More Order with Less Law: On Contract Enforcement, Trust, and Crowding

IRIS BOHNET *Harvard University*
BRUNO S. FREY *University of Zürich*
STEFFEN HUCK *University of London*

Most contracts, whether between voters and politicians or between house owners and contractors, are incomplete. "More law," it typically is assumed, increases the likelihood of contract performance by increasing the probability of enforcement and/or the cost of breach. We examine a contractual relationship in which the first mover has to decide whether she wants to enter a contract without knowing whether the second mover will perform. We analyze how contract enforceability affects individual performance for exogenous preferences. Then we apply a dynamic model of preference adaptation and find that economic incentives have a nonmonotonic effect on behavior. Individuals perform a contract when enforcement is strong or weak but not with medium enforcement probabilities: Trustworthiness is "crowded in" with weak and "crowded out" with medium enforcement. In a laboratory experiment we test our model's implications and find support for the crowding prediction. Our finding is in line with the recent work on the role of contract enforcement and trust in formerly Communist countries.

Trust can increase efficiency in the economic and political spheres. Recent studies using aggregate data suggest the existence of an efficiency-enhancing feature of trust for countries and organizations.[1] We attempt to provide a microfoundation for some of these findings by investigating whether trustworthiness can have an economic payoff at the individual level. Important domains for trust and trustworthiness include the relationship between representatives and their constituents, such as between politicians and voters, managers and shareholders, or attorneys and clients. In all these situations, principals have to decide whether they want to enter a contract that they know will be incomplete (i.e., agents might have an incentive to breach). Offering the contract is a matter of trust, and performing it, a matter of trustworthiness.

The problem of trust is more pronounced in large, anonymous societies than in small groups. In the latter case, participants frequently interact, and reputation matters. Therefore, according to the folk theorem type of argument, cooperation can be sustained even in the absence of genuine trustworthiness.[2] This kind of argument fails in the case of large groups, and institutions such as the law are needed to facilitate efficient outcomes.[3] The law, however, may affect behavior not only by creating incentives but also by influencing preferences. Whereas rational choice theory focuses on the first aspect, we propose a model that integrates both. We analyze how the enforceability of a contract affects individual performance in the short run with given preferences and in the long run when preferences adapt to the new environment. Our analysis builds on an evolutionary approach.[4] We present analytical results and test their implications in a laboratory experiment.

A contractual relationship is represented by a game in which the first mover has to decide whether she wants to enter a contract without knowing whether the second mover will perform. If the second mover breaches, a chance move decides whether he is held liable for the cost of breach. Standard economic analysis of law predicts that the higher the expected cost of breach, the more likely is the second mover to perform.

Iris Bohnet is Assistant Professor of Public Policy, Kennedy School of Government, Harvard University, Cambridge, MA 02138 (Iris_Bohnet@Harvard.edu). Bruno S. Frey is Professor of Economics, Institute for Empirical Economic Research, University of Zürich, Bluemlisalpstreet 10, 8006 Zürich, Switzerland (bsfrey@iew.unizh.ch). Steffen Huck is Senior Lecturer in Economics, Department of Economics, Royal Holloway College, University of London, Egham, Surry TW20 0EX, United Kingdom (s.huck@rhbnc.ac.uk).

The authors thank Linda Babcock, Sam Bowles, Uwe Dulleck, Ernst Fehr, Simon Gächter, Werner Güth, Ernan Haruvy, Owen Jones, Michael Mandler, Jörg Oechssler, Elinor Ostrom, Fred Schauer, Richard Zeckhauser, and three anonymous referees, as well as participants of the 1998 American Law and Economics Association meetings at Berkeley, the 1999 Economic Science Association meetings at Lake Tahoe, the 1999 Symposium on Behavioral Economics at the Levy Institute, the 1999 SPUDM meetings in Mannheim, the 2000 Public Choice Association meetings in Charleston, the 2000 meeting of the MacArthur Foundation Network on Economic Environments and the Evolution of Individual Preferences and Social Norms in Chicago, and seminars at Harvard, Royal Holloway, and the University of Zürich for many helpful comments and suggestions. Further thanks are due to Dorothea Kübler for her help in conducting the experiments, to Joanna Veltri for her editorial support, and to Michael Zürcher for his legal advice. Financial support of the German and the Swiss National Science Foundation, the Kennedy School of Government Dean's Fund, and the Russell Sage Foundation are gratefully acknowledged.

[1] See, for example, Fukuyama 1995; Knack and Keefer 1997; La Porta et al. 1997; or Putnam 1993.

[2] Folk theorems say that in infinitely repeated games any feasible payoff combination can be supported by an equilibrium.

[3] See, for example, Greif, Milgrom, and Weingast 1994 and Milgrom, North, and Weingast 1990. For a proof of the folk theorem in the prisoner's dilemma with large groups and anonymous interaction, see Ellison 1994. The behavioral relevance of repetition and anonymity have been studied in many laboratory experiments, e.g., Andreoni 1988 and Bohnet and Frey 1999. For a survey, see Ostrom 1998, who also discusses how experimental results relate to political science.

[4] See, for example, Axelrod 1984; Becker 1976; Bendor and Swistak 1997; Bowles 1998; Boyd and Richerson 1985; or Güth and Kliemt 1999.

We show that, when preferences are subject to change, this need not be the case. More specifically, we find that the probability with which a contract is enforced has a nonmonotonic effect on behavior: Performance rates of second movers are high not only when the expected cost of breach is sufficiently large but also when it is sufficiently small.

We focus on preferences for contract performance and assume that individuals may experience psychological costs when they breach. Such trustworthy people are said to have a preference for honesty.[5] Based on the idea that a specific preference is more likely to be maintained and to flourish if it proves to be economically successful, we study a dynamic process in which preferences can change over time. Legal rules can "crowd in" as well as "crowd out" preferences. We find that intermediate levels of contract enforcement lead to crowding out, but low levels induce crowding in. The intuition for this is rather straightforward. Suppose a first mover must decide whether to enter a contract. If she knows that the legal system is inefficient, that is, contracts are rarely enforced, she will be extremely cautious. Clearly, she would like to enter if she knew the other party could be trusted. If a signal is received about the partner's trustworthiness, it must be "sufficiently good." This caution not only protects first movers from being exploited too often but also makes trustworthy second movers more successful than others, because on average they will get more contracts than others. Hence, honesty will be crowded in.

Contractual relationships with weak enforcement are typical in many organizational settings. Some firms purposely create a low enforcement environment in which interactions are not guided by the expected cost of breach but by intrinsic motivation. At the same time, they heavily invest in screening of potential employees, stressing that character is more important than the possession of specific skills.[6] Similarly, most microfinance institutions (e.g., Grameen Bank or Accion) that lend money to poor clients without physical collateral focus on "character-based lending." In the absence of external enforcement mechanisms, the intrinsic trustworthiness of clients is a key variable that makes the contract between borrower and lender possible (Murdoch 1999).

The same pattern applies to many other domains: The more leeway agents have—whether these are employees, borrowers, legislators, judges, or executives—the more careful are principals when deciding who will be offered a contract. That the leeway for politicians can be considerable becomes clear in Rose-Ackerman's (1999) analysis of corruption. She points out that even in the United States the law is not strict enough to deter elected officials from being corrupt. "The criminal penalties are 'not more than three times the monetary equivalent of the thing in value (i.e., the bribe) or imprisonment for not more than fifteen years, or both' (18 USC § 201 (a)). This is appropriate for officials who receive bribes except that multiplying by three may be a poor measure of the risk of detection and punishment. The actual probability of catch is likely to be well below one-third" (p. 55). Whether this probability is low enough to induce crowding-in is an empirical question. If the probability of contract enforcement is higher but not high enough to deter all second movers from breaching, then we expect crowding-out.

Under medium enforcement, the expected payoff of entering is higher than the payoff of abstaining, even if the first mover knows that the contract will be breached. Accordingly, she will enter regardless of her beliefs about the partner's trustworthiness. This unconditional trust makes second movers who want to maximize expected monetary payoffs more successful than honest types who forsake profitable opportunities to breach. Therefore, honesty will be crowded out, causing the aforesaid nonmonotonicity. In sum, under high levels of enforcement, all second movers perform because they are deterred regardless of their preferences, and all first movers enter the contract; preferences are irrelevant and outcomes are efficient. At intermediate levels, honesty is crowded out; more second movers breach, and resources are wasted in trials. At low levels, trustworthiness is crowded in; more second movers perform even though they would have an incentive to breach without a preference for honesty, and efficiency increases.

Contract enforcement probabilities that are too small to deter breach may be due to a badly functioning legal system with weak state protection, corrupt governments and judiciaries, or high enforcement costs. So far, the literature has focused mainly on how informal institutions, such as social norms, may substitute for an ineffective legal system and whether shame and ostracism can replace imprisonment and fines.[7] Our model focuses on formal law and intrinsic dispositions and shows how the effectiveness of each depends on the other. By providing a specific legal enforcement regime, the state affects the degree of trust and trustworthiness in a society. By including formal institutions in the analysis we address an aspect rarely examined in the current debate on trust and social capital.[8]

[5] We are purposely vague here for two reasons. Our model does not depend on the specifics of the psychological costs incurred and our experiment does not test for different kinds of such costs. Our results are compatible with intrinsic motivation (see, e.g., Frey 1997), inequality aversion (see, e.g., Bolton and Ockenfels 2000 or Fehr and Schmidt 1999), reciprocity (see, e.g., Rabin 1993), and psychological contracts (see, e.g., Rousseau 1995).

[6] For a summary of such high-commitment human resource management practices, see Baron and Kreps 1999.

[7] Social norms confine minor crimes, such as trespassing (Ellickson 1991), or the overuse of common pool resources (Ostrom, Gardner, and Walker 1994). For the legal debate about "alternative sanctions" see Kahan 1996, and for a general discussion of how social norms and the law interact, see Cooter 1998 or Sunstein 1996.

[8] Tarrow (1996, 395) asks: "Can we be satisfied interpreting civic capacity as a home-grown product in which the state has played no role?" Schneider et al. (1997) are among the few who discuss the influence of institutions, namely, the extent to which parents can choose a public school, on social capital. Those interested in this influence mainly focus on the relationship between institutions and trust rather than between institutions and trustworthiness (see, e.g., Brehm and Rahn 1997; Norris 1999; and Nye, Zelikow, and King 1997).

Our findings may help understand two tendencies observed in many countries of the former Soviet block. On the one hand, there is a demand for "more law" in order to enforce contracts and secure property rights. When the state cannot provide levels of enforcement high enough to deter breach, the demand for protection is satisfied privately. This is one explanation for the rise of the Mafia in Sicily (Gambetta 1993) and may account for its thriving in Russia (Varese 1994). On the other hand, a reemergence of the demand for and supply of trust-based relationships can be observed in the very same countries. Wintrobe (1995, 46–7) writes: "The absence of enforceability generates a demand for trust. The costs of trust formation are lower, when the two parties share common traits, such as a common language, ethnicity, and so on." His analysis differs from ours, but his conclusions are similar: More order can be achieved by relying on trust-based relationships when each party can predict the other's likelihood of cooperation.

We suggest a rationale for the coexistence of these two tendencies. If trustworthiness has been crowded out, people cannot help asking for "more law," and if it has been crowded in, they can rely on trust-based interactions. According to Simis (1982) and Varese (1994), the former Soviet system was characterized by corruption affecting only specific sections of the population, namely, the "nomenclatura." In the absence of a credible state, these groups are unable to engage in trust-based interactions and, thus, demand private protection. For ordinary people, trust was and remains the basis of their contractual relationships (Wintrobe 1995). Overall, our model predicts a pattern encapsulated in a Latin American quip: For friends everything, for enemies nothing, for the stranger the law.[9] The model's empirical validity is tested in a laboratory experiment that implements the theoretical setup as closely as possible, and the results support the crowding predictions.

The article is organized as follows. We first describe how agents behave with fixed preferences. Next, we analyze the crowding dynamics and discuss the model's main implications. We then present the design and the experimental results. In the final section we offer conclusions.

THEORY

The Contract Game

We model a situation in which two players have the opportunity to produce a joint surplus, but the players are asymmetrical. Player 1 has to enter the contract without knowing whether player 2 will perform. Therefore, player 1's decision to enter is a matter of trust.[10] We denote her trusting move with T and her nontrusting move with $\bar{T}$. In the case where she trusts, player 2 can perform (move P) or breach ($\bar{P}$). The game ends either if player 1 does not enter, which yields zero payoffs for both parties, or if player 1's trust is rewarded by player 2's cooperation. This yields payoffs of 1 for both players.[11] If player 2 breaches, we assume a final chance move that captures a litigation process, with probability p that 2 will be held liable (move L). The surplus is divided, as in the case of performance, but the loser has to bear the costs of the trial $c > 0$. In other words, we assume that perfect expectation damages place player 1 in the position in which she would have been if player 2 had performed and that all legal fees are paid by player 2, the loser.[12] Thus, the payoffs are 1 for player 1 and $1 - c$ for player 2. If player 2 is not held liable ($\bar{L}$), he profits from breach and receives a payoff of $1 + b$ (with $b > 0$); player 1 bears the legal cost and is not compensated for any investments she made by entering the contract, so her payoff is $-a$ with $a \geq c$.

Breach is never efficient. The benefits from it are not large enough to compensate the first mover, that is, $b < 1 + a$. Figure 1 shows this game in its extensive form.

We assume that all payoffs are monetary, but in order to solve the game we need utilities associated with the various outcomes. To map outcomes into (cardinal) utilities, we assume two possible preferences for players. One type of player (M) is only interested in (expected) monetary payoffs, so for this type the monetary payoffs in Figure 1 represent utilities. The second type of player (H) has a preference for honesty and suffers a psychological cost when breaching a contract. These costs are so high that H types will never betray regardless of the monetary gain b.[13]

With this set of possible preferences $\{M, H\}$, we can transform the "money game" of Figure 1 into a standard game in which the payoffs are von Neumann-Morgenstern utilities. This is done by replacing the payoff of player 2 after path $T\overline{PL}$ by $1 + b - \delta$, where $\delta = 0$ for type M, and $\delta > b$ for type H.

Assuming that player 1 recognizes whether player 2 is M or H, we have, for each possible match of players, a well-specified standard game.[14] We solve this game

[9] "A los amigos todo, a los enimigos nada, al extraño la ley" (Rose-Ackerman 1999, 97).

[10] For related games on trust, see Berg, Dickhaut, and McCabe 1995; Burnham, McCabe, and Smith 1999; Glaeser et al. 2000; Güth and Kliemt 1999; or Van Huyck, Battalio, and Walters 1995.

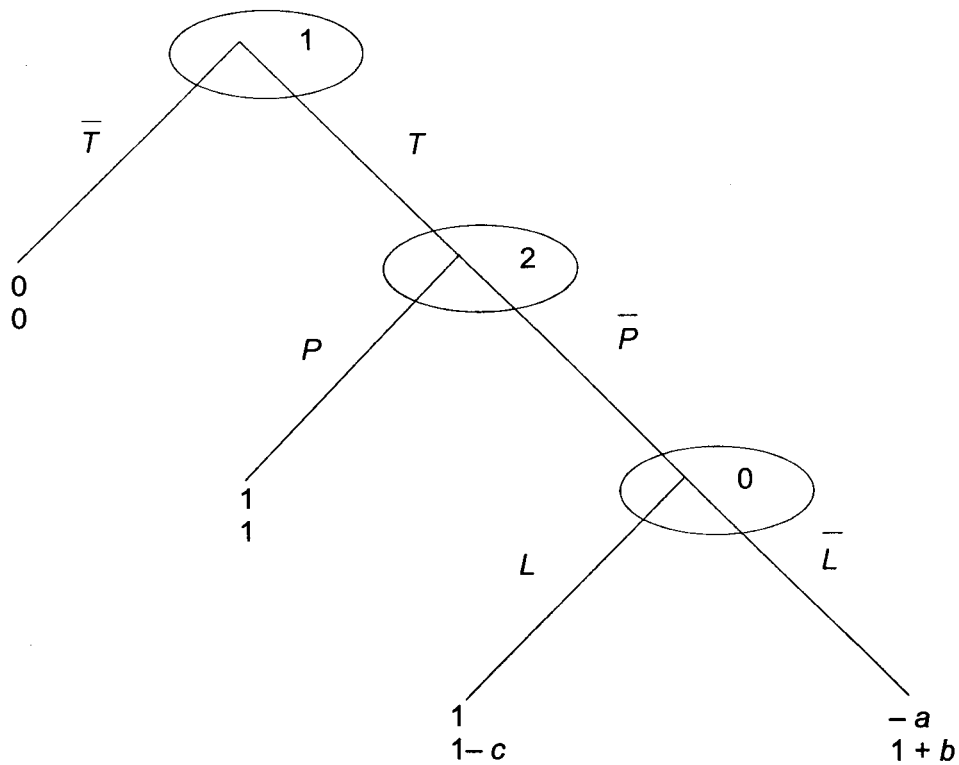
[11] This specification of the payoffs is a simple normalization without loss of generality.

[12] This corresponds to the English legal cost allocation rule.

[13] This assumption simplifies the analysis without altering the results, all of which would still hold if we allowed different levels of psychological cost. Notice that in the case of a continuous space—possibly ranging from infinite (psychological) costs to infinite gains—all that matters is whether the costs are larger or smaller than the monetary gain b.

[14] If player 2 is type M, the game is identical to that shown in Figure 1. If player 2 is type H, the payoff after path $T\overline{PL}$ is $1 + b - \delta < 1 + b$.

FIGURE 1. The Contract Game with Monetary Payoffs



Note: Player 1 chooses between T (her trusting move) and $\overline{T}$ (her nontrusting move). After T, player 2 chooses between P(erformance) and nonperformance ($\overline{P}$). After $\overline{P}$, chance decides whether player 2 is held L(iable) or not ($\overline{L}$). The end nodes show the two players' respective payoffs.

by backward induction.[15] Obviously, player 2 will breach if his expected payoff exceeds 1, that is, if

$$p(1 - c) + (1 - p)(1 + b - \delta) > 1. \qquad \textbf{(1)}$$

For type H this is never fulfilled, because a player 2 with a preference for honesty will always choose P. For type M, we can insert $\delta = 0$ into inequality 1 and rewrite it as

$$p < \frac{b}{b + c}. \qquad \textbf{(2)}$$

Player 2 of type M breaches if the probability of enforcement is smaller than the benefit of breach divided by the benefit plus the legal cost. Next, consider the decision of player 1. If she is confronted with an H type, she will surely trust. The same is true if she is confronted with an M type and

$$p > \frac{b}{b + c}$$

holds. But if she is confronted with an M type and (2) is fulfilled, she will enter only if her expected payoff exceeds her outside-option payoff, that is, if $p - a(1 - p) > 0$. This can be rearranged as

$$p > \frac{a}{1 + a}. \qquad \textbf{(3)}$$

From this a proposition follows.

PROPOSITION 1. *The (unique) subgame perfect equilibrium (SPE) is (T,P) if player 2 is of type H or if player 2 is of type M and*

$$p > \frac{b}{b + c};$$

this is the high-probability regime [High p]; (T,$\overline{P}$) if player 2 is of type M and

$$\frac{a}{1 + a} < p < \frac{b}{b + c},$$

called the medium-probability regime [Medium p]; and ($\overline{T}$, $\overline{P}$) if player 2 is of type M and

[15] Admittedly, the assumption of perfect type recognition is strong, but the results hold as long as players receive sufficiently informative signals about the opponent's type. This is also discussed later and illustrated in Appendix A. The main effect of assuming perfect signals is that it simplifies notation and analysis significantly. Recent experimental evidence supports the notion that type recognition is possible after preplay communication in a prisoner's dilemma game (Frank, Gilovich, and Regan 1993; Orbell and Dawes 1991). Kikuchi, Watanabe, and Yamagishi (1995) found that trustworthy types are better at predicting others' types than are nontrustworthy types. More generally, see Frank 1988.

$$p < \min\left\{\frac{a}{1+a}, \frac{b}{b+c}\right\},$$

the low-probability reigme [Low p].

We shall assume below that $ac < b$. Otherwise $(T,\overline{P})$ would never be an SPE, and the crowding analysis would be less rich and less interesting. Imposing this requirement means, informally speaking, that the loss player 1 incurs from uncompensated betrayal must be relatively small in comparison to the profit of player 2 and the legal costs.

Crowding

We have shown how individuals with given preferences behave under different legal regimes, and we now allow for preferences to adapt to the contractual situation. This enables us to study the implications of "preference crowding" for our model. Economically successful preferences are crowded in, and unsuccessful preferences are crowded out. Formally, the share of types with a certain preference grows faster than another share if and only if the average material earnings of the former exceed the average earnings of the latter. In the context of our model and in the absence of a fully fledged theory of preference formation, this assumption seems reasonable. Contracts are closed to secure material benefits, and the outcomes of our contract game are exclusively characterized by different resource allocations.[16]

The assumption that successful "traits" spread is often associated with models of genetic evolution, although it can be justified differently.[17] One justification is offered in Appendix B, which briefly illustrates a stochastic model of individual preference adaptation.[18] In our model, this implies the following: If honesty leads to forsaking profitable opportunities such that a typical H earns less than a typical M, honesty will be crowded out. If an environment favors honesty, such that, on average, H types earn more than M types, honesty will be crowded in.

In order to calculate the success of the two different types (of preferences), we assume a random match of players and enough individuals to ensure that the law of large numbers can be applied. This allows us to take expected values as a measure of success.

Proposition 1 shows that what happens when two players interact depends on the value of p, the probability that the contract is enforced. The proposition distinguishes three regimes, and in each case we can show how preferences are crowded in and out.

In the high-probability regime,

$$p > \frac{b}{b+c},$$

all individuals, regardless of type and matching, will play the SPE (T,P). Players of both types receive the same expected monetary payoff, so there will be no crowding. Regardless of the numbers of H and M types, the high enforcement probability ensures performance.

In the medium-probability regime,

$$\frac{a}{1+a} < p < \frac{b}{b+c},$$

behavior in a match depends on the type of player 2. If player 2 is an H, the SPE is (T,P). If he is an M, the SPE is $(T,\overline{P})$. Since M maximizes expected monetary payoffs but H does not, it follows that average earnings of M types exceed those of H types. In the role of player 2, M types earn on average $p(1-c) + (1-p)(1+b)$, which is strictly greater than 1, the payoff of an H type in the role of player 2. In the role of player 1, both types do equally well on average, since the expected payoff of player 1 is independent of her type. Thus, with medium p, M types always earn more than H types (regardless of their number). Accordingly, honest preferences will be crowded out. This implies, asymptotically for the long run, that H types will completely vanish and that all individuals will play equilibrium $(T,\overline{P})$.

In the low-probability regime,

$$p < \frac{a}{1+a},$$

we take into account that the SPE is $(\overline{T},\overline{P})$ if player 2 is an M and (T,P) if player 2 is an H. In this case, the earnings of H types exceed the earnings of M types. In the role of player 2, an H always receives 1, an M always 0. In the role of player 1, the average payoffs of both types are again identical. They do not depend on their own type. Thus, with low p, preferences for honesty are crowded in and, in the long run, type M will vanish, such that all individuals will play the trust-rewarding equilibrium (T,P).

We summarize our results in the following proposition.

Proposition 2. *In the high-probability regime, where*

$$p > \frac{b}{b+c},$$

there is no crowding; in the long run both types may be present in the society, and all individuals play (T,P).

[16] Other studies analyze endogenous preferences in a similar framework (e.g., Bester and Güth 1998; Huck and Oechssler 1999; To 1999), especially Güth and Kliemt 1999, who also deal with the issue of trust. Pioneering studies suggesting that endogenous preferences may result from payoff-driven (evolutionary) dynamics are Becker 1976 and Hirshleifer 1977. We do not wish to generalize too much, however, as different mechanisms may apply if one deals with attitudes toward drugs, religion, or other domains in which outcomes are less tangible. For a survey of alternative approaches, see Bowles 1998.

[17] For two such justifications, see Börgers and Sarin 1997 and Schlag 1998.

[18] We assume that the probabilities with which preferences are adapted depend on two factors, economic success and conformity. The more an individual's environment favors a certain preference and the more common is a certain preference, the more likely is an individual to develop the same preference. (Note, however, that adaptation is not deliberate. People cannot consciously decide about their preferences. For a different approach, see Cooter 1998.) We show that this process behaves, under certain additional regularity assumptions, like a growth-monotonic (evolutionary) process.

In the medium-probability regime, where

$$\frac{a}{1+a} < p < \frac{b}{b+c},$$

trustworthiness is crowded out; in the long run only type M will be present in the society, and all individuals play $(T,\overline{P})$.

In the low-probability regime, where

$$p < \frac{a}{1+a},$$

trustworthiness is crowded in; in the long run only type H will be present in the society, and all individuals play (T,P).

The long-run stable states also can be derived by analyzing evolutionary games in which the two types compete. For

$$p > \frac{b}{b+c}$$

this game is trivial, because both types behave identically and receive identical payoffs. For the other cases, the two matrix games in tables 1 and 2 emerge. (Note that the payoffs are based on the assumption that both types are equally likely to become player 1 or 2).

As

$$p < \frac{b}{b+c},$$

it is easy to see that the game has a unique evolutionary stable strategy (see Maynard Smith 1982). There is a unique equilibrium (M,M), and the equilibrium is strict. Hence, the unique evolutionary stable strategy is M. This mirrors the second result of proposition 2: In the long run only M types survive.

An even simpler matrix emerges in the low-probability regime, where

$$p < \frac{a}{1+a}.$$

Clearly, the unique equilibrium of this game is (H,H); because the equilibrium is strict, the unique evolutionary stable strategy is H. This mirrors the third result of proposition 2.

TABLE 1. The Evolutionary Game with Medium Probability *p*

$\frac{a}{1+a} < p < \frac{b}{b+c}$	Type M	Type H
Type M	$1 + b - a + p(1 - b + a - c)$	$2 + b - p(b + c)$
Type H	$1 - a + p(1 + a)$	2

TABLE 2. The Evolutionary Game with Low Probability *p*

$p < \frac{a}{1+a}$	Type M	Type H
Type M	0	1
Type H	1	2

Discussion

There are two main implications of proposition 2. The most fundamental one is that it is impossible to predict behavior in a group of agents playing the contract game without knowing their history. Individual preferences are subject to change, and outcomes in one round affect the distribution of types in the next. And because preferences depend on past regimes, so do actions. The longer a group plays in a low-probability environment, the more agents with a preference for honesty are present, and the less breach is observed. The opposite is true for a medium-probability environment. Also, it may be possible that groups have experienced regime changes in the past, in which case one needs to know not only how long the group has been playing under the current regime but also how long under the preceding one. The total crowding history matters.

The second significant implication is that the probability of contract enforcement exerts a nonmonotonic effect on behavior, which would not occur in standard models with fixed preferences.[19] The worst legal regime is not one in which contracts cannot be enforced but one with an intermediate level of enforceability. With an intermediate p, first movers do not care with whom they interact, because entering the contract is better than staying outside, even if the contract is breached. This lack of caution makes dishonest second movers economically successful, so the share of dishonest types will grow. There are then two alternatives: more order with more law or more order with less law.

With less law, first movers have to be extremely cautious. They have to think about their partners' trustworthiness, which makes honesty a successful preference. In our model, first movers receive a perfect signal about their opponents' type, and their decision rule is simple: "Only enter a contract if your opponent is trustworthy." With a stochastic signal the rule would be very similar: "Only enter if the signal is good enough." This illustrates how our results would extend to the more general case of imperfect but informative signals. With perfect signals, M types in the role of player 2 are *never* offered contracts when p is low, while H types *always* get contracts. With imperfect signals, some M types would get contracts, while some H types would not (namely, whenever the signal is wrong). If the signal is sufficiently informative, however, H types will get *more* contracts than M types, which is required

[19] To the best of our knowledge, the first authors to highlight the possibility of such nonmonotonicities are Akerlof and Dickens (1982). They study a model with cognitive dissonance that may induce players to reevaluate outcomes.

TABLE 3. Experimental Sessions

Session	Matching	Prob. 1–3	Prob. 4–9	Number of Subjects	Univ.
1 (RLLB)	Random	Low (p = .1)	Low (p = .1)	20	B
2 (RMLB)	Random	Medium (p = .5)	Low (p = .1)	28	B
3 (RHLB)	Random	High (p = .9)	Low (p = .1)	28	B
4 (FMLB)	Fixed	Medium (p = .5)	Low (p = .1)	16	B
5 (RLLH)	Random	Low (p = .1)	Low (p = .1)	34	H
6 (RMLH)	Random	Medium (p = .5)	Low (p = .1)	28	H

for crowding in. Appendix A elaborates on the case of imperfect signals further.

Comparing the two alternative policies for replacing a medium probability regime ("more or less law"), two differences can be observed. The first is due to the dynamic nature of our analysis. With "more law," more order is instantly achieved, since performance and entering becomes rational for everyone. This is different in the case of "less law" because after the change of the regime the crowding process needs some time. Though our experimental results indicate that adjustments can be fast in small groups, the behavior does not change instantly. This is an argument in favor of the standard law-and-order approach. "Less law," however, is less costly. In our model, we disregard all fixed costs of legal contract enforcement and variable costs being a function of p. Increasing p costs resources; decreasing p saves resources.

Here we do not wish to make any judgment about what is the better policy. Instead, we test our model's empirical validity in the laboratory.

EXPERIMENT

Design

In the experimental design we tried to implement our model as closely as possible. Subjects played a two-person contract game and were randomly matched. Six sessions with a total of 154 subjects were conducted, and the game was repeated several times. To examine crowding in of trust (more order with less law), we confronted subjects in all sessions with a low contract enforcement probability during the last rounds. In order to create different histories, we varied the legal regime in the first few rounds. If behavior is driven by incentives only, it should be independent of the history created in earlier rounds and should depend only on the current legal regime. In contrast, if preferences adapt to the legal regime, earlier history should affect the likelihood of performance in later rounds.

Experimental subjects were confronted with identical payoffs: 50 cents for each player whenever the first mover chose $\overline{T}$ (corresponds to choosing alternative A in the experimental payoff table in Appendix C); \$1.50 for each player in case of TP (corresponds to choosing alternative Y in the experimental payoff table); \$1.50 for the first and \$1.20 for the second mover in case of $T\overline{P}L$ (corresponds to the realization of state α in the experimental payoff table); and 20 cents for the first and \$2.50 for the second mover in case of $T\overline{P}L$ (corresponds to the realization of state β in the experimental payoff table). The normalized payoff variables were $a = .3$, $b = 1$, and $c = .3$. The probabilities were .1, .5, and .9.[20]

Subjects received payment for each round. Instructions (see Appendix C) were neutrally framed. After each round, aggregate information on outcomes was provided, that is, first and second movers knew how many contracts were offered and performed in the previous round. Providing information on the distribution of types only serves as a conservative test of our model, because individuals' types could not perfectly be detected. The crowding model assumes random matching, so we implemented a stranger treatment in five sessions and used a fixed-pair matching in one control session. Table 3 presents an overview of all sessions.

The experiments were conducted at the University of California, Berkeley, and at Harvard University (labelled B and H). Participation was voluntary, and students were paid a show-up fee of \$5. The experiments took approximately 45 minutes. Average earnings were approximately \$15.

Subjects were identified by code numbers only, and anonymity was fully preserved. After signing a consent form, participants were randomly assigned to the role of first and second mover; they were given written instructions, along with an envelope containing a code number sheet and nine decision sheets, all marked with the subject's code number. Instructions were repeated orally, which allowed subjects to ask questions and helped ensure that everyone understood the decision task. In all but session 4, they were truthfully assured that they would be randomly matched with a different person after each round. They were told that nine rounds would be played and that they would publicly be informed about the aggregate outcome after each round. Individual results were private information. Subjects could not anticipate the regime change before round 4,[21] but when informed about the new conditions, they were also told that they would play under the new regime for the remaining six rounds.

[20] Subjects carried out the chance moves themselves. After each round a randomly chosen participant picked a card from a pile of red and black cards.

[21] The instructions told them neither that the environment would remain constant nor that it would change.

TABLE 4. Data

Session	Behavior	Round 1	2	3	4	5	6	7	8	9
1 (RLLB)	No trust ($\bar{T}$)	4	8	10	8	3	2	3	1	5
	Trust and breach ($T,\bar{P}$)	5	2	0	0	2	2	1	2	0
	Trust and perf. (T,P)	1	0	0	2	5	6	6	7	5
2 (RMLB)	No trust ($\bar{T}$)	0	0	4	10	10	10	8	4	8
	Trust and breach ($T,\bar{P}$)	10	12	8	4	4	2	3	6	4
	Trust and perf. (T,P)	4	2	2	0	0	2	3	4	2
3 (RHLB)	No trust ($\bar{T}$)	0	0	0	5	8	2	7	6	4
	Trust and breach ($T,\bar{P}$)	2	2	2	7	2	8	4	2	2
	Trust and perf. (T,P)	12	12	12	2	4	4	3	6	8
4 (FMLB)	No trust ($\bar{T}$)	0	1	0	2	0	0	4	0	8
	Trust and breach ($T,\bar{P}$)	1	0	0	0	0	2	0	0	0
	Trust and perf. (T,P)	7	7	8	6	8	6	4	8	0
5 (RLLH)	No trust ($\bar{T}$)	6	3	9	6	7	7	8	9	10
	Trust and breach ($T,\bar{P}$)	5	6	2	5	3	3	3	3	2
	Trust and perf. (T,P)	6	8	6	6	7	7	6	5	5
6 (RMLH)	No trust ($\bar{T}$)	0	1	1	11	10	8	8	7	8
	Trust and breach ($T,\bar{P}$)	8	7	8	1	1	1	0	1	0
	Trust and perf. (T,P)	6	6	5	2	3	5	6	6	6

Note: The table summarizes behavior for all sessions over time. In RLLB, R stands for random matching, the first L for low probability in the first few rounds, the second L for low probability in the remaining six rounds, and B for Berkeley; analogously, for the other sessions.

Results

Our theory predicts history-dependent behavior. The longer individuals are confronted with a low-p environment, the more trustworthiness should be crowded in, and the higher performance rates should be. High enforcement probabilities are expected to be neutral with respect to crowding, and medium probabilities should crowd out trustworthiness.[22]

Table 4 presents the results for all sessions in all rounds. We briefly examine session 4 with fixed pairs, which is a simple control session because the requirements for crowding are not fulfilled. There is no random matching and therefore no interaction on the group level. Instead, subjects play a finitely repeated game. Experiments on games with cooperative gains (e.g., repeated public goods games or gift exchange games) reveal that cooperation rates are typically higher with this kind of matching than standard theory expects. Furthermore, cooperation rates seem to be relatively stable over time but break down toward the end of the game when the "shadow of the future" loses its power and reputation no longer plays a role. This strong drop in the last rounds has been called an "end game effect" (Selten and Stoecker 1983).[23] If the contract game is comparable to these games, we should observe a similar pattern. Table 4 confirms this expectation: We find that in the fixed-pairs session cooperation drops from 100% to 0% in the last round.

In contrast, our crowding theory predicts increasing cooperation over time and rules out an end-game effect. It predicts that trustworthy second movers perform because they receive less utility from breaching than from performing, even though breaching leads to a higher monetary payoff. Table 5 shows aggregate data for rounds 4 to 9 in all random-matching sessions. It suggests a trend toward more cooperation that does not break down. On the contrary, the performance rate (the number of contracts performed divided by the number of contracts offered) reaches its maximum in the last round, which is in line with the crowding prediction. We summarize by the following.

RESULT 1. *In the low probability environment with random matching, performance rates increase over time and there is no end-game effect.*

Our model assumes that it is most efficient if all second movers perform. We expect high efficiency rates instantly when enforcement is strong, and only slow increases in efficiency rates over time with low enforcement probabilities. Figure 2 presents the efficiency rates (the number of contracts performed divided by the number of all possible contracts), for

TABLE 5. Aggregate Random-Matching Data for Rounds 4 to 9

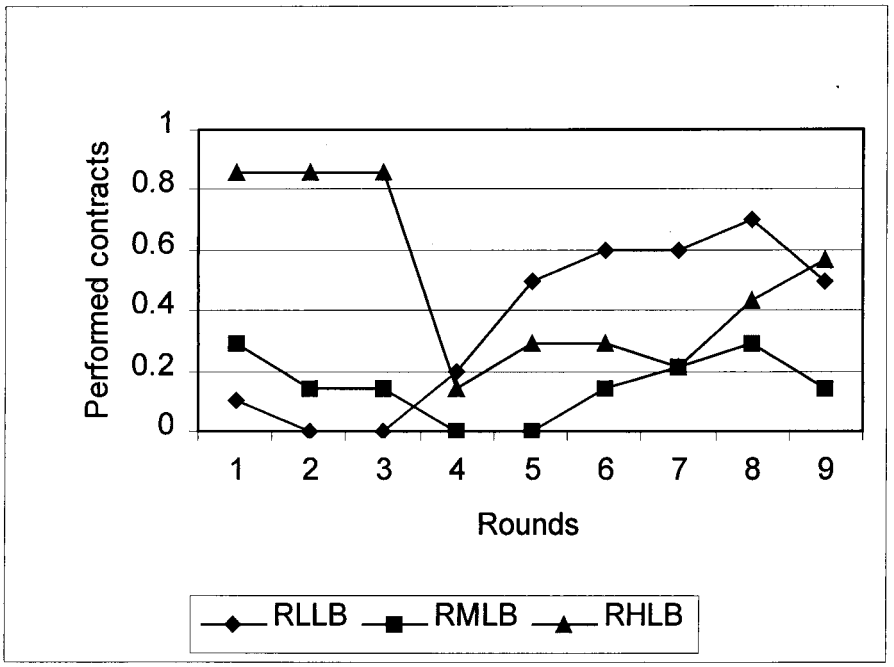
Aggregate Data	Round 4	5	6	7	8	9
No trust ($\bar{T}$)	40	38	29	34	27	35
Trust and breach ($T,\bar{P}$)	17	12	16	11	14	8
Trust and perf. (T,P)	12	19	24	24	28	26
Performance rate	.41	.61	.63	.69	.67	.77

Note: Numbers in the first three rows are absolute numbers; in the last row they are relative frequencies.

[22] These predictions cannot be viewed as deterministic, since the law of large numbers does not apply in the laboratory (see also Appendix B).

[23] For additional evidence, see Andreoni 1988 and Croson 1996, who differentiate between cooperation based on reputation and reciprocity.

FIGURE 2. Efficiency Rates over Time (Berkeley Only)



the sessions with random matching. In the short run, high enforcement probabilities lead to the most efficient outcomes. In the medium term, when the crowding dynamics start to become relevant, the low-p environment is most efficient. The longer subjects were confronted with a low-p environment, the less applicable were the differential effects of their respective crowding histories.

RESULT 2. *In the short run, efficiency rates are highest when enforcement is strong; in the medium term, they are highest in the low-probability environment; in the long run, the differential effects of enforcement and crowding tend to vanish.*

In order to analyze the data more thoroughly, we next estimated binary choice models for first movers' propensity to enter and second movers' propensity to perform, and we controlled for the relevance of crowding compared to economic incentives and for fixed effects of the university group.

In order to measure crowding, let

$$\gamma_j^t \begin{cases} 1 \text{ if } p \text{ is small in group } j \text{ in round } t; \\ 0 \text{ if } p \text{ is high in group } j \text{ in round } t; \text{ and} \\ -1 \text{ if } p \text{ is medium in group } j \text{ in round } t; \end{cases}$$

and let

$$CROWD_j^t = \sum_{h=1}^{t-1} \gamma_j^h.$$

The variable γ_j^t indicates whether our theory predicts crowding in (+1), crowding out (−1), or no crowding (0), and the variable $CROWD_j^t$ summarizes the "crowding history" of group j up to round t. If the theory is relevant, we would expect $CROWD_j^t$ to help explain the propensity of second movers to perform.[24]

In addition to *CROWD,* we include a number of variables as covariates.

- $PERFORM_j^{t-1}$ is the performance rate in group j in round $t - 1$, measured as the number of contracts performed divided by the number of contracts offered.
- $ENTER_j^{t-1}$ is the rate of entering in group j in round $t - 1$, measured as the number of contracts entered divided by the number of first movers in group j.
- $INCENT_j^t$ is a dummy variable indicating whether the first mover has an incentive to enter the contract if only monetary payoffs play a role, that is, $INCENT_j^t = 1$ if p is medium or high, 0 otherwise.
- $INCPERF_j^t$ is a dummy variable indicating whether the second mover has an incentive to perform the contract if only monetary payoffs play a role, that is, $INCPERF_j^t = 1$ if p is high, 0 otherwise.
- $UNIV_j$ is a dummy variable indicating the university group to which j belongs; $UNIV_j = 1$ for university H and 0 for university B.

[24] It can be argued that this definition is somewhat arbitrary, and we agree. By no means do we claim that *CROWD* captures the "true story," and certainly we do not claim that crowding is linear. Probably it is not. We think, however, that this a "Bayesian approach." In the absence of any specification that is a priori more rational than another, the above definition is the most simple and can be justified by taking expected values over equally probable alternatives. Furthermore, if we find a significant effect of *CROWD* in its current form, any "optimal" definition would increase its significance.

TABLE 6. Propensity to Enter (First Movers) and to Perform (Second Movers) a Contract (Logistic Regression)

	Without Subject Dummies				With Subject Dummies			
	First Movers							
	Coef.	S.E.	p	R	Coef.	S.E.	p	R
$CROWD_j^t$	.048	.035	.175	.000	−.103	.078	.189	.000
$PERFORM_j^{t-1}$	2.062	.447	.000	.160	5.441	.884	.000	.218
$ENTER_j^{t-1}$	.530	.458	.247	.000	−.263	.759	.729	.000
$INCENT_j^t$	2.553	.515	.000	.173	7.955	1.508	.000	.185
$INCPERF_j^t$	4.649	4.432	.684	.000	5.914	32.663	.856	.000
$UNIV_j$	−.518	.213	.015	−.072	−1.838	5.206	.724	.000
Constant	−1.296	.325	.000		−1.999	3.063	.514	
	Second Movers							
$CROWD_j^t$	.125	.058	.030	.077	.792	.198	.000	.183
$PERFORM_j^t$	1.493	.704	.034	.080	−.959	1.669	.566	.000
$ENTER_j^{t-1}$	−.864	.709	.223	.000	−1.731	1.738	.319	.000
$INCENT_j^t$	−.177	.538	.743	.000	−1.454	1.412	.303	.000
$INCPERF_j^t$	2.057	.749	.006	.115	15.225	40.668	.708	.000
$UNIV_j$	.594	.280	.034	.077	2.588	24.336	.915	.000
Constant	−.505	.557	.365		−.484	13.967	.972	

Note: $N = 552$ for first movers, $N = 312$ for second movers. R is the partial contribution measuring the relative importance of each variable.

Furthermore, we also ran (logistic) regressions including subject dummies S_{ji}.[25] The estimated model without subject dummies is

$$\ln\frac{q_{rji}^t}{1-q_{rji}^t} = \alpha + \beta_0 CROWD_j^t + \beta_1 PERFORM_j^{t-1} + \beta_2 ENTER_j^{t-1} + \beta_3 INCENT_j^t + \beta_4 INCPERF_j^t + \beta_5 UNIV_{j+}\varepsilon_{ji},$$

with q_{rji}^t indicating either first mover i's propensity to enter or second mover i's to perform. Table 6 presents the results for first movers and second movers. All sessions with random matching are included.

We find that first movers' behavior is mainly driven by the economic incentives they face and by the performance rate of the last round. In the estimation without subject dummies we also find a significant effect of the university dummy: On average, H students enter less often than B students. Since this dummy loses its significance when subject dummies are introduced, we are confident that the difference between H and B is not due to some differences in the experimental procedures (of which we were unaware). Rather, the difference is on the individual level.

In the case of second movers, we find that H students are more trustworthy than B students. Again, this effect disappears when we control for differences between individuals. In the estimation without subject dummies, three additional variables help explain second movers' decisions: economic incentives (less breaching if breach is deterred), last round's performance rate (such that there is some inertia), and crowding history (the longer subjects interact in a low-p environment, the more likely they are to be trustworthy). Only the last variable remains significant when we include subject dummies. We summarize as follows.

RESULT 3. *First movers' propensity to enter a contract (i.e., to select alternative B in the experimental payoff table) mainly depends on their monetary incentives and on second movers' previous performance rate. Second movers' propensity to perform mainly depends on their crowding history.*

The second part of result 3 is the key finding of this study because it confirms the qualitative predictions of the crowding theory. It is a result on the aggregate level, however, so it seems worthwhile to investigate whether the theory also predicts individual behavior. The main prediction of our model on the individual level is that, in the low-p environment, subjects who breach once are much more likely to switch to honest behavior than are honest subjects to switch to breaching. To analyze this hypothesis we simply count how many second movers switch from breaching to performing once p is low. We also count how many second movers switch in unpredicted directions and how many never switch once p is low. Table 7 presents the results. The picture is clear. In the low-enforcement regime 53.6% of all subjects changed their behavior, and 89.2% of these switched in the predicted direction. In other words, 33 second movers breached in the early phases of the experiment and then performed later, even though a money-maximizing strategy favored the

[25] We use nested effect coding (see Königstein 1998) because the subject dummies are nested in the university dummy.

TABLE 7. Individual Behavior in the Low-Probability Environment

Session	Breach→ Perform	Unpredicted Switches	Behavior Fixed	*N*
1 (RLLB)	7	—	3	10
2 (RMLB)	6	1	7	14
3 (RHLB)	9	—	5	14
4 (RLLH)	6	3	8	17
5 (RMLH)	5	—	9	14
Total	33	4	32	69

old behavior. None of these 33 subjects switched back to breaching. They became trustworthy, even though they started by breaching. There were very few unpredicted switches.[26] The null hypothesis that behavior switching is random can be rejected at a significance level of .01%. Thus, the crowding theory's predictions are confirmed by individual data.

CONCLUSION

In our model the (legal) rules of a game have not only short-run but also long-run effects on behavior because they affect preferences. In a contractual relationship, economic incentives have a nonmonotonic influence on contract performance. Our model complements recent work on the interaction of rules and preferences (e.g., Fehr and Schmidt 1999), which only allows for differences in preferences, whereas our model permits preferences to change. It suggests that the rules of the game determine which preferences dominate. More specifically, it predicts that low levels of legal contract enforcement increase trustworthiness. Because first movers cannot trust the legal system, they enter a contract only if they can trust the second mover. They are careful about the decision, which makes trustworthiness a successful trait.

Arguing from a different perspective, others come to very similar conclusions. Mansbridge (1999, 305), who discusses various ways of encouraging trustworthiness and trust, concludes:

> When the trustworthiness of a population is too low to sustain a general stance of initial trust, and when geographic and social mobility make reputational, kin and local sanctions less viable, the trustworthy members of a given population will benefit from finding ways of distinguishing themselves and other trustworthy individuals from the untrustworthy . . . the trustworthy would find it useful to train themselves to recognize subtle signs of trustworthiness in others and also to develop in themselves signs that could not easily be mimicked.

Differential incentives to learn about others' dispositions may account for some of the cross-cultural variation in behavior found in laboratory experiments. For example, Yamagishi, Cook, and Watanabe (1998) argue that Japanese subjects are less trusting and trustworthy than American subjects because contract enforcement mechanisms and assurance structures are more prevalent in Japan than in the United States. This corresponds to our high-p setting and our finding that when contracts are completely specified, interpersonal trust is replaced by institutional trust in the legal system. First movers enter a contract because second movers are deterred from breaching.

Previous work on crowding focuses on the relevance of preferences when contracts are complete.[27] We show that trustworthiness is crowded out not by complete contracts but by semispecified contracts. At intermediate levels of enforcement, second movers are not yet deterred from breaching, and first movers find entering a contract financially more attractive than remaining outside. Interpersonal trust is replaced by institutional trust in the legal system, and genuine trustworthiness is crowded out. Semispecified contracts cause nonmonotonic behavior: More order can result from less law, which yields a "motivation-compatible" environment (Bohnet and Frey 1997), and from more law, which yields an incentive-compatible environment.

Closely related to this finding are results by Huang and Wu (1994) and Huck (1998). Using psychological game theory,[28] Huang and Wu (1994) model games very similar to ours and show that if payoffs also depend on beliefs, then different levels of order may result from the same level of law; that is, there is a multiplicity of equilibria. Simply speaking, there is one equilibrium in which everyone believes society is functioning well and trust is rewarded, and this becomes self-fulfilling. If everyone believes the opposite, that becomes self-fulfilling. A crucial difference between the Huang and Wu approach and ours is that preferences (regarding payoffs and beliefs) are fixed in their model. Thus, no obvious dynamics lead from one state to another, and there is no straightforward link between institutional design and behavior. In contrast, Huck (1998) shows in the context of criminal law that if preferences are allowed to change, socially desirable behavior can be induced with lower levels of (monetary) punishments than one would conclude assuming fixed preferences.

In our experiment we tried to map the theoretical assumptions into a laboratory environment as precisely as possible, and the results support our qualitative predictions. Similar to Huck, we find that if there is enough time for the crowding dynamics to unfold, environments with low contract enforcement can produce outcomes as efficient as high levels of enforcement.

To the best of our knowledge, we provide the first empirical evidence of long-run effects of legal rules on behavior. Although experiments simplify reality, our contract game is informed by real-life institutions; it represents a situation in which legal enforcement leads

[26] Notice that, as the theory is stochastic, these instances are not entirely unpredicted.

[27] See Titmuss 1970 for policy examples; Frey 1997 for the crowding out of tax morale; Frey and Oberholzer-Gee 1997 for the crowding out of civic duty in a siting context; and Gneezy and Rustichini 1999 for the crowding out of parental discipline in a daycare center. For experimental evidence, see the extensive survey of psychology studies by Deci, Koestner, and Ryan 1999 as well as Falk, Gächter, and Kovacs 1999. For the disruption of "implicit agreements" in the organizational context, see Arrow 1974.

[28] See Geneakoplos, Pearce, and Stacchetti 1989.

to perfect expectation damages, and transaction costs are allocated according to the English legal cost allocation rule, where the loser pays all legal costs. The "long run" in our experiment is nine rounds (or 45 minutes), which we interpret as a conservative test of crowding. In fact, we were surprised that the dynamics unfolded so quickly and that subjects' inclination to trust and to be trustworthy changed in such a short time. The results support the view that institutional changes affect behavior, but they also reveal that, by affecting behavior, institutions affect preferences.

APPENDIX A. A SIMPLE CASE WITH IMPERFECT SIGNALS

Suppose first movers receive, before making their decision to enter a contract, a signal $s \in \mathbf{R}$. The signal technology is the following. If player 2's true type is H, then the signal is $1 + \varepsilon$; ε is normally distributed, with mean zero and variance σ^2. If player 2's true type is M, the signal is $0 + \varepsilon$, and ε comes from the same normal distribution.

How will player 1 decide whether to enter the contract? Nothing changes in the case of

$$p > \frac{a}{1+a}:$$

She will always enter. (This follows from the fact that she enters even when she knows for sure that player 2 is of type M.) In case of

$$p < \frac{a}{1+a},$$

player 1 has to update her beliefs about the type of player 2 by using the Bayes rule. She will enter if and only if the probability for player 2 being of type H is large enough. Hence, there will be a critical value $\bar{s}$, such that player 1 enters if $s \geq \bar{s}$ and stays out otherwise.

Given $\bar{s}$ and the signal technology, we can now compute the probabilities with which the two different types are offered contracts. Using these probabilities we can calculate the expected payoffs for both types or, in a population model, the average payoff for both types. Obviously, H types will always get more contracts than M types. (It is more likely that signal s exceeds $\bar{s}$ for H types.) Yet, M types benefit from profitable breach. The dominance of one of these effects depends critically on how noisy the signal is, that is, on variance σ^2.

There are two boundary cases. (i) $\sigma \to 0$ is the case of a perfect signal, and the first effect is stronger than the second (trustworthiness is crowded in). (ii) $\sigma \to \infty$ is the case without a signal, and the second effect is stronger (trustworthiness is crowded out). Obviously, we can now find a critical standard deviation, σ, that induces identical expected monetary payoffs for both types. If the standard deviation is above this level, M types will earn more than H types, and the third result of proposition 2 will be reversed. If the standard deviation is below this level, the original result is resurrected.

APPENDIX B. A STOCHASTIC MODEL OF INDIVIDUAL PREFERENCE ADAPTATION

Consider a large population of individuals who may be heterogeneous with respect to their preferences. Let Ω be the finite set of possible preferences, and let ω be a typical element of this set. When individuals interact, their behavior and earnings depend on the exact situation they face and on their type. For a given situation (e.g., a game that specifies only monetary payoffs) we can denote the monetary payoff of individual i as $\pi_i(\omega_i,\omega_{-i})$, where ω_{-i} denotes the types of all individuals with whom i is interacting. (In situations with multiple equilibria, this implies that a ready selection criterion is at hand.) Individuals are matched by a matching scheme S, that is, S maps the set of individuals itself. The average payoff earned by individuals of a certain type ω', depends on ω', S, and the current profile of types, denoted by the cumulative distribution function $F(\omega)$, which can be written as $\Pi(\omega',S,F(\omega))$ or as $\Pi(\omega')$. For convenience, we restrict the model to discrete time, t being the time index. Accordingly, let $f^t(\omega)$ denote the share of individuals of type ω at time t. The expression

$$\frac{f^{t+1}(\omega) - f^t(\omega)}{f^t(\omega)} = g(\omega)$$

reflects the growth rate of type ω.

We assume that preference changes occur stochastically and within individuals. Let $q(\omega',\omega'')$ be the probability that an individual's preference ω' changes to ω''. Obviously,

$$\sum_{\omega''} q(\omega',\omega'') = 1.$$

Furthermore, we assume that these probabilities depend on two factors, economic success and conformity.[29] To capture the role of conformity we assume the following.

ASSUMPTION 1 (conformity). *Ceteris paribus, $q(\omega',\omega'')$ is proportional to $f(\omega'')$.*

Assumption 1 implies that $q(\omega',\omega'')$ can be written as $f(\omega'')$ times some other function $Q(\omega',\omega'')$. In order to embed the role of economic success, we assume the following.

ASSUMPTION 2 (economic success). $Q(\omega',\omega'')$ *only depends on $\Pi(\omega'')$ and is strictly increasing in it.*

This gives rise to a stochastic process in which $F(\omega)$ develops over time. Below we show that, under certain additional (regularity) assumptions, such a process behaves like a growth-monotonic evolutionary process, that is, like a process assumed in our text: Shares of types grow according to their relative economic success.

ASSUMPTION 3 (regularity). *(a) The matching scheme S specifies random matching. (b) The population is large enough for the law of large numbers to apply.*

THEOREM 1. *Under assumptions 1 to 3 the dynamic process of individual preference adaptation behaves like a growth-monotonic evolutionary process, that is, $g(\omega') > g(\omega'') \Leftrightarrow \Pi(\omega') > \Pi(\omega'')$.*

Proof. With assumptions 1 and 3 we can write:

$$f^{t+1}(\omega') = \sum_{\omega} f^t(\omega) Q(\omega,\omega') f^t(\omega').$$

Therefore,

$$g(\omega') > g(\omega'') \Leftrightarrow \sum_{\omega} f^t(\omega) Q(\omega,\omega') > \sum_{\omega} f^t(\omega) Q(\omega,\omega'').$$

Due to assumption 2 $Q(\omega,\omega') > Q(\omega,\omega'') \Leftrightarrow \Pi(\omega') > \Pi(\omega'')$. Hence, the claim follows. *Q.E.D.*

The widely used replicator dynamics belong to the class of growth-monotonic evolutionary processes, and it is easy to see when a process of individual preference adaptation behaves like the replicator dynamics.

[29] For theories of conformity, see Akerlof 1997; Bernheim 1994; Bowles 1999; Boyd and Richerson 1985.

COROLLARY 1. *If*

$$q(\omega',\omega'') = \frac{\Pi(\omega'')}{\sum_{\omega} f(\omega)\Pi(\omega)} f(\omega''),$$

then the process of individual preference adaptation behaves like the replicator dynamics.

Proof. The proof is straightforward. *Q.E.D.*

APPENDIX C. SAMPLE INSTRUCTIONS

For the condition of random matching of subjects, medium probability, player 1, and the first phase of the experiment, instructions were as follows.

Welcome to this research project! You are participating in a study in which you have the opportunity to earn cash. The actual amount of cash you will earn depends on your choices and the choices of other persons. At the end of the study, the amount of cash earned will be added to your show-up fee and paid to you in cash.

What the study is about: The study is on how people decide. You are randomly matched with another person present in this room. You and the other person have to choose between two alternatives. The payoff table tells you how much money you earn depending on what you choose and what the other person chooses.

How the study is conducted: The study is conducted anonymously, without communication between the participants, and repeated 9 rounds. Participants are only identified by a letter or a number called "code number." Neither the other participants nor the researcher will ever know how you decide. You are randomly matched with another person after each round. You will never interact with the same person again.

You are person 1.

Start of the study.

Round 1: The payoff table reads as follows.

First you have to choose between A and B.

If you choose A, you and the other person receive 50 cents each.

If you choose B, person 2 gets to choose between Y and Z.

If person 2 chooses Y, you and person 2 receive 150 cents each.

If person 2 chooses Z, chance decides about your earnings.

You earn 150 cents with probability 0.5 (α) and 20 cents with probability 0.5 (β), that is, your expected earnings after a chance move are 85 cents.

The other person earns 120 cents with probability 0.5 (α) and 250 cents with probability 0.5 (β), that is, his or her expected earnings after a chance move are 185 cents.

Payoff Table

Who Decides	Alternatives	Earnings for 1	Earnings for 2
Person 1	A	50	50
	B	→	→
Person 2	Y	150	150
	Z	→	→
Chance	α (prob = .5)	150	120
	β (prob = .5)	20	250

Now, please open your envelope. It contains 9 decision sheets and a code number sheet. Please take everything out of the envelope. Keep the code number sheet. Then choose between A and B. Indicate your choice on the decision sheet marked "Round 1," put this decision sheet back into the envelope, and put it into the box which we will pass around. Keep all other decision sheets.

Persons 2 are randomly allocated an envelope and asked to look at your decision and—if they get to make a choice—indicate their choice of either Y or Z on the decision sheet. Decision sheets will be put back into the envelope and into the box.

We collect all decision sheets and count how many people in this room chose A, B, Y, and Z, respectively, and inform all of you of the aggregate outcome of the first round. We then give you the envelope back. Please take the decision sheet out. The information on the decision sheet is private. Please do not share it with anyone else.

Chance now decides whether α or β will be realized in this round. For this purpose we draw a card from a pile with five red and five black cards. Red implies α, black implies β.

We determine your earnings according to your choice and the choice of the other person after the study is over. Your earnings will be paid to you in cash.

End of round 1.

REFERENCES

Akerlof, George A. 1997. "Social Distance and Social Decisions." *Econometrica* 65 (September): 1005–27.

Akerlof, George A., and William T. Dickens. 1982. "The Economic Consequences of Cognitive Dissonance." *American Economic Review* 72 (3): 307–19.

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

Arrow, Kenneth. 1974. *The Limits of Organization.* New York: Norton.

Axelrod, Robert. 1984. *The Evolution of Cooperation.* New York: Basic Books.

Baron, James N., and David Kreps. 1999. *Strategic Human Resources.* New York: Wiley.

Becker, Gary S. 1976. "Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology." *Journal of Economic Literature* 14 (September): 817–26.

Bendor, Jonathan, and Piotr Swistak. 1997. "The Evolutionary Stability of Cooperation." *American Political Science Review* 91 (June): 290–307.

Berg, John, John Dickhaut, and Kevin A. McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10 (July): 122–42.

Bernheim, B. Douglas. 1994. "A Theory of Conformity." *Journal of Political Economy* 102 (October): 841–77.

Bester, Helmut, and Werner Güth. 1998. "Is Altruism Evolutionarily Stable?" *Journal of Economic Behavior & Organization* 34 (February): 193–209.

Bohnet, Iris, and Bruno S. Frey. 1997. "Rent Leaving." *Journal of Institutional and Theoretical Economics* 153 (3): 711–21.

Bohnet, Iris, and Bruno S. Frey. 1999. "Social Distance and Other-Regarding Behavior in Dictator Games: Comment." *American Economic Review* 89 (March): 335–39.

Bolton, Gary E., and Axel Ockenfels. 2000. "ERC: A Theory of Equity, Reciprocity, and Competition." *American Economic Review* 90 (March): 166–93.

Börgers, Tilman, and Rajiv Sarin. 1997. "Learning Through Reinforcement and Replicator Dynamics." *Journal of Economic Theory* 77 (November): 1–14.

Bowles, Samuel. 1998. "Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions." *Journal of Economic Literature* 36 (March): 75–111.

Bowles, Samuel. 1999. "Cultural Evolution and Human Social Structure." University of Massachusetts. Typescript.

Boyd, Robert, and Peter J. Richerson. 1985. *Culture and the Evolutionary Process.* Chicago and London: University of Chicago Press.

Brehm, John, and Wendy Rahn. 1997. "Individual-Level Evidence for the Causes and Consequences of Social Capital." *American Journal of Political Science* 41 (July): 999–1023.

Burnham, Terry, Kevin A. McCabe, and Vernon Smith. 2000. "Friend-or-Foe Intentionality Priming in an Extensive Form Trust Game." *Journal of Economic Behavior & Organization* 43 (1): 57–74.

Cooter, Robert. 1998. "Expressive Law and Economics." *Journal of Legal Studies* 27 (June): 585–608.

Croson, R. 1996. "Partners and Strangers Revisited." *Economics Letters* 53 (October): 25–32.

Deci, Edward L., Richard Koestner, and Richard Ryan. 1999. "A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation." *Psychological Bulletin* 125 (November): 627–68.

Ellickson, R.C. 1991. *Order Without Law: How Neighbors Settle Disputes.* Cambridge, MA: Harvard University Press.

Ellison, Glenn. 1994. "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching." *Review of Economic Studies* 61 (July): 567–88.

Falk, Armin, Simon Gächter, and Judith Kovacs. 1999. "Intrinsic Motivation and Extrinsic Incentives in a Repeated Game with Incomplete Contracts." *Journal of Economic Psychology* 20 (3): 251–84.

Fehr, Ernst, and Klaus Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114 (August): 817–68.

Frank, Robert H. 1988. *Passions within Reason: The Strategic Role of the Emotions.* New York: Norton.

Frank, Robert H., Thomas Gilovich, and Dennis T. Regan. 1993. "The Evolution of One-Shot Cooperation: An Experiment." *Ethology and Sociobiology* 14 (July): 247–56.

Frey, Bruno S. 1997. "A Constitution for Knaves Crowds Out Civic Virtues." *Economic Journal* 107 (July): 1043–53.

Frey, Bruno S., and Felix Oberholzer-Gee. 1997. "The Costs of Price Incentives: An Empirical Analysis of Motivation Crowding Out." *American Economic Review* 87 (September): 747–55.

Fukuyama, Francis. 1995. *Trust: Social Virtues and the Creation of Prosperity.* New York: Free Press.

Gambetta, Diego. 1993. *The Sicilian Mafia.* Cambridge, MA: Harvard University Press.

Geneakoplos, John, David Pearce, and Ennio Stacchetti. 1989. "Psychological Games and Sequential Rationality." *Games and Economic Behavior* 1 (1): 60–79.

Glaeser, Edward L., David I. Laibson, José A. Scheinkman, and Christine L. Soutter. 2000. "Measuring Trust." *Quarterly Journal of Economics* 115 (August): 811–46.

Gneezy, Uri, and Aldo Rustichini. 2000. "A Fine Is a Price." *Journal of Legal Studies* 29 (January): 1–18.

Greif, Avner, Paul Milgrom, and Barry R. Weingast. 1994. "Coordination, Commitment, and Enforcement: The Case of the Merchant Guild." *Journal of Political Economy* 102 (August): 745–76.

Güth, Werner, and Helmut Kliemt. 1999. "Evolutionarily Stable Co-operative Commitments." *Theory and Decision* 49 (November): 197–221.

Hirshleifer, Jack. 1977. "Economics from a Biological Viewpoint." *Journal of Law and Economics* 20 (April): 1–52.

Huang, Peter H., and H.-M. Wu. 1994. "More Order without More Law: A Theory of Social Norms and Organizational Cultures." *Journal of Law, Economics, & Organization* 10 (October): 390–406.

Huck, Steffen. 1998. "Trust, Treason, and Trials: An Example of How the Evolution of Preferences Can Be Driven by Legal Institutions." *Journal of Law, Economics, & Organization* 14 (April): 44–60.

Huck, Steffen, and Jörg Oechssler. 1999. "The Indirect Evolutionary Approach to Explaining Fair Allocations." *Games and Economic Behavior* 28 (July): 13–24.

Kahan, Dan M. 1996. "What Do Alternative Sanctions Mean?" *University of Chicago Law Review* 63 (September): 591–653.

Kikuchi, Masako, Yukio Watanabe, and Toshio Yamagishi. 1996. "Accuracy in the Prediction of Others' Trustworthiness and General Trust: An Experimental Study." *Japanese Journal of Experimental Social Psychology* 37 (1): 23–36.

Knack, Stephen, and Philip Keefer. 1997. "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation." *Quarterly Journal of Economics* 112 (November): 1251–88.

Königstein, Manfred. 1998. "Measuring Treatment-Effects in Experimental Cross-Sectional Time Series." Humboldt University. Typescript.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert W. Vishny. 1997. "Trust in Large Organizations." *American Economic Review* 87 (May): 333–8.

Mansbridge, Jane. 1999. "Altruistic Trust." In *Democracy and Trust,* ed. M. E. Warren. Cambridge: Cambridge University Press. Pp. 290–309.

Maynard Smith, John. 1982. *Evolution and the Theory of Games.* Cambridge: Cambridge University Press.

Milgrom, Paul, Douglass North, and Barry Weingast. 1990. "The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs." *Economics and Politics* 2 (1): 1–23.

Murdock, Jonathan. 1999. "The Microfinance Promise." *Journal of Economic Literature* 37 (December): 1569–615.

Norris, Pippa. ed. 1999. *Critical Citizens: Global Support for Democratic Governance.* Oxford: Oxford University Press.

Nye, Joseph S., Philip D. Zelikow, and David King, eds. 1997. *Why People Don't Trust Government.* Cambridge, MA: Harvard University Press.

Orbell, John, and Robyn M. Dawes. 1991. "A 'Cognitive Miser' Theory of Cooperators' Advantage." *American Political Science Review* 85 (June): 515–28.

Ostrom, Elinor. 1998. "A Behavioral Approach to the Rational Choice Theory of Collective Action: Presidential Address." *American Political Science Review* 92 (March): 1–22.

Ostrom, Elinor, Roy Gardner, and James Walker. 1994. *Rules, Games, and Common-Pool Resources.* Ann Arbor: University of Michigan Press.

Putnam, Robert D., with Robert Leonardi and Raffaella Y. Nanetti. 1993. *Making Democracy Work.* Princeton, NJ: Princeton University Press.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83 (December): 1281–302.

Rose-Ackerman, Susan. 1999. *Corruption and Government: Causes, Consequences and Reform.* Cambridge: Cambridge University Press.

Rousseau, Denise M. 1995. *Psychological Contracts in Organizations: Understanding Written and Unwritten Agreements.* Thousand Oaks, CA: Sage.

Schlag, Karl H. 1998. "Why Imitate, and If So, How? A Boundedly Rational Approach to Multi-Armed Bandits." *Journal of Economic Theory* 78 (January): 130–56.

Schneider, Maria, Paul Teske, Melissa Marschall, Michael Mintrom, and Christine Roch. 1997. "Institutional Arrangements and the Creation of Social Capital: The Effects of Public School Choice." *American Political Science Review* 91 (March): 82–93.

Selten, Reinhard, and Rolf Stoecker. 1983. "End Behavior in Finite Prisoner's Dilemma Supergames." *Journal of Economic Behavior and Organization* 7 (1): 47–70.

Simis, Konstantin. 1982. *USSR: The Corrupt Society.* New York: Simon and Schuster.

Smith, John M., and Paul A. Bell. 1994. "Conformity as a Determinant of Behavior in Resource Dilemma." *Journal of Social Psychology* 134 (April): 191–200.

Sunstein, Cass R. 1996. "Social Norms and Social Roles." *Columbia Law Review* 96 (May): 903–68.

Tarrow, Sidney. 1996. "Making Social Science Work Across Space and Time: A Critical Reflection on Robert Putnam's *Making Democracy Work*." *American Political Science Review* 90 (June): 389–97.

Titmuss, Richard M. 1970. *The Gift Relationship.* London: Unwin.

To, Ted. 1999. "Risk and Evolution." *Journal of Economic Theory* 13 (March): 329–43.

Van Huyck, John B., Ray C. Battalio, and Mary F. Walters. 1995. "Commitment Versus Discretion in the Peasant-Dictator Game." *Games and Economic Behavior* 10 (July) 143–70.

Varese, Frederico. 1994. "Is Sicily the Future of Russia? Private Protection and the Rise of the Russian Mafia." *Archives Européennes de Sociologie* 35 (2): 224–58.

Wintrobe, Ronald. 1995. "Some Economics of Ethnic Capital Formation and Conflict." In *Nationalism and Rationality,* ed. Albert Breton, Gianluigi Galeotti, Pierre Salmon, and Ronald Wintrobe. Cambridge: Cambridge University Press. Pp. 43–70.

Yamagishi, Tashio, Karen S. Cook, and Motoki Watabe. 1998. "Uncertainty, Trust and Commitment Formation in the United States and Japan." *American Journal of Sociology* 104 (July): 165–94.