



**The Coevolution of
Behavior and Normative
Expectations**

Customary Law in the Lab

Christoph Engel
Michael Kurschilgen





The Coevolution of Behavior and Normative Expectations

Customary Law in the Lab

Christoph Engel / Michael Kurschilgen

December 2011

The Coevolution of Behavior and Normative Expectations Customary Law in the Lab

Abstract: Customary law has been criticized from very different angles. Rational choice theorists claim that what looks like custom is nothing but self-interest. Positivists doubt that anything beyond consent assumes the force of law. In this paper, we adopt an experimental approach to test these claims. We show that the willingness to overcome a dilemma transcends self-interest. Cooperation is significantly higher in the presence of a meta-rule for the formation of customary law. Yet only if it is backed up by sanctions, law is significantly more effective than mere comity. Customary law guides behaviour into the normatively desired direction as normative expectations and behavioural patterns coevolve.

Keywords: Customary Law, Normativity, Crowding Out, Public Good, Experiment

JEL: C14, C91, D03, D62, D63, H41, K10

I. Introduction

Conceptual rigour can be a dangerous weapon. From several directions, this weapon has been directed to customary law, with lethal effect as aggressors pretend, or leaving the target badly wounded but able to recover as defenders believe. This article objects that neither position gives customary law the credit it deserves. It uses a laboratory experiment to show the power of customary law, and the behavioural forces driving it, thereby contributing to the nascent experimental law and economics literature (characteristic contributions include Croson and Johnston 2000; Arlen, Spitzer et al. 2002; Loewenstein and Moore 2004; McAdams and Nadler 2005; Nadler and Seidman Diamond 2008; Grechenig, Nicklisch et al. 2010; Zeiler 2010).

In almost all legal orders of the world, customary law is acknowledged as a valid source of law. Law need not originate in legislation or precedent, but may be created by the behaviour and the will of those supposed to abide by it. Yet in most national legal orders, customary law in the strict sense has become rare. Within nation states, if there is need for a new rule, alternative law making procedures are readily available. Society need not wait until custom has formed. Since the alternative sources of law rely on an explicit decision in a formalized procedure, usually there is also less debate about the precise contents of the rule. Substitutes for customary law are less easily employed in the dealings of sovereign states with each other. The main reason for this difference is, of course, the absence of a sovereign ruler who could ordain reluctant states to subdue to the common will or good. This explains why, these days, most of the legal debate on customary law is conducted by international lawyers.

In this paper, we use a lab experiment to test the main claims made in the legal debate over customary law. This creates an obvious tension. The debate originates in a field of law where entire states are the main actors. We test these claims with individuals. Since it is impossible to bring states to the lab, or to engage them in a field experiment, we have to trade off a loss in context specificity for a gain in experimental control and thus causal inference. We do of course not mean to claim that states essentially behave the same way as individuals. States are highly aggregate, institutionally and historically embedded corporate actors. We abstract from all of this. Yet we believe this price for experimental control to be justified for three reasons: First, the legal debate does not posit that customary international law is conceptually different from customary law in national law. Insights that are valid for customary national law, which does engage and address individuals, are therefore in principle also valid for international law. Second, the arguments brought forward in the international law debate over customary law, which we review in the next section, do not rely on the character of states as corporate actors, but invoke mechanisms that hold for any actor, and for individuals as well. Third, a very similar legal debate is indeed concerned with the normative force of custom on the behaviour of individuals. While this is not

customary law in the strict sense, there are many instances where the law acknowledges the normative relevance of custom, short of assuming the force of a source of law. Classic illustrations are *lex mercatoria*, trade practice, or codes of conduct.

We argue that the critics of customary law use an overly narrow and therefore inappropriate concept of normativity. Norms do not only direct behaviour if norm violation is against the addressee's self-interest. On the other hand, the duty to abide by the law in force is not the exclusive motivating force either. Norms matter because they provide guidance. Most actors are most of the time willing to follow the norms prevailing in their context, or at least to be not too far off the mark, and most actors expect other actors to be thus guided.

The article takes issue with the furthest reaching claim: customary law is "epiphenomenal"; what looks like an effect of law actually is nothing but an act of self-interested behaviour. The proponents of this claim rely on game theory. They model states as unitary actors, i.e. as if they were individuals. In the experiment, we test a situation where game theory unequivocally predicts total defection: a public good game. In line with a rich literature in experimental economics, we refute the claim even in our *Baseline*, where normativity plays no (explicit) role. In our *Law* treatment, we introduce a meta-rule for the formation of customary law. In game-theoretic terms, this meta-rule is totally irrelevant. Yet it turns out to have a pronounced positive effect on people's cooperativeness.

Positivists will not be surprised by this result. Since we have explicitly invoked the legal order, this is what they would expect. Yet for them, the motivational force rests in the legal order. Actors abide by the law since this is their duty. Consequently, if we do not invoke the law but only ask participants whether they believe a norm to exist, the effect should vanish. This is what we test with our *Comity* treatment. Contrary to the positivist prediction yet in line with the position of Maurice Mendelson (Mendelson 1998), this turns out to be at least as effective at enhancing cooperation as the *Law* treatment.

Based on these findings, we suggest an alternative conceptualisation of customary law: The function of customary law consists of guiding behaviour into the normatively desired direction as normative expectations and behavioural patterns coevolve. We back this claim by further data analysis. What our participants have (privately) stated in the previous period significantly explains their behaviour in the following period; what the group has done in the previous period significantly explains their statements about norm existence and minimum required contributions in the following period.

Thus far, our results seem to suggest that the power of customary law collapses with the power of normativity. Law would be immaterial, not

because it has no behavioural effect, but because one does not need the legal order to bring the effect about. To test this hypothesis, we rerun the experiment in a setting where participants can sanction each other (*Baseline-S*). Also in this setting, the introduction of the meta-rule for the formation of customary law (treatment *Law-S*) has a strong positive effect on cooperation compared to *Baseline-S*. However, if we only privately ask participants whether there is a norm (treatment *Comity-S*), cooperation is even lower than in *Baseline-S*. If participants do not perceive the norm to be legal, sanctions “crowd out” some of the beneficial effect of the norm. By contrast sanctions and normative expectations corroborate each other if normative behaviour is required in law.

In the next section, we develop the legal research question from the lively debate over customary law in public international law, and derive the hypotheses to be tested in the experiment. Section 3 presents the design of our first experiment. Section 4 reports the results. Section 5 investigates the driving forces. Section 6 reports findings from the additional three treatments with sanctioning opportunities. Section 7 concludes.

II. The Legal Debate

In an influential, provocative paper, Jack Goldsmith and Eric Posner have argued that customary international law is a mere epiphenomenon. What looks as if states were abiding by international law effectively is nothing but an exercise of self-interest. States are rational maximisers of utility. They cooperate if this is the optimal strategy given the (expected) behaviour of other states. Cooperation may even occur if states face a dilemma. Through the folk theorem, if their interaction is repeated and the end is uncertain, cooperation may occur as long as neither of them is too impatient (Aumann and Shapley 1994). Yet cooperation is much less likely to occur if the group is large. In an n-person prisoner’s dilemma, the conditions for cooperation are close to heroic (Goldsmith and Posner 1999).

“A nation's ‘compliance’ with the cooperative strategy in the bilateral prisoner's dilemma has nothing to do with following a norm from a sense of legal obligation. Nations do not act in accordance with a norm that they feel obliged to follow; they act because it is in their interest to do so” (Goldsmith and Posner 1999: 1132).

Many have taken issue with this claim (for a survey see Norman and Trachtman 2005). They have argued in particular that the authors give too little credit to reputation, retaliation and segmentation (Chinen 2001), that international law changes the payoffs of the game (Guzman 2008), and that the strategies of grim trigger and penance make it possible to sustain cooperation even if the number of actors is large (Norman and Trachtman 2005). Our approach differs in that we take exactly a situation

Goldsmith and Posner claim makes cooperation impossible, and test experimentally whether it nonetheless occurs. We thus test:

H₁: In a multi-person dilemma game, there is no cooperation. A meta-rule requiring persons to abide by the rules of customary law is immaterial.

The positivist tradition stands in sharp contrast to the ideas of Goldsmith and Posner. For them the force of customary law rests in the individual addressee's consent, either to a specific rule on the issue at hand (Triepel 1899; Anzilotti 1955), or to a meta-rule stipulating the conditions under which a new rule comes into being (Kelsen 1952; Morelli 1967); (Elias 1995). Consequently, the scope for customary law that positivists are willing to grant is rather small.

Positivists are not primarily interested in extra-legal effects. Their main contribution is to the doctrine about rules on rules. Positivists define the conditions under which an utterance of words assumes the force of law. Yet the exercise has a natural corollary. If the law's addressees care about the law at all, they should care much more if this utterance of words is actually law, rather than a mere statement about desirability, on whatever non-legal grounds.

From this angle, the positivist position is related to a facet of the debate over an "expressive" function of law. While some contributions narrowly conceive law as a mere sanction (Bohnet and Cooter 2001; Tyran and Feld 2006; Galbiati and Vertova 2008; Bernasconi, Corazzini et al. 2010), others adopt a richer concept of law, arguing that the law serves as a focal point (McAdams 2000; McAdams and Nadler 2005; McAdams and Nadler 2008), informs people about behaviour others will approve (McAdams 2000), induces people to change their beliefs about the consequences associated with an action (Geisinger 2002), or changes the perception of underlying social norms (Feldman and Nadler 2006). Empirical findings have been mixed. The willingness to disregard the copyright protection of electronic works was not significantly affected by either making the illegality salient, nor by also pointing to informal sanctions, like a loss in reputation (Feldman and Nadler 2006). By contrast, in Switzerland voter turnout went down once a canton had abolished the legal obligation to vote, although enforcement had only been symbolic (Funk 2007).

As long as there is no enforcement, in psychological parlance the difference between mere comity and actual law boils down to a frame. It makes people see the issue in a different light. The requested behaviour is not only desirable; it is mandated by the legal order. Frames have been shown to strongly influence behaviour. Whether context makes people see a choice as involving a gain or a loss has a heavy impact, despite the fact that the task can easily be reframed by manipulating the reference point

(Tversky and Kahneman 1981). Also if a choice is contrasted with another, irrelevant outside option, choices change substantially (Tversky and Simonson 1993). When an agency aims at eliciting voters' willingness to pay for a public venture, responses heavily depend on how the issue is presented (Kahneman, Ritov et al. 1999). In all these tasks, the frame activates people's world knowledge. By the same token, we expect that people's attitudes toward law in general are activated if it is made salient that custom can be binding law (for a more elaborate theory of what this implies see Engel 2008). Most legal rules are meant to tame egoism and to make the law's subjects see the issue in the light of what is socially desirable. We therefore derive the following positivist prediction:

H₂: Participants are significantly more likely to overcome a dilemma if they are made aware that this is their duty in law.

Rational choice theorists have a hard time with customary law since it seems unclear why self-interested actors should contribute to the formation of a rule that will prevent them from acts of selfishness. Positivists have a hard time with customary law since it seems unclear how a new norm of customary law could ever come into being. Must those who originally claim the norm to exist actually have been mistaken (Elias 1995: 503; Stern 2001: 97)? Are the proponents of a new rule of customary law actually only making a proposal, which must be accepted by other actors (Elias 1995: 508)? Must one have recourse to some external authority, like divine emanation, natural law, or social necessity, to explain the formation of new customary law (Stern 2001: 92)?

It seems that both the positivist and the rational choice conceptualisations of customary law miss a key ingredient. Practice turns into law since behaviour and normative expectations coevolve. If nearly everybody behaves in a certain way, this not only shapes beliefs about others' future behaviour. If the pattern has been repeated for a while, the *behavioural* belief turns into a *normative* expectation. It no longer is purely cognitive. A motivational component is added to it. If one actor deviates from established practice, she violates others' normative expectations. Others regard such behaviour as not only anti-social but illegitimate.

In his Hague lectures, Maurice Mendelson has offered a related explanation (Mendelson 1998). For him, the textbook approach to customary law misses a central feature of public international law. The international legal order is "semi-anarchic" (166), embryonic, and in a deep way incomplete. "Whilst modern domestic societies are characterised by highly centralised and compulsory systems of law-making and adjudication, not to mention enforcement, international society is not like that" (168). Therefore a "formalistic approach" (168) is misplaced. It is not possible to state in an abstract way the conditions that must be fulfilled for a new rule of customary international law to come into being (172). "The characteristic of this kind of law is that it is not just unwritten,

it is informal" (172). "The customary process is in fact a continuous one, which does not stop when the rule has emerged [...]. Even after the rule has 'emerged', every act of compliance will strengthen it, and every violation, if acquiesced in, will help to undermine it" (175). Customary international law rests on the conviction that "states should comply with the legitimate expectations of the international community" (185), where the ambiguity of the term "expectation" is deliberate: "If, within a social group, people habitually behave in a certain way, then, particularly if others rely on the continuation of this conduct, the sentiment may develop within that society that one is obliged to continue so to act. In other words, a norm emerges from what is normal [...]. If the generality of states has regularly behaved in certain ways [...], then a legitimate expectation arises that they will continue to do so" (185 f.).

In essence, this is a claim about normativity in general. Yet it can be combined with the reasons expressive law theorists give for the specific behavioural effects of law, which we have reported above. If this were to hold true, we would have to see

H₃: Normative expectations and behaviour coevolve. Coevolution is faster and more robust if the normative expectation originates in the legal order.

III. Experimental Design

To test these hypotheses, we run a standard linear four person public goods game. In this game, payoffs are given by (1):

$$\pi_i = e - c_i + m \sum_{j=1}^N c_j \quad (1)$$

Every round, each participant receives the same endowment e . She is free to keep it, or to use a fraction, or the entire endowment, as her contribution c_i to the joint project. Contributions by all members are multiplied by marginal per capita rate $m < 1$. As long as $mN > 1$, in this game the society of all group members is best off if everybody contributes everything. However individually, each member is best off if others contribute while she freerides. This constitutes an n-person prisoner's dilemma with continuous action space. As is standard in the experimental literature on public goods (for overviews see Ledyard 1995; Zelmer 2003; Chaudhuri 2011), we have announced after how many periods the game ends. Under standard game theoretic assumptions, players reason backwards. Since it is rational to defect in the last round, to preempt being the sucker in this round, a rational player defects in the penultimate

round, and so forth until the first round (Selten 1978; Rosenthal 1981). Hence the unique subgame perfect Nash equilibrium is all players defecting from the beginning. Note that, by this design, we give the rational choice critique of customary law its best shot. Since the number of periods is announced, game theory would even predict that customary law is pointless if there are only two players (states).

In our *Baseline* treatment participants play this game for fixed groups of 4, interacting over 30 announced periods, with an endowment of 20 experimental currency units (ECU), and a marginal per capita rate $m = .4$. These parameters are standard in the experimental literature on public goods, but for the number of repetitions. We have replaced the usual duration of 10 periods by a longer spell since we want more scope for the coevolution of norms and behaviour. With these parameters, if all participants contribute their entire endowments to the joint project, all receive 32 ECU. If all keep their entire endowments, all receive 20 ECU. If three participants contribute fully, while the fourth keeps everything, the former have 24 ECU, while the latter has 44 ECU. If one contributes fully, while three freeride, the former has 8 ECU, while the latter each have 28 ECU.

The first stage of the *Comity* treatment is exactly the same as the baseline. However, *Comity* has one additional stage: In each period, after participants have made their contribution decisions, but before giving them feedback, we ask them the following two questions:

- “1. Do you believe in your group exists a general norm regarding an adequate minimum contribution to the project (yes/no)?
2. If so, which is the generally expected minimum contribution (number from 0 to 20)?”

The instructions make it clear that participants will not get feedback about other participants' statements.

The *Law* treatment is identical to *Comity* but for the fact that, in addition, participants read the following paragraph in the instructions:

“For new law to originate, it is not necessary that the legislator pass a statute, or that the parties agree on an explicit contract. Customary law is equally valid and binding. Customary law comes into being if the large majority of those affected for a sufficiently long period behave in a sufficiently similar way. The fact that some contribute even more to the joint project does not prevent that a norm of customary law originates. Hence a rule of customary law may prescribe a minimum standard. Customary law may originate here in the lab as well.”

Note that we have deliberately kept this paragraph procedural. Participants are informed about the conditions under which *they themselves* are able to make new law. This is not only in line with customary law doctrine, but also a safeguard against the risk of imposing any specific material norm on participants.

The experiment was run in the Bonn EconLab and programmed in zTree (Fischbacher 2007). Participants were invited using ORSEE (Greiner 2004). After being seated in individual cubicles, participants received experimental instructions and answered a set of control questions (see the Appendix). We had 20 participants, interacting in 5 groups in the *Baseline*,¹ and 24 participants, interacting in 6 groups, in each of the two treatments. Participants were randomly drawn from a pool of some 3500 subjects. They held various majors. Approximately half of them were female. In the *Baseline*, participants on average earned 12.42 € (15.73 \$), in the *Comity* treatment earnings were on average 14.07 € (17.82 \$), and in the *Law* treatment 13.80 € (17.48 \$).

IV. Results

A. Treatment Effects

From Figure 1 one directly sees that both our treatments have a pronounced effect on cooperation. Whereas the *Baseline* displays the characteristic decay of contributions over time (see for example Fehr and Gächter 2000), both in *Comity* and in *Law*, average contributions even increase in the first periods, and they are much higher than in the *Baseline* until the endgame effect kicks in. Descriptively, *Comity* has a slightly stronger effect on contributions than *Law*.

¹ The planned 6th group could not be filled since participants did not show up.

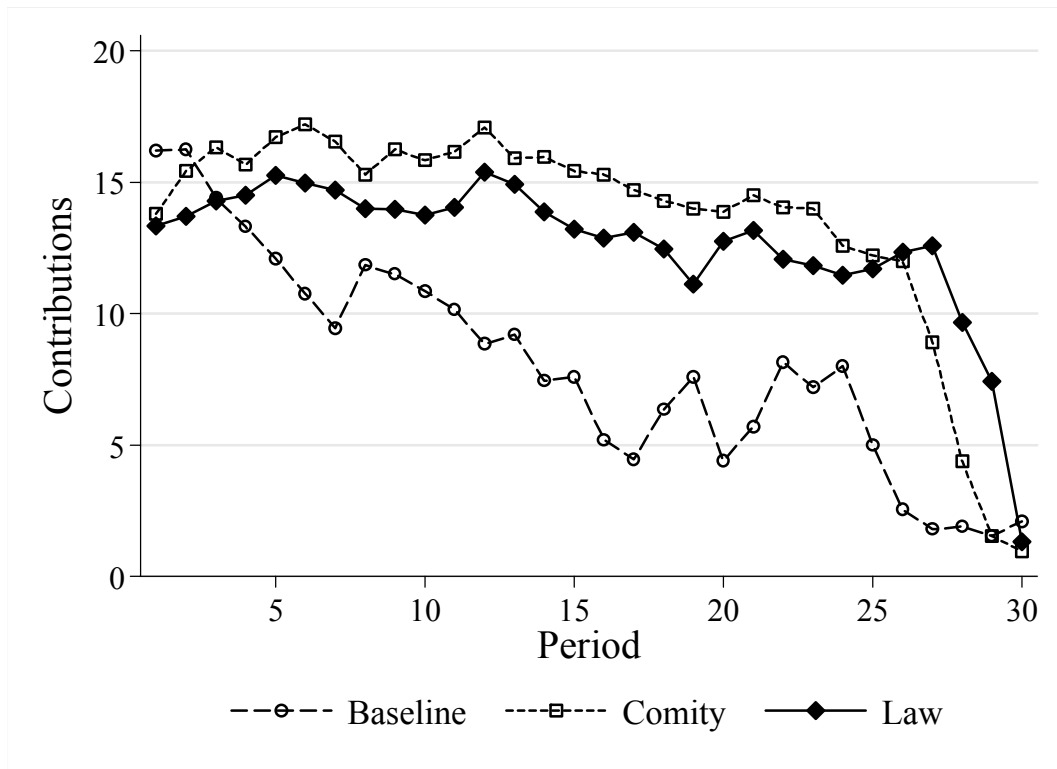


Figure 1
Contributions to the Public Good

The visual impression is confirmed by statistical analysis. In a conservative non-parametric test over means per group, the difference between the *Baseline* and *Comity* is significant at conventional levels (Mann Whitney, $N = 11$, $p = .0446$). In this test, the difference between the *Baseline* and *Law* is insignificant ($p = .1441$).

We analyse the data parametrically using a random effects Tobit estimator.² Using this procedure, in the regressions of Table 1 we establish a significant treatment effect for both *Comity* (model 1) and *Law* (model 2), in comparison with the *Baseline*. By contrast the difference between *Comity* and *Law* is clearly insignificant (model 3, p -value *Law* = .535).

Our results clearly refute H_1 , which was derived from the rational choice critique of customary law. In a way, already our *Baseline* speaks against the claim that there is no cooperation in a multi-person dilemma. For

² The random effect captures the dependence at the level of individuals and the Tobit functional form accounts for the fact that our dependent variable is both left and right censored (many participants contribute their entire endowment of 20 tokens, while others contribute nothing). Participants interact in the same group of four over 30 periods. This creates a second source of dependence. Ideally we would therefore want to estimate a mixed effects model with two random effects: one for the group and another for the individual. Yet unfortunately there is no generally acknowledged mixed effects estimator for censored data. As a substitute, to make sure standard errors are not deflated, we bootstrap the estimator, with random draws of entire groups.

sure, in the reality of international law, there are many more than just four actors. Yet note that a four person game excludes stabilising cooperation by threatening a defector with Nash reversion. Since there is more than one partner, players cannot use their own contributions, in the subsequent period, as a sanctioning technology. They would not only hit free riders, but also those who have faithfully contributed to the joint project. On the conceptual grounds on which the rational choice critique of customary law is built, a four actor dilemma is no different from a 192 actor dilemma.

Moreover, as model 2 shows, if we only compare the *Baseline* with the *Law* treatment, law clearly matters. The fact that behaviour is required by law has a big, positive effect. Note that our *Law* treatment is very subtle. The legal norm not only lacks sanctions, there is not even communication among those expected to abide by the legal rule, neither about its existence nor about its contents.

By contrast, we do not find support for **H₂**, which we derived from legal positivism, and from expressive law theory. We do not find any statistical difference between *Comity* and *Law*. In our data, customary law boils down to normativity. The fact that the rule is embedded in the legal order does not help participants overcome the dilemma even better. Descriptively they even fare slightly worse. We will revisit this issue in our second experiment, but first turn to our process hypothesis **H₃**.

	model 1	model 2	model 3
	Baseline vs. Comity	Baseline vs. Law	Comity vs. Law
Comity	13.661*		
Law		9.597*	-3.150
Period	-.813***	-.604***	-.517***
Cons	18.175***	15.217***	26.483***
N	1320	1320	1440
left censored	311	327	198
right censored	469	380	595
p model	<.001	<.001	<.001

Table 1
Treatment Effect on Contributions

random effects Tobit, $ll(0) ul(20)$, bootstrapped with random draws of entire groups, 50 reps
reference category: models 1 and 2: *Baseline*, model 3: *Comity*
*** p < .001, ** p < .01, * p < .05

B. The Coevolution of Behaviour and Normative Expectations

Figure 2 shows for both the *Comity* and the *Law* treatment that normative expectations and behaviour are indeed synchronous. As long as many participants believe there is a norm, contributions are high. If this belief erodes, ultimately behaviour is affected. There is also a clear relationship

between statements about the minimum expected contribution and actual contributions.³

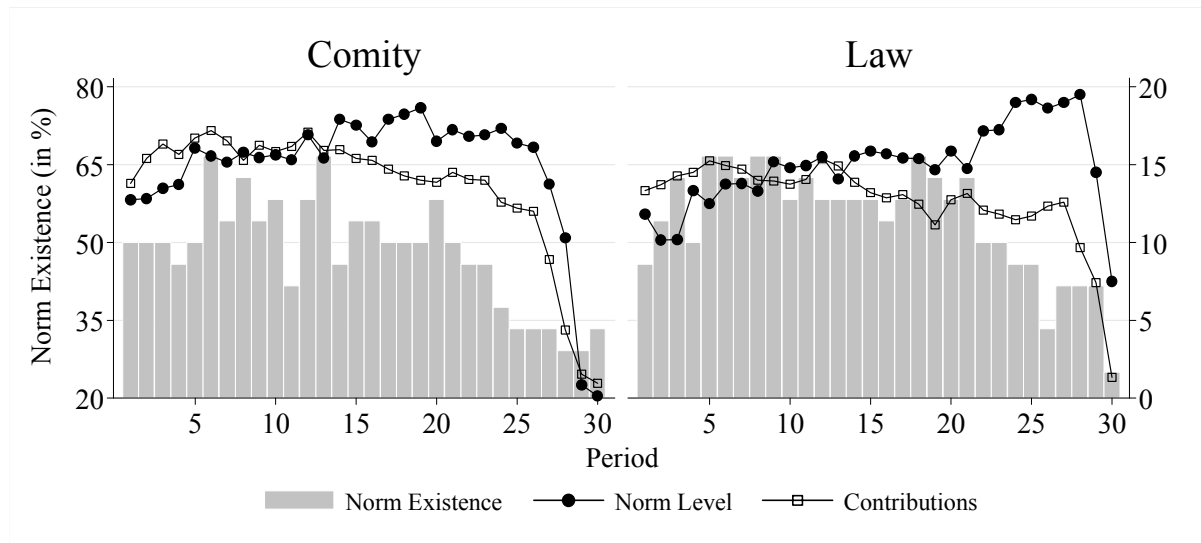


Figure 2
Coevolution of Behaviour and Normative Expectations
 left axis: % of participants who say there is a norm
 right axis: actual and expected contribution

We can exploit the panel structure of our data to identify effects in both directions. If a participant has claimed in the previous period that there is a norm, this significantly and substantially increases how much she contributes to the joint project in the subsequent period (Table 2, model 1). Likewise, if one only looks at those participants who claim norm existence, the higher a participant has claimed the minimum contribution to be, the more she contributes in the subsequent period (model 2). This finding suggests a desire for self-consistency between one's stated normative expectations and one's contribution behaviour. Note that this would not be different from normativity, but would be the mechanism by which subjectively accepted norms guide behaviour. In the reverse direction, the higher mean contributions in the participant's group in the previous period⁴, the more this participant is likely to state that there is a norm (model 3).⁵ Likewise, the higher mean contributions in the previous period, the more demanding the statements about the required minimum contribution (model 4).

³ The apparent kink in statements about the required contribution level by the end in the *Law* treatment is due to a selection effect. If we analyse first differences of these statements, in neither treatment do they go up over time. Hence in the final periods, those remaining faithful to the norm do not believe it to be even stronger. Rather those who always believed the norm to be more demanding are more likely to still think it exists.

⁴ We work with the lag of average contributions, although participants made their statements after contribution choices since they did not yet have feedback about this period's contributions.

⁵ In a logit model, both events are equally likely if the predicted coefficient is 0.

	model 1	model 2	model 3	model 4
Dependent Variable:	contribution	contribution	statement of norm existence	statement of norm level
Lagged statement of norm existence	6.465**			
Lagged statement of norm level		1.699***		
Lagged mean contribution in group			.202***	1.189***
Law	-3.530	-1.231	.784	-.914
Period	-.510***	-.445*	-.030**	.123 ⁺
cons	23.408***	3.803	-2.588***	-3.583
N	1392	724	1392	715
left censored	198	54		23
right censored	574	424		303
p model	<.001	<.001	<.001	<.001

Table 2
Coevolution of Behaviour and Normative Expectations

models 1,2,4: random effects Tobit, $\text{II}(0)$ $\text{ul}(20)$
standard error from bootstrap, sampled at group level, 50 reps
model 3: mixed effects logit, period nested in individual nested in group
*** p < .001, ** p < .01, * p < .05, + p < .1

Figure 3 separately draws contributions of participants who say and who do not say there is a norm, together with the lagged mean contribution. The left panel displays the *Comity* treatment, and the right panel the *Law* treatment. As one sees, in all periods those who claim norm existence on average contribute more than the mean in the previous period. By contrast, those who do not say there is a norm contribute less than the lagged mean in all but the very first periods. Moreover, until the endgame effect kicks in, contributions of those who say there is a norm are almost stable, while contributions of the remaining subjects visibly decay over time. We thus support the main claim of **H₃**: customary law matters since normative expectations and behaviour coevolve. These observations hold likewise for the *Comity* and the *Law* treatment. We do not find any additional effect of law over mere social expectations. The next section further examines this result within a slightly modified experimental paradigm.

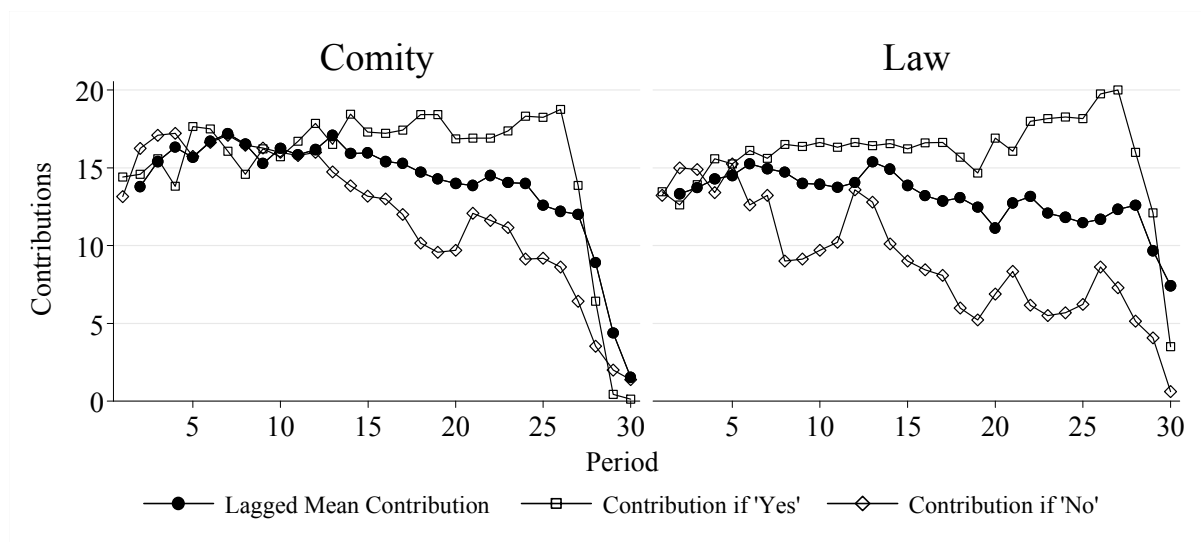


Figure 3
Contributions by Claim of Norm Existence

C. Law beyond Normativity?

The results reported above strongly suggest that customary law matters. It substantially improves behaviour, as we have clearly seen in Figure 1. This is welcome news for public international law, and for law in general. Yet it is troubling that, seemingly, an equally strong behavioural effect could be brought about without explicit reference to law. Seemingly, customary law only matters because it makes normativity salient. In the terminology of international law doctrine we might thus conclude that customary law is not more effective than mere comity. If that was the end of the story, why has (customary) law evolved? Why do states seem to care? Why do they exert considerable effort to show that what might appear to be in violation of customary international law actually is in harmony with it, for instance because the rule has an exception? Within international law doctrine, the answer is straightforward: If and only if the state that has violated an obligation from customary international law, it must tolerate that the victim retaliates by itself violating another rule of international law (within the limits of the law of reprisals, for sure). Yet is this doctrinal distinction of any behavioural relevance?

To answer these questions, we have run a second experiment. In this experiment we repeat our three treatments from above with a slight modification. In every period we now add an additional stage in which participants have the possibility to sanction each other (cf. Fehr and Gächter 2000; Nikiforakis and Normann 2008). Specifically, to destroy one ECU of another group member, a participant must spend one ECU of her own period income. The first stage payoff function remains as in (1). If we write π_1 for the first stage income, total payoff is given by (2):

$$\pi = \pi_1 - \sum_{j \neq i}^N s_{ij} - \sum_{j \neq i}^N s_{ji} \quad (2)$$

First stage income is reduced by all sanctioning points this player gives to another group member (s_{ij}), and by all sanctioning points other group members inflict on her (s_{ji}). Another 72 students have participated in these three treatments with sanctioning opportunities (“S-treatments”). They have on average earned 14.45 € (18.44 \$) in the *Baseline-S*, 13.78 € (17.58 \$) in *Comity-S*, and 15.19 € (19.38 \$) in *Law-S*.

Of course, the caveat made when introducing the first experiment applies here too: we test students, and the reactions of states to the availability of a sanctioning mechanism may well be driven by other forces. In particular, the cost or risk involved in enforcing international law against a powerful state may be very high. Yet note that apart from that our design is very close to the situation of customary international law. There is no central enforcement agency. The cost of enforcement is borne by those who engage in it. If the rule is valid *erga omnes* (for an economic analysis see Engel 2009; Posner 2009), each state is entitled to enforce it, even if another state has been violated. Since the rule originates in state practice, it need not be explicit. Frequently, whether there is a rule or not is disputed. Also, using a 1:1 fine-to-fee ratio, we have made decentral punishment as costly as it reasonably can be made in the lab; typically in public good experiments the leverage of punishment is much higher.

Figure 4 compares cooperation in the three treatments without (left panel) and with (right panel) decentral sanctioning. There are two messages. The first is straightforward. Sanctions are very effective at improving cooperation, even if they are weak. For the rational choice theorists of customary international law this again should be surprising. In our game every player is best off if others bear the cost of sanctioning, which is why theory predicts zero sanctions (Yamagishi 1986; Heckathorn 1989). Consequently adding the sanctioning option should not change behaviour. Yet empirically people, and states for that matter, are willing to engage in costly sanctioning. Potential addressees rightly expect sanctions, and react by changing their behaviour (key contributions to this literature are Fehr and Gächter 2000; Fehr and Gächter 2002). Both the difference between *Comity* and *Law-S* ($N = 12$, $p = .0163$) and the difference between *Law* and *Law-S* ($N = 12$, $p = .0104$) is already significant in a simple Mann-Whitney test over group means.

The more important message of Figure 4 is more subtle, but also more relevant for understanding the difference between comity and customary law. While contributions were slightly lower in *Law* than in *Comity* when there was no sanctioning option, with sanctioning the order reverses. Now *Law-S* outperforms all other treatments, while *Comity-S* performs even

poorer than the *Baseline-S*. There is a well-understood behavioural explanation for the latter effect. Extrinsic interventions crowd out intrinsic motivation (Bolton and Katok 1998; Fehr and Gächter 2001; Frey and Jegen 2001; Nyborg and Rege 2001; Chan, Godby et al. 2002; Janssen and Mendys 2004; Eckel, Grossman et al. 2005; Borges and Irlenbusch 2007). Yet interestingly the socially detrimental effect of sanctions disappears if behaviour is not only socially but legally required. Now, to the contrary, sanctions and intrinsic motivation are no longer substitutes. They become complements. Law makes sanctions behaviourally robust.

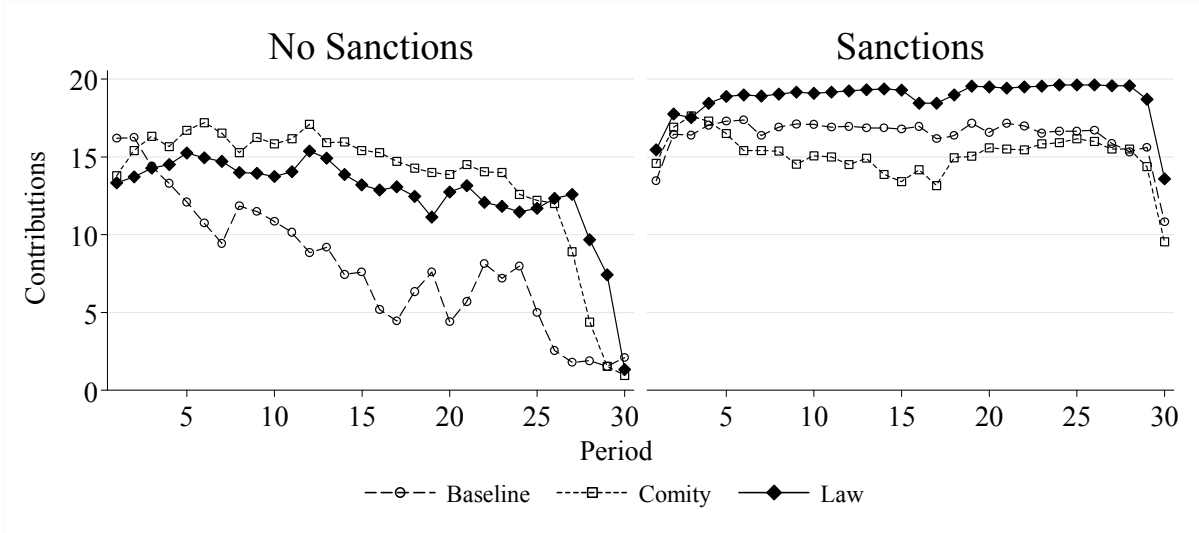


Figure 4
Contributions with and without Sanctions

Table 3 analyses the complementarity between law and sanctions in further depth. Specifically, we want to see how people change their contribution behaviour as a result of receiving a sanction in the previous period. Consequently, we do not look at contribution levels, but at contribution changes, i.e. first differences.

There is a strong main effect of sanctions: each sanctioning point a participant has received in the previous period induces her to increase her contributions by .659 points. More importantly even, from the negative interaction of sanction with *Comity-S* we see that the beneficial effect of punishment is significantly less pronounced when the normative expectation does not originate in law. In the *Comity-S* treatment, one punishment point only induces the recipient to adjust her contributions by $.659 - .217 = .442$ points.

lagged received sanction	.659***
Comity	-.161
lagged received sanction* Comity	-.217**
period	-.008
period 30	-5.418***
cons	-.026
N	1392
p model	<.001

Table 3
Sensitivity to Sanctions

depvar: first differences of contributions
mixed effects model, period nested in individual nested in group
Reference group: *Law* treatment
*** p < .001, ** p < .01, * p < .05, + p < .1

As Figure 5 shows, the main difference between *Comity-S* and *Law-S* originates in the perception of norm existence. If we make the possibility of the formation of customary law salient, and if participants have the possibility to sanction norm violations, after a small number of periods almost all participants believe a norm to exist. And this norm closely corresponds with actual behaviour. One way of explaining this effect is game theoretic. Our *Law-S* treatment can be interpreted as turning the meta-rule, i.e. the possibility that a normative expectation emerges, into perceived common knowledge.⁶ Our experiment thus also points to an additional option for explaining the effect of customary law when keeping the rational choice assumption that actors just maximise their payoffs.

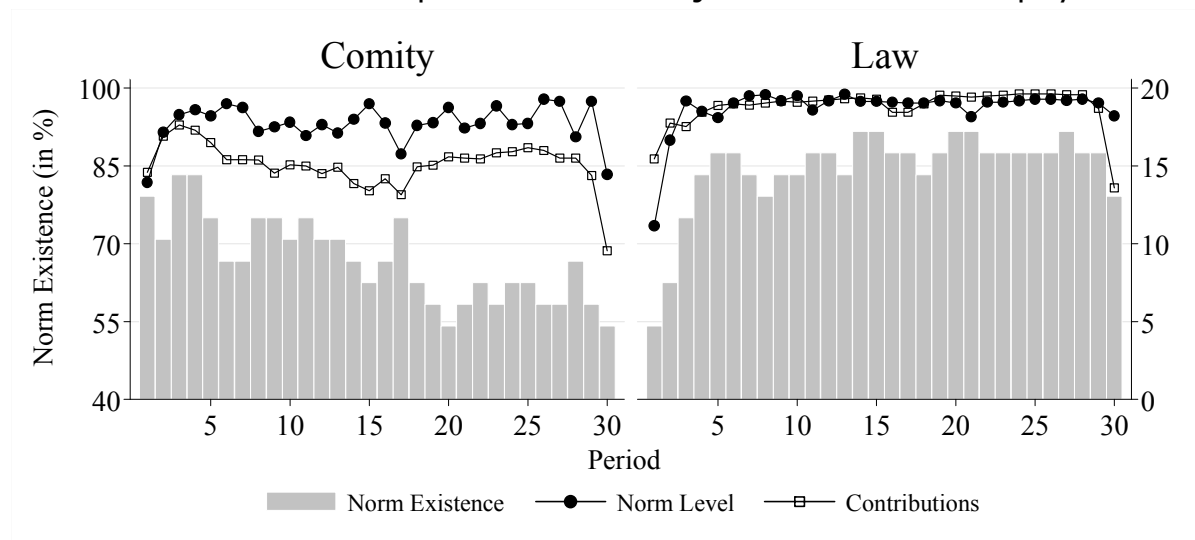


Figure 5
Norm Existence and Norm Level with Sanctioning

Table 4 shows that the dynamics of norm perception are substantially different in *Comity* and *Law*. The positive and significant time trend indicates that participants become more and more likely over time to state there is a norm when this norm originates in law. However, they are less

⁶ We are grateful to Carlos Alós-Ferrer for this interpretation. Descriptively, we already see a similar effect without sanctions, Figure 2, yet it is not significant, Table 2.

likely to say so in the absence of law (.100 - .217 = -.117). This qualifies our finding from the first experiment. While we do not find support for **H₂** in the absence of sanctions, once the law is backed up by a sanction, law is indeed more powerful than mere comity. In our second experiment we thus support **H₂**.

Comity	2.059
Period	.100***
Comity*period	-.217***
Cons	2.280*
N	1440
p model	<.001

Table 4
Statement of Norm Existence with Sanctions

depar: statement of norm existence (yes/no)
 mixed effect logit, statement nested in individual nested in group
 data from *Comity* and *Law* treatments only (since norm question is not asked in the *Baseline*)
 reference category: *Law* treatment
 *** p < .001, ** p < .01, * p < .05

V. Conclusion

Using a standard paradigm from experimental economics, in this paper we show that, in comparison with a situation free from institutional guidance, customary law has a strong beneficial effect. It helps experimental participants overcome a social dilemma. If there are no sanctions, the effect basically coincides with the behavioural effect of what public international law calls comity. In essence, customary law governs behaviour since normative expectations and behaviour coevolve. Whether the rule invokes the authority of the law is at best immaterial, if not detrimental. Yet the authority of the law becomes instrumental as soon as there are sanctions. If combined with comity, sanctions crowd out some of the beneficial effect. If the rule originates in law, however, the authority of the law and the threat of sanctions reinforce each other.

There is an obvious gap between the behaviour of students in a computer lab and state practice. It could well be that states are not affected by normativity the same way as individuals. But as laid out in section 2, as it stands, the debate in public international law does not rest on factors that are specific to states being the actors. The underlying claims address actors in general, and should therefore also be valid for individuals. It is these claims we address.

To gain further insights about the coevolution of behaviour and normative expectations in public international law, it would be attractive to explore normativity in international relations (cf. Risse 1999; Risse 2000) qualitatively. In one respect, the motivation of states as corporate actors is even easier to study than the motivation of individuals. While the formation of the individual will occurs in her *forum internum*, the

formation of the corporate will is open to public scrutiny. Through freedom of information legislation, even internal government deliberation is made accessible. While the ultimate proof may not be provided in this paper, it seems plausible that normativity in international relations is not fundamentally different from normativity in personal relations. Normativity is a key ingredient of human sociality (Wyman, Rakoczy et al. 2009). It seems likely that it is scaled up in international law, but that it is not a substantially different force. On that assumption, the experimental findings reported in this paper should help international lawyers to gain a better understanding of the oldest source of law: custom.

References

- ANZILOTTI, DIONISIO (1955). Corso di Diritto Internazionale. Padova,, Cedam.
- ARLEN, JENNIFER, MATTHEW L. SPITZER, et al. (2002). "Endowment Effects within Corporate Agency Relationships." Journal of Legal Studies **31**: 1-37.
- AUMANN, ROBERT J. and LLOYD S. SHAPLEY (1994). Long Term Competition - A Game Theoretic Analysis. Collected Papers I. R. J. Aumann. Cambridge, MIT Press: 395-409.
- BERNASCONI, MICHELE, LUCA CORAZZINI, et al. (2010). 'Expressive' Obligations in Public Good Games. Crowding-in and Crowding-out Effects. Working Papers.
- BOHNET, IRIS and ROBERT D. COOTER (2001). Expressive Law. Framing or Equilibrium Selection?
- BOLTON, GARY E. and ELENA KATOK (1998). "An Experimental Test of the Crowding Out Hypothesis. The Nature of Beneficent Behavior." Journal of Economic Behavior & Organization **37**: 315-333.
- BORGES, GEORG and BERND IRLBUSCH (2007). "Fairness Crowded Out by Law. An Experimental Study of Withdrawal Rights." Journal of Institutional and Theoretical Economics **163**: 84-101.
- CHAN, KENNETH S., ROB GODBY, et al. (2002). "Crowding-out Voluntary Contributions to Public Goods." Journal of Economic Behavior & Organization **48**(3): 305-317.
- CHAUDHURI, ANANISH (2011). "Sustaining Cooperation in Laboratory Public Goods Experiments. A Selective Survey of the Literature." Experimental Economics **14**: 47-83.
- CHINEN, MARK A. (2001). "Game Theory and Customary International Law. A Response to Professors Goldsmith and Posner." Michigan Law Review **23**: 143-189.
- CROSON, RACHEL T.A. and JASON SCOTT JOHNSTON (2000). "Experimental Results on Bargaining Under Alternative Property Rights Regimes." Journal of Law, Economics and Organization **16**: 50-73.
- ECKEL, CATHERINE C., PHILIP J. GROSSMAN, et al. (2005). "An Experimental Test of the Crowding Out Hypothesis." Journal of Public Economics **89**: 1543-1560.
- ELIAS, OLUFEMI (1995). "The Nature of the Subjective Element in Customary International Law." International and Comparative Law Quarterly **44**: 501-520.
- ENGEL, CHRISTOPH (2008). "Learning the Law." Journal of Institutional Economics **4**: 275-297.
- ENGEL, CHRISTOPH (2009). "Erga omnes. Why does Public International Law Ignore Privity of Contract?" Journal of Institutional and Theoretical Economics **165**: 24-28.
- FEHR, ERNST and SIMON GÄCHTER (2000). "Cooperation and Punishment in Public Goods Experiments." American Economic Review **90**: 980-994.

- FEHR, ERNST and SIMON GÄCHTER (2001). Do Incentive Contracts Crowd Out Voluntary Cooperation?
- FEHR, ERNST and SIMON GÄCHTER (2002). "Altruistic Punishment in Humans." Nature **415**: 137-140.
- FELDMAN, YUVAL and JANICE NADLER (2006). "The Law and Norms of File Sharing." San Diego Law Review **43**: 577-618.
- FISCHBACHER, URS (2007). "z-Tree. Zurich Toolbox for Ready-made Economic Experiments." Experimental Economics **10**: 171-178.
- FREY, BRUNO and RETO JEGEN (2001). "Motivation Crowding Theory. A Survey of Empirical Evidence." Journal of Economic Surveys **15**: 589-611.
- FUNK, PATRICIA (2007). "Is there an Expressive Function of Law? An Empirical Analysis of Voting Laws with Symbolic Fines." American Law and Economics Review **9**(1): 135-159.
- GALBIATI, ROBERTO and PIETRO VERTOVA (2008). "Law and Behaviours in Social Dilemmas. Testing the Effect of Obligations on Commitment." Games and Economic Behavior **64**: 146-170.
- GEISINGER, ALEX (2002). "A Belief-Change Theory of Expressive Law." Iowa Law Review **88**(1): 35-73.
- GOLDSMITH, JACK and ERIC A. POSNER (1999). "A Theory of Customary International Law." University of Chicago Law Review **66**: 1113-1177.
- GRECHENIG, KRISTOFFEL, ANDREAS NICKLISCH, et al. (2010). "Punishment Despite Reasonable Doubt. A Public Goods Experiment with Sanctions Under Uncertainty." Journal of Empirical Legal Studies **7**(4): 847-867.
- GREINER, BEN (2004). An Online Recruiting System for Economic Experiments. Forschung und wissenschaftliches Rechnen 2003. K. Kremer and V. Macho. Göttingen: 79-93.
- GUZMAN, ANDREW T. (2008). How International Law Works. A Rational Choice Theory, Oxford University Press, USA.
- HECKATHORN, DOUGLAS D. (1989). "Collective Action and the Second-Order Free-Rider Problem." Rationality and Society **1**: 78-100.
- JANSSEN, MAARTEN C.W. and EWA MENDYS (2004). "The Price of a Price. On the Crowding Out of Social Norms." Journal of Economic Behavior & Organization **55**: 377-395.
- KAHNEMAN, DANIEL, ILANA RITOV, et al. (1999). "Economic Preferences or Attitude Expressions? An Analysis of Dollar Responses to Public Issues." Journal of Risk and Uncertainty **19**: 203-235.
- KELSEN, HANS (1952). Principles of International Law. New York, Rinehart.
- LEDYARD, JOHN O. (1995). Public Goods. A Survey of Experimental Research. The Handbook of Experimental Economics. J. H. Kagel and A. E. Roth. Princeton, NJ, Princeton University Press: 111-194.
- LOEWENSTEIN, GEORGE and DON A. MOORE (2004). "When Ignorance Is Bliss. Information Exchange and Inefficiency in Bargaining." Journal of Legal Studies **33**: 37-58.
- MCADAMS, RICHARD H. (2000). "An Attitudinal Theory of Expressive Law." Oregon Law Review **79**: 339-390.

- MCADAMS, RICHARD H. (2000). "A Focal Point Theory of Expressive Law." Virginia Law Review **86**: 1649-1729.
- MCADAMS, RICHARD H. and JANICE NADLER (2005). "Testing the Focal Point Theory of Legal Compliance. The Effect of Third-Party Expression in an Experimental Hawk/Dove Game." Journal of Empirical Legal Studies **2**: 87-123.
- MCADAMS, RICHARD H. and JANICE NADLER (2008). "Coordinating in the Shadow of the Law: Two Contextualized Tests of the Focal Point Theory of Legal Compliance." Law and Society Review **42**: 865-898.
- MENDELSON, MAURICE H. (1998). "The Formation of Customary International Law." Recueil des Cours **272**: 155-410.
- MORELLI, GAETANO (1967). Nozioni di Diritto Internazionale. Padova,, CEDAM.
- NADLER, JANICE and SHARI SEIDMAN DIAMOND (2008). "Eminent Domain and the Psychology of Property Rights: Proposed Use, Subjective Attachment, and Taker Identity." Journal of Empirical Legal Studies **5**: 713-749.
- NIKIFORAKIS, NIKOS S. and HANS-THEO NORMANN (2008). "A Comparative Statics Analysis of Punishment in Public Good Experiments." Experimental Economics **11**: 358-369.
- NORMAN, GEORGE and JOEL P. TRACHTMAN (2005). "The Customary International Law Game." American Journal of International Law **99**: 541-580.
- NYBORG, KARINE and MARI REGE (2001). Does Public Policy Crowd Out Private Contributions to Public Goods?
- POSNER, ERIC A. (2009). "Erga Omnes Norms, Institutionalization, and Constitutionalism in International Law." Journal of Institutional and Theoretical Economics **165**: 5-23.
- RISSE, THOMAS (1999). "International Norms and Domestic Change. Arguing and Communicative Behaviour in the Human Rights Area." Politics & Society **27**: 529-559.
- RISSE, THOMAS (2000). "'Let's Argue!'. Communicative Action in World Politics." International Organization **54**: 1-39.
- ROSENTHAL, ROBERT W. (1981). "Games of Perfect Information, Predatory Pricing and the Chain Store Paradox." Journal of Economic Theory **25**: 92-100.
- SELTEN, REINHARD (1978). "The Chain Store Paradox." Theory and Decision **9**: 127-159.
- STERN, BRIGITTE (2001). "Custom at the Heart of International Law." Duke Journal of Comparative and International Law **11**: 89-108.
- TRIEPEL, HEINRICH (1899). Völkerrecht und Landesrecht. Leipzig,, C.L. Hirschfeld.
- TVERSKY, AMOS and DANIEL KAHNEMAN (1981). "The Framing of Decisions and the Psychology of Choice." Science **211**: 453-458.
- TVERSKY, AMOS and ITAMAR SIMONSON (1993). "Context-Dependent Preferences." Management Science **39**: 117-185.

- TYRAN, JEAN-ROBERT and LARS P. FELD (2006). "Achieving Compliance when Legal Sanctions are Non-Deterrent." Scandinavian Journal of Economics **108**: 135-156.
- WYMAN, EMILY, HANNES RAKOCZY, et al. (2009). "Normativity and Context in Young Children's Pretend Play." Cognitive Development **24**: 146-155.
- YAMAGISHI, TOSHIO (1986). "The Provision of a Sanctioning System as a Public Good." Journal of Personality and Social Psychology **51**: 110-116.
- ZEILER, KATHRYN (2010). "Cautions on the Use of Economics Experiments in Legal Analysis." Journal of Institutional and Theoretical Economics **166**: 178-193.
- ZELMER, JENNIFER (2003). "Linear Public Goods. A Meta-Analysis." Experimental Economics **6**: 299-310.

Appendix

In the following, we present the experimental instructions as well as the control questions for Treatment *Law-S*. All other treatments are reduced forms of *Law-S*:

- Omitting the box “Customary Law:...” we obtain *Comity-S*.
- Additionally omitting the shaded areas we obtain *Baseline-S*.
- Finally, by omitting Stage 3 of the instructions and Question 5 of the questionnaire, we obtain the 3 treatments without sanctions: *Law*, *Comity*, and *Baseline*.

General instructions for the participants

Welcome to our experiment!

If you read the following explanations carefully, you will be able to earn a substantial sum of money, depending on the decisions you make. It is therefore crucial that you read these explanations carefully.

During the experiment there shall be absolutely no communication between participants. Any violation of this rule means you will be excluded from the experiment and from any payments. If you have any questions, please raise your hand. We will then come over to you.

During the experiment we will not calculate in euro, but instead in taler. Your total income is therefore initially calculated in taler. The total number of taler you accumulate in the course of the experiment will be transferred into euro at the end, at a rate of

1 Euro = 60 Taler

At the end you will receive from us the **cash** sum, in euro, based on the number of taler you have earned.

The experiment consists of **30 periods**, and each period consists of **3 stages**. Participants are randomly divided into groups of four. Apart from yourself, your group therefore has 3 further members. During these 30 periods, the constellation of your group of four remains unchanged. **Hence, you are with the same people in the same group for 30 periods.** At the beginning, each group member is allocated a random number between 1 and 4. This number remains unchanged for the entire 30 periods.

Stage 1:

At the beginning of each period, each participant is given **20 taler** to work with, referred to henceforth as **endowment**. Your task is to decide upon how to use your endowment.

You must decide how many of the 20 taler you wish to pay into a common **project**, and how many you wish to keep for yourself. The consequences of this decision are explained in more detail below.

Your **endowment** hence consists of **20 taler in each period**. You make a decision on your payments by typing whole numbers between 0 and 20 in the input field on your screen. Once you have keyed in your amount, press **Continue**. As soon as you have done this, you may no longer reverse your decision for this period.

Once all group members have made their decisions, you are told how much each individual group member has contributed to the project.

Your total income (in taler) therefore consists of two parts: (1) the taler income from the common project and (2) the taler you have retained.

Total income (in taler)	=	Income from the common project	+	Taler retained
--------------------------------	----------	--------------------------------	----------	----------------

The **income from the common project** is calculated as the total sum of all contributions to the project (within your group of four) times 0.4.

Income from the common project	=	total sum of all contributions to the project (within your group of four)	×	0.4
---------------------------------------	----------	---	----------	-----

Example:

If the sum of contributions from all group members to the common project is 60 taler, you and each other group member receive an income from the project of $0.4 \times 60 = 24$ taler. If the group members have contributed a total of 9 taler to the project, you and each other group member receive a taler income from the project of $0.4 \times 9 = 3.6$.

If you contribute one taler from your endowment to the group project, the sum of contributions to the common project increases by 1 taler, and your income from the project increases by $0.4 \times 1 = 0.4$ taler. However, this also means that each individual other group member's income increases by 0.4 taler, so that the total income of the group increases by $0.4 \times 4 = 1.6$ taler. The other group members therefore also earn something from your contribution to the project. On the other hand, you profit from the contributions made by the other group members. For each taler contributed to the project by another group member, you earn $0.4 \times 1 = 0.4$ taler. Hence, if each member of your group of four contributes 1 taler to the project, each of you receives $0.4 \times 1 \times 4 = 1.6$ taler as income from the project.

Stage 2

In Stage 2, you will see a screen requesting you to answer the following questions:

1. Do you believe that there is a general **norm** in your group on an **appropriate minimum contribution** to the project? (Yes/No)
2. If yes, how high can this minimum contribution be expected to be? (Number between 0 and 20).

Customary Law:

For new law to originate, it is not necessary that the legislator pass a statute, or that the parties agree on an explicit contract. Customary law is equally valid and binding. Customary law comes into being if the large majority of those affected for a sufficiently long period behave in a sufficiently similar way. The fact that some contribute even more to the joint project does not prevent that a norm of customary law originates. Hence a rule of customary law may prescribe a minimum standard. Customary law may originate here in the lab as well.

From the second period onwards, you will receive information on the behavior of individual group members in past periods. In order to receive this, you will have to click on an appropriate **button** on your screen. This can be done as often as you like.

- Button "**contributions**": how much have the individual group members contributed to the common project?

Stage 3

In Stage 3 you learn how much the other group members have contributed to the common project in this period. You then have the possibility of reducing the other players' income by **distributing points**. Each point you distribute to the other players costs you 1 Taler and reduces the other players' income also by 1 Taler.

Do you have any further questions?

Control Questionnaire

- 1.** Each group member has an endowment of 20 taler. Nobody (including you) contributes any taler to the project. What is:
 - a. Your income from the common project?
 - b. Your total income?.....

- 2.** Each group member has an endowment of 20 taler. You contribute 20 taler to the project. All other group members contribute 20 taler each to the project. What is:
 - a. Your income from the common project?
 - b. Your total income?.....

- 3.** Each group member has an endowment of 20 taler. You contribute 0 taler to the project. The other three group members contribute together a total of 30 taler to the project. What is:
 - a. Your income from the common project?
 - b. Your total income?.....

- 4.** Each group member has an endowment of 20 taler. You contribute 15 taler to the project. The other three group members contribute together a total of 5 taler to the project. What is:
 - a. Your income from the common project?
 - b. Your total income?.....

- 5.** After Stage 1 you have a total income of 30. Then you distribute 2 points to group member 1 and 3 points to group member 2. You also receive from the members of your group a total of 4 points. What is your total income now?