ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Ahn, T. K. et al.

Working Paper

Trust and communication in a property rights dilemma

Working Paper, No. 2016-5

Provided in Cooperation with:
Department of Economics, Brown University

Suggested Citation: Ahn, T. K. et al. (2016) : Trust and communication in a property rights dilemma, Working Paper, No. 2016-5, Brown University, Department of Economics, Providence, RI

This Version is available at:
https://hdl.handle.net/10419/162755

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Trust and Communication in a Property Rights Dilemma

T.K. Ahn[‡], Loukas Balafoutas[§], Mongoljin Batsaikhan[¶], Francisco Campos-Ortiz[+],
Louis Putterman[†] and Matthias Sutter[#]

Abstract. We study a laboratory social dilemma game in which incentives to steal from others lead to the socially inefficient diversion of resources from production unless the members of a given mini-society can abide by norms of non-theft or engage in low cost collective protection of their members' wealth accumulations. We compare two treatments in which subjects have opportunities to exchange free-form messages to one without such opportunities, finding that most subjects allocate far less to theft and most groups achieve much greater efficiency in the presence of communication. Ease of identifying who has engaged in theft varies across the two communication treatments, but is of minor importance to the outcome. We find several coding-amenable elements of message content to be statistically significant predictors of group and individual outcomes.

KEYWORDS: Property Rights; Theft; Efficiency; Experiment; Communication
JEL CODES: C91, C92, D03, H41, P14

[‡] Department of Political Science and International Relations, Seoul National University. Email: tkahn@snu.ac.kr
[§] Department of Public Economics, University of Innsbruck. Email: Loukas.Balafoutas@uibk.ac.at
[¶] School of Foreign Services in Qatar, Georgetown University. Email: mb1712@georgetown.edu
[+] Prudential Fixed Income – Global Macroeconomic Research, francisco.camposortiz@prudential.com
[†] Corresponding author. Department of Economics, Brown University. Email: Louis_Putterman@Brown.Edu
[#] Department of Economics, University of Cologne, and Department of Public Economics, University of Innsbruck, Email: matthias.sutter@uni-koeln.de

"[N]o one can have property in anything except as others acknowledge it. … my right of property in a thing depends not upon my claim to it but precisely upon your readiness to admit my claim as privileged" – John Taylor, 1993.

1. Introduction

Secure property rights, which are considered crucial for incentives to invest in production and accumulation of productive assets, can be said to rest on varying combinations of private vigilance, collective protection, and voluntary adherence to normative constraints against theft. While laboratory decision experiments have shown that the behaviors of many individuals are significantly impacted by concerns for fairness or reciprocity (Cooper and Kagel, forthcoming), experimental economists have also confirmed that when able to do so, many will transfer potential payoffs from others to themselves in line with traditional assumptions of self-interest (Abbink *et al*., 2000). Under what conditions subjects can establish norms to deter one another from stealing, when group efficiency depends on secure property, remains a relatively under-explored topic within the experimental and behavioral literatures on social dilemmas.

A few studies, including Durham *et al*. (1998) and Duffy and Kim (2005) have used laboratory decision experiments to investigate endogenous property rights or appropriative conflict. The present study more closely resembles Ahn *et al*. (forthcoming), whose set-up emphasizes trust and other social dimensions of the securing of property rights, partly by including a collective enforcement activity. Notably absent in that study, however, is the possibility of communication, which Ostrom (2010) argues is a key element in many real world successes in addressing social dilemmas. The present paper brings the experimental study of property rights and of communication together by adding communication opportunities to treatments otherwise identical to a core treatment in Ahn *et al*.

Standard economic theory which assumes rationality, strict self-interest, and common knowledge that all share and know one another to know of those features, predicts no improvement in efficiency from adding communication in our setting. In sharp contrast with this prediction, we find that the introduction of communication causes average theft to fall by more than 50%, and causes allocations to production to rise and those to private protection to fall by about 70%. Desistance from theft and the resulting rise in efficiency are significantly greater in our treatments with communication than they are in treatments with availability of mandatorily-financed property protection studied by Ahn *et al*., contrary to predictions for rational, selfish subjects who are expected to select dramatically efficiency-enhancing institutions when on offer. This result suggests that trust-building can overcome obstacles assumed in standard economic theory, whereas an incentive-compatible institutional option that requires no more than selfish rationality to implement may fail to achieve comparable results, absent opportunities to first build trust.

The increase in efficiency brought about by opportunities for communication could in principle be driven by simple adherence to norms of cooperation or by reluctance to break pledges made during message exchange (lying- or guilt-aversion), but for some subjects, the threat of being punished by other group members engaging in coordinated retaliatory theft can in principle have played a crucial role. To gauge the importance of punishment threats versus norms and pledges acting independently of such

threats, we ran two communication treatments, in one of which communication could easily operate through both channels because subjects were provided with information on the sources of one another's earnings, and in the other of which punishment threats are less credible due to the absence of this information. In the event, outcomes in our "low information chat" treatment are statistically indistinguishable from those in its higher information counterpart, although a quantitatively small and statistically insignificant difference in the expected direction (slightly less cooperation in the information-poor environment) is observed. Moreover, as detailed below, threats of punishment can have been effective only if the large majority of members of the affected groups believed that others would abide by their agreements. Regression analysis using coded chat content shows that groups in which the idea of punishing thieves by stealing from them was articulated more in pre-play communication did not achieve better outcomes than ones in which few or no subjects suggested this, whereas the more members agreed on a strategy not mentioning punishment, the less theft occurred and the more production took place.

The remainder of our paper proceeds as follows. In Section 2, we describe our experimental design and theoretical predictions, also briefly discussing past experimental findings especially with respect to communication. We report and discuss subjects' decisions in the aggregate in Section 3. Section 4 reports a coding exercise in which we translate the subjects' chat room messages into a small set of categorical and quantifiable variables so that we can investigate econometrically the predictive power of message contents. We find several dimensions of message content to be statistically significant predictors of behaviors. In Section 5, we discuss the effects of communication in some illustrative groups, providing further insights into the interplay between messages and decisions over the course of experimental sessions. Section 6 summarizes and concludes the paper.

2. Setting, Experimental Design, and Predictions

We conduct a laboratory decision-making experiment that shares with many others the fact that there are well-defined privately optimal choices for rational, self-interested individuals with common knowledge of type, and that past observation suggests that actual behaviors will diverge from such predictions in a manner potentially consistent with norm-abiding or social preferences (Camerer and Fehr, 2004, Sobel, 2005). In these respects, the experimental game played by our subjects can be classed along with simpler games such as the prisoners' dilemma, the public goods or voluntary contributions game, and common pool resource games, as a *social dilemma*: a situation in which the social optimum is at odds with private incentives to defect, free ride or over-extract (Ostrom, 2010; Van Lange, Rockenbach and Yamagishi, eds., 2014). The experiment's focus on acts of taking, or refraining from taking, is also shared by experiments such as Abbink et al. (2000)'s "moonlighting game" (see also List, 2007), and it fits more specifically into a literature on the trade-offs between investments in production, protection and predation (Grossman and Kim, 1995, Duffy and Kim, 2005, Kimbrough, Smith and Wilson, 2010).

The decision-making game played in our experiment involves five individuals assigned to a group whose composition stays fixed for 24 interaction periods, seated in decision lab sessions of 15 or 20 participants from amongst whom one's partners cannot be identified. Every individual receives

(virtually) an endowment of 10 "effort tokens" at the start of each period. Effort tokens can be used for (1) *production*, a private activity generating "wealth tokens" with diminishing returns shown in Table 1; (2) *theft*, an activity in which each token member *i* allocated to stealing from the accumulation of member *j* gives *i* 10 of *j*'s wealth tokens with probability (1 - P_{jt}), where P_{jt} reflects the total property protection *j* enjoys in period *t*; (3) *private protection*, an activity that raises *j*'s protection level P_{jt} by 0.1 (10%) for each token *j* allocates to it; or (4) *collective protection*, an activity that raises the protection level P_{kt} of all group members (*k* includes *i* and *j*) by 0.06 (6%) for every token assigned to it up to a maximum of 12 tokens (72% protection). An in dividual's property protection level cannot exceed 100%, to which it can be brought by a combination of collective and private protection. For example, 12 tokens allocated to collective protection by some combination of group members' contributions and 3 tokens allocated to private protection by *j* brings P_{jt} to 100%, while if *j* lowers her allocation to private protection by 1 token in the otherwise identical situation P_{jt} is 92% ([12*.06]+[2*.10]). A given period's attempt by any *i* to steal from any *j* either fully succeeds or fully fails, the outcome being determined by an independent random draw with the stipulated probability. For example, if *i* directs 3 tokens to stealing from *j*, *j* loses and *i* receives 30 tokens with probability (1 - P_{jt}) and 0 tokens with probability P_{jt}. Each subject begins period 1 with 100 wealth tokens, and her wealth token accumulation can never fall below 0.[1] Whereas effort tokens are given the subject in limited quantity each period, must be divided among the four activities, and have no monetary value in their own right, each subject's final wealth token accumulation is converted to cash at the end of the session.[2]

We can denote *i*'s budget constraint as

$$10 = m_{it} + t_{it} + p_{it} + c_{it}$$

where *m*, *t*, *p* and *c* denote allocations to making, taking, privately protecting, and collectively protecting wealth, and using a reduced form that suppresses details on who directs theft at whom we can represent *i*'s wealth token earnings in period *t* as

$$w_{it} = f(m_{it}) + g(t_{it},P_{-it}) - l(t_{-it},P_{it})$$

where *f* is the wealth production function (see again Table 1), *g* a function mapping *i*'s theft, the protection levels of those it is targeted at, and the outcomes of the relevant random draws, into *i*'s gains from theft, and *l* is the corresponding function mapping others' theft, *i*'s protection level, and chance into *i*'s losses from theft. It is straightforward to show that total earnings are maximized, with w_{it} = 70 for all *i* when each player chooses m_{it} = 10. However, were others to devote all tokens to production, one could profit by devoting tokens to theft (since each token would return 10 wealth tokens versus the

[1] A thief's "takings" are limited, if necessary, to make sure that this constraint is observed. The restriction never had to be imposed in any treatment, however.

[2] Each wealth token converted to $0.014 in the U.S., €0.0125 in Austria, $0.16 pesos in Mexico, ₮11 in Mongolia, and ₩14 in South Korea, with conversion rates scaled to yield payoffs of similar value relative to local student cost-of-living. Subjects received a fixed show-up fee of $5 in the U.S. and comparable amounts (relative to local earnings) in the other countries, with total earnings from wealth tokens averaging an additional $21.40, €20,3 (≈ $25.86), $239.5 (≈ $18.21 [USD]), ₮13,933.9 (≈ $10.92), and ₩21,034.1 (≈ $17.44), respectively (at average 2010 exchange rates).

smaller returns from units above the third put into production, see Table 1). Private payoff maximizing players who assume one another to be of the same type can thus gain by deviating until each chooses m_{it} = 3, t_{it} = 7, p_{it} = 0 and c_{it} = 0 and each earns w_{it} = 39, making clear that the environment is a social dilemma. If subjects were perfectly rational, strictly self-interested, and believed one another to be so (common knowledge), it is this (3, 7, 0, 0) decision vector that is to be expected of them.[3] But the social optimum would have each individual use all tokens for production, thus earning w_{it} = 70 each period. Individual rationality accordingly leads to the forfeiting of a little over 44% of potential earnings.

A detail omitted from the description above is that in each period of interaction, subjects first select their allocations to collective protection and the aggregate amount so allocated is announced to group members, who then decide the assignment of their remaining tokens among the other three activities. This temporally segmented process is made available to subjects so that we can investigate the possibility that cooperatively inclined group members would invest in collective protection as a way of signaling cooperative intentions while simultaneously discouraging theft. Notice that if each member were to invest 3 tokens in private protection in a given period, their protection levels would be only 30% and also unknown to one another, leaving expected gains from each theft token at 7 units (if the protection level were accurately anticipated), so that theft would remain more profitable on the margin than production. Investment by each of 3 tokens in collective protection, in contrast, yields a 72% protection level for each, hence expected gains of only 3.8 units per token assigned to theft, leaving production the most profitable option for the remaining units of endowment. This sequential decision and information-sharing process changes nothing under the triple assumptions of strict self-interest, rationality, and common knowledge: allocating 0 to collective protection, 0 to private protection, and 7 and 3 respectively to theft and production remains privately optimal. However, the opportunity to signal cooperative intent offered by it is of potential interest under alternative assumptions, as discussed below.

Another detail of relevance to our later discussion is that at any point during the decision process, subjects are free to consult a statistics screen on which is displayed the current wealth accumulation of their fellow group members along with each one's net accumulation of wealth tokens in the most recent past period. In two treatments, this "stats window" also breaks each group member's accumulations down by origin: output from production, gain from theft, loss from theft. Conceivably, strategic or psychological factors might lead to targeting of theft with particular attention to how wealth was obtained—e.g., subjects might seek retribution for past thefts committed against them, or in the treatment having both this high level of information and communication opportunities, group members could reach and attempt to implement agreements to punish any individual thief by collaborative

[3] The reasoning behind this solution to the private payoff maximization problem is the same as in Ahn *et al*. Briefly, if subject *i* assumes others to follow the indicated decision profile, she cannot achieve expected benefit by deviating from the same behavior, since diverting a token from theft to private protection yields an additional expected payoff of 7 (by lowering expected loss to theft from 70 to 63 tokens) while lowering her expected payoff by 10. If *i* deviated by instead allocating the token to collective protection, her expected gain would be an even smaller 4.2 tokens, with the same expected loss of 10. Likewise, a one-unit deviation towards production yields an expected gain of 9 tokens, versus the 10 available from theft. From any other decision profile, in contrast, it is possible to demonstrate that deviations exist that can increase expected payoff.

retaliatory theft. As with the detail that elements c_{it} of the decision vector are chosen and made known in the aggregate before each group member allocates her remaining tokens, the availability of the stats window does not change predicted decisions under self-interest, rationality, and common knowledge, which still should adhere to the (3, 7, 0, 0) allocation pattern.[4] In our third treatment, however, we lower the information level in the stats window to display of accumulations without information about source, in order to explore whether having the source information plays a major role in facilitating cooperation when communication opportunities are available.

In all, we study three treatments that share the above structure. The 24 interaction periods are divided into six phases of four periods each, with short breaks between phases. In our No Communication (hereafter **No-CHAT**) treatment, subjects learn of one another's simultaneously-made decisions but are not permitted to communicate in any other way, with the breaks between phases serving only for purposes of parallelism with other treatments.[5] In the remaining two treatments, members of each group are invited to communicate with each other using typed messages in an online chat room immediately before commencing the first four periods of play and during the breaks between phases. The duration of chat opportunities is fixed in advance and becomes shorter as the experiment proceeds: 4 minutes during the first opportunity (before period 1), then 3.5 minutes, 3 minutes, 2.5 minutes, and 2 minutes for each of the final two chat meetings. We started with relatively long chat periods so that all issues could be aired, then successively shortened them to reduce the likelihood of boredom.[6] Between-phase breaks in **No-CHAT** last for 1 minute each.

As alluded to previously, our two treatments with communication differ from each other in that the structure of information during play in one, which we'll call simply **CHAT**, is identical to that in **No-CHAT**, whereas in the other, which we'll call **L-CHAT** (for chat treatment with low information), subjects are given less information with which to identify who in their group might be engaging in theft. Whereas in **No-CHAT** and **CHAT**, the stats window shows not only each group member's current accumulation of wealth tokens, but also what part of that accumulation was obtained by production and theft, as well as how many wealth tokens the individual has lost due to theft, in the **L-CHAT** treatment, the stats window shows only the number of wealth tokens gained in the most recent period and the accumulation thus far of wealth tokens, giving no breakdown of acquisition by production versus theft and losses from theft. As explained above, we added the **L-CHAT** treatment to check whether, and if so how much, the benefit of permitting communication in **CHAT** was due to ease of implementing threats to punish thieves by group coordination on stealing from any thief.

---

[4] Technically, the availability of the statistics may be pertinent insofar as subjects are unable to coordinate about whom to aim each period's 7 theft tokens at, hence net accumulations might randomly vary over time and one can't rule out *a priori* that it would sometimes be profitable to direct theft tokens at group members revealed by this screen to have a higher accumulation. However, if subjects would otherwise have spread their theft tokens roughly evenly over other group members, the likelihood of random bankruptcies would be quite small.

[5] A concern in experiments of this kind is that after even pre-announced breaks, there can be "restart effects," characterized by surges in cooperation (see e.g. Andreoni 1988). Having breaks in the no communication treatment helps us to distinguish the effect of breaks from the effect of communication.

[6] Having all of a session's groups adhere to the same chat schedule helped to preserve anonymity of group membership.

It is worth noting that the subject numbers by which group members were identified to one another during the chat periods were not the same as those used in the decision rounds, nor were they uniform across group members. Thus, a given subject could be identified as 2 to one group member but as 1 to a second one and 3 or 4 to a third. And there was no way for subjects to link communicators thus identified to them during chat with decisions inferable from the stats screens during rounds of play. A subject might accordingly complain that there is a thief in the group, in his messages, while in fact being the group's main thief. So long as few were deviating, this disconnect between messages and actions would not in itself prevent a group from implementing the strategy "put all in production unless someone steals, and if someone steals use several tokens to steal from that person." Nevertheless, it raised the hurdle faced by verbal interaction, increasing the importance of having a group's discussions lead as nearly as possible to full "buy in" by group members, if they were to have a strong effect.[7]

Predicted impact of communication

We observed above that our subjects faced a social dilemma in that the equilibrium prediction for strictly selfish, rational agents with common knowledge of type entails earnings of 39 wealth tokens each period, whereas earnings of 70 per period are available if all engage in production only. If subjects are unable to refrain from stealing, they might also achieve a socially second-best outcome by each allocating 3 tokens to collective protection, thereby making it individually rational to put 7 rather than 3 tokens into production, and bringing earnings to 64 tokens per period—an efficiency loss of under 9% compared to the loss of 44% in the sub-game perfect equilibrium. But this improvement is also unavailable if strict self-interest, rationality and common knowledge prevail.

Under those same assumptions, opportunities to communicate can help to select better equilibria when the problems at issue are ones of coordination (Farrell and Rabin, 1996), but social dilemmas involve outright conflicts of interest, and no message exchange can change the fact that one is always privately better off allocating more to theft than is socially optimal, taking others' actions as given. So while common intuition may suggest that players of our game who are given the opportunity to communicate will attempt to reach agreement on a non-theft pact, whether aided by the second-best step of allocations to collective protection or going directly for the full social optimum, traditional economic theory says that any such agreements amount to cheap talk only and can have no impact on behaviors.

A considerable body of experimental evidence seems at odds with this conclusion, however. Subjects have often been found to achieve considerably higher levels of cooperation in voluntary contribution experiments (Isaac and Walker, 1988, Brosig, Ockenfels and Weimann, 2003, Bochet, Page and Putterman, 2006) and trust games (Charness and Dufwenberg, 2006, Ben-Ner and Putterman, 2009, Ben-Ner, Putterman and Ren, 2011) when there are opportunities to communicate than when they are

[7] Permitting pre-play communication without having an identifiable link to individual actors' decisions has been common in other experiments, for example Brosig *et al*. (2003) and Bochet *et al*. (2006), where subjects engage in communication, then retain anonymity with respect to their individual decisions. In the present experiment, having fixed identifiers within chat periods helps group members to follow the flow of remarks, while the IDs displayed in interactions themselves allow targeting of theft at suspected thieves, although without identifiable responsibility for statements made in the chat room.

absent. Similar findings are reported in other settings (Sally, 2005; Ellingsen and Johannesson, 2004; Ellingsen *et al*., 2009). The mode of communication also seems to matter, face-to-face communication generally being most powerful (Brosig *et al*. 2003, Bochet *et al*. 2006), free-form written communication that preserves anonymity also being powerful but a little less so (Bochet *et al*., 2006), and choices from a message menu, conveying an action under consideration by choice of numbers, or similarly constrained pre-written message communication, being much less effective (Bochet *et al*., 2006, Charness and Dufwenberg, 2010, Ben-Ner *et al*., 2011). We let our subjects engage in text communication while preserving their anonymity because the method seemed likely to have significant effects, it is easily implemented by the computer interface used in all treatments, and it rules out effects due to ongoing personal relationships beyond the lab.[8]

Because communication has been found to increase cooperation in other experiments, we thought it likely that allocations to production would rise and those to theft fall in **CHAT** and **L-CHAT** relative to **No-CHAT**. We could not, however, predict the size of this impact, and we were also unsure of the degree to which groups would use agreements to allocate tokens to collective protection rather than rely on simpler non-theft pacts, nor could we anticipate the frequency with which explicit plans to punish thieves would be agreed on. Finally, we expected at least a small drop in cooperation in **L-CHAT** as compared to **CHAT**, due to the difficulty of implementing any decision to punish thieves, but were unsure of how big a drop to expect.

Note that even if the possibility of using punishment turns out to be a factor helping propel the higher cooperation and efficiency reached in the **CHAT** (and perhaps **L-CHAT**) treatment(s), a significant positive effect of communication on efficiency would still require at least some suspension of joint assumptions of rationality, self-interest, and common knowledge. In the game's last period of play, that is, it is selfishly rational for each player to devote seven tokens to theft. Since selfish, rational players with common knowledge would know to expect this behavior in period 24, no plan to punish thieves could be made rational to adhere to in period 23, and so forth back to the first period of play. At a minimum the assumption of common knowledge must be abandoned, to derive a prediction that coordinating on punishment strategies, or on an agreement to only produce, can be effective.

However, once one or more of the three traditional assumptions is dropped, the number of potential explanations for cooperation is large. Most subjects could, for example, assign positive probabilities to the existence of conditionally cooperative preferences on the parts of others—preferences such that if others are believed to be desisting from theft, one's utility payoff is higher when also desisting from theft oneself than when stealing, although one's monetary payoff is higher. Even a selfish payoff maximizer may calculate that she will earn more by encouraging others to assign high probability to her being a conditional cooperator for most, potentially up to 23, of the 24 interactions (cf. Kreps *et al.*, 1982). Beliefs in prevalence of aversion to lying may function similarly. Beyond beliefs, some subjects may actually be conditionally cooperative, lying averse, and so on. Some combination of belief in the possible presence, or actual presence, of preferences including aversion to lying (Vanberg,

[8] Although subjects were drawn from the broad student bodies of universities having thousands of students, we cannot rule out that a few in a given session might know one another outside the lab.

2008; Ellingsen *et al*., 2009), guilt aversion argument (Charness and Dufwenberg, 2006), or conditional cooperation (Fischbacher and Gächter, 2010), can explain success of communication and punishment opportunities in inducing cooperation in mutual respect of property.

Subject pools and sessions

For ease of participant recruiting and to have subjects comfortable with the computer interface on which such an experiment must rely, we use as subjects university students recruited from the general student populations of universities at five sites.[9] Ahn *et al*. (forthcoming) selected five sites in culturally and historically diverse regions of the world with the aim of studying difference of response to alternative institutional options. Countries represented are: Austria, Mexico, Mongolia, South Korea, and the United States. While difference of cultural response is not a theme of this paper, we draw on the same subject pools, with experiments conducted at University of Innsbruck (Austria), ITAM (Mexico), Mongolia University of Science and Technology (Mongolia), Korea University (S. Korea) and Brown University (U.S.). [10]

A total of 580 subjects, all university students, participated in sessions taking 90 to 120 minutes including reading of instructions, guided practice for familiarization with the experiment software and the sequence of decisions, and 24 decision rounds. Roughly equal numbers of sessions of each treatment were conducted in each of five countries, with an overall total of 37 subject groups in No-CHAT, 40 in CHAT and 39 in L-CHAT treatment, and a total of 22 to 24 groups all told in each of the five countries.

3. Results

The four quadrants of Figure 1 show the evolution of allocations to production, collective protection, private protection, and theft over the course of the 24 periods, while Table 2 shows the overall average allocation to each activity and, for comparison, the allocation predicted by standard theory. Average allocations to production and theft in **No-CHAT** lie in between the levels that are privately payoff-maximizing assuming common knowledge of the associated preferences, and those that are socially optimal. That is, subjects allocate more than predicted to production and less than predicted to theft. As in public goods experiments, there is also a trend over time away from a more cooperative action—in this case, contributing to collective production—and towards a more private payoff-seeking one—here, theft. A plausible interpretation is that many subjects initially explore the possibility of sustaining cooperation with their group-mates, since such cooperation can be advantageous even to selfish subjects if sustained for many periods and if reciprocation is observed, a sign that others are either genuinely of conditional cooperative disposition (i.e., are not strict

[9] We share critics' concern that university students are not necessarily representative of broader populations. Nevertheless, much of the literature on the topic suggests that student and non-student subject behaviors are similar, with students if anything displaying somewhat less influence of social concerns than older, non-student adults.

[10] No subject participated in more than one treatment of the experiments of either Ahn *et al*. or the present paper. The **No-CHAT** treatment reported in this paper is reported in both the present paper and Ahn *et al*. where it is referred to as **VCP** treatment.

maximizers of own payoff), or that they are mimicking such preferences for as long as this proves beneficial (Reuben and Suetens, 2012; Embrey, Frechette and Yuksel, 2015). Cooperation gradually decays because some are seen to be less cooperative than others, lowering conditional cooperators' expectations of cooperation (Fischbacher, Gächter and Fehr, 2001; Fischbacher and Gächter, 2010). Cooperation in our context is likely to have particular appeal to some due to the normative principle of not stealing, but the typical subject's sense of obligation to adhere to that norm may already be low thanks to the game-like setting, and may quickly wane if she sees others steal, especially from herself.

Private and collective protection allocations are always above the traditional theory prediction—i.e., zero—in **No-CHAT**. Ahn *et al*. discuss several possible explanations for this, including the possibility that subjects who are normatively disinclined to steal engage in constrained optimization: an individual with moral reluctance to steal but expecting others to do so would find it constrained-payoff-maximizing in our design to place some tokens in private protection.[11] Insofar as any tokens are to be used for protection, placing them in collective rather than private protection is more socially efficient, but such allocations are seen to decline rapidly, closely resembling the decay of cooperation in traditional public goods experiments.

Our main focus is on how (if at all) opportunities to discuss the dilemma in a chat room affect levels of cooperation. Figure 1 and Table 2 show that subjects in **CHAT** and **L-CHAT** come far closer to socially optimal behaviors than those in **No-CHAT**. For example, whereas allocations to production hover around half of the available ten tokens in **No-CHAT**, averaging 4.83 overall, they average around 8 of 10 tokens in the chat treatments for the first twenty periods, thereafter declining but retaining a substantial efficiency edge over **No-CHAT** production allocations even in the last period. The average production allocation in **CHAT** is 8.12, a 68% increase over that in **No-CHAT**, making the gap from the efficient allocation of 10 only about a third that which **No-CHAT** subjects display. Conversely, allocations to theft and private protection in **CHAT** and **L-CHAT** average less than half those in **No-CHAT**, except in the last periods, in which the gap narrows but does not close.

Considering outcomes on a group by group basis is illuminating because behaviors can evolve differently in different groups facing the same treatment due to the impacts subjects can have on fellow group members' subsequent decisions. As we illustrate in Section 5, a single opportunistic subject determined to exploit the cooperative inclinations of his group mates may suffice to steer a group which might otherwise have cooperated into cycles of theft and counter-theft. Figure 2, which displays overall outcomes by group and treatment, shows that only in the two treatments with communication were

[11] Up to 40% of subjects, depending on site, invest nothing in theft in the first period, and doing so is significantly correlated with lower theft investments in later periods at the level of the individual subject, consistent with a normative or social preference difference. If others are expected to allocate seven of their ten tokens to theft and direct their theft attempts randomly, a subject disinclined to stealing will lose seventy wealth tokens to theft on average each period if she allocates nothing to protection, hence a token used for private protection prevents an average of seven wealth tokens loss to theft. By comparison to the returns from production shown in Table 1, at least the last four tokens available to allocate each period have a higher return in private protection. Other explanations for allocations to private protection include loss (but not risk) aversion, and investment in protection when others' theft is expected to be asymmetrically aimed at oneself due to past success at stealing from them. Discussion appears in Appendix E of Ahn *et al*., forthcoming.

very high levels of cooperation ever sustained. In this figure, each group is represented by four bars indicating average token allocation to production, collective protection, private protection and theft during its session's 24 periods, with groups ordered from most to least cooperative. No group in the **No-CHAT** treatment allocates more than 7 tokens on average to production, whereas more than half of **CHAT** and **L-CHAT** groups do so. Even in **No-CHAT**, to be sure, allocations are far from the sub-game perfect equilibrium prediction, with the average allocation to theft rarely reaching even half of the 7 tokens that are predicted. Still, the difference in production levels in communication treatments versus **No-CHAT** is dramatic.

Statistically, Mann-Whitney tests using group-level observations find the differences of average production, private protection, and theft, both when comparing **No-CHAT** to **CHAT** and comparing **No-CHAT** to **L-CHAT**, are significant at the 1% level (see Table 3). Average earnings are also significantly higher in the two communication treatments, doubling the share of the potential gain from cooperation that **No-CHAT** subjects achieve.[12] Differences between the two communication treatments themselves are never statistically significant.

For collective protection, the story is slightly different. Similar to private protection, allocations to collective protection are significantly lower in **CHAT** than in **No-CHAT**, suggesting that most groups in the **CHAT** treatment achieved more socially-optimal outcomes not by using the second-best cooperation route of agreed collective protection to signal cooperation at the outset of each interaction period and thus render theft unprofitable, but by the more direct and efficient route of simply desisting from theft despite its private profitability. Allocations to collective protection are more similar to those of **No-CHAT** in the **L-CHAT** treatment. Indeed, in this case the average allocation is higher in **L-CHAT** than in **No-CHAT**. But while that difference is statistically significant at the 5% level, it is economically quite small.

Our most important result thus far is therefore:

*Result 1: There is substantially and statistically significantly more production, higher earnings, less theft, and less allocation of tokens to private protection in both **CHAT** and **L-CHAT** than in **No-CHAT**.*

The fact that opportunities to communicate in a rich verbal space increase cooperation is not a surprise given past results on communication in social dilemmas. However, the magnitude of the improvement, especially in the subject pools of the countries scoring highest on trust and norm abidance—the U.S. and Austria—is surprisingly large.[13] In the U.S., for example, contributions,

[12] Table 2 reports average efficiency gain as the percentage of the prospective 31 wealth token gain achievable by moving to perfect cooperation from the 39 wealth tokens of earnings at sub-game perfect equilibrium. While **No-CHAT** subjects already achieve 36.6% of that potential efficiency gain, on average, those in **CHAT** and **L-CHAT** achieve 77.2% and 70.5% of it, respectively—a result that is all the more remarkable given that there are diminishing returns to production.

[13] Ahn *et al*. report that of the five countries from which our subject pools are drawn, the U.S. and Austria score highest on trust in the World Values Survey and parallel regional surveys using the same survey question, and that theft by experimental subjects in **No-CHAT** (there referred to as **VCP**) and other treatments studied by them is corresponding lower, with the trust measure and within-experiment theft being well aligned across the five countries.

averaging 8.72 tokens per period in **CHAT**, are 88% higher than the 4.63 tokens per period in **No-CHAT**. In Austria, average allocations to theft are 85% lower, at 0.25 per period in **CHAT** versus 1.66 tokens per period in **No-CHAT**. Indications of the pronounced impact of communication include the facts that whereas not a single group among the 37 in **No-CHAT** treatment achieved average allocations of 8 or more out of 10 tokens to production, 25 of 40 groups achieved or exceeded this benchmark in **CHAT** and 18 of 39 groups did so in **L-CHAT**. Indeed, 20 and 14 groups respectively exceeded 9.5 of 10 tokens on average going to production in those treatments. Given diminishing returns to production allocations, the 8 of 10 benchmark secures over 90% of the efficiency gain possible by moving from the private payoff maximizing profile of 3 out of 10 tokens going to production to the socially optimal profile of all 10 being so allocated. Putting 9.5 tokens into production realizes almost 97% of that potential gain.

Another way to put in perspective the magnitude of the jump in cooperation when communication is possible is to compare outcomes in **CHAT** and **L-CHAT** to those in the **VOTE** treatment of Ahn *et al*. In that treatment, if a majority of group members voted to make contributions to collective protection mandatory, the median proposed allocation to collective protection was implemented. With this institution available, it is privately payoff-maximizing to mandate that three tokens be assigned to collective protection by each member, raising expected earnings and achieving 81% of maximum potential efficiency gain.[14] In actuality, however, **VOTE** treatment subjects on average realized only 48.0% of the potential efficiency gain, comparing poorly to this paper's **CHAT** and **L-CHAT** subjects, who achieve 77.2% and 70.5% respectively. Thus, subjects are more successful at building trust and cooperation by means of communication—a device that ought not to affect incentives if individuals remain strictly selfish, rational, and convinced of sharing and having common knowledge of that type—than at achieving efficiency through a binding mechanism requiring only clear-headed selfishly rational choices.

Turning to the differences between the **CHAT** and **L-CHAT** treatments, recall that because cooperation was so high in many groups in **CHAT** treatment, with fully half of the treatment's groups putting 95% or more of tokens into production, we felt it important to check whether this might be due mainly to the availability of a punishment technology rather than to simple inclinations to abide by agreements reached by exchange of messages. We therefore designed and conducted as a robustness test the **L-CHAT** treatment, in which discovering who engaged in theft is considerably more difficult unless thefts are quite rare. How much do outcomes in **CHAT** and **L-CHAT** differ? Figure 1 shows behaviors in the two treatments to be closely parallel. Differences in average allocations to production and to the two forms of protection are consistent with intuition, in that there is usually slightly less production and slightly more allocated to both kinds of protection in **L-CHAT** than in **CHAT**. But none of these differences are statistically significant, according to the Mann-Whitney test results shown in Table 3. Nor is the difference in earnings significant. As for theft, the curves for average allocation in the two treatments cross at several points in Figure 1, and the Mann-Whitney tests again find no statistically significant difference. We conclude:

[14] Still higher efficiency is unattainable on a mandatory basis mainly due to the production sacrificed to achieve maximum collective protection; see Ahn *et al*.

*Result 2: There are no statistically significant differences in allocations between the **CHAT** and **L-CHAT** treatments. Hence, ease of catching and punishing those who steal does not seem to be the main factor at work in achieving high cooperation when communication is available.*

It is important to acknowledge that the lack of significant differences between **CHAT** and **L-CHAT** behaviors does not rule out a role for the joint punishment of thieves. To use the terminology of Ostrom *et al*. (1992), the presence of a "sword" could have lent efficacy to "covenants" in both of these treatments. However, the threat of punishment can function as a deterrent to theft only if a given theft event is expected to stand out, hence only for those who found the verbally made commitments of others to be credible. In fact, as the next section will show, explicitly articulated strategies of punishment are associated with worse, not better outcomes, in our communication treatments.

4. Coded early communication contents help to predict outcomes

Economists sometimes view free-form communication as an unpromising research tool, because exactly how such communication affects decisions is difficult to investigate using quantitative methods. In this section, we discuss how the messages sent by subjects in our experimental chat rooms can be represented by numerical codes, and we report the insights these coded chat contents yield when subjected to econometric analysis.

To keep the exercise manageable, we code only the statements subjects made during their groups' first chat opportunity (which took place before the first allocation decisions and lasted four minutes). Our goal is to see whether rendering their statements into coded form yields statistically significant predictions of how groups would fare in their 24 periods of payoff-relevant interaction. We code for eight aspects of chat content that we conjectured might predict later actions. At the individual level, each takes the value 1 if it holds, 0 if not. The coded contents are:

1. Agreement on a plan. The individual endorses a strategy or plan endorsed by at least two other group members and not disavowed by those three before the chat period ends. (In other words, the individual can be considered part of a majority coalition that has agreed on a strategy.)
2. All to production. 1. holds for the individual AND the strategy in question is that subjects will put all of their tokens into production each period.
3. Collective protection. 1. holds for the individual AND the strategy in question entails subjects putting some of their tokens into collective protection each period. Note that 2. and 3. are mutually exclusive.
4. Punish thieves. At some point during the first chat, the individual states that those who engage in theft can or should be punished by stealing from them, or agrees with another's advocacy of this idea. Unlike 1., 2. and 3., an individual can be coded 1 for this variable without requiring that any other also expressed the idea.
5. Trust. The individual mentions trust or echoes such mention by one or more others. This includes affirmative statements ("I trust you guys"), negative ones ("we can't trust each other"), and statements of doubt ("can we trust each other?").
6. Alternative plan. The individual proposes, raises the possibility of, or agrees with a plan or strategy that was ultimately not embraced by the group's majority in its first chat period. For

example, the individual alone or as one of two or more group members proposes asking all to contribute to collective protection, but the majority ends up rejecting this in favor of an all to production strategy.

7. Profit maximization. The individual explicitly appeals to the idea that an action increases the earnings of all or of the group as a whole.
8. Example 1. The individual explicitly mentions Example 1 of the instructions. Example 1 is one of several examples offered as to how tokens could end up being allocated; in it, all tokens are used for production and each group member accordingly earns 70 wealth tokens. 8. differs from 2. ("all to production") in that the example as such must be referenced by the subject before we set this variable's value to 1. For instance, the subject writes "let's follow example 1".

For each site, two research assistants fluent in the local language independently coded the first round chat of all participants in the **CHAT** and **L-CHAT** treatments, then reconciled any differences to produce the coding on which our analysis is based.[15] We also counted the total number of words each individual wrote during the chat period as an indicator of her level of participation, giving us nine indicators of chat content overall.

We estimate multivariate regressions using both group and individual level data to explore the correlations between chat content and outcomes. To aggregate chat content to group level, we simply add up the number of group members for whom the individual level variable takes value 1.[16] We estimate a separate regression for each activity to which tokens can be allocated (theft, collective protection, private protection, production). The dependent variables in the regressions using group level data are measured as group total allocations to the activity in question averaged over the 24 periods of play. Because non-negligible numbers of groups had no contribution at all to theft, collective protection, or private protection, there are numerous observations at the minimum value of 0, while for production there are values at the maximum value of 10, so we estimate Tobit regressions. We use country fixed effects to net out those differences of outcome attributable to subject pool as such, and also cluster standard errors by country. The results are shown in Table 4.

Five chat items show statistically significant effects in at least two of the regressions. First, groups in which agreement was reached to engage in *collective protection* exhibit more theft, more private protection, and less production, all significant at the 1% level. While the sign on the *collective protection* variable is positive in the regression for collective protection allocations (column (2)), the coefficient is insignificant there. A possible interpretation is that agreements on collective protection were more likely in groups whose members trusted one another less and perhaps correctly anticipated substantial theft allocations. Members of such groups may thus have allocated more to protection—but ultimately more to private than to collective protection, due to the social dilemma logic—and also more to theft, hence less to production.

[15] This is endorsed as standard procedure by Cooper and Kagel, 2005.

[16] Since variables 1., 2. and 3. can have non-zero values for some individuals only if at least three of five group members agree on a plan, the corresponding group level variables can have values 0, 3, 4 or 5 only. For the remaining variables, the group level variable can take any of those values as well as values 1 or 2.

Second, there is a diametrically opposite impact of references to *example 1*, which may be similarly if inversely explained: groups having more such appeals may have been ones having more trusting members, associated behaviorally with less theft and more production (both significant at the 1% level) as well as less private protection (significant at the 10% level). Third, achieving *agreement* as such displays a similarly negative association with theft and production, significant at the 1% level.[17] Fourth, fewer tokens are allocated to collective protection (significant at the 1% level) when there is agreement on *all to production*, which may reflect a straightforward influence of such an agreement, although that variable is also associated with greater allocations to theft.[18] Fifth, when there are more mentions of trust, there is significantly less theft and significantly more production. While not associated with significant coefficients, it is worth noting, finally, that more references to the idea of punishing group members who steal shows no significant effects other than a marginally negative impact on allocations to collective protection.

Result 3: Agreement on a plan by more group members, and more references to trust and to the most cooperative and efficient of the illustrative scenarios in the instructions, are significantly associated with lower allocations to theft and higher allocations to production, while endorsements of plans that include allocation to collective protection and allocations only to production are associated with more theft and (in the case of collective protection agreements) less production, at the group level.

We also investigate whether individuals' allocation decisions, as opposed to average allocations in their groups, are significantly predicted by their own statements in the initial chat stage, holding other group members' statements constant by using group fixed effects. The dependent variables are the proportions of total within-group expenditure on each activity—theft, collective protection, private protection, and production—accounted for by the given individual's allocations, over the 24 periods of interaction as a whole. We find substantial numbers of observations at both the 0 and 1 limits for theft, and substantial numbers at the 0 limit for collective protection and private protection, so we report Tobit regression estimates for those dependent variables, but use linear regressions for production, where our share measure is never at either limit. In addition to group fixed effects, we cluster standard errors by country.

Table 5 displays the relevant regressions. No single variable is as successful in explaining outcomes at the individual level as was the case with the group level regressions. Nonetheless, of the 36 coefficients estimated, four are statistically significant at the 1% level, three at the 5% level, and four at the 10% level. The most successful variable in terms of statistical significance across specifications is the proposal to punish thieves: those making or endorsing that proposal account for a smaller share of their groups' private protection allocations, significant at the 1% level, and for a larger share of their

[17] We note that *collective protection* takes positive values only if there is some *agreement*, hence the net effect of agreements on collective protection, measured by the sum of those coefficients, is still favorable to production and unfavorable to theft. The coefficients on collective protection should be interpreted as measures of the difference in effect of an agreement that includes collective protection versus one that does not.

[18] As with agreements on *collective protection*, the net effect on theft of agreements remains positive, though smaller (the net effect can be approximated as -1.304 + 0.830 = -0.474 provided that the numbers of group members joining agreements in general and agreements of the "all tokens to production" variety are approximately the same).

groups' theft allocations, significant at the 5% level. The most straightforward explanation is that advocates of using theft to punish those who steal were indeed active in doing so, and that their more retaliatory stances also manifested as less protective strategies.

Two variables are significant at the 1% level in one regression and at the 10% level in a second one. First, if there is *agreement* on a strategy of any kind by at least three group members, those who explicitly supported the agreement in the first chat account for significantly more of their group's allocations to collective protection and to (marginally) significantly less of the allocations to private protection; 'committing' verbally thus seems to be associated with more cooperation on the second-order social dilemma of using collective protection, if not in the first-order dilemma of refraining from theft and engaging in production. Second, when the agreement is on "all to production," those committing to it account for significantly less of their group's allocations to private protection—reinforcing the effect of the *agreement* variable—and for engaging marginally more in production. Relatedly, perhaps, when the agreement entails *collective protection*, the individuals committing to it account for significantly less of the group's theft allocations.[19] Finally, those referencing *example 1* engage in somewhat more stealing (significant at the 5% level) and marginally more private protection than others, while those contributing more words to the chat interaction contribute significantly less than others to collective protection.[20]

We summarize that choice of message is significantly associated with relative allocations within groups as follows:

Result 4: Individual group members who express support for agreements including ones requiring that all tokens be allocated to production or that tokens should go to collective protection show signs of actual cooperation in the form of more allocation to collective protection, less allocation to private protection, and less allocation to theft, relative to other group members. References to an efficient example and relative activeness in chat are associated with relatively uncooperative behaviors—more allocation to theft and less allocation to collective protection, respectively.

To conclude this section, coded chat contents show some significant associations with subject behaviors, and some of these associations—e.g., the associations at group level between having agreements and mentioning trust, on the one hand, and engaging in production and avoiding theft, on the other—are intuitive. Others seem rather less so, and we see little reason to expect them to be replicated if the study were repeated with other participants. While communication was resoundingly successful in allowing many groups to solve the social dilemma our experiment posed to them, our

[19] It need not be contradictory that groups having a majority agreeing to engage in collective protection exhibit more theft (Table 4) whereas the specific individuals who joined in the agreement engage in less of it. Such a pattern would be consistent with the interpretation above (that including collective protection in the agreement is more likely when there is less trust) along with a disproportionate role in stealing by those not expressing support for the agreement.

[20] These last two results may be related to the kinds of cases, illustrated in Section 5, in which a group member seeming to be the most active promoter of cooperation in the chat room is the least cooperative group member in terms of actions.

simple attempt to represent chat content with codable indicators displays modest success, only, in explaining which groups succeeded and which individuals within groups were most cooperative.

5. A look at some illustrative groups

An alternative way to try to "break open" the opaque (if not black) box of communication is to look at the messages and behaviors over time of subjects group by group, paying attention to which subjects wrote what messages and how that correlates with the costly decisions that they and other group members made. A few examples help provide a sense of how some groups were able to achieve quite high cooperation with the aid of communication, others less so. We briefly discuss four groups that provide interesting examples as well as contrasts to one another.

Quite a few groups achieved nearly complete cooperation in the treatments with chat. In all, 3 and 7 groups achieved perfect cooperation, using every token for production and no token for theft, in the **CHAT** and **L-CHAT** treatments, respectively. Another 19 and 10 groups, respectively, had average contribution levels of at least 9 (but less than 10) tokens. No group in No-CHAT achieved either benchmark. We select one group to illustrate achievement of sustained high cooperation, with a second group illustrating success at achieving high cooperation for part (about half) but not all of its session. We then discuss two groups that struggled and did only modestly better or no better than average groups without communication opportunities. Their examples help to illustrate the kinds of interactions between group members that are associated with such struggles.

Near-perfect cooperation

The illustrative high cooperation group tracked by Figure 3a is taken from the **CHAT** treatment and the U.S. subject pool.[21] As the figure shows, its average allocation to production began above 9 and reached the perfect cooperation level of 10 in period 4, staying there without exception through period 23 before experiencing a small end-game drop due to one member's deviation. Of the 13 out of 1200 total effort tokens available that were not used for production, 4 went to collective and 9 to private protection.

Initial chat remarks indicate some differences among the group members with respect to appreciation of the dilemma and anticipated behaviors. The differences in first messages correlate nicely with differences in later observed behaviors. In particular, the two who ultimately devoted every token to production, Subject 10 and 17, appeared most certain of the desirability of an all-to-production strategy from the outset.[22] In contrast, Subject 1, who ended up allocating tokens to private or collective protection during each of the first three periods, initially advocated for each putting three tokens into collective protection. Subject 10 responded that "we don't need protection if nobody steals" then led the others into agreement with the remarks "listen, nobody steal, devote everything to

[21] Note that because the present paper does not focus on the cross-country differences emphasized by Ahn *et al*., the groups selected are not intended to be representative of behaviors in their country, per se.
[22] To differentiate the subjects, we use the same subject numbers as were randomly assigned by the computer to participants in each session.

production" and "but if someone steals, then everyone steal from that person" and finally "everything to production every time".

Remaining chat periods in this group were mostly devoted to the members congratulating themselves on their success in cooperating, and to making small talk. During the final chat period, the two leaders made efforts to head off final defections, with Subject 10 writing "so, last round, lets finish off strong" and Subject 17 writing "I just want to say that you all did a great job out there today".

An amazing recovery

A remarkable case in which high cooperation was achieved following a string of failures occurred in an **L-CHAT** groups in Mongolia. After a struggling first phase and an effort to cooperate that ended disastrously in the second, the group's members achieved almost perfect efficiency for three of the six phases. One member resumed stealing at the end of Phase 5, and the final phase then saw cooperation decline toward its previous low levels. Figure 3b tells the story.

Interestingly, the chat that preceded the group's remarkable Phase 3 recovery does not look very different in content from the one before phase 2. Evidently, the large Phase 2 theft outbreak and heavy punishment received by the group's biggest thief (Subject 12) caused the pledges of cooperation this time to be matched by more truthful intentions. The third phase began with all putting 10 into production except Subject 12, who put 3 into private protection. Thereafter, each of the five group members put all 10 of their tokens into production for ten consecutive periods, with 4 subjects doing so for another period as well, allowing all to earn more than in the initial phases.

The decision record suggests that the group had two members strongly inclined towards cooperation, one of whom played a leadership role in this direction during chat, also being aggressive in retaliation when stolen from. The other was quieter in the chat room, but never engaged in theft. The other three members expressed doubt from the outset that theft could be avoided, and each played a role in starting off rounds of theft—Subject 12 in Phase 1, joined by Subject 10 in Phase 2, and Subject 16 at the end of Phase 5 and in Phase 6. The presence of these opportunistic members clearly helps explain the difference between this group's and the previous group's outcomes, overall.

Good, bad, and in between

For our third illustration, shown in Figure 3c, we consider an **L-CHAT** treatment group from Austria that achieved only modestly greater cooperation and earnings with communication than did most groups in **No-CHAT**. As with the previous group, this group nicely illustrates behavioral heterogeneity among group members. In this case, a single member's decision to ignore the chat room agreement to cooperate may account for the difference from the several highly cooperative groups at the same site.

In their first chat period, group members followed a frequently seen pattern of agreeing that allocating all tokens to production would be ideal. It seems possible this might have worked and that the group might have enjoyed relatively cooperative outcomes, since three of the members in fact allocated all tokens to production in period 1, while a fourth allocated six to production and four to private

protection. However, one member—Subject 9—provocatively allocated all of his tokens to theft in period 1, then allocated 5, 0, and 10 to theft in the remaining three periods of the phase. By that time, two other subjects were engaging in theft at more modest levels (2 tokens and 1 token, respectively). Tellingly, Subject 9 had been the only group member to express doubt about the all production strategy in the initial chat, writing "With 5 people, there is always someone that deviates".

For the remainder of their session, the group seems to be struggling to agree on and implement a strategy that includes a mix of private and collective protection. The decision record shows the level of collective protection in phases 2 through 6 to be ranging between 12% and 48%, with an average of 30.6%, while the average group member is investing an additional 1.8 tokens into private protection, yielding average protection levels of 51%. Theft also continues, however, most often initiated and invested in most amply by Subject 9. While 9's earnings ended up exceeding the group average by 6%, even he earned 9% below the social optimum.

"Do not think like Mexicans"

We lastly discuss a group in the **CHAT** treatment in Mexico—see Figure 3d—that resembles the Austrian group in that group members initially agree on all-to-production, one member deviates by stealing at the outset, and the conversation then shifts to coordinating on collective protection. This group's average allocation to production ends up at 5.8 tokens, slightly below the 6.1 tokens in the Austrian group. Note that while both groups are underperformers in their national subject pools for the respective treatments, they both nonetheless manage to use the chat opportunity to some advantage (the average group in **No-CHAT** put a little under 5.5 tokens into production in Austria and 4.4 tokens into production in Mexico).

The group's "bad boy" was Subject 1. When three of the five members were following the agreement to put all ten tokens into production in round 1 with the fourth deviating by putting three into private protection, Subject 1 put 3 tokens into theft. Two others then quickly join him in stealing, and by the end of the first phase all group members have stolen at least once. Similar to the Austrian group's Subject 9, Subject 1 of the Mexican group tends to "set up" the others in his society in order to plunder them. For example, after the all-to-production agreement falls apart in Phase 1 beginning with Subject 1's round 1 theft, Subject 12 writes "what happened to our strategy, huh?" and Subject 1 pretends outrage, writing "very bad" and "didn't we say 'no stealing' and stuff?" When the group's members debate whether to use private or collective protection to counter theft, Subject 1 comes down strongly for the latter: "public protection is the way to go". When the agreement to allocate to collective protection unravels in the following phase, with Subject 1 among the first to ignore it, he writes "what happened with public protection?" and amusingly "Mexicans you are …the best countries, like the USA…they all have money…because nobody steals…and everybody looks after each other".

Members of this group do not seem entirely incapable of cooperating. After the average level of collective protection falls to only 4.5% in each of phases 3 and 4 from 22.5% in Phase 2, they manage to revive collective protection to an average of 24% in Phase 5, then congratulate themselves and encourage still more collective protection during their final chat. Subject 6, the original booster of collective protection, writes "did you see the difference?" and Subject 19 responds "it turned out

better". Subject 1 continues to hide behind his chat room anonymity, pretending to be one of the cooperators: "I keep putting in the public account" "…but nobody else does" "…I only lose ☹". Other group members nevertheless manage to raise the collective protection level to 48% and average tokens to production rise correspondingly to 6.4 in rounds 21 to 23 before a mild last period drop in protection and rise in theft.

6. Conclusions

We conducted laboratory experiments in which subjects can generate wealth through a private production technology but face a dilemma because allocating most of their effort tokens to the unproductive wealth-transferring activity of theft rather than production is privately payoff-maximizing under traditional assumptions. In our **No-CHAT** treatment, we find that levels of production and theft are almost always substantially higher and lower, respectively, than predicted by models of strictly selfish agents choosing rationally under common knowledge of type. However, substantial numbers of tokens are diverted in that treatment to equally unproductive activities of property protection, the bulk of it in the socially inferior form of private rather than collective protection. In all, a little over a third of the potential efficiency gain from theft avoidance is achieved in that treatment.

Our paper focuses on the effects of introducing in this setting opportunities for text communication immediately before and at periodic breaks between sets of interaction periods. According to traditional economic theory, these opportunities do not alter incentives and hence should leave equilibrium play unaffected. In the event, compared to **No-CHAT**, we find that adding communication approximately doubles the proportion of the maximum efficiency gain that subjects attain (where potential efficiency gain is defined as the wealth increase associated with a shift from the sub-game perfect equilibrium to the social optimum). In particular, allocations both to theft and to protective activities fall by well over half their levels in the no communication baseline, when there is communication, and allocations to production rise by a little under two-thirds of their baseline level.

Because the increase in cooperation associated with text communication in our experiment is so pronounced, we also investigated whether it might mainly be explained by availability of a way to punish violators of anti-theft agreements by banding together to steal from them. We did this by implementing a second communication treatment with reduced information about others' stealing. The added treatment results in slightly lower but statistically indistinguishable cooperation levels, suggesting that trust plays at least as large and arguably a larger role than punishment opportunities in explaining the impact of communication. Econometric analysis using coded chat content also fails to show a relationship between proposing to use theft to punish thieves, and more cooperative outcomes.

At a more general level, it appears that communication leads to higher levels of cooperation because some participants are able to identify and express commitment to cooperative plans of action, because they find one another's expressions of commitment sufficiently credible to justify conditional adherence to those plans, and also perhaps because many group members dislike going back on their word if others don't do so, dislike behaving selfishly when others cooperate, or both. The evidence includes estimated regressions using coded indicators of chat content, and four brief case studies of group dynamics. Our regressions confirm that groups in which the majority of members agreed on a

plan of action, in which fewer members advocated for collective protection, in which more addressed the issue of trust and more referenced an example corresponding to the efficient allocation, achieved significantly more efficient outcomes. Our analysis of chat messages and behaviors in the illustrative groups suggested that groups of individuals who were sufficiently trusting and who expressed commitment to non-theft agreements with straightforward intentions could reach agreements that they then adhered to, while less successful groups seem to contain one or more opportunistic individuals who use the chat periods and the anonymity afforded by our design to repeatedly set up fellow group members as targets of theft. Even these groups, however, seem able to benefit from the opportunity to communicate, since subsets of their members often succeed in cooperating with one another.

Acknowledgements: We are grateful to Iñaki Arbeloa, Nicholas Deak, Cheewadhanaraks Matanaporn, Nicole Chen, and Ben Chowdhury for their assistance with the organization and analysis of our data and results, Karla Henning for assistance with literature review, and Arjun Bansal for programming all experimental treatments. We thank the U.S. National Science Foundation for funding the experiments reported in this paper with grant SES-0921733. We also wish to thank the following, who assisted with conducting the experiments: Jorge Tarraso in Mexico City, Benjamin Furlan in Innsbruck, Moon-Sun Kang, Sang-Hoon Ahn and Namun Cho in Seoul, and Dulamzaya Batjargal, Munkherdene Gochoo and especially Amarsanaa Dashdavaa in Ulaanbaatar.

References

Abbink, Klaus, Bernd Irlenbusch and Elke Renner, 2000, "The Moonlighting Game: An Empirical Study on Reciprocity and Retribution," *Journal of Economic Behavior and Organization,* 42: 265-77.

Ahn, T.K., Loukas Balafoutas, Mongoljin Batsaikhan, Francisco Campos-Ortiz, Louis Putterman, Matthias Sutter, forthcoming, "Securing Property Rights: A Dilemma Experiment in Austria, Mexico, Mongolia, South Korea and the United States," *Journal of Public Economics* (in press).

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

Ben-Ner, Avner and Louis Putterman, 2009, "Trust, Communication and Contracts: An Experiment," *Journal of Economic Behavior and Organization* 70 (1-2): 106-121, 2009.

Ben-Ner, Avner, Louis Putterman and Ting Ren, 2011, "Lavish Returns on Cheap Talk: Non-binding Communication in a Trust Experiment," *Journal of Socio-Economics* 40 (1): 1 – 13, 2011.

Bochet, Olivier, Talbot Page and Louis Putterman, 2006, "Communication and Punishment in Voluntary Contribution Experiments," *Journal of Economic Behavior and Organization* 60: 11-26.

Brosig, Jeannette, Axel Ockenfels and Joachim Weimann, 2003, "The Effect of Communication Media on Cooperation," *German Economic Review* 4: 217-242.

Camerer, Colin and Ernst Fehr, 2004, "Measuring Social Norms and Preferences using Experimental Games: A Guide for Social Scientists," pp. 55 – 95 in J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr and H. Gintis, eds., *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford University Press.

Charness, Gary and Martin Dufwenberg, 2006, "Promises and Partnerships," *Econometrica* 74 (6): 1579-1601.

Charness, Gary and Martin Dufwenberg, 2010, "Bare Promises," *Economics Letters* 107 (2): 281-83.

Cooper, David and John Kagel, 2005, "Are Two Heads Better Than One? Team Versus Individual Play in Signaling Games," *American Economic Review* 95: 477 – 509.

Cooper, David and John Kagel, forthcoming, "Other-Regarding Preferences: A Selective Survey of Experimental Results," in J. Kagel and A. Roth, eds., *Handbook of Experimental Economics, Vol. 2*. Princeton University Press.

Duffy, John and Minseong Kim, 2005, "Anarchy in the Laboratory (and the Role of the State)," *Journal of Economic Behavior and Organization*, 56: 297-329.

Durham, Yvonne, Jack Hirshleifer and Vernon L. Smith, 1998, "Do the Rich Get Richer and the Poor Poorer? Experimental Tests of a Model of Power." *American Economic Review* 88: 970-83.

Ellingsen, Tore, Magnus Johanneson, Jannie Lilja and Henrik Zetterqvist, 2009, "Trust and Truth," *Economic Journal* 119 (534): 252 – 276.

Ellingsen, Tore and Magnus Johannesson, 2004, "Promises, Threats and Fairness," *Economic Journal* 114: 397-420.

Embrey, Matthew, Guillaume Frechette and Sevgi Yuksel, 2015, "Cooperation in the Finitely Repeated Prisoner's Dilemma," Working Paper, NYU Dept. of Economics.

Farrell, Joseph and Matthew Rabin, 1996, "Cheap Talk," *Journal of Economic Perspectives* 10 (3): 103 – 118.

Fischbacher, Urs, Simon Gächter, and Ernst Fehr, 2001, "Are people conditionally cooperative? Evidence from a public goods experiment," *Economics Letters*, 71: 397-404.

Fischbacher and Gächter, 2010, "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments," *American Economic Review* 100: 541 – 556.

Grossman, Herschel and Minseong Kim 1995, "Swords of Plowshares? A Theory of the Security of Claims to Property," *Journal of Political Economy,* 103: 1275-88.

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

Janssen, Marco, Robert Holahan, Allen Lee and Elinor Ostrom, 2010, "Lab Experiments for the Study of Social-Ecological Systems," *Science* 328, 30 April, 613 – 617.

Kimbrough, Erik, Vernon Smith and Bart Wilson, 2010, "Exchange, Theft, and the Social Formation of Property," *Journal of Economic Behavior and Organization,* 74: 206-229.

Kreps, D., Milgrom, P., Roberts, J. and Wilson, R., 1982, "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory*, vol. 27(2), pp. 245-52.

List, John, 2007, "On the Interpretation of Giving in Dictator Games," *Journal of Political Economy* 115 (3): 482 – 493.

Ostrom, Elinor, James Walker and Roy Gardner. 1992, "Covenants with and without a Sword: Self Governance is Possible," *American Political Science Review* 86(2): 404-416.

Ostrom, Elinor, 2010, "Beyond Markets and States: Polycentric Governance of Complex Economic Systems," *American Economic Review* 100 (3): 641 – 672.

Reuben, Ernesto and Sigrid Suetens, 2012, “Revisiting Strategic versus Non-Strategic Cooperation,” *Experimental Economics* 15: 24 – 43.

Sally, David, 1995, “Conversation and cooperation in social dilemmas: a meta-analysis of experiments from 1958 to 1992,” *Rationality and Society* 7, 58-92.
Sobel, Joel, 2005, “Interdependent Preferences and Reciprocity,” *Journal of Economic Literature* 43: 392 – 436.

Taylor, John, 1993, “The Ethical Foundations of the Market,” in Vincent Ostrom, David Feeny and Hartmut Picht, eds., *Rethinking Institutional Analysis and Development: Issues, Alternatives, and Choices*. San Francisco: Institute for Contemporary Studies Press.

Vanberg, Christoph, 2008, “Why Do People Keep Their Promises? An Experimental Test of Two Explanations,” *Econometrica* 76 (6): 1467 – 1480.

Van Lange, Paul, Bettina Rockenbach and Toshio Yamagishi, eds., 2014, *Reward and Punishment in Social Dilemmas*. New York: Oxford University Press.

Figures and Tables

Table 1. Wealth production schedule

# Effort Tokens	# Wealth Tokens Produced
1	15
2	28
3	39
4	48
5	55
6	60
7	64
8	67
9	69
10	70

Table 2. Predicted and actual average token allocations and outcomes per period by treatment

	Collective Protection	Production	Private Protection	Theft	Earnings	% of Max. Efficiency Gain*
Prediction[+]	0	3	0	7	39	0%
No-CHAT[+]	0.43	4.83	2.74	2.01	50.35	36.6%
CHAT	0.20	8.12	0.82	0.87	62.93	77.2%
L-CHAT	0.45	7.64	1.04	0.87	60.85	70.5%

*Percentage of maximum efficiency gain is the fraction of the 31 wealth token difference between earnings predicted in all three treatments (39) and socially optimal earnings (70). [+] Data from Ahn *et al.*, forthcoming.
[+] Predicted allocations in finitely repeated play on assumption that all group members are rational maximizers of own payoff and have common knowledge of this, as demonstrated in Ahn *et al.*, forthcoming.

Table 3. *p*-values of Mann-Whitney tests of difference in allocations across treatments

Collective protection	**CHAT**	**L-CHAT**
No-CHAT	< 0.01	0.03
CHAT	-	0.13
Production	**CHAT**	**L-CHAT**
No-CHAT	< 0.01	< 0.01
CHAT	-	0.38
Private protection	**CHAT**	**L-CHAT**
No-CHAT	< 0.01	< 0.01
CHAT	-	0.36
Theft	**CHAT**	**L-CHAT**
No-CHAT	< 0.01	< 0.01
CHAT	-	0.90
Earnings per period	**CHAT**	**L-CHAT**
NO-chat	< 0.01	< 0.01
CHAT	-	0.41

Note: Tests use values of group-level averages. Hence, the numbers of observations are the same as the numbers of groups, namely 37 in No-CHAT, 40 in CHAT, and 39 in L-CHAT.

Table 4. Regressions Exploring Chat Content Impacts with Group Level Data

Model	(1) Tobit	(2) Tobit	(3) Tobit	(4) Tobit
Dependent variable	**Theft**	**Collective protection**	**Private protection**	**Production**
agreement	-1.304 *** (0.281)	-0.123 (0.392)	-0.644 (0.488)	2.006 *** (0.662)
all tokens to production	0.830** (0.416)	-0.503 *** (0.129)	-0.006 (0.475)	-0.405 (0.424)
punish thieves	0.150 (0.320)	-0.233 * (0.118)	0.319 (0.594)	-0.291 (0.469)
alternative plan	0.529 (0.366)	-0.024 (0.033)	0.175 (0.290)	-0.609 (0.545)
trust	-1.194 *** (0.393)	-0.222 (0.282)	-1.022 (1.026)	1.947 ** (0.948)
collective protection	0.538 *** (0.186)	0.507 (0.324)	0.610 *** (0.206)	-1.422 *** (0.401)
profit maximization	-0.049 (0.696)	0.104 (0.154)	-0.300 (1.144)	0.210 (1.502)
example 1	-1.247 *** (0.373)	-0.527 (0.473)	-1.796 * (0.908)	2.513 *** (0.585)
total number of words	0.002 (0.004)	-0.000 (0.001)	0.007 (0.007)	-0.008 (0.009)
N	77	77	77	77
Number of censored observations	17	20	20	10

Notes:

1. The dependent variables refer to total group allocations towards a given activity per period on average (i.e., total group allocations during the experiment divided by 24)
2. Standard errors (shown in parentheses) are clustered by country.
3. Country fixed effects in all specifications.
4. Although there were a combined total of 79 groups in the **CHAT** and **L-CHAT** treatments, it proved impossible to match chat contents to behaviors due to a computer problem affecting two groups at the Austrian site. This brings the number of includable groups down to 77.

Table 5. Regressions Exploring Chat Content Impacts with Individual Level Data

Model	(1) Tobit	(2) Tobit	(3) Tobit	(4) Linear Regression
Dependent variable	**Theft share**	**Coll. prot. Share**	**Private prot. share**	**Production share**
agreement	-0.062 (0.046)	0.132 *** (0.044)	-0.149 * (0.083)	-0.003 (0.007)
all tokens to production	-0.071 (0.062)	0.155 (0.149)	-0.116 *** (0.031)	0.027 * (0.012)
punish thieves	0.199 ** (0.091)	-0.099 (0.130)	-0.112 *** (0.039)	-0.008 (0.006)
alternative plan	-0.056 * (0.031)	0.130 (0.095)	0.049 * (0.026)	0.001 (0.003)
trust	0.018 (0.056)	0.058 (0.068)	0.033 (0.021)	-0.003 (0.004)
collective protection	-0.108 *** (0.030)	-0.020 (0.095)	0.080 (0.049)	0.007 (0.007)
profit maximization	0.021 (0.056)	0.059 (0.117)	0.025 (0.064)	0.002 (0.004)
example 1	0.215 ** (0.104)	0.267 (0.191)	0.076 * (0.039)	-0.001 (0.001)
total number of words	-0.000 (0.000)	-0.001 ** (0.000)	-0.000 (0.000)	0.000 (0.000)
N	280	265	265	365
Number of censored observations	87	96	61	n/a

Notes:

1. The dependent variables refer to the proportion of total group allocations towards a given activity accounted for by an individual (i.e., total allocations by *i* divided by total allocations during the experiment by all in *i*'s group, *i* included).
2. Standard errors (shown in parentheses) are clustered by country.
3. Group fixed effects in all specifications.
4. Sample size is limited by availability of data matches for 73 groups only (see Table 4, note 4.), yielding a maximum of 365 observations, as seen in column (4). Since the dependent variable would have a denominator of zero, 17 groups that saw no tokens allocated to theft, 20 groups that saw no tokens

allocated to collective protection, and 20 groups that saw no tokens allocated to private protection, are excluded from the analysis of columns (1), (2) and (3) respectively.

5. If column (4) is estimated using the Tobit command, it yields exactly the same coefficients but different standard errors, so that the coefficient on *all tokens to production* attains significance at the 1% level. Because there are no censored observations of the dependent variable, however, we show a linear regression estimate.

Figure 1: Average token allocations by period and treatment

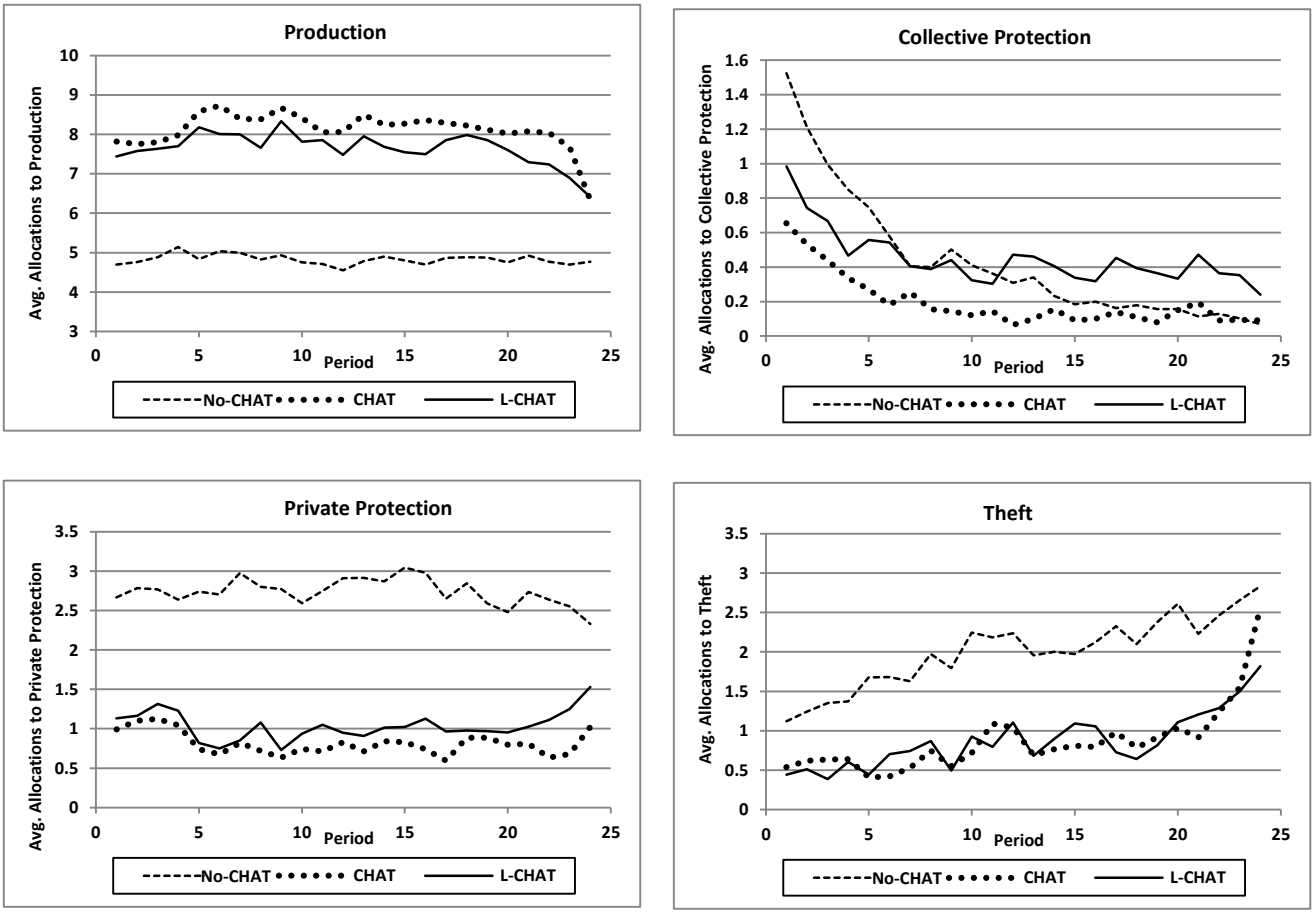
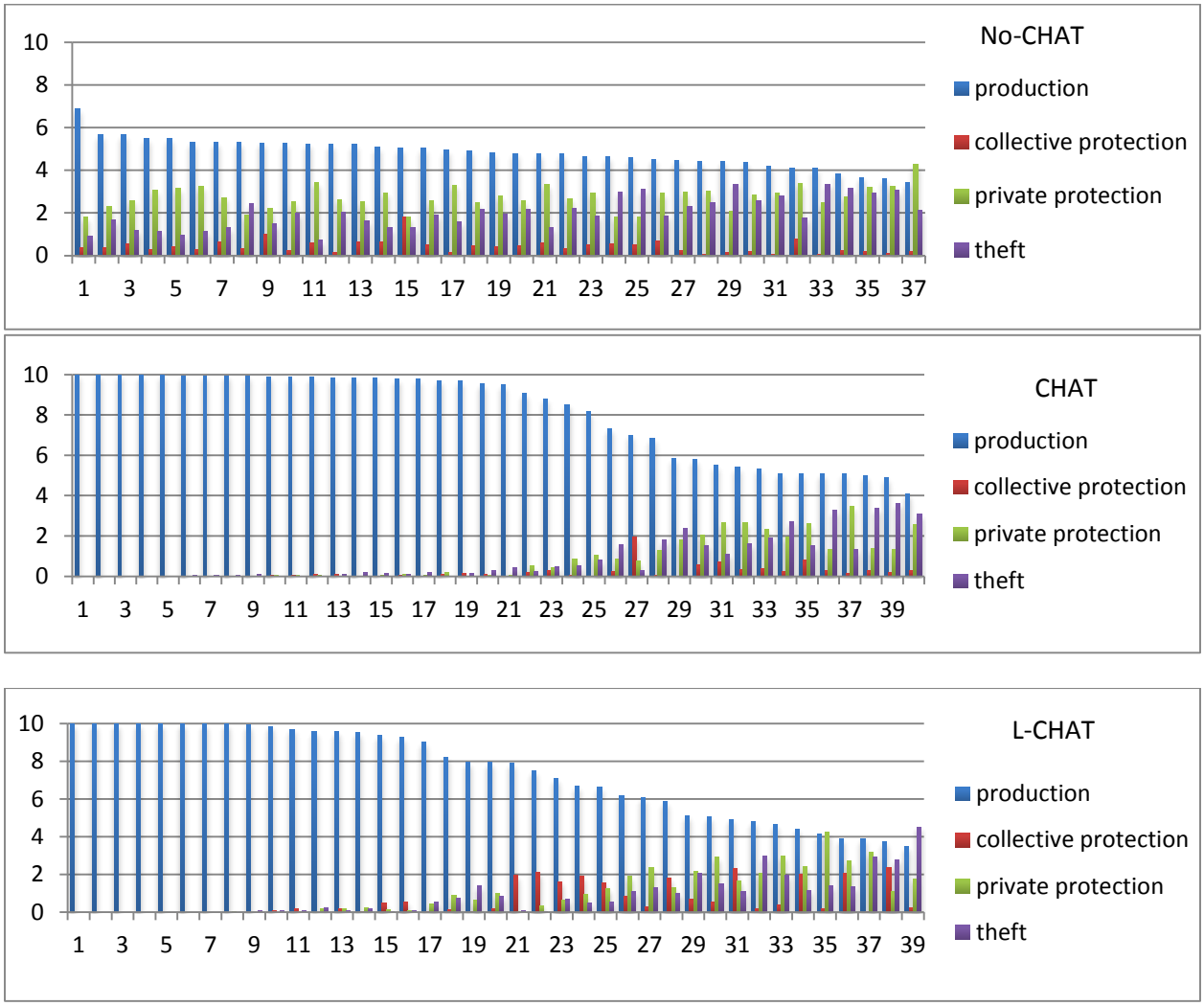


Figure 2. Average token allocations by group and treatment, ordered by efficiency



Note: Within each treatment, groups are ordered from those with highest average allocation to production to those with lowest average allocation to production, breaking ties through ordering by average allocations to collective protection, then to private protection, and finally to theft.

Figure 3a

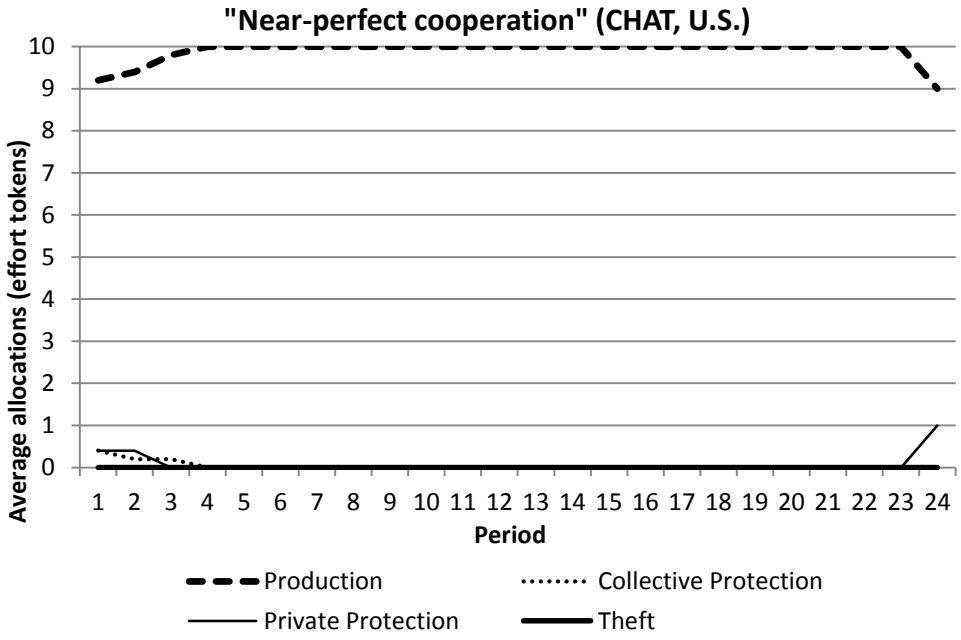


Figure 3b

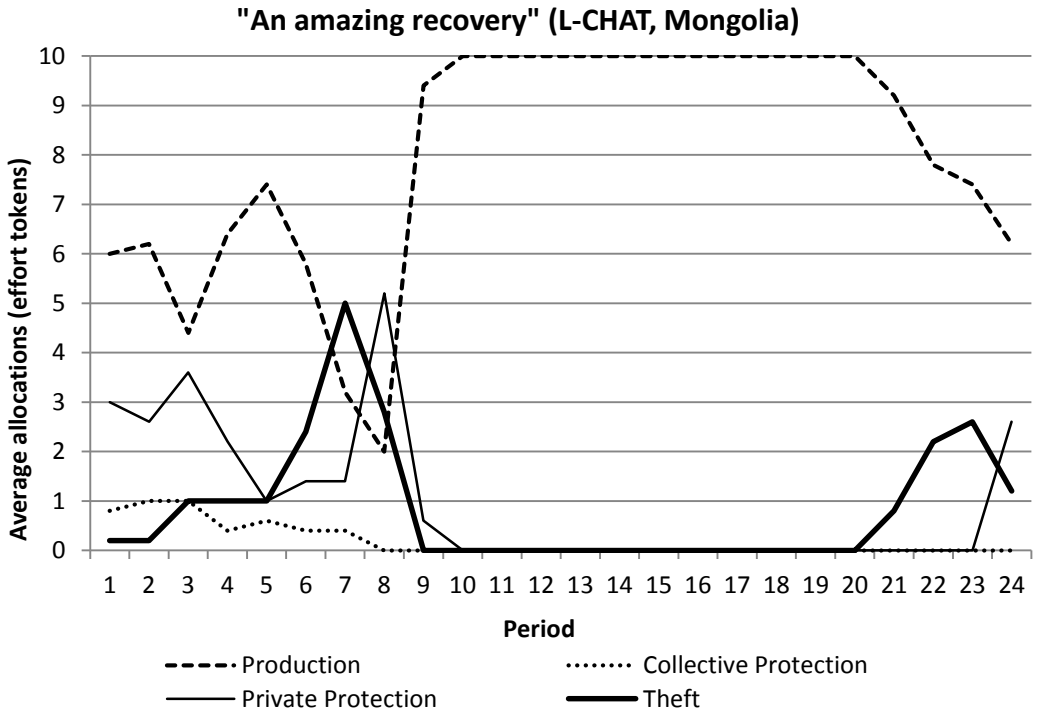


Figure 3c

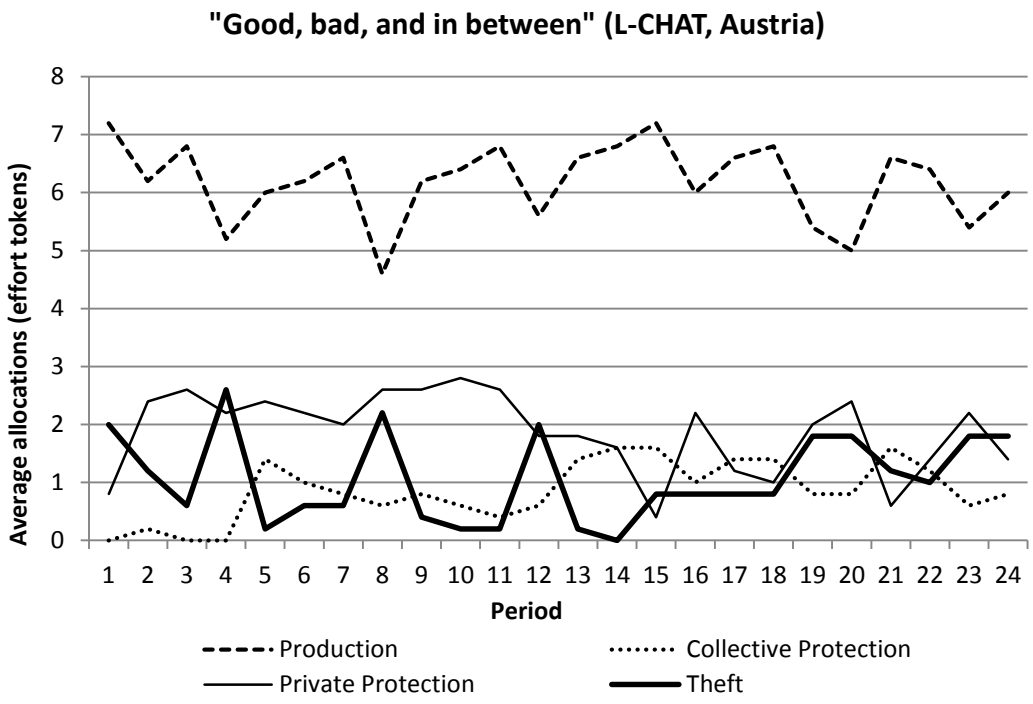
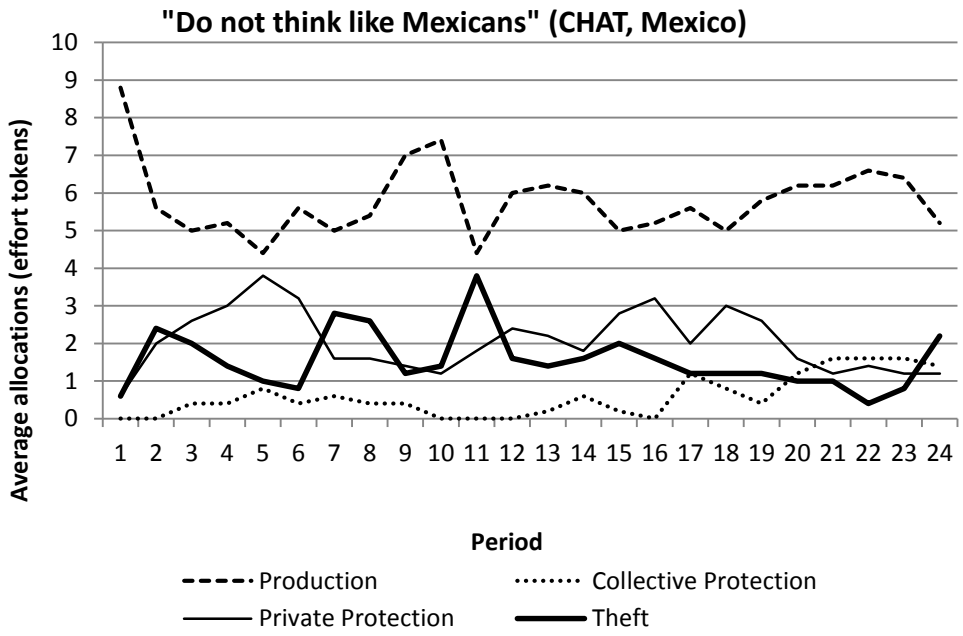



Figure 3d