JENA ECONOMIC RESEARCH PAPERS

2011 – 018

The emergence of norms from conflicts over just distributions

by

Luis Miller
Heiko Rauhut
Fabian Winter

www.jenecon.de

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact markus.pasche@uni-jena.de.

Impressum:

Friedrich Schiller University Jena
Carl-Zeiss-Str. 3
D-07743 Jena
www.uni-jena.de

Max Planck Institute of Economics
Kahlaische Str. 10
D-07745 Jena
www.econ.mpg.de

The emergence of norms from conflicts over just distributions

Luis Miller[1], Heiko Rauhut[2], and Fabian Winter[3,†]

[1] CESS, Nuffield College
New Road, OX1 1NF, Oxford, Great Britain, Luis.Miller@nuffield.ox.ac.uk

[2] ETH Zurich, Swiss Federal Institute of Technology
Clausiusstrasse 50, CLU E6, 8092 Zurich, Switzerland, rauhut@gess.ethz.ch

[3] Max Planck Institute of Economics
Kahlaische Straße 10, 07745 Jena, Germany,
Phone: +49-3641-686-673, Fax: +49-3641-686-667 winter@econ.mpg.de

[†] corresponding author

April 4, 2011

7366 words, 5 figures, 3 tables

JEL-Classification: Z113, C91, D63
Keywords: social norms, normative conflict, bargaining, cooperation, experiment

Abstract

Why is it that well-intentioned actions can create persistent conflicts? While norms are widely regarded as a source for cooperation, this article proposes a novel theory in which the emergence of norms can be understood as a bargaining process in which normative conflicts explain the finally emerging norm. The theory is tested with a

dynamical experiment on conflicts over the consideration of equality, effort or efficiency for the distribution of joint earnings. Normative conflict is measured by the number of rejected offers in a recursive bargaining game. The emerging normative system is analyzed by feedback cycles between micro- and macro-level. It is demonstrated that more normative cues cause more normative conflict. Further, under the structural conditions of either simple or complex situations, the convergence towards a simple and widely shared norm is likely. In contrast, in moderately complex situations, convergence is unlikely and several equally reasonable norms co-exist. The findings are discussed with respect to the integration of sociological conflict theory with the bargaining concept in economic theory.

1 Introduction

Why is it that well-intentioned actions can create persistent conflicts? What is the reason that norm adherence is often not enough for the creation of social order and that even the overcoming of self-interest does not suffice for the establishment of lasting cooperative relations?

For example, wage schemes depend heavily on social norms employers and employees adhere to (Elster 1989). For some people, fixed, equal wages are appropriate, while others call for the consideration of individual effort (e.g. piece rate wages), or regard the individually created added value as important. Also combinations of these input and output measures are possible and can create normative conflicts among holders of more complex rules with those of simpler rules.

Also everyday examples support the notion that well-intended, cooperative actions can create normative conflicts: Consider a Ph.D. student has a common lunch with her supervisor and both want to signal their cooperative intentions. However, they may experience tensions or even conflicts if they want to split their bill based on different rules, be it an equal split, based on their individual consumption, status based (the professor pays), or even combinations of these rules (e.g. the professor pays the drinks but the meal is paid individually).

We believe that the *emergence* of norms should be understood as a process of conflict in which actors bargain over different rules regarding how mutual cooperation should be achieved. In this perspective, norms are typically subject to negotiations and continual bargaining. The emerging macro-level system of norms therefore heavily depends on the micro-level dynamics of how actors reach agreements concerning which norm ought to be followed. The inherent dynamical nature and the mechanisms of normative conflict have not been fully addressed yet in the literature on social norms.

The dynamics of normative conflict and the resulting emergence of normative systems can be illustrated by examples from the sociology of science: While most academic disciplines have agreed on what makes a "good" publication, two norms co-exist in sociology: publishing in books and publishing in peer-reviewed journals (Clemens et al. 1995). Publications have enormous influence on the field, as reputation, grant acquisitions, or getting a job largely rests on publication records. Especially in mixed genre departments, however, finding the right merit metric may lead to conflicts and prolonged faculty meetings due to conflicting opinions about what constitutes a good

way of communicating scientific results. Time-consuming bargaining can go on, wastes resources and persist, because actors adhere to different norms.

Similarly, multiple norms for receiving credit for authorship co-exist between and within different academic fields (Zuckerman 1968, Hudson 1996, Tscharntke et al. 2007). Alphabetic ordering, bracketing (with the main contributor as the first-author and the supervisor in the end), ordering based on merits (in a decreasing order based on relative contributions), or even multiple first-authors (with a footnote identifying all authors who contributed equally to the work). If norms clash and authors cannot agree on which norm should apply for the ordering of names, fruitful collaborations do not come about and even finished manuscripts can remain unpublished.

Our contribution investigates the *evolution* of norms from normative conflict. While previous analyses focused on demonstrating the existence of normative conflict (Winter, Rauhut, and Helbing forthcoming), it remains open how ongoing bargaining over norms at the micro-level leads to convergence of which norms at the population level. We develop a measurement of the persistence of normative conflict by using a bargaining experiment in which we introduce different normative cues, resulting in the adherence to different norms and expectations. Our design builds up on experimental work by Gantner, Güth, and Königstein (2001), Winter et al. (forthcoming) and Rauhut and Winter (2010). The novel experimental design proposed here allows to study the dynamical interplay between micro-level bargaining over norms and the emergence of normative systems at the macro-level.

2 A conflict theory of norm emergence

We illustrate our theory by referring to the example of normative conflict among co-authors. For the sake of simplicity, suppose two scientists co-author an academic article. Before submitting the article for publication, they have to decide who the first author of the paper should be. Let us further assume that the situation is ambiguous in the sense that there is no predominant norm in the field regarding the order of authors.

In such situations, several cues may be relevant. A *cue* is defined as a variable which determines the kind of behavior that is prescribed or proscribed in a norm-relevant situation. For example, if the co-authors are equal in every way, they may consider an alphabetical order of their names to signal equal contributions. However, if the co-authors differ in their effort, their

status or other variables, these cues may determine the order of their names (cf. Tscharntke et al. 2007).

A different number of considered cues may lead to the emergence of different norms. In this respect, we distinguish between less complex or *unconditional norms* and more complex or *conditional norms*. Unconditional norms do not require information about the context of the decision, for instance information about individual contributions to the article. A paradigmatic example is the equality norm (Winter et al. forthcoming). Equality norms aim at ex post equality, which would relate to an alphabetical or random ordering of authorship in our example. In contrast, conditional norms do require information about the context of the decision. An example of a conditional norm is the equity norm, which states that "the value of what a member of a group receives from other members should be proportional to his investments" (Homans 1961:237). In our example, credit for authorship would be received for effort or in other cases for status.

The process of normative conflict and the emerging normative system can be understood as feedback cycles between micro- and macro-level: A larger number of normative cues at the macro-level makes the situation more ambiguous, leading to more normative conflict at the micro-level which eventually results in the emergence of different norms at the macro-level. This mechanism is depicted in Figure 1.

Our idea is that the multiplicity of cues leads to heterogeneity in beliefs and in norms among individuals (the *meso level*). For example, co-authors who agree on a conditional norm could still have different beliefs whether to condition on status (seniority), on actual contributions (effort put in the article) or on the achieved level of improvement of the article (generated outcome from the effort). This heterogeneity in beliefs may lead to conflict at the meso level, which we call normative conflict. *Normative conflict* is thus defined as conflict due to the adherence to different normative beliefs. Of course, a larger number of normative cues does not necessarily lead to more ambiguity. It is particularly likely that normative conflict arises from ambiguous situations, if the actor's norms are not readily on display. Sometimes, we can quite reliably infer an individual's norms from specific signals, which may reduce the extent of normative conflict. A monk wearing a cowl signals rather unambiguously his religious norms, whereas a random person's beliefs may be harder to assess.

Normative conflict is instantiated at the meso level, i.e. among interacting individuals. We can understand the structure of conflict from bottom-up

by investigating the micro level. More precisely, we distinguish between initial beliefs and norms of individuals and their resulting actions. In our example, individuals may hold different beliefs regarding how much effort, status or improvement different co-authors accomplished and which of these cues are valid for the respective norm they hold.[1] The interplay of norms and beliefs then determines the ordering of names a person considers appropriate. Normative conflict emerges at the meso level if the co-authors reach different conclusions regarding which order of authors should be applied.

In a nutshell, normative conflict emerges in ambiguous situations in which multiple cues lead to heterogeneity in beliefs and actions at the individual level. If there is no commonly shared hierarchy of cues, individuals will not *a priori* agree on one focal norm and conflict emerges at the meso level. Therefore, our first hypothesis concerns the purported relationship between ambiguity of the situation and the extent of normative conflict.

> **Hypothesis 1** *The more normative cues, the stronger the normative conflict.*
>
> We conjecture that more normative cues will increase the strength of normative conflict as each cue triggers different norms. In our experiment, the number of normative cues is varied in the treatments and the strength of normative conflict is measured by the length of the negotiation (the number of bargaining rounds).

Normative conflict is essentially a dynamical concept and contributes to the understanding of the emergence of norms. In most situations, the bargaining process triggers an updating of the beliefs and norms. In our example, an initial conflict may be followed by a change of the co-authors' views regarding the appropriate order of their names. This is illustrated by the right-hand side of Figure 1. After conflict at the interaction level, beliefs and norms are updated at the micro level, and a new round of bargaining begins. Back again at the micro level, individuals decide once more. In our example they reconsider the order of co-authors. Theoretically, the process of beliefs-actions-updates could be infinitely repeated, but in the real world and in experiments there usually exists a "natural" end to this process.

[Figure 1 about here.]

We term the feedback loops between individual and interaction level as the *negotiation of norms* and the ones between interaction and population

level as *emergence of norms*. The relationship on the population level regarding how the ambiguity of the situation affects the complexity of the emerging norm can be understood by the dynamical negotiation of norms at the interaction and individual level (Figure 1).

Taking into account the previous bargaining process, parties may finally agree or disagree on which norm to follow. It is important to note that this process of normative negotiations does not necessarily end up in agreement. The long-term result of this process may also be persistent conflict. In contrast, if parties reach an agreement, a specific norm has emerged. We understand with norm emergence a process of convergence such that a sufficiently large number of interaction partners repeatedly update their beliefs and consequently adopt the same norm as the solution to their social interaction problem. Therefore, our hypothesis about the relationship on the population level considers how the ambiguity of the situation affects the complexity of the emerging norm. A large number of cues may lead to many conflicts and finally to a simpler emerging norm than a lower number of cues.

> **Hypothesis 2** *The more ambiguous the situation, the less complex the emerging norm.*
>
> We conjecture an inversely u-shaped relationship. More ambiguous situations trigger a larger extent of normative conflicts, which eventually results in less complex (unconditional) norms. In contrast, simpler situations trigger less conflict and allow the emergence of more complex (conditional) norms.

3 Application on distributive justice and definition of fairness norms

To test our two hypotheses about normative conflict, a laboratory study in the specific domain of distributive justice was conducted. At least since Homans (1961), sociologists have investigated the normative principles of distributive justice. Homans (1961) and a number of other social scientists (Adams 1965, Selten 1976, Cook and Hegtvedt 1983, Güth 1988) proposed an *equity norm*, for which "the received benefits of a group member should be proportional to her investments" (Homans 1961:237). The dilemma is that people "differ in their ideas of what legitimately constitutes investment, re-

ward, and cost, and how these things are to be ranked." (Homans 1961:246). This may create transaction failures when several parties disagree due to their different conceptions of what they consider as valid measures of equity or proportionality.

Recent experimental studies demonstrate that people adhere to different normative standards. Gantner et al. (2001) show that several equity standards are applied in bargaining situations. Frohlich, Oppenheimer, and Kurki (2004) find that, besides selfishness, two main behavioral norms determine individual decision making in distributive justice problems; *egalitarianism* and *equity*. Equity is by far the most often observed behavioral pattern in their experiments. Cappelen et. al (2007) report similar results. They propose a pluralism of fairness ideals, where the concept of equity is decomposed into the liberal egalitarian and the libertarian ideal. *Liberal egalitarianism* denotes input equity, for which actors are rewarded according to their previous effort. *Libertarianism* denotes output equity, for which actors are rewarded according to their achievements for the joint good, regardless of the effort with which this achievement was produced.

Our operationalization uses the set of normative principles proposed by Cappelen et al. (2007). More specifically, three fairness norms will be considered.

1. The *equality norm* states that the common good should be divided equally and actors' contributions to the good should be ignored. This ensures ex post equality of outcomes.

2. The *equity norm* demands that the common good should be distributed in proportion to actors' efforts into providing the common good. Equity refers to the input of efforts and not to actual contributions to the good.

3. The *libertarian norm* considers actual contributions to the good. This principle refers to output equity, meaning actors' contributions to the common good. Accordingly, the good should be distributed in proportion to participants' actual contributions, which can be a function of individual effort, but also of other factors for which the individual is not accountable for. Common examples include luck, genetic disposition, or parents' socio-economic status.

An unintended consequence of people following different normative standards at the micro level is *normative conflict* at the macro level. Actors can

have the best intentions and do their best, but nevertheless, their behavior can be perceived as improper. In this view, in a context where several standards are possible, social norms are not only promoting cooperation and social welfare, as previous theoretical and empirical literature has posited; they can also undermine it.

4 Method

4.1 Procedure and participants

The experiment was conducted using the *z-Tree* software developed by Fischbacher (2007). The subjects were 72 male and 116 female undergraduate students from a large European university, recruited from a wide range of academic disciplines with the online recruiting system ORSEE (Greiner 2004).[2] The average earnings per subject were 14.19 Euro ($\sigma^2 = 3.28$) and ranged from a minimum of 2.56 Euro to a maximum of 25.11 Euro, including a show-up fee of 2.50 Euro.

4.2 Experimental design

Our measurement of the persistence of normative conflict uses a recursive random bargaining protocol. It introduces a larger number of different *normative cues* as in Gantner et al. (2001) and Winter et al. (forthcoming). Our design consists of two stages. In stage 1, the subjects earn money in a real effort task, in stage two they bargain with changing partners about how to split the joint earnings.

4.3 The real effort task

At the beginning of each session, the subjects were randomly seated in cubicles with computer terminals. Some general instructions regarding the procedure were given on the computer screen (Schmelz 2010) and read out loud to ensure that everybody understood them and to demonstrate that everybody had received the same instructions. In the following real effort task, subjects had to answer 20 questions on a seven-page long text of a *Wikipedia* entry on the Westminster Palace, which they had received five days in advance by email.[3] There were five answer categories, one of which

was correct. For each correct answer, subjects had the opportunity to earn experimental currency units (ECU), which were later transformed into Euros at an exchange rate of 100 ECU = 1 Euro. An accompanying letter informed the subjects that their preparation of the text will influence their possible earnings in the experiment. Thus, every subject could decide on her own as to invest spare time in order to earn more money later on.

4.4 The bargaining game

In the bargaining part, the joint earnings of two randomly drawn players were pooled to form the pie. This procedure was designed to induce a feeling of personal effort and inherent monetary earnings. In particular, the effort was real in the sense that subjects could spend their own spare time.

After completing the quiz, the subjects received the instructions for the bargaining game experiment sketched in figure 2. In our game, two players bargained over several periods over the sum of money jointly produced in the quiz. In each period, both players could offer a share of the pie to the other player. At the same time, they entered an acceptance threshold, stating what the other player had to offer at least so that the offer was acceptable. As soon as both players submitted their respective offers and acceptance thresholds, the computer randomly chose one of the offers with equal probability. This offer was compared to the other player's threshold and was accepted if it was at least as high as the other player's threshold. In case of acceptance, the accepting player yielded the offer and the offering player received what was left on the joint account. If the offer was below the responder's threshold, the game proceeded to the next period and the pie shrinked by $2^{(t-1)}$ ECU in each period t. Note that the costs of delay are an increasing function of the length of bargaining, meaning that disagreement is cheap in the beginning and becomes increasingly expensive over time. Players continued this bargaining process until they either reached agreement or no more money was left on the account.

[Figure 2 about here.]

The bargaining game was repeated 10 times with one initial unpaid trial period with random matching in matching groups of 16 subjects. Before the beginning of the experiment, subjects had to answer test questions regarding the game. This allowed us to verify that the participants understood the

rules. For exclusion of wealth effects, one of the 10 potentially payment-relevant rounds was randomly selected for payment. The experiment started when there were no further questions to the experimenter.

4.5 Treatments and measures

For the measurement of the dynamics of normative conflict, we introduced cues which were designed to trigger norms of equity, equality or libertarianism.

The different normative cues were implemented in three different treatments, ranging from only one cue up to three cues. Equality was implemented in the C-treatment, where we transferred 1200 ECU for each player to the joint account. In this case, all norms prescribed an equal split. In the E-treatment, we implemented heterogeneity in efforts by composing the "pie" from the individual earnings in the quiz. Only if subjects were sensitive to equity or efficiency norms, they would condition their behavior on the other player's effort. In addition to heterogeneous efforts, the F-treatment implemented efficiency by randomly assigning subjects to different efficiency factors, which subjects kept throughout the experiment. These factors were used as multipliers of the subjects' contributions to the common pool, such that the contributions of half of the subjects in treatment F were tripled, while the other player's gains from the quiz were counted only once. In this respect, some subjects were more efficient than others. Here, only those subjects who were sensitive to the libertarian norm would take this factor into account, while subjects adhering to the equity norm would only rely on the other player's relative earnings from the quiz. Those interested in equality would split the pie equally. Table 1 gives an overview over the treatments and the respective composition of the pie.

[Table 1 about here.]

Offers and acceptance thresholds were submitted on a single screen, where the provided information depended on the treatment. In all treatments, subjects were informed about the remaining size of the pie, the initial transfers to the joint account, and a table describing the pie size for all bargaining rounds. In addition to that, the E-treatment provided the absolute and relative contributions from correct quiz answers of both players. In the F-treatment, subjects' also learned about the efficiency of their contributions,

i.e. their earnings from the quiz multiplied by the subjects' respective efficiency factors.

Subjects submitted their offers and acceptance thresholds with sliders on the computer screen, which automatically calculated and displayed the resulting absolute and relative distribution. After submitting the offers and thresholds, subjects proceeded to the information screen. This screen displayed both players' offers, the own acceptance threshold, which offer was selected and whether this offer was accepted.

Table **??** relates the theoretical constructs from figure 1 to the operationalization of the dependent and independent variables and to the statistics in the results section.

[Table 2 about here.]

5 Results

5.1 The bargaining process

Figure 3 gives an exemplary impression of the bargaining process. The figure depicts the decisions of two players (black and grey) until they reach an agreement in the F-treatment. Both players answered the same number of questions correctly, however, the black player's contribution was tripled, such that this player effectively put in 75 %, while the other player contributed only 25 %. A black/grey circle depicts the relative offer of the black/grey player towards the other player in shares of the pie, while a black/grey cross depicts the respective player's acceptance threshold.

[Figure 3 about here.]

The grey player offers and accepts around 50 % for quite some time, which corresponds to our definition of an equality player. The black player, on the other hand, adheres to the libertarian norm by offering only slightly less than 25 % and accepting about 75 %. This fits what the player effectively contributed to the pie.

We chose the display of this case due to the long duration of the bargaining process, which exemplifies severe normative conflict. No player wanted to give in first and their hassle would probably have continued if there would have been money left on their accounts after the 12th round.

As a manipulation check, we video-taped subjects in our video lab during a pre-test session of this experiment. The conversations among the bargainers often referred to fairness arguments and would become increasingly angry when the counterpart made unappropriate offers.

5.2 Consequences of normative conflict: Costly delays

Our hypothesis 1 states that the strength of normative conflict increases with the number of normative cues. Figure 4 gives a first "eye-ball"-test of this claim, plotting the number of rounds until agreement was reached for all treatments. In all treatments, a considerable fraction of interactions already ended in the first period because agreements were immediately reached. However, from period two onwards, the bold line representing treatment C is always above the two other lines, meaning that subjects in this treatment found an agreement faster than those in the other two treatments.

In addition to the cumulative density plot in figure 4, we estimated a random effects poisson regression model in order to test our claim more systematically (see the table in figure 4).[4] We chose a random effects model in order to adjust for correlated errors in multiple decisions of single individuals. Because rounds of bargaining are discrete count-data, we used a poisson model and not a linear random effects model.

[Figure 4 about here.]

The regression results generally confirm hypothesis 1. The strength of conflict (measured by bargaining rounds) is significantly higher in the F-treatment involving three cues compared to the C-treatment involving just a single cue ($t = 4.20$, $p < 0.001$ if controlled for subject specific errors and $t = 2.00$, $p < 0.05$ if controlled for subject and group). In addition, the difference between the E-treatment (two cues) and the C treatment is significant ($t = 2.30$, $p < 0.05$ and $t = 1.76$, $p < 0.10$, respectively). However, the difference between the E-treatment and the F-treatment is not statistically significant, meaning that the average time until subjects find an agreement is about equal in the two treatments.

5.3 Macro emergence of norms (mixture model)

Our second hypothesis states that the complexity of a situation has an inverse u-shaped effect on the emerging norm. To test this hypothesis, we

statistically extract the prevalent norms at the macro-level from micro-level data. We assume that the data are generated by three different types of players, corresponding to the fairness norms introduced above. By using this assumption, we can estimate the evolving population composition over time. Technically speaking, we estimate the proportion of the different types in a finite mixture model. These models have been originally introduced by Pearson (1894), but have only recently become popular by the increase in computational power.[5] In sociology and social psychology, these models have also been known as "latent class" or "structural equation" models (Harrison and Rutström 2009).

Our analysis is based on a simplified version of the random behavioral approach proposed by Conte and Moffatt (2010). To estimate the parameters of interest, we define utility functions for up to three different types of decisions and estimate the respective proportions of normative types $\lambda_j, j \in \{$*equality, equity, libertarian*$\}$ in the population.

As a first step, we formally define the three norms $N_i, i \in \{$*equality, equity, libertarian*$\}$ which we expect to find in the population. Let π denote the size of the pie, e_i subject i's relative effort, and ϕ_i subject i's efficiency factor. Than the respective norms N_k are given by

$$\begin{aligned} N_{equality} &= \frac{\pi}{2} \\ N_{equity} &= \frac{e_i}{e_i + e_j}\pi \\ N_{libertarian} &= \frac{\phi_i e_i}{\phi_i e_i + \phi_j e_j}\pi. \end{aligned} \tag{1}$$

Moreover, decisions are usually a balance between norm adherence and individual interests, meaning that normative expectations are often undercut. We formally account for that observation by introducing a global "norm-adherence term" δ, describing how much the average offer falls below the respective fairness norm.[6] Finally, decisions are often prone to little errors, which we capture by an error term $\epsilon_i \sim N(0, \sigma^2)$.

The desired share of the pie $\tilde{x}$ is a function of the player's norm, her

adherence to the norm and a random error term:

$$\begin{aligned}(\tilde{x}|N_{equality},\delta,\sigma) &= \frac{\pi}{2}+\delta+\epsilon,\\ (\tilde{x}|N_{equity},\delta,\sigma) &= \frac{e_i}{e_i+e_j}\pi+\delta+\epsilon,\\ (\tilde{x}|N_{libertarian},\delta,\sigma) &= \frac{f_i e_i}{f_i e_i+f_j e_j}\pi+\delta+\epsilon.\end{aligned} \tag{2}$$

As a matter of fact, our experimental design does not allow us to discriminate between all types in every treatment. Consequently, we can only estimate the share of those types we can clearly identify in a given treatment, resulting in different models for different treatments. From the different utility functions introduced above, we can construct the following models, where the likelihood contribution of a single decision x_i is given by:

$$L(\lambda_{equality},\lambda_{equity},\delta,\sigma|X)=\sum_{j=equality}^{g=equity}\lambda_j\frac{1}{\sigma}\phi\left(\frac{(\pi-x_i)-N_i-\delta}{\sigma}\right) \tag{3}$$

for the E- treatment with the two distinguishable types equity and equality, and

$$L(\lambda_{equality},\lambda_{equity},\lambda_{libertarian},\delta,\sigma|X)= \sum_{j=equality}^{g=libertarian}\lambda_j\frac{1}{\sigma}\phi\left(\frac{(\pi-x_i)-N_i-\delta}{\sigma}\right) \tag{4}$$

for the F-treatment, where we can identify the three types equality, equity and libertarians.

Given our experimental data matrix X, we can maximize the likelihood of the functions 3 and 4 and jointly obtain estimates for the population shares $\lambda_{equality},\lambda_{equity}$ (and $\lambda_{libertarian}$ in the F-treatment), as well es the norm-adherence δ and the variance of the errors σ.[7]

[Figure 5 about here.]

Figure 5 displays the offers and thresholds in the C treatment and furthermore the estimated mixing proportions of normative types λ_j in the

other two treatments. In the C-treatment, we can observe a convergence towards the equal split, both in offers and even more pronounced in acceptance thresholds. In the E-treatment, the equity and the equality norm co-exist throughout the whole experiment.[8] In the E-treatment, both norms for offer behavior co-exist without convergence over time. For acceptance thresholds, there is also co-existence of both norms. In the F-treatment, the complex situation leads to the emergence of a simple equality norm, as stated by hypothesis 2. Though already prevalent in the beginning, the estimated share of equality offers approaches roughly 80 % in the end. The same is true for acceptance thresholds: The equality line is above the equity line most of the time. Furthermore, it is constantly increasing from period four onwards.

To connect the two analyses in figures 4 and 5, we estimate the probability of severe conflict in a logit model. It follows from our theory that the convergence towards a single norm (as observed in treatments C and F, see figure 5) reduces the probability that the complete pie is lost after ongoing, persistent conflict. Although the number of these "worst-case" outcomes is small compared to the overall number of interactions, a respective random intercept logit model can still be estimated. This model adjusts for correlated errors within individuals and matching groups. In addition, we correct for the fact that one interaction consists of one interaction among two players, which means that our sample size is half compared to an individual-level analysis.

[Table 3 about here.]

The results show that there are no significant differences between treatment C and treatment F, but the likelihood of severe conflict is significantly higher in the E-treatment.[9] This confirms the robustness of our mixture model: The likelihood of conflict decreases if a population converges towards a norm. This means that the collective convergence to a shared norm reduces the costs of normative conflict. Such convergence can be easier achieved in simple or complex situations and is more difficult for moderately complex situations.

6 Discussion and open questions

We have outlined a new theory of norm emergence. Our perspective draws on the idea that the content of norms is negotiated in social interactions. If

actors base their behavior on different normative premisses, their interactions can generate serious transaction failures, which we refer to as normative conflict. Once an interaction does not work as expected and conflict emerges, "the right course of action" has to be negotiated in a dynamical process. Our theory introduces macro-level variables, like the complexity of a situation, in order to predict the content of the emerging norm. Therefore, our theory opens the black box of "path dependence", a term often referred to in the theoretical debate about the emergence of norms.

The concept of bargaining is essential in our theory of norm emergence. Our theory bridges bargaining concepts in economics and sociology. In economics, bargaining is a rigorous and therefore rather narrow concept which can be represented in simple formal models (Nash 1950, Rubinstein 1982). In sociology, the concept of bargaining is applied to a much greater variety of social phenomena, for instance, to the meaning of certain words (i.e. sociolects), rules of proper behavior or dress codes. Our general idea of norm emergence certainly refers to the broader sociological concept. Our theory can be generalized to a whole range of areas such as how we dress, talk, eat or even die. In order to test our hypotheses, however, we have to operationalize our theory into a tractable model of normative conflict. Therefore, we chose a specific formal bargaining model which can be experimentally tested.

The experimental results give a first corroboration of our proposed theory of norm emergence. Under the structural conditions of either simple or complex situations, the population converges towards a simple, widely shared norm. In contrast, in moderately complex situations, convergence is unlikely and several equally reasonable and prominent norms co-exist. The robustness of our findings should be further tested in the the lab and in the field: Which norms emerge, if the information about individual claims is unreliable or fuzzy? Does this change the bargaining process and if so, how? Our theory could be applied to very different kinds of norms and tested with a number of different research designs, which could substantiate our general theory.

Endnotes

[1]Note that two important forces can influence the decision. The *self-serving bias* can shape individual beliefs in a way that boosts own efforts and minimizes contributions of others (Babcock, Wang, and Loewenstein 1996). In a similar fashion, the *Thomas Theorem* (Merton 1995) highlights to focus rather on subjective beliefs than objective contributions

for the investigation of the roots of normative conflict.

[2]We ran 7 sessions with 3 sessions involving 32 subjects, two session involving 30 subjects, 2 sessions involving 16 subjects. Matching groups consisted of 16 subjects with two exceptions of matching groups of 14 subjects.

[3] Wikipedia contributors, "Westminster Palace," Wikipedia: The Free Encyclopedia, http://de.wikipedia.org/wiki/Palace_of_Westminster (accessed May 04,2008 14:40)

[4]All statistical tests were performed using STATA 10.1.

[5]See McLachlan and Peel (2000) for an introduction and a survey of the development of mixture models, and Aitkin and Rubin (1985) for an application to hypothesis testing.

[6]Conte and Moffatt (2010) call it the " selfishness premium".

[7]Non-convergence is a common problem for the estimation of mixture models with a numeric maximum-likelihood procedure if the proportion of one or more types in the population is small (i.e. $\exists \Lambda_i : \Lambda_i \rightarrow 0$). This is the case for the offers in period 9 and the acceptance thresholds in period 6 and 10 in treatment F (see the right side of figure 5). In this case, we restrict $\Lambda_{libertarian}$ to 0 and estimate equation 3 instead of equation 4. This is reflected in Figure 5 by disconnected symbols at the respective periods.

[8]Note that the share of "equity-offers" is sometimes as high as 76 % (period 1), or as low as 23 % (period 2).

[9]Note that the estimate of the E-treatment is significantly different from zero if we estimate a two-level model and only control for dependence of errors on the subject level. It is only marginally different from zero if we additionally control for the dependence of errors on the matching group level.

References

Adams, J. Stacy. 1965. "Inequity in Social Exchange." *Advances in Experimental Social Psychology* 2:267–269.

Aitkin, Murray and Donald B. Rubin. 1985. "Estimation and Hypothesis Testing in Finite Mixture Models." *Journal of the Royal Statistical Society. Series B (Methodological)* 47:pp. 67–75.

Babcock, Linda, Xianghong Wang, and George Loewenstein. 1996. "Choosing the Wrong Pond: Social Comparisons in Negotiations That Reflect a Self-Serving Bias." *The Quarterly Journal of Economics* 111:1–19.

Cappelen, Alexander W., Astri D. Hole, Erik Ø. Sørensen, and Bertil Tungodden. 2007. "The pluralism of fairness ideals: An experimental approach." *American Economic Review* 97:818–827.

Clemens, E.S., W.W. Powell, K. McIlwaine, and D. Okamoto. 1995. "Careers in print: Books, journals, and scholarly reputations." *American Journal of Sociology* 101:433–494.

Conte, Anna and Peter G. Moffatt. 2010. "The econometric modeling of social Preferences." Jena Economic Research Papers in Economics 2010-042, Friedrich-Schiller-University Jena, Max-Planck-Institute of Economics.

Cook, Karen S. and Karen A. Hegtvedt. 1983. "Distributive Justice, Equity, and Equality." *Annual Review of Sociology* 9:217–241.

Elster, Jon. 1989. *The Cement of Society: A Study of Social Order*. Cambridge: Cambridge University Press.

Fischbacher, Urs. 2007. "Ztree- Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics* 10:334–352.

Frohlich, Norman, Joe Oppenheimer, and Anja Kurki. 2004. "Modeling Other-Regarding Preferences and an Experimental Test." *Public Choice* 119:91–117.

Gantner, Anita, Werner Güth, and Manfred Königstein. 2001. "Equitable choices in bargaining games with joint production." *Journal of Economic Behavior & Organization* 46:209–225.

Greiner, Ben. 2004. "An Online Recruitment System for Economic Experiments." In *GWDG Bericht 63*, edited by Kurt Kremer and Volker Macho, pp. 79–93. Göttingen: Ges. für Wiss. Datenverarbeitung.

Güth, W. 1988. "On the behavioral approach to distributive justice: A theoretical and experimental investigation." *Applied behavioral economics* 2:703–717.

Harrison, Glenn and E. Rutström. 2009. "Expected utility theory and prospect theory: one wedding and a decent funeral." *Experimental Economics* 12:133–158. 10.1007/s10683-008-9203-7.

Homans, George Caspar. 1961. *Social Behavior*. New York: Harcourt, Brace & World.

Hudson, John. 1996. "Trends in Multi-Authored Papers in Economics." *The Journal of Economic Perspectives* 10:pp. 153–158.

McLachlan, Geoffrey J. and David Peel. 2000. *Finite mixture models*. Wiley-Interscience.

Merton, Robert K. 1995. "The Thomas Theorem and the Matthew Effect." *Social Forces* 74:pp. 379–422.

Nash, John F. 1950. "The Bargaining Problem." *Econometrica* 18:155–162.

Pearson, Karl. 1894. "Contributions to the Mathematical Theory of Evolution." *Philosophical Transactions of the Royal Society of London. (A.)* 185:71–110.

Rauhut, Heiko and Fabian Winter. 2010. "A sociological perspective on measuring social norms by means of strategy method experiments." *Social Science Research* 39:1181–1194.

Rubinstein, Ariel. 1982. "Perfect Equilibrium in a Bargaining Model." *Econometrica* 50:97–109.

Schmelz, Katrin. 2010. "E-nstructions: Using electronic instructions in laboratory experiments." *Jena Economic Research Papers* p. 047.

Selten, Reinhard. 1976. "The Equity Principle in Economic Behavior." In *Decision Theory and Social Ethics*, edited by Hans-Werner Gottinger and Werner Leinfellner, Theory and Decision Library, pp. 289–302, Dordrecht. D. Reidel Publishing Company.

Tscharntke, Teja, Michael E. Hochberg, Tatyana A. Rand, Vincent H. Resh, and Jochen Krauss. 2007. "Author sequence and credit for contributions in multiauthored publications." *PLoS biology* 5.

Winter, Fabian, Heiko Rauhut, and Dirk Helbing. forthcoming. "How norms can generate conflict: An experiment on the failure of cooperative micro-motives on the macro-level." *Social Forces* .

Zuckerman, Harriet A. 1968. "Patterns of Name Ordering Among Authors of Scientific Papers: A Study of Social Symbolism and Its Ambiguity." *The American Journal of Sociology* 74:pp. 276–291.

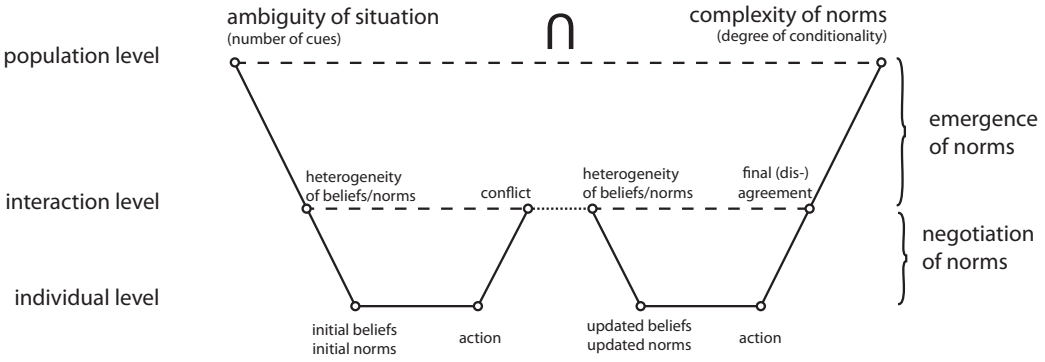


Figure 1. A multilevel theory of the dynamics of normative conflict

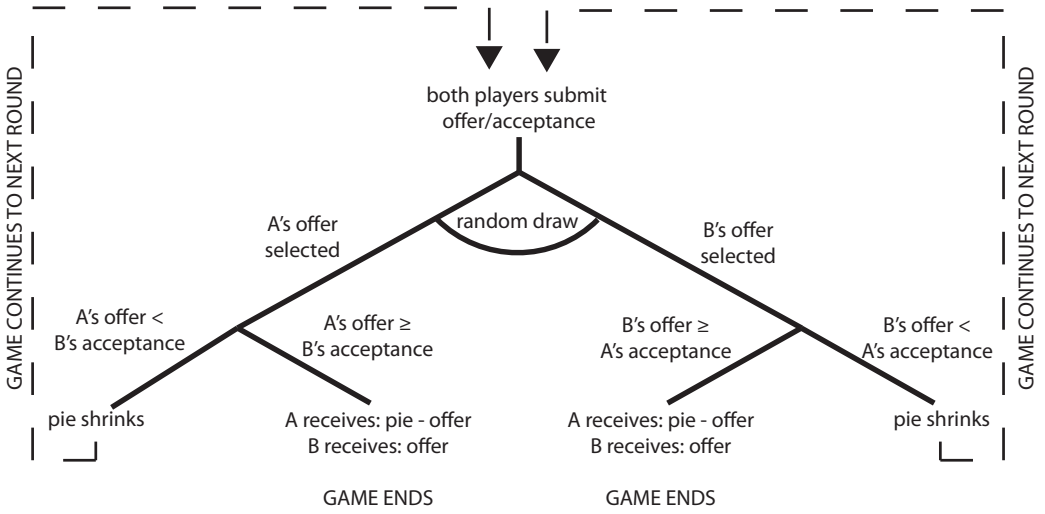


Figure 2. The bargaining game among two players A and B

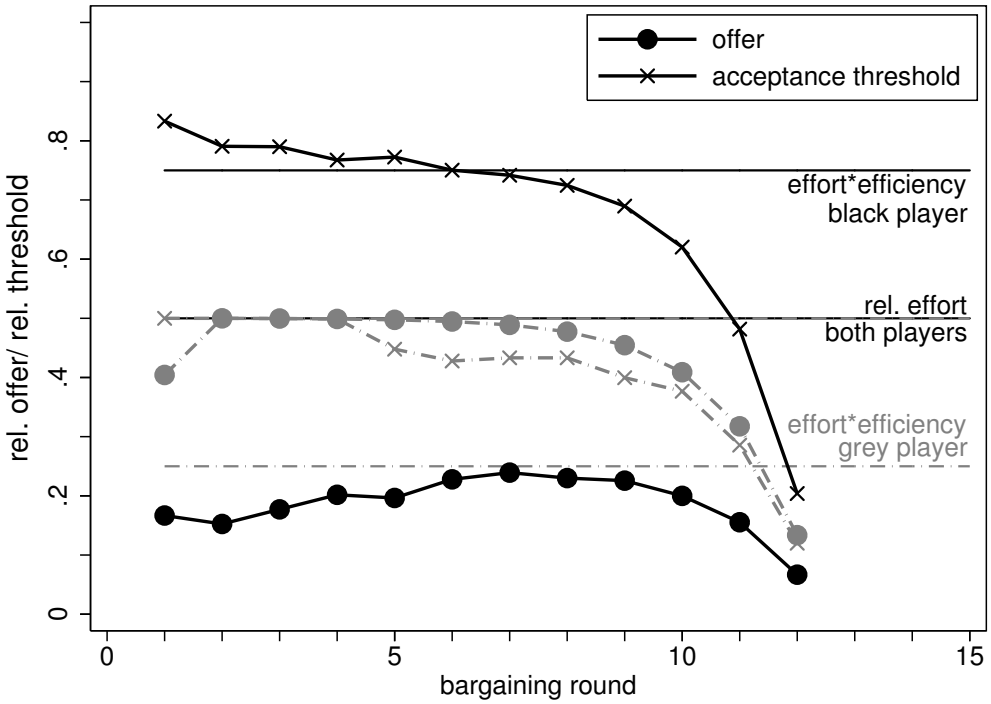
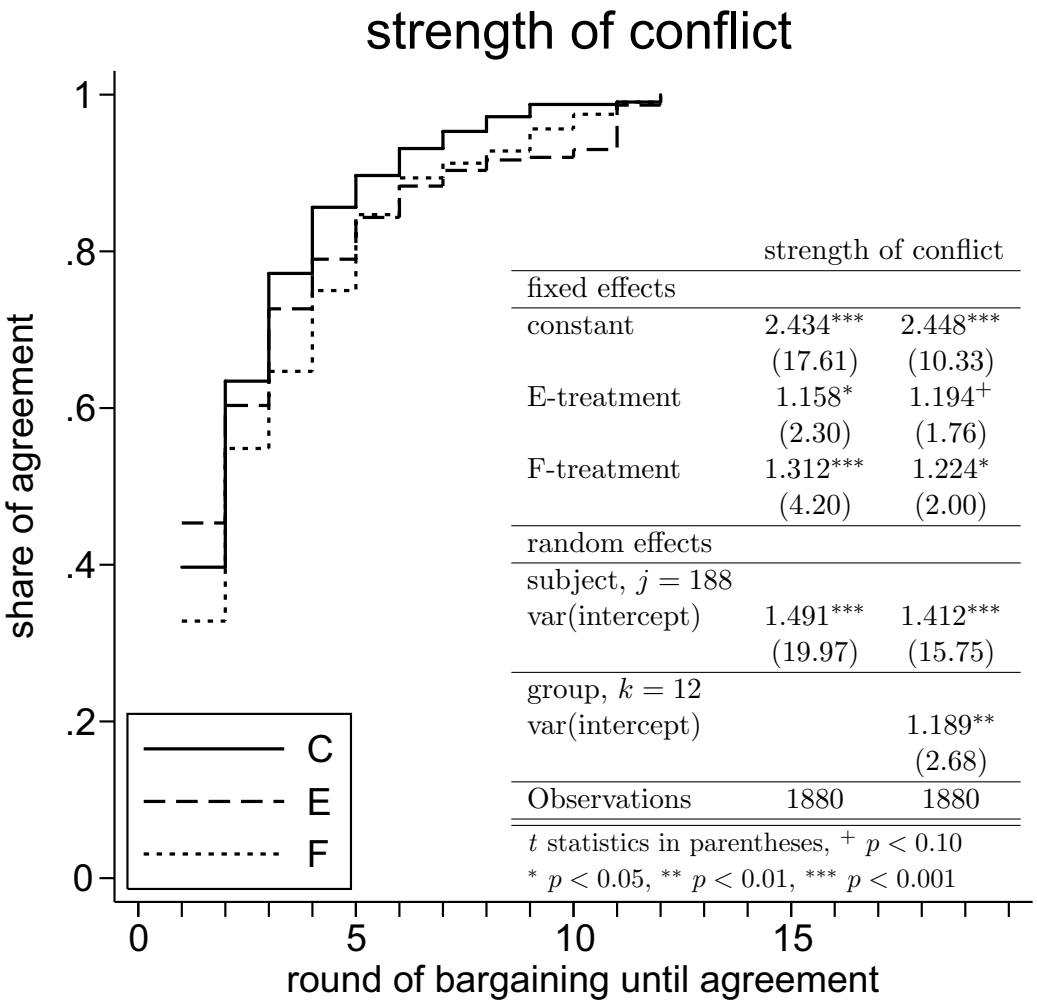


Figure 3. An exemplary interaction between two players (grey and black). The dot (•) represents the respective player's relative offer, the cross ($\times$) the relative acceptance threshold.



	strength of conflict	
fixed effects		
constant	2.434^{***}	2.448^{***}
	(17.61)	(10.33)
E-treatment	1.158^{*}	1.194^{+}
	(2.30)	(1.76)
F-treatment	1.312^{***}	1.224^{*}
	(4.20)	(2.00)
random effects		
subject, $j = 188$		
var(intercept)	1.491^{***}	1.412^{***}
	(19.97)	(15.75)
group, $k = 12$		
var(intercept)		1.189^{**}
		(2.68)
Observations	1880	1880

t statistics in parentheses, $^{+}$ $p < 0.10$
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 4. Cumulated density plot of the strength of conflict in different treatments. The horizontal lines represent the share of interactions in which an agreement has been found in the respective period displayed on the X-axis. The table shows two random-effects poisson models. Model (1) controls for subject specific errors, model (2) additionally for group specific errors. They demonstrate a significantly higher strength of conflict in the F-treatment ($p < 0.001$ and $p < 0.05$, respectively) and in the E-treatment ($p < 0.05$ and $p < 0.10$, respectively) compared to the C-treatment. The dependent variable strength of conflict is measured by rounds of bargaining.

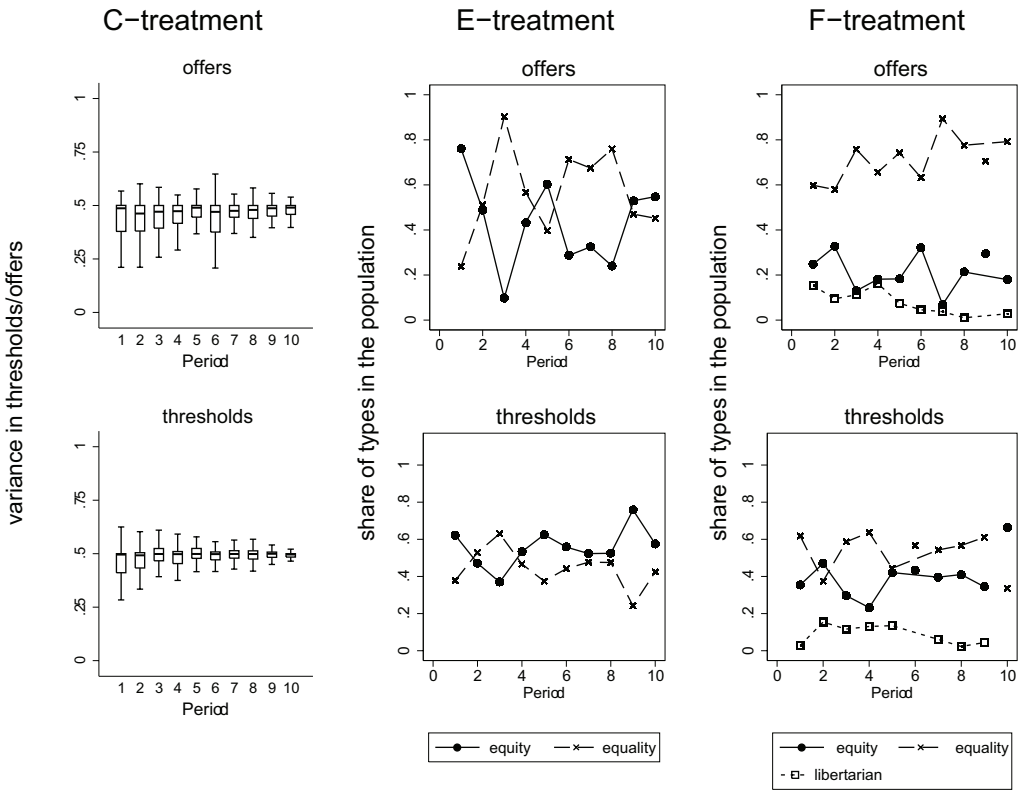


Figure 5. Distribution of different types of proposer decisions (top) and responder decisions (bottom) as estimated in model 3 and 4. The horizontal axis denotes the 10 repetions of the bargaining game, the vertical axes the relative share of types in the population. Types in the C-treatment are not estimated but decisions are represented by box plots (outliers omitted). The model assumes two types in the E-treatment (equity and equality), and three types in the F-treatment (equity, equality and libertarian). In the treatments C, behavior converges towards the equal split. In the E-treatment, there is co-existence of equity and equality norms. In the F-treatment, the equality norm becomes increasingly prominent over time and the libertarian norm vanishes.

treatment	pie	ECU per answer	efficiency factor	pie size
C (control)	given	10	–	2400
E (effort)	produced	100	–	$e_1 + e_2$
F (efficiency)	produced	50	$\phi_l : 1$, $\phi_h : 3$	$\phi_l e_1 + \phi_h e_2$

	normative cue		
	egalitarian	equity	libertarian
C (control)	X		
E (effort)	X	X	
F (efficiency)	X	X	X

Table 1. Overview of treatments and number of normative cues in a given treatment. The number of normative cues increases from C to E to F. ϕ_l refers to the low efficiency factor and ϕ_h to the high efficiency factor.

theoretical construct	operationalization	statistics
ambiguity of situation	treatment variables (number of cues)	(not applicable)
heterogeneity of norms	heterogeneity of offers/thresholds	mixture-model (figure 5)
initial beliefs initial norms	not measured directly	evidence from the video lab
actions	offer / threshold	exemplary interactions (figure 3)
normative conflict	rounds of bargaining until agreement	cdf-plot, random effects poisson model (figure 4)
final (dis-)agreement	no agreement until the end	logit model (table 3)
complexity of prevailing norm	offer/threshold	mixture model (figure 5)

Table 2. Overview of theoretical constructs, their operationalization and corresponding statistics

	no solution	
fixed effects	(1)	(2)
constant	-5.686***	-5.573***
	(-6.67)	(-5.65)
E-treatment	2.016*	1.799+
	(2.48)	(1.70)
F-treatment	-0.168	-0.0510
	(-0.16)	(-0.04)
random effects		
subject, j=94		
var(intercept)	1.545***	1.135*
	(3.51)	(2.24)
matching group, k=12		
var(intercept)		0.960*
		(2.04)
Observations	940	940

t statistics in parentheses, $^{+}$ $p < 0.10$

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3. Random effects logistic regression. The dependent variable represents whether a bargaining couple did not find a solution in the bargaining process until the end where all money was gone due to bargaining costs. Model (1) controls for subject specific errors, model (2) additionally for group specific errors.