(\alpha_i = \alpha_l | m_i = 1) &= p \cdot U(\alpha_l, \alpha_h, c_l, c_h) + (1 - p) \cdot U(\alpha_l, \alpha_l, \hat{c}_l, \hat{c}_l) \\ U(\alpha_i = \alpha_l | m_i = 0) &= p \cdot U(\alpha_l, \alpha_h, c_l, c_h) + (1 - p) \cdot U(\alpha_l, \alpha_l, c_l, c_l) \\ U(\alpha_i = \alpha_l | m_i = 1) - U(\alpha_i = \alpha_l | m_i = 0) &= (1 - p) \cdot \{U(\alpha_l, \alpha_l, \hat{c}_l, \hat{c}_l) - U(\alpha_l, \alpha_l, c_l, c_l)\} - p \cdot mc\end{aligned}$$

There exists a $p^*_{\alpha_l}$ st. $p < p^*_{\alpha_l}$ $U(\alpha_i = \alpha_l | m_i = 1) - U(\alpha_i = \alpha_l | m_i = 0)$ is positive.

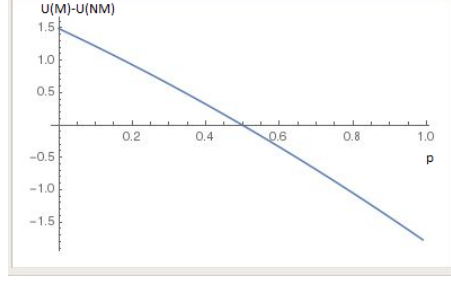


Figure 23 Difference in utility between voting for a monitor or not $U(m_i = 1) - U(m_i = 0)$

For type: $\alpha_i = \alpha_h$

$$U(\alpha_i = \alpha_h | m_i = 0) = p \cdot U(\alpha_h, \alpha_h, c_h, c_h) + (1 - p) \cdot U(\alpha_h, \alpha_l, c_h, c_l)$$

$$U(\alpha_i = \alpha_h | m_i = 1) = p \cdot U(\alpha_h, \alpha_h, c_h, c_h) + (1 - p) \cdot U(\alpha_l, \alpha_l, \hat{c}_h, \hat{c}_l)$$

$$U(\alpha_i = \alpha_h | m_i = 1) - U(\alpha_i = \alpha_h | m_i = 0) = (1 - p) \cdot \{U(\alpha_h, \alpha_l, \hat{c}_h, \hat{c}_l) - U(\alpha_h, \alpha_l, c_h, c_l)\} - p \cdot mc$$

In calculating the difference, we find that $U(\alpha_h | m_i = 1) - U(\alpha_h | m_i = 0) < 0$, therefore a high type would always vote for no monitor.

Combining the two results above, there exists an interval of p where type α_l would choose a monitor whereas type α_h would choose no monitor. Therefore there exists a separating equilibrium with p st. $0 < p < p^*_{\alpha_l}$ where the above equilibrium holds true. If $p > p^*_{\alpha_l}$ then there is no separating equilibrium and both types vote for no monitor.

Appendix E: Model extension with three agents

We expand the two agent model to three agents for it to be more representative of the interaction. Let us first consider the exogenous case with no signalling. Since the monitoring technology is randomly assigned and not chosen by the group, there is no update of the prior μ_{0i} . The voting rule m^* is slightly different with no tie possible.

For type α_h

$$U(m^* = 1) = p(1 - p) \cdot [U(\hat{c}_h^{exo}, \hat{c}_h^{exo}, \hat{c}_l^{exo}) + U(\hat{c}_h^{exo}, \hat{c}_l^{exo}, \hat{c}_h^{exo})] +$$

$$p^2 \cdot [U(\hat{c}_h^{exo}, \hat{c}_h^{exo}, \hat{c}_h^{exo}) + U(\hat{c}_h^{exo}, \hat{c}_l^{exo}, \hat{c}_l^{exo})]$$

$$U(m^* = 0) = p(1 - p) \cdot [U(\hat{c}_h^{exo}, \hat{c}_h^{exo}, \hat{c}_l^{exo}) + U(\hat{c}_h^{exo}, \hat{c}_l^{exo}, \hat{c}_h^{exo})] +$$

$$p^2 \cdot [U(\hat{c}_h^{exo}, \hat{c}_h^{exo}, \hat{c}_h^{exo}) + U(\hat{c}_h^{exo}, \hat{c}_l^{exo}, \hat{c}_l^{exo})]$$

For type α_l

$$\begin{aligned} U(m^* = 1) &= p(1-p) \cdot [U(\hat{c}_l^{exo}, \hat{c}_h^{exo}, \hat{c}_l^{exo}) + U(\hat{c}_l^{exo}, \hat{c}_l^{exo}, \hat{c}_h^{exo})] + \\ &\quad p^2 \cdot [U(\hat{c}_l^{exo}, \hat{c}_h^{exo}, \hat{c}_h^{exo}) + U(\hat{c}_l^{exo}, \hat{c}_l^{exo}, \hat{c}_l^{exo})] \\ U(m^* = 0) &= p(1-p) \cdot [U(\hat{c}_l^{exo}, \hat{c}_h^{exo}, \hat{c}_l^{exo}) + U(\hat{c}_l^{exo}, \hat{c}_l^{exo}, \hat{c}_l^{exo})] + \\ &\quad p^2 \cdot [U(\hat{c}_l^{exo}, \hat{c}_h^{exo}, \hat{c}_h^{exo}) + U(\hat{c}_l^{exo}, \hat{c}_l^{exo}, \hat{c}_l^{exo})] \end{aligned}$$

where \hat{c}_i^{exo} denotes the contribution of player i when there is the monitor and c_i^{exo} when there is no monitor. Solving for each contribution level c_l^{exo} , c_h^{exo} , \hat{c}_l^{exo} , \hat{c}_h^{exo} we get for the exogenous assignment of monitoring technology,

$$\begin{aligned} \hat{c}_l^{exo} &= \frac{2\alpha_l p - 1}{4} + \frac{\delta}{2\theta} & c_l^{exo} &= \frac{2\alpha_l p - 1}{4} \\ \hat{c}_h^{exo} &= \frac{2\alpha_h p - 1}{4} + \frac{\delta}{2\theta} & c_h^{exo} &= \frac{2\alpha_h p - 1}{4} \end{aligned}$$

Given the calculated level of contribution under the exogenous monitor, we reiterate Proposition 1 where

$$\hat{c}_i^{exo} > c_i^{exo}$$

for $i = h, l$

In the endogenous case we have to take into account the election rule and now the monitor outcome (m^*) becomes a signal according to which players update their belief about player j 's type. We show in Proposition 2 that in both dense and sparse groups, average contribution in two rounds is higher if endogenous is played first and no monitor ($m^* = 0$) is elected by the group. In the event of a monitor ($m^* = 1$) being elected by the group, contribution decreases more in dense groups for a small p , where $p < \frac{2}{3}$. Given the updated priors, we can write the utility function for type α_h

$$\begin{aligned} U(\alpha_i = \alpha_h, \cdot, m^* = 1) &= U(\hat{c}_h^{end}, \hat{c}_l^{end}, \hat{c}_l^{end}) \\ U(\alpha_i = \alpha_h, \cdot, m^* = 0) &= p(1-p)[U(c_h^{end}, c_h^{end}, c_l^{end}) + U(c_h^{end}, c_l^{end}, c_h^{end})] + p^2 U(c_h^{end}, c_h^{end}, c_h^{end}) \end{aligned}$$

For type α_l

$$\begin{aligned} U(\alpha_i = \alpha_l, \cdot, m^* = 0) &= U(\hat{c}_l^{end}, \hat{c}_h^{end}, \hat{c}_h^{end}) \\ U(\alpha_i = \alpha_h, \cdot, m^* = 0) &= p(1-p)[U(c_l^{end}, c_h^{end}, c_l^{end}) + U(c_h^{end}, c_l^{end}, c_h^{end})] + p^2 U(c_l^{end}, c_l^{end}, c_l^{end}) \end{aligned}$$

Solving for each contribution level c_l^{end} , c_h^{end} , \hat{c}_l^{end} , \hat{c}_h^{end} we get,

$$\begin{aligned} \hat{c}_l^{end} &= \frac{4\alpha_l - 3}{12} + \frac{\delta}{2\theta} & c_l^{end} &= \frac{2\alpha_l - 1}{4} \\ \hat{c}_h^{end} &= -\frac{1}{4} + \frac{\delta}{2\theta} & c_h^{end} &= \frac{4\alpha_h - 3}{12} \end{aligned}$$

To establish the separating equilibrium,

For type: $\alpha_i = \alpha_l$

To show that voting for a monitor is optimal we calculate ex-ante utility of player i using prior beliefs via ,

$$\begin{aligned} U(m_i = 1) &= p(1-p) \cdot [U(\hat{c}_l^{exo}, \hat{c}_h^{exo}, \hat{c}_l^{exo}) + U(\hat{c}_l^{exo}, \hat{c}_l^{exo}, \hat{c}_h^{exo})] + \\ &\quad p^2 \cdot [U(c_l^{exo}, c_h^{exo}, c_h^{exo}) + U(\hat{c}_l^{exo}, \hat{c}_l^{exo}, \hat{c}_l^{exo})] \\ U(m_i = 0) &= p(1-p) \cdot [U(c_l^{exo}, c_h^{exo}, c_l^{exo}) + U(c_l^{exo}, c_l^{exo}, c_h^{exo})] + \\ &\quad p^2 \cdot [U(c_l^{exo}, c_h^{exo}, c_h^{exo}) + U(\hat{c}_l^{exo}, \hat{c}_l^{exo}, \hat{c}_l^{exo})] \end{aligned}$$

There exists a $p^*_{\alpha_l}$ st. $p < p^*_{\alpha_l}$ $U(\alpha_i = \alpha_l | m_i = 1) - U(\alpha_i = \alpha_l | m_i = 0)$ is positive.

For type: $\alpha_i = \alpha_h$

$$\begin{aligned} U(m_i = 1) &= p(1-p) \cdot [U(\hat{c}_h^{exo}, \hat{c}_h^{exo}, \hat{c}_l^{exo}) + U(\hat{c}_h^{exo}, \hat{c}_l^{exo}, \hat{c}_h^{exo})] + \\ &\quad p^2 \cdot [U(c_h^{exo}, c_h^{exo}, c_h^{exo}) + U(\hat{c}_h^{exo}, \hat{c}_l^{exo}, \hat{c}_l^{exo})] \\ U(m_i = 0) &= p(1-p) \cdot [U(c_h^{exo}, c_h^{exo}, c_l^{exo}) + U(c_h^{exo}, c_l^{exo}, c_h^{exo})] + \\ &\quad p^2 \cdot [U(c_h^{exo}, c_h^{exo}, c_h^{exo}) + U(\hat{c}_h^{exo}, \hat{c}_l^{exo}, \hat{c}_l^{exo})] \end{aligned}$$

In calculating the difference, we find that $U(\alpha_h | m_i = 1) - U(\alpha_h | m_i = 0) < 0$, therefore a high type would always vote for no monitor.

Combining the two results above, there exists an interval of p where type α_l would choose a monitor whereas type α_h would choose no monitor. Therefore there exists a separating equilibrium with p st. $0 < p < p^*_{\alpha_l}$ where the above equilibrium holds true. If $p > p^*_{\alpha_l}$ then there is no separating equilibrium and both types vote for no monitor.

5 Summary

This thesis combines experiments and theoretical models with data analysis to answer questions on the role of social network and aggregated information. Chapter 1 studies a multi-state binary choice experiment in which in each state, one alternative has well understood consequences whereas the other alternative has unknown consequences. Subjects repeatedly receive feedback from past choices about the consequences of unfamiliar alternatives but this feedback is aggregated over states. Varying the payoffs attached to the various alternatives in various states allows us to test whether unfamiliar alternatives are discounted and whether subjects' use of feedback is better explained by similarity-based reinforcement learning models (in the spirit of the valuation equilibrium, Jehiel and Samet 2007) or by some variant of Bayesian learning model. Our experimental data suggest that there is no discount attached to the unfamiliar alternatives and that similarity-based reinforcement learning models have a better explanatory power than their Bayesian counterparts.

Chapter 2 studies a natural follow up to the question to understand how these findings would change in the face of the feedback being individual specific in the spirit of learning by doing. The follow up experiment allows subject to experiment and generate individual level feedback unlike the endogenous group feedback in the original one. Our experimental data suggest that there is not much difference in learning and the choice in the proportion of Bayesian learners.

Chapter 3 studies the demand for monitoring and its effectiveness across different group compositions. In developing countries where formal institutions are often weak or non-existent, the community is responsible to enforce local agreements. Peer monitoring represents a natural mechanism for the enforcement of social norms and local agreements in such a setting. In this paper we collect original network data in 19 villages in rural Nepal and conduct an experiment to study who is elected as a monitor in a public good game. Individuals play in groups of three, both with their close friends and with people socially distant in the network. They receive the opportunity to anonymously choose their preferred "institution". We combine a theoretical model and a unique lab-in-the-field experiment to show that closely knit groups are significantly more likely to not choose any monitor, while sparse groups tend to prefer a monitor who is highly central in their network. Low central monitors are seldom chosen. Endogenous election of the high central monitor ensures higher cooperation compared to an exogenous assignment, but only in sparse groups.

Field: Economics

Key words: Social network, Peer monitoring, Experiment, Public good game, Learning, Bounded Rationality, Valuation equilibrium

References

- Acemoglu, D., & Wolitzky, A. (July, 2015). Community enforcement vs. specialized enforcement. *MIT Department of Economics Working Paper No. 15-06*.
- Ali, S. N., & Miller, D. A. (2016). Ostracism and forgiveness. *American Economic Review*, 106.
- Attanasiy, G., & Nagelz, R. (2008). Psychological games. In *Games, rationality and behaviour. essays on behavioural game theory and experiments*. Houndmills: Palgrave Macmillan.
- Badev, A. (2013). Discrete games in endogenous networks: theory and policy. *PSC Working Paper Series Population Studies Center*.
- Baldassarri, D., & Grossman, G. (2010). The effect of group attachment and social position on prosocial behavior. evidence from lab-in-the-field experiments. *PLOS ONE*, 8(3).
- Ballester, C., Calvo-Armengol, A., & Zenou, Y. (2006). Who's who in networks. wanted: The key player. *Econometrica*, 74(5).
- Banerjee, A., Chandrashekhar, A., Duflo, E., & Jackson, M. (2015). Gossip: Identifying central individuals in a social network. *NBER working paper 20422*.
- Banerjee, A., Duflo, E., Jackson, M. O., & Chandrashekhar, A. (2013). The diffusion of microfinance. *Science*, 341(6144).
- Barron, S., Kai Huck, & Jehiel, P. (2019). Everyday econometricians: Selection neglect and overoptimism when learning from others. *Working paper*.
- Beersma, B., & Van Kleef, G. A. (2011). How the grapevine keeps you in line: gossip increases contributions to the group. *Social Psychology and Personality Science*, 2.
- Bjorkman, M., & Svensson, J. (2009). Power to the people: Evidence from a randomized field experiment on community-based monitoring in uganda. *The Quarterly Journal of Economics*, 124(2).
- Bowles, S. (2008). Economic experiments undermine "the moral sentiments": Evidence from policies designed for self-interested citizens. *Science*, 320(1605).
- Camerer, C., & Ho, H.-T. (1999). Experience-Weighted Attraction Learning in Normal form games. *Econometrica*, 67.
- Cason, T. N., & Gangadharan, L. (2016). Swords without covenants do not lead to self-governance. *Journal of Theoretical Politics*, 287(1).

- Cason, T. N., Sheremeta, R. M., & Zhang, J. (2012). Communication and efficiency in competitive coordination games. *Games and Economic Behavior*, 76, 26-43.
- Chandrashekhkar, A., & Breza, E. (2018). Social networks, reputation and commitment: evidence from a savings monitors experiment. *Econometrica*.
- Chandrashekhkar, A., Breza, E., & Larreguy, H. (2016). Network centrality and informal institutions: evidence from a lab experiment in the field. *NBER working paper 20309*.
- Chandrashekhkar, A., Kinnan, C., & Larreguy, H. (2018). Social networks as contract enforcement: Evidence from a lab experiment in the field. *American Economic Journal: Applied Economics*.
- Charness, G., & Levin, D. (2005). When optimal choices feel wrong: a laboratory study of bayesian updating, complexity, and affect. *The American Economic Review*, 95(4), 1300-1309.
- Charness, R., Gary Cobo Reyes, & Jimenez, N. (2008). An investment game with third-party intervention. *Journal of Economic Behavior and Organization*, 68.
- Cheung, Y.-W., & Friedman, D. (1997). Individual learning in normal form games: some laboratory results. *Games and Economic Behavior*, 19, 46-76.
- Cownden, D., Eriksson, K., & Strimling, P. (2018). The implications of learning across perceptually and strategically distinct situations. *Synthese*, 195, 511-528.
- Dal Bo, A., Pedro Foster, & Putterman, L. (2010). Institutions and behavior: experimental evidence on the effects of democracy. *American Economic Review*, 100.
- Dufwenberg, M. (2008). Psychological games. In *Palgrave macmillan (eds) the new palgrave dictionary of economics*. London: Palgrave Macmillan.
- Elliott, M., & Golub, B. (2017). A network approach to public goods. *Journal of Political Economy*.
- Ellsberg, D. (1961). Risk, ambiguity and the savage axioms. *The Quarterly Journal of Economics*, 643-661.
- Enke, B. (2019). What you see is all there is. *Working paper; Harvard University*.
- Enke, B., & Zimmermann, F. (2019). Correlation neglect in belief formation. *The Review of Economic Studies*, 86, 313-332.
- Epstein, L., & Halevy, Y. (2019). Hard-to-interpret signals. *Working Paper 634 University of Toronto*.

- Epstein, L., & Schneider, M. (2007). Learning under ambiguity. *Review of Economic Studies*, 74, 1275-1303.
- Erev, I., & Roth, A. E. (1998). Predicting how people play games : Reinforcement learning in experimental games with unique mixed strategy. *The American Economic Review*, 88(4), 848-881.
- Esponda, I., & Vespa, E. (2019). Contingent preferences and the sure-thing principle: Revisiting classic anomalies in the laboratory. *Working paper*.
- Etang, D., Alvin Fielding, & Knowles, S. (2011). Does trust extend beyond the village? experimental trust and social distance in cameroon. *Experimental Economics*, 14(1).
- Fehr, D., & Sutter, M. (2019). Gossip and the efficiency of interactions. *Games and Economic Behavior*, 113.
- Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review*, 90(4).
- Ferraro, P. J., & Vossler, C. A. (2010). The source and significance of confusion in public goods experiments. *The B.E. Journal of Economic Analysis Policy Contributions*, 10(1).
- Fonseca, M. A., & Peters, K. (2018). Will any gossip do? gossip does not need to be perfectly accurate to promote trust. *Games and Economic Behavior*, 107.
- Fryer, R., & Jackson, M. O. (2008). A categorical model of cognition and biased decision making. *The B.E. Journal of Theoretical Economics*, 8(1), 1-44.
- Fryer, R. G., Harms, P., & Jackson, M. O. (2019). Updating beliefs when evidence is open to interpretation: implications for bias and polarization. *Journal of European Economic Association forthcoming*.
- Fudenberg, D., & Levine, D. K. (1998). *The theory of learning in games*. MIT Press.
- Gelade, W., & Guiringer, C. (2018). The enforcement advantage of external monitoring: Lessons from an experiment with joint-liability groups in burkina faso:. *Journal of Economic Behavior and Organization*, 151.
- Gilboa, I., & Schmeidler, D. (1995). Case based decision theory. *The Quarterly Journal of Economics*, 110, 605-639.
- Glaeser, E., Laibson, D., Scheinkman, J. A., & Soutter, C. (2000). Measuring trust. *Quarterly Journal of Economics*, 115(3).
- Glockner, A., Kube, S., & Nicklisch, A. (2018). The joint benefits of observed and unobserved social sanctions. *Journal of Behavioral and Experimental Economics*, 75.

- Goeree, J., McConnell, M., Mitchell, T. D., Tromp, T., & Yariv, L. (2007). The 1/d law of giving. *American Economic Journal: Microeconomics*, 2(1), 183–203.
- Greif, A. (1989). Reputation and coalitions in medieval trade: Evidence on the maghribi traders. *Journal of Economic History*, 49(4).
- Greif, A. (2006). *Institutions and the path to the modern economy: Lessons from medieval trade*. Cambridge University press.
- Grimm, V., & Mengel, F. (2012). An experiment on learning in a multiple games environment. *Journal of Economic Theory*, 147(6), 2220–2259.
- Grossman, G., & Baldassarri, D. (2012). The impact of elections on cooperation: Evidence from a lab-in-the-field experiment in uganda. *American Journal of Political Science*, 56(4).
- Guillen, P., Schwieren, C., & Staffiero, G. (2006). Why feed the leviathan? *Public Choice*, 130.
- HeB, S. H., Jaimovich, D., & Schundeln, M. (2019). Development projects and economic networks: Lessons from rural gambia. *Working paper*.
- Hoffman, K., Elizabeth McCabe, & Smith, V. L. (1996). Social distance and other-regarding behavior in dictator games. *The American Economic Review*, 3(86).
- Hopkin, E. (2002). Two competing models of how people learn in games. *Econometrica*, 70, 2121–2166.
- Isaac, R. M., & Walker, J. M. (1988). Group size effects in public goods provision: The voluntary contributions mechanism. *The Quarterly Journal of Economics*, 103(1).
- Jackson, M. O., Tomas, R. B., & Tan, X. (2012). Social capital and social quilts: Network patterns of favor exchange. *American Economic Review*, 102(5).
- Jackson, M. O., & Wolinsky, A. (1996). A strategic model of social and economic networks. *Journal of Economic Theory*, 71(2).
- Jehiel, P. (2005). Analogy-based expectation equilibrium. *Journal of Economic Theory*, 123(2), 81–104.
- Jehiel, P., & Koessler, F. (2008). Revisiting games of incomplete information with analogy based equilibrium. *Games and Economic Behavior*, 62(2), 533–557.
- Jehiel, P., & Samet, D. (2005). Learning to play extensive games by valuation. *Journal of Economics Theory*, 121(557), 129–148.

- Jehiel, P., & Samet, D. (2007). Valuation equilibrium. *Theoretical Economics*, 163-185.
- Kandori, M. (1992). Social norms and community enforcement. *The Review of Economic Studies*, 59.
- Karlan, D., Mobius, M., Rosenblat, T., & Szeidl, A. (2009). Trust and social collateral. *The Quarterly Journal of Economics*, 124(3).
- Ketz, P. (2018). Subvector inference when the true parameter vector may be near or at the boundary. *Journal of Econometrics*, 207(2), 285-306.
- Kosfeld, M., & Rustagi, D. (2015). Leader punishment and cooperation in groups: experimental field evidence from commons management in ethiopia. *American Economic Review*, 105(2).
- Kranton, R. E. (1996). Reciprocal exchange: A self-sustaining system. *The American Economic Review*, 86(4).
- Lageras, A., & Seim, D. (2016). Strategic complementarities, network games and endogenous network formation. *International Journal of Game Theory*, 3.
- Leider, S., Mobius, M. M., Rosenblat, T., & Quoc-Anh, D. (2009). Directed altruism and enforced reciprocity in social networks. *Quarterly Journal of Economics*, 124(4).
- Martinez-Marquina, M., Alejandro Niederle, & Emanuel, V. (2019). Probabilistic states versus multiple certainties: The obstacle of uncertainty in contingent reasoning. *forthcoming American Economic Review*.
- McKelvey, R. D., & Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10, 6-38.
- Mengel, F. (2002). Learning across games. *Games and Economic Behavior*, 74, 601-619.
- Merkle, E. C., You, D., & Preacher, K. J. (2016). Testing non nested structural equation models. *Psychological Methods*, 21.
- Nepal, A. K., Mani Bohara, & Berrens, R. P. (2007). The impacts of social networks and household forest conservation efforts in rural nepal. *Land Economics*, 83(2).
- Olken, B. A. (2007). Monitoring corruption: Evidence from a field experiment in indonesia. *Journal of Political Economy*, 115(2).
- Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge University press.
- Ozbay, E. F., & Ozbay, E. Y. (2014). Effect of an audience in public goods provision. *Experimental Economics*, 17.

- Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American Economics Review*, 83(5).
- Rustagin, D., Engel, S., & Kosfeld, M. (2010). Conditional cooperation and costly monitoring explain success in forest commons management. *Science*, 330(1).
- Ryan, B., & Gross, N. (1943). Acceptance and diffusion of hybrid corn seed in two iowa communities. *Rural Sociology*, 8, 15-24.
- Salmon, T. C. (2001). An evaluation of econometric models of adaptive learning. *Econometrica*, 69, 1597-1628.
- Samuelson, L. (2001). Analogies, adaptations and anomalies. *Journal of Economic Theory*, 97, 320-367.
- Sommerfeld, R., Krambeck, H.-J., Semmann, D., & Milinski, M. (2007). Gossip as an alternative for direct observation in games of indirect reciprocity. *Proceedings of the National Academy of Sciences of the U.S.A*, 104.
- Sutter, M., Haigner, S., & Kocher, M. G. (2010). Choosing the carrot or the stick? endogenous institutional choice in social dilemma situations. *The Review of Economic Studies*, 77(4).
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge Press.
- The World Bank . (2004). *World development report: Making services work for poor people*. The World Bank and Oxford University Press.
- Train, K. (2003). *Discrete choice methods and simulations*. Cambridge University Press.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, New Series*, 185(4157), 1124-1131.
- Tversky, A., & Kahneman, D. (1979). Prospect theory, an analysis of decision under risk. *Econometrica*, 47(2), 263-291.
- Vespa, E., & Wilson, A. J. (2016). Communication with multiple senders: An experiment. *Quantitative Economics*, 7, 1-36.
- Wilcox, N. (2006). Theories of learning in games and heterogeneity bias. *Econometrica*, 74, 1271-1292.
- Wu, J., Balliet, D., & Van Lange, P. (2015). When does gossip promote generosity? indirect reciprocity under the shadow of the future. *Social Psychology and Personality Science*.

Wu, J., Balliet, D., & Van-Lange, P. (2016). Gossip versus punishment: The efficiency of reputation to promote and maintain cooperation. *Scientific reports*.

Young, H. P. (2009). Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning. *American Economic Review*, 99(5), 1899-1924.

Zimmermann, F. (2019). The Dynamics of Motivated Beliefs. *CRC TR 224 Discussion Paper Series; University of Bonn and University of Mannheim, Germany*.

List of Figures

1	Feedback structure for treatment sessions	33
2	Set up of the different treatment sessions	33
3	Evolution of choice across treatments with aggregated feedback	34
4	Evolution of choice across treatments with dis-aggregated feedback	35
5	Simulated choices using the reinforcement model	36
6	Results for Montecarlo simulations for Model 1.	41
7	Improving the fit of simulated data in T1 with a lower noise parameter λ . .	42
8	Feedback structure for treatment sessions	52
9	Set up of the different treatment sessions	52
10	Evolution of choice across treatments with aggregated feedback: followup design	53
11	Evolution of choice across treatments with aggregated feedback: original design	53
12	Imbalance in the choice of Red across states	54
13	Evolution of choice across treatments with aggregated feedback: followup design	59
14	Evolution of choice across treatments with dis-aggregated feedback: Control	60
15	Example of formation of groups	80
16	Experimental Design	80
17	Percentage of individuals voting in Sparse and Dense groups	81
18	Average contribution endogenous v/s exogenous monitors with selection . .	81
19	Contribution with endogenous v/s exogenous monitors without selection . .	82
20	Order of endogenous and exogenous monitors	83
21	Distribution of average path length across groups	89
22	Variation in individual choice within a group	90
23	Difference in utility between voting for a monitor or not $U(m_i = 1) - U(m_i = 0)$	93

List of Tables

1	Information about the relative payoff of urns Blue and Green after 100 random draws is reported at the start of each session.	37
2	Parameters for similarity-based reinforcement learning model	37
3	Parameter for Bayesian model with bounds	37
4	BIC values for the two competing models	37
5	Parameters for variant1 of similarity-based reinforcement learning model . .	37
6	Parameters for variant2 of similarity-based reinforcement learning model . .	38
7	Parameter for Bayesian model with Blue and Green updating	38
8	Parameter for elaborate Bayesian model	38
9	Information about the relative payoff of urns Blue and Green after 100 random draws is reported at the start of each session.	54
10	Optimal choice under the two solution concept.	54
11	Parameters for similarity-based reinforcement learning model	55
12	Parameters for variant1 of similarity-based reinforcement learning model . .	55
13	Parameter for Bayesian model with bounds	55
14	Variation in voting within individual across different groups	83
15	Average contribution in the exogenous treatment	84
16	Multilogit regression on monitor choice	84
17	Fixed effect regression on average contribution under exogenous monitors .	85
18	Endogenous v/s Exogenous Contribution Regression FE	85
19	Effect of order on Contribution	86
20	Summary statistics of the participants	90
21	OLS regression for monitor choice behavior	91