RETALIATION, REMEDIES, AND CONTRACTS

Sergio Rubens Mittlaender Leme de Souza [*]

An award of damages for breach of contract is justified because of the incentives it creates for parties to behave in socially optimal manners, and hence by its contribution to social welfare. In order to induce efficient levels of performance, damages shall impose a cost on breach equivalent to the loss of expectancy endured by the promisee. In legal scholarship and positive law, in contrast, they shall provide relief for victims of breach to redress breach. In doing so, damages can substitute private for public redress efficiently, and compensation can thereby be justified on social welfare grounds. The article reports results from an economic experiment that provides evidence for the capacity of expectation damages to induce performance if and only if performance is socially efficient as well as to crowd out promisees' observed tendency to retaliate to breach in socially costly manners. Both functions of damages deliver their own net contribution to social welfare, and under the parameters of the implemented trade game, gains of welfare provided by the crowding out function outweigh gains of welfare provided by incentives for promisors to perform efficiently.

INTRODUCTION

An award of remedies for breach of contract is justified, on social welfare grounds, because of the ex ante incentives it creates for parties to behave in socially optimal manners. Promisors anticipate the prospect of the payment of damages – the cost of breach – and change their behavior accordingly, responding rationally to the incentives created by law. Most importantly, by imposing a cost on breach that is equal to the promisee's loss of expectancy, expectation damages are apt to induce promisors to perform when performance is, in any contingency, socially efficient, thereby providing for maximal social welfare.[1]

---

[*] Rotterdam Institute of Law and Economics, European Doctorate in Law and Economics. The experiment was implemented in April and May 2013. [Acknowledgments].

[1] *See* John Barton, *The Economic Basis of Damages for Breach of Contract*, 1 JOURNAL LEGAL STUDIES 277 (1972); Robert Birmingham, *Breach of Contract, Damage Measures, and Economic Efficiency*, 24 RUTGERS LAW REVIEW 273 (1969). Remedies for breach further provide incentives for individuals to enter into mutually profitable contracts, for parties to rely on promises and to take precautions, among others. *See* Steven Shavell, *Damage Measures for Breach of Contract*, 11 BELL JOURNAL OF ECONOMICS 466 (1980); Steven Shavell, *The Design of Contracts and Remedies for Breach*, 99 QUARTERLY JOURNAL OF ECONOMICS 121 (1984) (introducing incentives for reliance); Robert Cooter, *Unity in Torts, Contracts, and Property: The Model of Precaution*, 73 CALIFORNIA LAW REVIEW 1 (1985) (introducing incentives for precautions).

Economic theories of contract law have, however, largely ignored one specific effect of an award of compensatory damages upon parties' behavior. Among the several incentives created by damages identified in the literature, no reference is made to the effect of compensatory remedies on the victims' tendency to reciprocate to perceived wrong in breach of contract. In the absence of reliance investments, optimal remedies for breach require only that promisors internalize all losses created by breach. They do not need to provide relief for the promisee to redress breach, and economic theories have difficulties providing a justification for compensation for the victim.[2]

For legal scholars, on the contrary, compensation is without doubt the goal pursued by an award of damages for breach.[3] In existing positive law, the goal is also to provide compensation, and legal relief aims at placing the victim of breach in the position she would have been in had the promisor performed.[4] Accordingly, compensation is the fundamental principle in the enforcement of contracts, both in legal theories and in the law.[5]

---

[2] *See, e.g.*, Richard Craswell, *Instrumental Theories of Compensation: A Survey*, 40 SAN DIEGO LAW REVIEW 1135, 1139 (2003) ("From the standpoint of modern economics, the analysis is instrumental all the way down, so the concept of compensation does no meaningful work.") *Id*., at 1178 ("In short, in economic theories the concept of compensation can be dispensed with entirely, whereas in corrective justice theories that concept is absolutely crucial.") Steven Shavell mentions two different justifications for why there could be value in compensation for promisees. Firstly, the promisee must earn something through a lawsuit in order to report breach. This however does not explain why the promisee is entitled by law to recover all lost expectancy. Secondly, compensation could provide implicit insurance for promisees. This, however, requires that promisees be risk-averse, more risk-averse than promisors, that the risk to be bear is detrimental and monetary, and that insurance markets concomitantly fail. *See* STEVEN SHAVELL, FOUNDATIONS OF ECONOMIC ANALYSIS OF LAW 311 (2003) (in the presence of functioning insurance markets, "the need for damages to compensate the victim is negated, and damages have a role mainly as an incentive device.")

[3] *See* 2 WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND 226-227 (George Sharswood ed., Philadelphia, J.B. Lippincott 1893) (1765-1769) ("upon all contracts or promises … just as for all infringements of the natural or relative rights of another, the law gives redress by action against the wrongdoer by an action to recover the damage sustained"); 1 THEODORE SEDGWICK, A TREATISE ON THE MEASURE OF DAMAGES 34 cmt. a (7th ed., New York, Baker, Voorhis & Co. 1880) (1847) ("The general principle undoubtedly at common law, both in actions for breach of contract and …, is to give compensation for pecuniary injury"); WILLIAM ANSTON, PRINCIPLES OF THE ENGLISH LAW OF CONTRACTS AND OF AGENCY IN ITS RELATION TO CONTRACT 377 (8th ed., New York, Banks Law 1899) ("Damages for breach of contract are by way of compensation"); 3 SAMUEL WILLISTON, THE LAW OF CONTRACTS 2392-2393 (1st ed. 1920) ("Compensation is the fundamental principle. In fixing the amount of these damages [for breach of contract], the general purpose of the law is, and should be, to give compensation"); Max Radin, *Fundamental Concepts of the Roman Law*, 12 CALIFORNIA LAW REVIEW 481, 485 (1924) ("At the common law, the nature of the obligation created by a wrongful act is in general undisputed. The purpose is compensation."); E. ALLAN FARNSWORTH, CONTRACTS 757 (4th ed. 2004) ("The basic principle for the measurement of those damages [for breach of contract] is that of compensation based on the injured party's expectation"); Robert Cooter & Melvin Eisenberg, *Damages for Breach of Contract*, 73 CALIFORNIA LAW REVIEW 1434, 1434 (1985) ("the goal, compensation, and the means, expectation damages, are so ingrained in contract law as to seem self-evident.")

[4] *See* RESTATEMENT (SECOND) OF CONTRACTS, introductory note to chap. 16, at 100 (1981) ("The traditional goal of the law of contract remedies has not been compulsion of the promisor to perform his promise but compensation of the promisee for the loss resulting from breach"); U.C.C. § 1-305 cmt. 1 (explaining that the central purpose of contract damages is to provide compensation for disappointed promisees).

[5] At common law, the principle of compensation shaped the legal rule according to which damages must put the promisee in the position she would have been in if the contract had been performed, and thereby the protection of the expectation interest. *See* Robinson v. Harman 1 Exch. 850, 855 (1848), per Baron Parke ("The rule of the common law is, that where a party sustains a loss by reason of a breach of contract, he is, so far as money can do it, to be placed in the same situation, with respect to damages, as if the contract had been performed."); Wertheim (Sally) v. Chicoutimi

Its goal and function is arguably to *substitute private for public redress*, thereby providing its own contribution to the welfare of society.[6] As advanced by Corbin,

> "[t]here is more than one purpose underlying the rules of law that provide for the giving of damages for breach of contract. One of the ends to be obtained is, without doubt, the keeping of the peace. The party injured by the breach has a sense of grievance. In the absence of a public remedy, he would do his best to redress his own wrong. This means private war, with all of the resulting harm that it entails to the interest of other people. A second purpose in the giving of damages, however, one that is equally important as the first, is the prevention of similar harms in the future. The fact that damages must be paid tends directly to the prevention of breaches of contract"[7]

In effect, the human tendency to retaliate to perceive wrongful and unfair behavior is a well-established type of behavior repeatedly documented in a large body of empirical literature, and often with its own beneficial effect, for it is apt to induce cooperation by the potential victim.[8]

Pulp Co. A.C. 301, 307 (1911), per Lord Atkinson (P.C.) ("And it is the general intention of the law that, in giving damages for breach of contract, the party complaining should, so far as it can be done by money, be placed in the same position as he would have been in if the contract had been performed ... That is a ruling principle. It is a just principle.") The principle and the legal rule it defines is equally found in the American law of contracts. *Cf.* Jaquith v. Hudson, 5 Mich. 123, 133-134 (1858) ("The law, following the dictates of equity and natural justice … adopts the principle of just compensation for the loss or injury actually sustained"); Hawkins v. McGee, 84 N.H. 114, 146 A. 641 (N.H. 1929) ("By 'damages,' as that term is used in the law of contracts, is intended compensation for a breach.")

6 *Cf.* RUDOLPH VON JHERING, GEIST DES RÖMISCHEN RECHTS 113 (Leipzig, Breitkopf und Härtel 1852) ("*Die ersten unausbleiblichen Regungen des verletzten Rechtsgefühls bestehen in der gewaltsamen Reaktion gegen das zugefügte Unrecht, in der Selbsthülfe und Rache; mit Selbsthülfe und Rache hat daher ein jedes Recht begonnen.*") ("The first inevitable impulse of the injured sense of justice consist in a violent reaction against the inflicted wrong, in self-help and revenge; every right thus began with self-help and revenge.") (own transl.); O. W. HOLMES, THE COMMON LAW 37 (Toronto, Typographical Society 2011) (1881) ("My aim and purpose has been to show that the various forms of liability known to modern law spring from the common ground of revenge"); MAX WEBER, WIRTSCHAFT UND GESELLSCHAFT 421 (Tübingen, Paul Siebeck 1922) ("*Die ökonomische Rationalisierung des Rechts begünstigte die Entstehung der Vorstellung, daß die Sühnehaftung nicht sowohl Abkauf der Rache (die ursprüngliche Auffassung) wie Ersatz des Schadens sei. Nichterfüllung eines Kontrakts konnte nun ebenfalls als sühnepflichtige Schädigung qualifiziert werden.*") ("The economic rationalization of the law favored the rise of the conception that liability for composition was not only buying off of vengeance (the original conception) but also compensation for harm. Nonperformance of contract could now be qualified as harm requiring compensation.") (own transl.).

7 5 ARTHUR L. CORBIN, A COMPREHENSIVE TREATISE ON THE RULES OF CONTRACT LAW 30-31 §1002 (1st ed. 1951). *See also* JOHN CALAMARI & JOSEPH PERILLO, CONTRACTS 6 (6th ed. 2009) ("Anthropology and history have proven that a basis of contract law is the desire to keep the public peace.")

8 Retaliation and costly punishment induce cooperation in social dilemmas, as observed in experiments with public good and prisoners' dilemma games. *See* Ernst Fehr & Simon Gächter, *Cooperation and Punishment in Public Goods Experiments*, 90 AMERICAN ECONOMIC REVIEW 980 (2000) (costly punishment of free-riders lead to almost universal cooperation, and maintained it stable in repeated interactions); Fehr & Gächter, *Altruistic Punishment in Humans*, 415 NATURE 137 (similar); Anna Dreber et al., *Winners don't punish*, 452 SCIENCE 348 (2008) (costly punishment more than doubled the amount of cooperation in prisoners' dilemmas); Bettina Rockenbach & Manfred Milinski, *The Efficient Interaction of Indirect Reciprocity and Costly Punishment*, 444 NATURE 718 (2006) (interaction between punishment and reputation building further increases cooperation in public good games); Marco Casari & Luigi Luini, *Cooperation Under Alternative Punishment Institutions: An Experiment*, 71 JOURNAL OF ECONOMIC BEHAVIOR & ORGANIZATION 273 (2009) (punishment under consensus by members of the group induced higher levels of cooperation than individual punishment).

In fact, the role of retaliation to deter socially undesirable behavior (as crimes) was intuited and well explained by Richard Posner in *Retribution and Related Concepts of Punishment*, 9 JOURNAL OF LEGAL STUDIES 71, 78 (1980)

In case of contracts, retaliation can possibly deter violations of the moral norm of keeping promises, or of the principle of *pacta sunt servanda*, if promisors fear and anticipate acts of retaliation by promisees. It is, however, most often unable to provide a net contribution to welfare, since whenever it is implemented, retaliation entails costs for the individual retaliating as well as for its victim. As several empirical studies consistently identify, its social costs most often outweigh the benefits, and the result is a loss of welfare.[9]

Moreover, in the presence of legal enforcement, retaliation is not necessary to induce promisors to perform, for the prospect of the payment of damages fulfills that very same function. Damages for breach consist in a transfer of resources from the party in breach to the breachee, and thus do not create the same loss of resources that retaliation creates when they are implemented.[10] Legal relief is thereby well suited to substitute private for public redress efficiently, and to induce performance of contracts while crowding out the tendency to retaliate to breach of contract, thereby avoiding the loss of welfare it occasions.[11]

This conjecture has not yet been subject to empirical investigation. The existing literature has not examined whether disappointed promisees in fact retaliate to breach of bargained-for promises responsible for no more than loss of expectancy. Moreover, no experimental evidence exists either for the effect of expectation damages to induce performance if and only if performance

---

("A desire to retaliate not motivated by hope of compensation or by desire to establish a reputation that will deter future wrongdoing directed against oneself is a form of negative altruism"). On the evolution of liability in ancient law, *see* Francesco Parisi, *The Genesis of Liability in Ancient Law*, 3 AMERICAN LAW AND ECONOMICS REVIEW 82 (2001).

9 *See* Elinor Ostrom, James Walker & Roy Gardner, *Covenants with and without a Sword: Self-Governance is Possible*, 86 AMERICAN POLITICAL SCIENCE REVIEW 404 (1992) (sanctions alone decreased group earnings because of the costs of sanctioning); Martin Sefton, Robert Shupp & James Walker, *The Effect of Rewards and Sanctions in Provision of Public Goods*, 45 ECONOMIC INQUIRY 671 (2007) (direct costs of sanctioning outweighed the beneficial effect of increased contributions); Martijn Egas & Arno Riedl, *The Economics of Altruistic Punishment and the Maintenance of Cooperation*, 275 PROCEEDINGS OF THE ROYAL SOCIETY: BIOLOGICAL SCIENCES 871 (2008) (costly punishment lead to an overall loss in individual and group welfare); Anna Dreber et al., *Winners don't Punish*, *supra* note 8 (while punishment increased cooperation, it did not increase average payoff in prisoners' dilemmas). *But see* Simon Gächter, Elke Renner & Martin Sefton, *The Long-Run Benefits of Punishment*, 322 SCIENCE 1510 (2008) (punishment decreased welfare in a public good game repeated 10 times, but increased it when the game was repeated 50 times). Additionally, spiteful or anti-social punishment aimed at those that cooperate also exists and is responsible for lower average rates of cooperation, and hence for a loss of welfare. *See* Benedikt Herrmann, Christian Thöni & Simon Gächter, *Antisocial Punishment across Societies*, 319 SCIENCE 1362 (2008) (the higher antisocial punishment in a society, the lower the rate of increase in cooperation and overall welfare).

10 *Cf.* Louis Kaplow & Steven Shavell, *Fairness versus Welfare*, 114 HARVARD LAW REVIEW 961, 998 fn. 73 (2001) ("the bare fact that money may change hands in a lawsuit in certain circumstances, thereby changing how a loss is divided between the two parties, is of no consequence under welfare economics.")

11 Under expectation damages, breach is predicted to occur whenever it is, in the realized contingency, socially efficient. When it occurs, the victim might feel aggrieved and tempted to retaliate if she is not entitled to a legal remedy. Therefore, even if promisors anticipate the prospect of payment of damages perfectly, and only efficient breaches are committed, the function of legal relief to crowd out retaliation is still necessary for maximal social welfare.

Moreover, under imperfect information, imperfect cognitive abilities, or imperfect legal enforcement, breach of contract might arise even when it is not socially efficient (as it in fact happened in the experiment), and is hence expected to be more often committed in reality than in theory. Under any of those imperfections, the need to crowd out retaliation for a maximal welfare is even more pervasive.

is socially efficient, or for the role of legal relief to crowd out retaliation to breach of promises given with consideration.

There are uncountable manners through which promisees can retaliate to breach in modern legal systems that, in general, prohibit the use of violence by the citizens. These can always tell others about breach by the promisor, thereby harming her reputation in the market, or refuse to transact with the promise-breaker in the future even when doing so would be mutually profitable, thereby withholding cooperation.[12] They may file frivolous suits, name and shame the promise-breaker or, as mentioned by Hart and Moore, shade on performance.[13] As noted by Richard Posner, since "retaliation may sometimes operate as a constraint on market activity, even economists who take a narrow view of the proper scope of economics might include retaliation within that scope."[14]

The article seeks to fill in this gap in the literature by developing an experimental study into these two functions of damages for breach, and it attempts to provide empirical evidence for the welfare benefits of legal relief through both channels. To achieve those goals, it investigates, firstly, the function of expectation damages to induce efficient levels of performance. Secondly, it investigates the function of compensatory remedies to crowd out retaliation to breach by the victim. Lastly, it compares the estimated gains of social welfare provided by each of these two functions of damages for breach in order to discuss how far, and under what circumstances, compensation is necessary for the maximization of social welfare.

In order to investigate those questions empirically, the experiment distinguishes between different situations in which the promisor can breach. These take the form of contingencies not covered by the parties' agreement, which is, in the experiment, incomplete and does not condition upon those, as most real-life agreements. The different contingencies distinguish between socially efficient and inefficient breaches, and between breaches committed to avoid losses resulting from an increase in costs of production ("unfortunate contingencies," in the "loss-avoidance paradigm") from breaches committed to achieve higher profits from a substitutive transaction ("fortunate contingencies," in the "overbidder paradigm").[15]

---

[12] Similar to the forms of retaliation discussed by David Charny in commercial relationships. *Cf.* David Charny, *Nonlegal Sanctions in Commercial Relationships*, 104 HARVARD LAW REVIEW 373, 392f. (1990).

[13] *See* Oliver Hart & John Moore, *Contracts as Reference Points*, 73 QUARTERLY JOURNAL OF ECONOMICS 1, 9-10 (2008) ("It may be useful to give some examples of shading. There are many ways one trading partner can hurt another. A seller can shade by cutting quality: in the wedding example, she can stint on some of the ingredients of the wedding cake. Or the seller may withhold cooperation … A third example would be 'working to rule': the seller abides by the strict terms of the contract and offers no more. Buyers can also shade. Although it is harder to imagine a buyer cutting back on quality, it is easy to think of situations where a buyer refuses to make minor concessions or to cooperate… The buyer can also make life difficult for the seller by quibbling about details of performance, by delaying payment, or by giving a bad reference").

[14] *Retribution and Related Concepts of Punishment*, 9 JOURNAL OF LEGAL STUDIES 71, 73 fn. 4 (1980).

[15] The distinction between the "loss-avoidance paradigm" and the "overbidder paradigm" is made by Melvin Eisenberg in *Actual and Virtual Specific Performance, the Theory of Efficient Breach, and the Indifference Principle in Contract Law*, 93 CALIFORNIA LAW REVIEW 975, 997 (2005). *See also* Melvin Eisenberg, *The Disgorgement Interest in Contract Law*, 105 MICHIGAN LAW REVIEW 559, 571 (2006). The distinction between "fortunate" and "unfortunate" contingencies is made in ROBERT COOTER & THOMAS ULEN, LAW AND ECONOMICS 238, 241 (3rd ed. 2000).

Moreover, both functions of damages for breach operate inside a relationship in which parties are bound to each other by the contractual obligation. Damages provide monetary incentives for promisors to perform when the promisor is already committed to perform. This commitment, established by the giving of a promise with consideration, can already induce promisors to perform, independently of the material incentives provided by the legal remedy. At the same time, breach of that commitment provides a reason and motive for the promisee to feel aggrieved and thereby tempted to retaliate. The experiment does not neglect those, but, quite on the contrary, explicitly studies how far promissory commitment, independent of damages for breach, can by itself induce performance by the promisor, and how far breach of promise can trigger retaliation by the promisee.

In sum, the experiment investigates the role of promises given with consideration to (i) induce performance by the promisor and (ii) instigate retaliation to breach by the promise. It then studies the role of expectation damages to (iii) induce efficient performance by the promisor and (iv) crowd out retaliation to breach by the promisee. All these four effects are captured in different types of contingencies that distinguish between the efficiency and inefficiency, and the fairness and unfairness of breach. With that, the experiment provides estimates for the behavioral and welfare effects of expectation damages to induce performance only when performance is socially efficient, and of compensatory remedies for breach to crowd out retaliation in circumstances where breach of promise may in fact induce retaliation.

The next section presents the individual contractual behavior examined in the experiment and the related existing literature. The third section explains the design of the experiment, describes the trade game that subjects played in the different treatments, the content of each treatment, predictions from rational choice theory, alternative hypotheses, and the details of the experimental procedure. The fourth section reports the obtained results and the statistical analysis, and the last section concludes.

I. Parties' Behavior under Investigation and Related Literature

The behavior of contractual parties that is subject to investigation involves the promisor's decision to perform and the promisee's decision to retaliate. They can both be influenced by the two duties created by the contractual obligation: the primary duty to perform and to keep one's part of the deal, and the secondary duty to pay damages and to compensate the promisee in case of breach. Both duties can influence the behavior of both parties, thus delivering four different effects under empirical scrutiny.[16]

The primary duty to perform arises with the giving of a promise with consideration, and creates an obligation for the promisor to undertake the promised act. In establishing that obligation,

[16] Additionally, the experiment studies the effect of the prospect of retaliation by the promisee upon the promisor's decision to perform. This effect is described in the alternative hypotheses in section II.D *infra*.

promissory commitment is apt to induce, by itself, performance by the promisor. In case it is breached, a wrong is committed, which is apt to instigate retaliation by the promisee.

There is evidence that individuals often keep their promises even at a cost for themselves.[17] There is still, however, a wide and unresolved debate on the causes of such behavior, or on why individuals keep promises. Two main explanations put forward in the literature are guilt aversion and moral commitment. According to the theory of guilt aversion, individuals feel guilt when they let the other individual's expectations down and, in doing so, suffer a loss of utility because of that act.[18] According to the theory of moral commitment, individuals have preferences for keeping promises per se, and derive disutility from behaving inconsistently.[19] While there is some evidence for the existence of guilt aversion, there is more evidence in favor of the commitment explanation.[20] More recently, Ederer and Stremitzer provide evidence that promisors can be motivated to keep their promises in order not to disappoint the expectations they create upon promisees.[21]

Existing evidence that individuals keep promises relies, however, on a modified version of the trust game, and is thus obtained where the promisee incurs real monetary losses in case of breach by the promisor. These studies do not implement an incomplete contract where breach is individually unprofitable only in certain contingencies, but rather a game where breach is always profitable. They do not implement contingencies where performance is the dominant strategy and both parties profit from the transaction, as under the status quo, and contingencies where the dilemma indeed emerges, as in case of breach of wholly executory and incomplete contracts studied in the experiment.[22]

---

17 *See* Gary Charness & Martin Dufwenberg, *Promises and Partnerships*, 74 ECONOMETRICA 1579 (2006); Christoph Vanberg, *Why do People Keep their Promises? An Experimental Test of Two Explanations*, 76 ECONOMETRICA 1467 (2008); Charness & Dufwenberg, *Bare Promises: An Experiment*, 107 ECONOMIC LETTERS 281 (2010); Tore Ellingsen et al., *Testing Guilt Aversion*, 68 GAMES AND ECONOMIC BEHAVIOR 95 (2010); Charness & Dufwenberg, *Participation*, 101 AMERICAN ECONOMIC REVIEW 1213 (2011).

18 *See* Charness & Dufwenberg, *Promises and Partnerships*, *supra* note 17, at 1583 (stressing that the motivation for this type of behavior is the individual's beliefs about the beliefs of the other, that is, an individual only experiences guilt if she believes her behavior falls short on the expectation of the other). For the formal model, *see* Pierpaolo Battigalli & Dufwenberg, *Guilt in games*, 97 AMERICAN ECONOMIC REVIEW: PAPERS AND PROCEEDINGS 170 (2007); Battigalli & Dufwenberg, *Dynamic Psychological Games*, 141 JOURNAL OF ECONOMIC THEORY 1 (2009).

19 *See* Tore Ellingsen & Magnus Johanesson, *Promises, Threats and Fairness*, 114 ECONOMIC JOURNAL 397 (2004) for a formal model and initial empirical evidence, and Ying Chen et al., *Selecting Cheap-Talk Equilibria*, 76 ECONOMETRICA 117 (2008) for a model of belief-independent costs of inconsistency.

20 *Compare* Charness & Dufwenberg, *Promises and Partnerships*, *supra* note 17 (evidence in favor of guilt aversion) with Vanberg, *Why do People Keep their Promises*, *supra* note 17 (evidence that favors the commitment explanation) and Charness & Dufwenberg, *Bare Promises: An Experiment*, *supra* note 17 (limited support for both explanations) and Ellingsen et al., *Testing Guilt Aversion*, *supra* note 17 (scarce support for the existence of guilt aversion in dictator and trust games, with complete and incomplete information) and Charness & Dufwenberg, *Participation*, *supra* note 17, at 1234 (concluding that "since, on balance, both theories capture many but not all aspects of the data, we shall not declare a winner.")

21 *See* Florian Ederer & Alexander Stremitzer, *Promises and Expectations*, Cowles Foundation Discussion Paper n. 1931 (2014) (also proposing a lexicographic theory of promise-keeping). *See* further Stephan Tontrup, *Pacta Sunt Servanda – Testing Voluntary Compliance to Contracts*, Max-Planck Institute of Economics Working Papers (2015).

22 *See* note 47 *infra* and the accompanying text.

There is pervasive evidence that individuals engage in costly punishment against those that violate a norm of cooperation in social dilemmas, or a fairness norm in bargaining games.[23] Quite remarkably, there is very little evidence on the existence of retaliation to breach of promise, and, similarly to experiments investigating promise-keeping, it is obtained in games where the promisee suffers material losses because of breach, and not only a loss of expectancy.[24]

In the present study, the question examined experimentally is whether promisees retaliate to breach of promise when no more than the expectation interest is at stake, and why they may do so: because of (i) the loss of expectancy they endure, (ii) the unfairness of the resulting distribution, or (ii) the inefficiency of the result. In distinguishing between different contingencies where breach of promise arises, and where only some lead to an unfair or inefficient result, the experiment inquiries into the circumstances where retaliation to breach of promise is most likely to arise.

The secondary duty to pay damages, in turn, arises only if the promised event does not come to pass, and it creates an obligation for the promisor to compensate the promisee. In imposing a monetary cost on breach, expectation damages are apt to induce efficient levels of performance by the promisor. In providing compensation for the promisee, they are apt to crowd out retaliation.

To the author's best knowledge, there is no precise experimental evidence that expectation damages in effect induce performance of contractual promises if performance is socially efficient, and not otherwise.[25] Expectation damages make the promisor internalize the negative externality of breach, given by the amount the promisee does not earn because of breach, and are predicted to induce only socially efficient performance. This clear theoretical prediction concerning parties' individual behavior further implies that social welfare will be maximal under expectation damages and is subject to empirical examination.

Expectation damages further provide relief for promisees to redress breach, and can thereby provide their own contribution to social welfare by crowding out the victim's tendency to retaliate to breach, in effect fulfilling the function to substitute private for public redress, thereby avoiding a loss of welfare. If promisees perceive breach of contractual promises as a wrong in need of

---

[23] *See supra* note 17 (evidence that individuals retaliate in social dilemma games). Punishment in ultimatum games, for example, is pervasive across societies and cultures. *See* Joseph Heinrich et al., *In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies*, 91 AMERICAN ECONOMIC REVIEW 73 (2001).

[24] *But see* Olivier Bochet & Louis Putterman, *Not just Babble: Opening the Black Box of Communication in a Voluntary Contribution Experiment*, 53 EUROPEAN ECONOMIC REVIEW 309, 322 (2009) (providing some evidence, in their result 9 of higher levels of costly punishment in case of breach of promise than in the absence of a promise, in a public good game, and thus where the individual that promised induced others to contribute to the public good, and hence to incur a real monetary loss – and not just loss of expectancy – in case of breach).

[25] *Cf.* Iris Bohnet, Bruno Frey & Steffen Huck, *More Order with Less Law: On Contract Enforcement, Trust, and Crowding*, 95 AMERICAN POLITICAL SCIENCE REVIEW 131 (2001) (experiment with a trust game that included expectation damages but where breach could never be efficient, and where the promisee could not decide whether to claim damages or not, since the promisor was held liable according to some exogenous probabilities); Ben Depoorter & Stephan Tontrup, *How Law FramesMoral Instuitions: The Expressive Effect of Specific Performance*, 54 ARIZONA LAW REVIEW 673 (2012) (experiment that included, among others, the possibility for parties to claim the equivalent of expectation damages, but with very different research questions).

redress, and in effect tend to retaliate against it, then damages for breach are well apt to crowd out retaliation. This requires that legal enforcement of contracts go beyond the imposition of a price on breach and further provide compensation for disappointed promisees, as it in fact does.

Lastly, the experiment investigates whether legal relief to promisees to redress breach has its own beneficial impact upon social welfare not only by inducing efficient performance, but also by reducing welfare losses from retaliation. In order to estimate the welfare gains from inducing promisors to perform, there is the need to subtract from those gains the possible gains that promises themselves deliver, independent of damages, in inducing performance. On the other hand, in order to estimate the welfare gains from reducing retaliation by promisees, there is the need to subtract the gains created by retaliation in deterring breach from the social costs of retaliation, so that the total net effect upon social welfare captures both the benefits and costs of retaliation.[26] By comparing those net estimated social welfare effects, the welfare functions of legal redress are examined. If legal redress indeed provides net gains through both channels, then compensation for breach of promise can be, on social welfare grounds, justified in the law of remedies for breach.

II. THE EMPIRICAL STUDY

The experiment attempts to investigate the effects of contractual promises, expectation damages, and retaliation on individual behavior in a controlled environment. This approach permits a more careful analysis of causal relationships between those institutions and behavior in a game where subjects' decisions and rewards are real and material. It attempts to abstract from, and to control for, other factors that can influence parties' contractual behavior, with the goal of providing evidence for the effect of the institutions under study, and their interactions, in isolation.[27]

This does not mean that other factors present in real contractual relationships are irrelevant for the behavior of the parties. Trust, custom, commercial good manners, reputation, and several others elements all play their role upon contractual behavior. Empirical research, however, must control for the factors outside the scope of the study in order to investigate the effect of any specific institution upon individual behavior, and this guided the design of the study.

The experimental design attempts to establish such a controlled environment first through the implementation of a control group. It aims at capturing the influence of other factors on parties' behavior, providing estimates of the effect of promises and damages net of those. Moreover, the experiment followed the orthodox procedure in economics, ensuring anonymity, the absence of external communication between subjects, monetary rewards for individual choices (incentive-compatibility), no deception, and the precise same amount of information provided to all subjects.

---

[26] As explained before, retaliation induces cooperation and induces other forms of socially desirable behavior, but it does so only at a cost, and this is the reason why its net effect upon social welfare must consider both.

[27] On the experimental method in economics, *see, e.g.*, Vernon Smith, *Experimental Economics: Induced Value Theory*, 66 AMERICAN ECONOMIC REVIEW 274 (1976); Armin Falk & James Heckman, *Lab Experiments are a Major Source of Knowledge in the Social Sciences*, 326 SCIENCE 535 (2009).

A. The Implemented Trade Game

Subjects played, in all treatments, a trade game that resembles the interaction between a seller and buyer. In it, sellers take the role of promisors and buyers of promisees. The seller could produce and deliver one unit of a good to the buyer, who in return makes the payment of its price. The trade game consisted of four different phases, or "dates": at date 1, parties met and could enter into an agreement to trade (except in the control group, where parties could not communicate). At date 2, the state of the world was realized, and both parties were informed of it. At date 3, the seller decided to perform or to breach. At date 4, and only in case of breach by the seller, the buyer could retaliate (in the treatments with retaliation) and concomitantly claim damages for breach (only in the treatment with contracts).

After the realization of the state of the world, at date 2, both parties observed the realized state, and the buyer observed the decision of the seller at date 3. Consequently, the seller was always well aware of the consequences of her decision upon her own earnings and upon the earnings of the buyer, as well as of the efficiency and fairness of the result of her decision. The seller knew that the buyer also observed the realized state and her decision to perform or breach, and that the buyer could thus take the motives and consequences of breach into consideration before deciding to claim damages and to retaliate (in the treatments including those actions).

Figure 1. Timeline of events

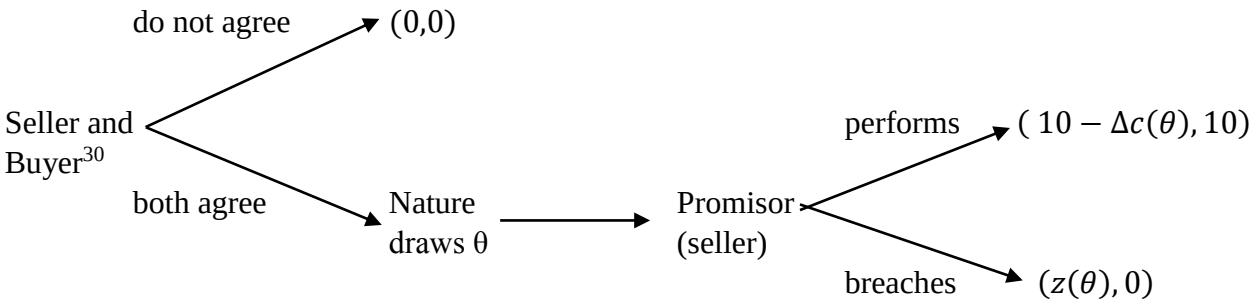
Date 1	Date 2	Date 3	Date 4
Agreement stage	*Resolution of risk*	*Performance stage*	*Enforcement stage*
Subjects are matched and can enter into an agreement	The state of the world is realized, determining the seller's costs	Seller decides to perform or breach; Buyer pays in case of performance [28]	Buyer decides, in case of breach, to retaliate and to claim damages

The buyer's valuation of the good was given by $V = 30$ and the price in the agreement was $P = 20$. The buyer's expectancy in the performance of the agreement was hence always equal to 10. The seller's costs of performance, in contrast, depended on her costs of production $c(\theta)$ and on her net gains $z(\theta)$ from breach to profit from a substitutive transaction, where θ_n denotes the realized state of the world, with $n = \{0,1,2,3,4\}$.

[28] The buyer could not actively decide to pay or to default on payment, as detailed below, since this would create uncertainty in the seller, who could then decide to breach because of fear of default by the buyer. Since this is not the object of the study, there is the need to control for such effect, and therefore payment by the buyer was implemented automatically whenever the seller delivered the good to the buyer, and automatically *not* implemented whenever the seller decided not to deliver.

Under the status quo θ_0, the seller's costs of production c_0 were equal to 10, and the seller had no outside option. She thereby earned, just as the buyer, 10 through performance of the agreement. In all other states, either the seller's costs rose by $\Delta c(\theta)$, or the seller was offered an outside option and could breach in order to make net profits of $z(\theta)$.[29]

Figure 2. The trade game



There were five different possible contingencies that could be drawn at date 2. The status quo (state 0) was maintained in half of the interactions. In the other half, Nature selected one out of four possible contingencies (states 1 to 4). These included a low and a high increase in costs of production, and an outside price offer that was slightly or considerably higher than the price in the original contract. They thereby distinguish socially efficient from socially inefficient breaches, and breaches committed to achieve higher profits (the ones that create inequality, or unfair breaches) from breaches committed to avoid incurring losses (the ones that avoid inequality, or fair breaches).

Table 1. Variations in the seller's costs of performance across states

Higher costs of production (fair breach)	Higher outside offer (unfair breach)
STATE 1 (θ_1) *Inefficient breach* that is *fair* (*avoids inequality*)	STATE 2 (θ_2) *Inefficient breach* that is *unfair* (*creates inequality*)
STATE 3 (θ_3) *Efficient breach* that is *fair* (*avoids inequality*)	STATE 4 (θ_4) *Efficient breach* that is *unfair* (*creates inequality*)

[29] Note that the third-party was not, in the experiment, a real person but a mere offer that the seller could accept instead of trading with the buyer. This was reiterated to participants: "this offer is hypothetical and not done by another participant in the lab" (as in the instruction in the experiment). Moreover, it is assumed that the third-party, who "arrives later" than the first buyer, makes an offer corresponding to her whole valuation of the good.

[30] More precisely, if either the buyer or the seller do not agree to exchange promises to perform, then they cannot trade and realize any gain from trade or loss. The seller would not incur costs to produce the good if the buyer does not promise to come back to pick it up and pay for it, and the buyer will not come back to pick up the good if the seller does not promise she will produce it. If they both agree to exchange promises, or on a contract, then the game proceeds.

More precisely, the contingencies were as follows:

θ_0, the status quo, with $c(\theta_0) = 10$ and no outside offer

θ_1, an increase in costs of production by 15 and no outside offer, with $c(\theta_1) = 25$

θ_2, an outside price of 25 and cost of production remain as in θ_0, with $z(\theta_2) = 15$

θ_3, a high increase in costs of production by 25 and no outside offer, with $c(\theta_3) = 35$

θ_4, a high outside price of 35 and cost of production remain as in θ_0, with $z(\theta_4) = 25$

For convenience, the final payoffs of the seller and of the buyer, in each contingency, are all presented below, together with the predictions on the seller's behavior.

B. Experimental Treatments

The experimental design consisted of three main treatments (trade, promises and contract), all implemented with and without the possibility of retaliation, and hence in six treatments. In all of them, subjects took the role of a buyer or a seller and played a series of single anonymous trade games. The parameters of the game described above and the different possible contingencies were constant across all treatments, which differed only in the possibility to exchange promises before playing the game and in the possibility to retaliate or to claim damages after the game.

In the *first treatment "trade"* (the control group), subjects played the trade game without any previous communication between them, and hence in the absence of any form of promissory commitment. They were, in each interaction, only randomly allocated with another participant in the opposite role at date 1, and took no decision or action at that moment. They subsequently observed the state of the world at date 2, and at date 3 the seller decided to produce the good and deliver it to the first buyer in exchange for payment of the price or not.

In the *second treatment "promises,"* subjects were matched at date 1 and could enter into an exchange of promises. It stated that "the seller promises to produce the good and trade it with the current buyer, while the buyer promises to pay the price of 20 points to the seller for the good". If the seller and the buyer both promised, then the state of the world was realized at date 2, and the seller decided to keep the promise at date 3. Apart from the promises, all other remaining elements of the game were identical to the control group.

In the *third treatment "contract,"* subjects were matched at date 1 and could enter into a contract. The contract consisted in an exchange of promises with the precise same content as in treatment promises, but also included a clause allowing the buyer to claim damages in case the seller decided "not to deliver the good to the buyer." The amount of damages the buyer was entitled to claim in case of breach was measured by her loss of expectancy, equal to 10. Apart from this

clause and the related possibility to claim damages in case of breach, all remaining elements were identical to treatment promises.

Each main treatment was implemented with and without the possibility of retaliation by the buyer. Retaliation was implemented in the traditional form of costly punishment, meaning that the buyer could spend 2 points to deduct 10 points from the earnings of the seller. Each subject played in just one main treatment (trade, promises, or contract), in the presence and absence of possible retaliation (the details of the procedure are presented below). The experimental design is depicted in table 2:

Table 2. Implemented treatments

	Seller's decision to perform	*Buyer's decision to retaliate*	
	Treatment trade no commitment, no enforcement	*Treatment trade with retaliation* no commitment, retaliation	
Effect of promises to induce performance	*Treatment promises* promissory commitment, no enforcement	*Treatment promises with retaliation* promissory commitment, retaliation	Effect of breach to trigger retaliation
Effect of damages to induce performance	*Treatment contract* promissory commitment, damages	*Treatment contract with retaliation* promissory commitment, retaliation, and damages	Effect of damages to crowd out retaliation

Although subjects actively decided to enter into the exchange of promises, or into the contract, as well as, only in treatment contract, to claim damages for breach, the treatment variables under study are *performance* and *retaliation*.[31] In treatments without retaliation, depicted in the first column of table 2, the focus lies on the effect of promissory commitment (between treatments trade and promises) to induce performance by the promisor, and of the material and pecuniary incentives created by expectation damages to induce performance *net* of the effect of promises (between treatments promises and contract). In treatments with retaliation, depicted in the second column, the focus lies on the effect of breach of promise to trigger retaliation by the promisee

[31] All three treatments equally involved a seller and a buyer of a good, who only in treatments "promises" and "contract" can accurately be denoted promisor and promisee (because they entered into the exchange of promises). In the present text, these terms are used interchangeably. In the experiment, differently, parties were always called seller and buyer in all treatments in order to avoid possible demand effects, as explained below in the section that explain the experimental procedure.

(between treatments trade and promises), and of expectation damages to crowd it out (between treatments promises and contract).

The function of treatment trade in the experiment is to control for the effect of other factors beyond promissory commitment and damages on the behavior of the parties.[32] For example, the seller may want to keep the promise because of her individual preferences against the inequality that would result from breach. The buyer may, in turn, desire to retaliate to breach because of the resulting inequality, independent of the presence of promissory commitment. With the control group, the estimated *difference* in behavior across treatments excludes those effects, which are captured inside the control group, and can then be attributed solely to the effect of promises and damages for breach.

Considering all treatments, the complete version of the game is described as follows.

Let

$q_s \in \{0,1\}$ be the decision variable of the seller to produce the good and deliver it to the buyer, equal to unity in case of performance;

$r_b \in \{0,1\}$ be the decision variable of the buyer to retaliate against the seller by imposing a loss of $R = 10$ at own personal costs $K = 2$, and

$e_b \in \{0,1\}$ be the decision variable of the buyer to enforce the contract and claim the equivalent of expectation damages, given by $D = V - P = 10$.

The profits of sellers and buyers, represented respectively by π_s and π_b, are as follows:

Seller's payoff: $\pi_s = q_s\big(P - c_0 - \Delta c(\theta)\big) + (1 - q_s) \cdot (z(\theta) - e_b D - r_b R)$

For the parameters of the game, the seller's payoffs are given by

$$\pi_s = \begin{cases} 20 - 10 - \Delta c(\theta) & if\ seller\ performs \\ z(\theta) - 10e_b - 10r_b & if\ seller\ breaches \end{cases}$$

The seller earns the price of 20 minus costs of production $c_0 = 10$ if she decides to perform, minus possible increases $\Delta c(\theta)$ in those. If she breaches, then she makes net profits $z(\theta)$ from the possible outside transaction ($z(\theta) = 0$ if no outside offer materializes), loses 10 if the buyer decides to claim damages, and loses another 10 if the buyer decides to retaliate.

Buyer's payoff: $\pi_b = q_s(V - P) + (1 - q_s) \cdot (-r_b K + e_b D)$

For the chosen parameters, the buyer's payoffs are given by

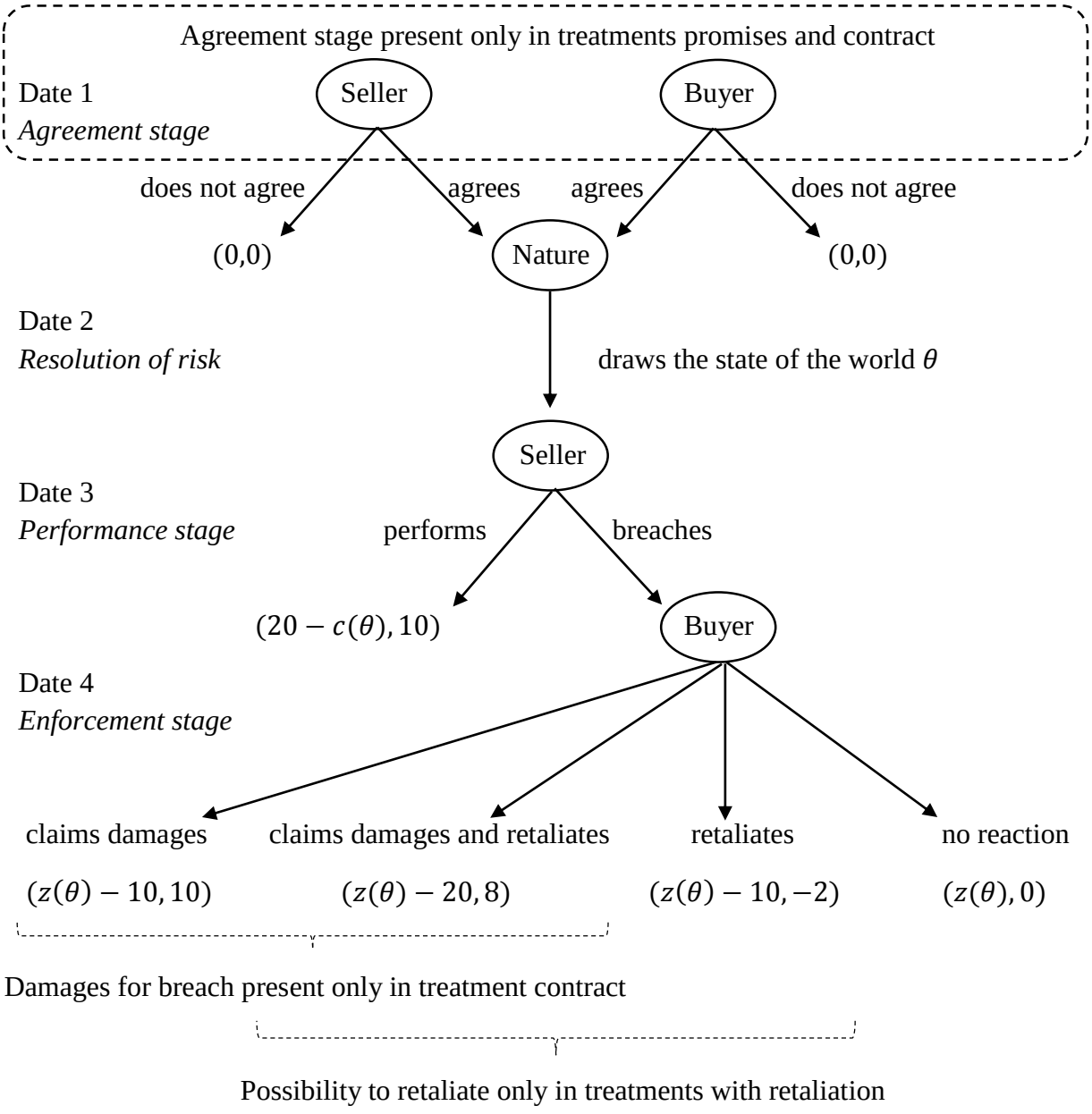
$$\pi_b = \begin{cases} 10 & if\ seller\ performs \\ -2r_b + 10e_b & if\ seller\ breaches \end{cases}$$

[32] All three treatments equally involved a seller and a buyer of a good, who only in treatments "promises" and "contract" can accurately be denoted promisor and promisee (because they entered into the exchange of promises). In the present article, these terms are used interchangeably. In the experiment, differently, parties were always called seller and buyer in all treatments in order to avoid demand effects, as explained below in section III.E.

The buyer always earns her expectancy of 10 in case of performance by the seller. In case of breach, the buyer neither earns nor loses anything in the absence of retaliation and damages. Only her expectancy is at stake in the game. In the treatments that include enforcement possibilities, the buyer incurs the costs of retaliation equal to 2 if she decides to retaliate, and earns damages for breach equal to 10 if she decides to enforce the contract. For convenience, the parties' payoffs in each possible contingency, depending on the seller's decision, are resumed in table 3 further below, together with the predictions from strict rational choice.

The complete game-tree takes the following form:

Figure IV.3. Complete game-tree



Entering into the agreement was individually profitable for both parties independent of the promisor's individual understanding of the morality of promises and contracts. A promisor that for moral reasons planned to keep promises under all circumstances had an expected utility of 5 by entering into the agreement and accordingly performing in all contingencies:

$$EU(q_s, \theta) \sum_{\theta=0}^{\theta=4} p_\theta \pi_s\,(q_s = 1) = 0.5 \cdot 10 + 0.125(-5) + 0.125(-15) + 0.125 \cdot 10 \cdot 2 = 5$$

C. Strict Rational Choice Predictions

The payoffs and further consequences of the seller's decision, in each possible contingency, are resumed in table 3 below. In the absence of retaliation and damages, the seller's decision at date 3 completely determined both parties' earnings, which are depicted in the first and second columns. The seller's individually optimal decision is there denoted with the star.

Table 3. Payoffs and consequences of the seller's decision in treatments trade and promises (stars indicate equilibrium behavior and outcomes)

	Seller's decision and related payoffs (seller,buyer)	Consequences of breach *for the seller*	Consequences of breach *for the buyer*	Consequences of breach *upon SW*	Consequences of breach *upon inequality*
State 0	*perform* *(10,10)* breach (0,0)	does not earn 10	does not earn 10	− 20	none
State 1	perform (-5,10) *breach* *(0,0)*	avoids loss of 5	does not earn 10	− 5	avoids inequality of 15
State 2	perform (10,10) *breach* *(15,0)*	gains extra 5	does not earn 10	− 5	creates inequality of 15
State 3	perform (-15,10) *breach* *(0,0)*	avoids loss of 15	does not earn 10	+ 5	avoids inequality of 25
State 4	perform (10,10) *breach* *(25,0)*	gains extra 15	does not earn 10	+ 5	creates inequality of 25

The seller is predicted to perform, in the absence of damages, only under the status quo (state 0). In all other states, it is individually optimal for the seller not to trade, but rather to breach, since by doing so the seller maximizes her individual gains. The third column summarizes the consequences of breach for the seller herself. These are calculated with respect to the gains or losses that the seller would make through performance, and reflect the distinction between gain-

seeking and loss-avoiding breaches. Although there is no difference, in economic terms, between breach in states 1 and 2, and between breach in states 3 and 4, the distinction is of relevance for the creation or avoidance of an unequal final distribution between the parties.

The last three columns summarize the consequences of the seller's decision that are, under strict individual self-interest, immaterial for the seller's maximization of profits. The fourth column reports the consequences of the seller's decision for the buyer, given by the buyer's loss of gains from trade, which is equal to 10 in all contingencies. The fifth column reports the consequences of breach for aggregate social welfare, and the last column the consequences of breach for the disadvantageous inequality in the final payoffs experienced by the buyer.

In the absence of promises and damages for breach, strict rational choice predictions are clear and univocal. As explained in the following paragraphs, these do not change in the presence of promises nor in the presence of retaliation, but only in the presence of expectation damages. Additionally, the buyer is never predicted to retaliate, either in the absence or presence of promises.

Consider first the *seller's decision to keep promises*. According to standard rational-choice models, promises, in the absence of an effective enforcement mechanism, are simply "cheap talk" because they do not directly affect payoffs.[33] They cannot induce a change in the behavior of the seller between treatments trade and promises. Sellers are hence predicted not to change their behavior because of a prior promise, but are rather, on the contrary, *predicted to behave equally in treatment trade and in treatment promises,* keeping their promises only under the status quo, and breaking them in all other states.

Rational-choice models also do not predict *buyers to retaliate,* since costly punishment is a strictly dominated strategy in single and anonymous interactions. It cannot deliver any material gain to the individual, either present or prospective. The same prediction holds for the buyer's decision to retaliate to breach of promise. The perceived wrong in breach by the seller is irrelevant for self-interested behavior that requires *no* investment in retaliation by the buyer. *Buyers are hence predicted to behave equally in the absence or presence of prior promises*, never retaliating against the seller's decision.[34]

---

[33] The only equilibrium in any game with cheap talk is a "babbling" equilibrium where communication and promises do not alter individual behavior. *Cf.* Joseph Farrell & Matthew Rabin, *Cheap Talk*, 10 JOURNAL OF ECONOMIC PERSPECTIVES 103, 107-108 (1996) (because of common knowledge of rationality, cheap talk is ignored and irrelevant in search for equilibria). A player's utterances need not be correlated with her private information or true intentions, such that the other player must ignore the other one's utterances, which are just "babble" and meaningless. The refined argument for talk being cheap is in Robert Aumann, *Nash Equilibria are not Self-Enforcing*, in ECONOMIC DECISION MAKING: GAMES, ECONOMETRICS AND OPTIMISATION 201 (J. Gabszewicz et al. eds, Amsterdam, North-Holland 1990) (communication will not affect the result of a game if the signaler's preferences depend on the other player's strategy choice). *See also* Ying Chen et al., *Selecting Cheap-Talk Equilibria*, 76 ECONOMETRICA 117 (2008) ("every cheap-talk game has a degenerate, 'babbling' equilibrium outcome in which the Sender's message contains no information, and, on the equilibrium path, the Receiver's response is equal to her ex ante optimal choice").

[34] Promises, in the present game and experiment, cannot, according to refined economic theory, *signal* anything meaningful for the other individual about one's true type, morality, or true intentions because they are, first of all, costless. Secondly, even if they were costly utterances, they would further need to be *negatively* correlated with what

The conjunction of both predictions from strict rational-choice models delivers the prediction concerning the *seller's decision under retaliation*. Rational sellers must anticipate the buyer's decision not to retaliate, both in treatment trade and promises. Sellers are hence predicted never to alter their behavior because of the possibility of costly punishment. They are predicted to take the very same decisions in treatments *trade*, *promises*, *trade with retaliation* and *promises with retaliation*, never deciding to trade except in state 0, the only case where trade and promise-keeping is individually profitable.

Predictions do change when the agreement is legally enforceable. In the treatment contract, damages for breach do change the equilibrium of the game, as the contract allows buyers to claim damages whenever the seller breaches the contract. In the experiment, this action was a real choice of the participant in the role of a buyer, and one that did not entail any cost for her. The buyer is hence predicted to claim damages whenever the seller breaches, since this maximizes the buyer's earnings. Sellers, accordingly, must anticipate the choice of buyers when making their ex ante optimal decision to perform or breach in the first place.

Table IV.4. Payoffs and consequences of the seller's decision to breach in treatment contract (stars indicate equilibrium behavior and outcomes)

	Seller's decision and related payoffs (seller,buyer)	Consequences of breach *for the seller*	Consequences of breach *for the buyer*	Consequences of breach *for total SW*	Consequences of breach *upon inequality*
status quo	*perform* *(10,10)* breach (-10,10)	loses 20	-	- 20	create inequality of 20
State 1 c = 25	*perform* *(-5,10)* breach (-10,10)	loses 5	-	- 5	increases inequality by 5
State 2 z = 25	*perform* *(10,10)* breach (5,10)	loses 5	-	- 5	creates inequality of 5
State 3 c = 35	perform (-15,10) *breach* *(-10,10)	*avoids losing 5*	-	+ 5	avoids (disadv.) inequality of 5
State 4 z = 35	perform (10,10) *breach* *(15,10)*	*earns 5 extra*	-	+ 5	creates (disadv.) inequality of 5

they would be signaling, otherwise everybody would invest in the signal equally and the signal would then not distinguish between types (such as moral, immoral, and amoral individuals). In precise terms, promises here do not fulfill Spence's critical assumption in his canonical model of signaling (in the labor market). *See* Michael Spence, *Job Market Signaling*, 87 QUARTERLY JOURNAL OF ECONOMICS 355, 358 (1973) ("a signal will not effectively distinguish one applicant from another, unless the costs of signaling are negatively correlated with productive capability").

Incentives created by expectation damages induce performance under the status quo but also in states 1 and 2, as reported in the second column of table 4. In the game, expectation damages induce performance if and only if performance is, in the realized contingency, socially efficient. In contingencies where performance is socially inefficient, i.e. in states 3 and 4, expectation damages do not induce performance, but rather give incentives for the seller to breach and pay damages. The prediction is that sellers will perform in states 1 and 2 and will breach in states 3 and 4, when breach and payment of damages is the individual strictly dominant strategy.

The buyer, as reported in the fourth column, never suffers any loss in treatment contract, independent of the seller's decision. Expectation damages make the buyer whole and compensate her for lost expectancy, the buyer's sole interest at stake in the game. The fifth column reports the consequences of breach for overall social welfare, and the last column reports the consequences of breach for inequality between the parties. The only contingency where breach of contract followed by the payment of expectation damages still creates disadvantageous inequality (of 5) for the buyer is state 4, as reported in the last column of table 4.

Note that although damages for breach do alter the *equilibrium* of the game, they do not alter in any manner the maximal possible gains of social welfare that parties could realize in any contingency – as well as the overall maximal gains from trade in the game as a whole. That is, the same total social welfare is equally achievable in treatments trade, promise and contract, with and without retaliation.

Table 5 resumes the predictions from strict rational choice models for the seller's decision to perform and for the buyer's decision to retaliate in all three treatments:

Table 5. Predictions from strict rational choice models

	Seller	Buyer
Treatment trade	Perform only under the status quo θ_0, refuse to trade in other contingencies	Never retaliate against any type of refusal to trade
Treatment promises	Perform only under the status quo θ_0, break promises in other contingencies	Never retaliate against any type of breach of promise
Treatment contract	Perform when efficient $(\theta_0, \theta_1, \theta_2)$, breach the contract otherwise (θ_3, θ_4)	Never retaliate against any type of breach of contract, always claim damages

D. Alternative Hypotheses

The *first hypothesis is that promisors will perform more often in the presence of promissory commitment than in its absence*. According to the design, sellers are hypothesized to perform more often in treatment promises than in treatment trade, where there was no obligation for sellers to take any specific course of action. Except in the status quo, breach is the action that maximizes the seller's profits and is hence her strictly dominant strategy. However, individuals may feel morally obliged to undertake the promised act even when circumstances change and performance becomes individually unprofitable, and hence to keep their promises for moral reasons.

The *second hypothesis is that disappointed promisees will retaliate more often in the presence of promises than in their absence*. In the design, it corresponds to hypothesized higher rates of retaliation in treatment promises than in treatment trade, where the seller, in deciding not to implement the exchange, does not break any promise.[35] Moreover, this effect may depend or not on the type of the realized contingency.

There are three possible causes, distinguished by the different contingencies, for the buyer to retaliate. If loss of expectancy causes retaliation, then buyers are expected to retaliate more often in all contingencies because they equally lost their expectation in all of them. If inefficiency from breach causes retaliation, then buyers are expected to retaliate more often only in the contingencies where breach is inefficient, and hence only in states 1 and 2. If the realization of inequality from breach of promise causes retaliation, then buyers shall retaliate more often only in states 2 and 4, where the seller creates inequality through her deliberate decision to breach.

The *third hypothesis is that expectation damages induce performance by the promisor, and that this effect is restricted only to those contingencies where performance is socially efficient*. In the design, it corresponds to the inducement of performance only in the contingencies in which performance is socially efficient. These are, as detailed above, states 1 and 2, or those in which performance generates overall gains of welfare. In states 3 and 4, in contrast, expectation damages are predicted to induce breach.

The *fourth hypothesis is that the availability of expectation damages reduces retaliation by the promisee*, in effect fulfilling the function to substitute private for public redress. Promisees are expected to retaliate less often in treatment contract, where they can claim expectation damages in case of breach, than in treatment promises, where this possibility is absent. If promisees perceive breach of bargained-for promises as a wrong in need of redress, and in effect retaliate more often against breach of promise than mere refusals to trade, as in hypothesis 2, then damages for breach are predicted to crowd out retaliation to breach.

---

[35] Individuals with social preferences may have a tendency to retaliate against the decision of the seller because of the inequality or inefficiency it creates. Between treatments trade and promises, however, inequality and inefficiency are held constant by design, and therefore cannot explain any difference in behavior between treatments, and in the same contingency. If buyers retaliate more often against the seller in treatment promises than in treatment trade, then this is caused by the presence of promissory commitment.

A last and *fifth hypothesis is that promisors will perform more often whenever promisees can retaliate*. This hypothesis relates the impact of retaliation to breach, if existent, upon the seller's behavior and decision to perform. Promisors may well anticipate irrational retaliation and decide to keep promises and contracts in order to avoid possible losses from retaliation by buyers. Sellers are hypothesized to perform more often in treatments where the buyer has the option to retaliate than in those where the buyer does not have it.

E. Experimental Procedure

There were six experimental sessions, and each session included three parts. Participants were randomly allocated to the role of either buyer or seller at the beginning of the session and kept the same role throughout the whole session. Each subject participated in only one session. In each, subjects played only one of the main three treatments, and in the different parts of the session, that main treatment was implemented with and without retaliation, as depicted in table 6 below.

Table 6. Order of implementation of treatments

Session	*Main Treatment*	*Part 1*	*Part 2*	*Part 3*
1	Trade	Trade	Trade + retaliation	Trade + retaliation
2	Trade	Trade + retaliation	Trade	Trade + retaliation
3	Promises	Promises	Promises + retaliation	Promises + retaliation
4	Promises	Promises + retaliation	Promises	Promises + retaliation
5	Contract	Contract	Contract + retaliation	Contract + retaliation
6	Contract	Contract + retaliation	Contract	Contract + retaliation

In each part, participants played a series of *eight trade games*. They played four games in which the status quo was maintained and one game where each of the four different contingencies materialized.[36] The order in which the contingencies materialized was unknown to participants, but the same for all of them in each period. Sellers received feedback on whether they were punished or not only at the end of each treatment. Each complete session involved hence 24 games, which were simple, brief, and easily understandable, and each session lasted only about one hour.

[36] The order in which the contingencies materialized was not known to participants, but pseudo-randomized. It was not simply random only because of the need to have each contingency materialize once in each treatment, and a random ordering could lead to a certain contingency never being implemented. This was the reason for the pseudo randomization of the order of realization of the different states of the world. Subjects were informed that the status quo would occur in half of the eight interactions, and that each of the other contingencies would occur in one out of every eight interactions, as in fact implemented.

After each game, subjects played the next game with another random subject, and hence single interactions were implemented throughout the whole experiment. It is worth emphasizing that only single interactions are analyzed and studied. By abstracting from repetition and other factors, and controlling for its effects upon parties' behavior, the experiment analyzes the relationship between promissory commitment, retaliation, and damages for breach in isolation, developing a *ceteris paribus* analysis. Therefore, prospective gains from cooperation that could be induced by punishment in the form of tit-for-tat or grim-trigger strategies, reputational concerns, signaling, and self-selection all *cannot* explain parties' behavior in the experiment. All such factors are isolated, abstracted from, and kept constant *by design*.

Written instructions were distributed at the beginning of each part and subjects were not informed about the content of each part until the beginning of that part. For example, subjects in session 1 first played treatment trade without any knowledge of the content of the subsequent parts, and only in part 2 received new instructions explaining that they would participate in another part that included the possibility of retaliation. Another treatment with retaliation was implemented in part 3 of each session in order to obtain more observations on the buyer's decision to retaliate, which could only be observed when the seller in fact breached, and were expected to be fewer.[37]

In the control group "trade," parties did not have any option to enter into an agreement to trade, and therefore were only randomly matched at date 1. Accordingly, in the experiment, subjects in the control group read on the first screen:

> "You have been randomly matched by the computer to another (buyer/seller) with whom you may trade one unit of a good."

In treatment "promises," subjects read the exact same sentence above, but were furthermore presented with the possibility to agree to enter into an exchange of promises at date 1 with the following content:

> "In the exchange of promises, the seller promises to produce the good and trade it with the current buyer, while the buyer promises to pay the price of 20 points to the seller for the good."

In treatment "contract," the agreement was complemented with a damages clause that

> "allows the buyer to claim 10 points of compensation in case the seller does not fulfill it."[38]

---

[37] The reader that worries with ordering effects and with the possibility that the treatment implemented in part 3 could not deliver unbiased observations in that part can relax. Tests of hypotheses are implemented with all data as well as without data from part 3 (the last ones are presented in footnotes since results are mostly consistent). Moreover, regressions considering data collected only in part 1, in part 1 and part 2, and in all three parts are all developed and concomitantly presented in the tables of results, and controls for ordering effects are also included.

[38] The only difference in the way both agreements were stated was in how they were called: an exchange of promises in treatment promises, and a contract in treatment contract. All the remaining wording was exactly the same. Therefore, in treatment contract, the agreement consisted of: "In the contract, the seller promises to produce the good and trade it with the current buyer, while the buyer promises to pay the price of 20 points to the seller for the good. The contract allows the buyer to claim 10 points of compensation in case the seller does not fulfill it."

Both parties were then informed of the realized state in that game. The seller then had to decide whether "to produce the good and deliver it to the buyer," and was always informed of the earnings that she and the buyer would make depending of her decision. The seller was also informed of the possibilities that the buyer would have depending on her decision. With that, subjects did not have to engage in the (very simple) calculations needed in order to make a well-informed decision in each state of the world, and also did not have to remember the buyer's possible ex post reactions to their decision.[39]

In all three treatments, sellers equally read:

"Do you want to produce the good and trade it with the original buyer?
If you choose yes, you earn < $20 - c(\theta)$ > points and the buyer earns 10 points."

Since payoffs of not doing so depended on the realized contingency, sellers read in the status quo and in all contingencies involving increases in costs of production (states 0, 1 and 3) that

"If you choose no, you earn 0 points and the buyer earns 0 points."

In contingencies involving an outside option (states 2 and 4), sellers read instead

"If you choose no, you sell the good to the third party and earn < $z(\theta)$ > points while the buyer earns 0 points."[40]

In treatment contract, sellers were always informed that, if they chose not to deliver the good to the original buyer, then

"in this case, the buyer can claim 10 points of compensation."

In all treatments with retaliation, sellers were informed that if they made that choice, then

"the buyer can spend 2 points to deduct 10 points from your earnings."

The experiment was implemented in the Experimental Laboratory for Sociology and Economics at Utrecht University between April and May 2013. A total number of 160 participants participated in the experiment. All interactions were anonymous and participants did not know the identity of other participants they interacted with. The computer program was written with the software z-Tree and the implementation of the experiment was fully computerized.[41] Average earnings per participant were of 12 Euros. There was no instance of any anomaly and subjects understood the game extremely well, as documented by several types of decision subjects consistently made, as described in the outset of the next section.

---

39 Moreover, subjects had in the written instructions a table with payoffs of both decisions in each state.

40 Note that the third party was not, in the experiment, a real person but a mere offer that the seller could accept instead of trading with the buyer. This was explicitly explained to participants: "this offer is hypothetical and not done by another participant in the lab" (as in the instruction in the experiment). Moreover, it is assumed that the third party, who "arrives later" than the first buyer, makes an offer corresponding to her whole valuation of the good, in the hope to induce the seller to sell for her instead of for the buyer that "arrived earlier".

41 *See* Urs Fischbacher, *z-Tree: Zurich Toolbox for Ready-made Economic Experiments*, 10 EXPERIMENTAL ECONOMICS 171 (2007).

III. RESULTS

A. Descriptive Results

Initial aggregate findings, presented in table 7 below, reveal that subjects understood the game well. Firstly, there were 994 observations of the decision to perform under the status quo in all different treatments, as reported in the first row. In 993 cases, sellers did choose to implement the exchange, to keep the promise, or to fulfill the contract under the status quo (state 0), as always expected.

Table 7. Initial aggregate findings

	Trade	Trade with ret.	Promise	Promise with ret.	Contract	Contract with ret.
Average performance in the status quo (θ_0)	1	1	0.99	1	1	1
N	104	208	96	194	112	230
Proportion of decisions to enter into agreement			0.985	0.98	0.98	0.99
N			400	800	464	928
Proportion of decisions to enforce the contract					0.99	0.98
N					136	220

Secondly, the number of successful agreements is extremely high in all treatments that included them (promises and contract), as reported in the second row. This was predicted since entering into the exchange of promises, enforceable or not, was always profitable even for an individual who planned to keep her word under all circumstances. There were 2553 observed instances of participants deciding to enter into the agreement out of 2592 observations.

Thirdly, there were 356 instances where buyers could claim damages, following breach by the seller, as seen in the last row. In 350 of them, disappointed buyers actively chose to do so, as predicted from theory, since this decision could deliver only benefits for the buyer.

In sum, whenever there were no reasons or arguments for deviations from strictly rational behavior, subjects behaved far and away as predicted in theory. In virtually all instances, (i) sellers chose to perform whenever doing so was individually profitable and circumstances did not change (in 99.89% of the cases), (ii) subjects opted to enter into the mutually profitable agreement (in 98.49% of the cases), and (iii) buyers decided to claim compensation for breach (98.31%).

Observed results concerning the behavior of sellers, or the decision to perform and to keep promises and contracts, in the absence of retaliation, are largely in accordance to the predictions delivered by strict rational choice. Table 8 below presents the frequency of breach in the three main treatments implemented without retaliation.

Table 8. Frequency of breach (treatments without retaliation)

	Characteristics of breach	Trade	Promises	Contract
State 0		0 % (N=104)	1 % (N=96)	0 % (N=112)
State 1	Socially inefficient, avoided inequality	96 % (N=26)	96 % (N=23)	28 % (N=29)
State 2	Socially inefficient, created inequality	88 % (N=26)	84 % (N=25)	28 % (N=29)
State 3	Socially efficient, avoided inequality	100 % (N=26)	100 % (N=25)	100 % (N=28)
State 4	Socially efficient, created inequality	92 % (N=26)	96 % (N=25)	92 % (N=26)
Total (*excluding status quo*)		94 % (N=104)	94 % (N=98)	61 % (N=112)

The only difference between treatments trade and promises is the presence of promissory commitment, and observed differences in rates of breach between those, in the same contingency, are attributed to the behavioral effect of promises, providing estimates for the test of hypothesis 1. Analogously, the only difference between treatments promises and contract is the availability of damages for breach, and observed differences in rates of breach between those treatments are attributed to the material incentives created by expectation damages, net of the effect of promise-keeping, hence providing estimates for the test of hypothesis 2.

Firstly, there is clearly no difference in observed rates of breach between treatments trade and promises. Descriptive results thus deliver *no evidence that promises induce performance* when performance becomes individually unprofitable, and when breach caused no more than loss of expectancy for the promisee. Promissory commitment alone, in states 1 to 4, did not lead sellers to perform, and rates of breach in those states in treatments trade and promises are virtually identical.

Secondly, there is a pronounced difference in observed rates of breach between treatments promises and contract, and *promisors performed, under expectation damages, significantly more often than in its absence*. In total, 94% of them breached in the absence of damages while only 61% did so in their presence. Moreover, *expectation damages induced performance only when it was socially efficient, and not otherwise*. These differences, observed only in states 1 and 2, are not only very pronounced but also statistically extremely significant in both cases (p=0.000, Fisher's exact test). Still, expectation damages did not induce all sellers to perform whenever performance

was socially efficient. In fact, 28% of them (8 out of 29) behaved irrationally and breached when breach was individually and socially unprofitable.

Table 9 below presents frequencies of breach in the main treatments implemented with the possibility of retaliation by the buyer. As the comparison of average rates of breach without retaliation, in table 8, with average rates of breach with retaliation reveal, the *possibility* of irrational retaliation can be a powerful force driving the decision of the seller, inducing sellers to perform, to avoid retaliation, in situations where sellers would not be willing to trade.

Table 9. Frequency of Breach (treatments with retaliation) [42]

	Characteristics of breach	Trade	Promises	Contract
State 0		0 % (N=208)	0 % (N=194)	0 % (N=230)
State 1	Socially inefficient, avoided inequality	64 % (N=52)	50 % (N=48)	5 % (N=56)
State 2	Socially inefficient, created inequality	50 % (N=52)	49 % (N=49)	14 % (N=57)
State 3	Socially efficient, avoided inequality	100 % (N=52)	96 % (N=48)	89 % (N=57)
State 4	Socially efficient, created inequality	90 % (N=52)	91 % (N=45)	84 % (N=57)
Total (*excluding status quo*)		76 % (N=208)	71 % (N=190)	48 % (N=227)

[42] Considering only data collected in parts 1 and 2 of the experiment, and excluding all data collected in part 3, rates of breach were quite consistent, as table IV.9.b reveals.

Table 9.b. Frequency of Breach (excluding data from part 3)

	Characteristics of breach	Trade	Promises	Contract
State 0		0 % (N=104)	0 % (N=95)	0 % (N=115)
State 1	Socially inefficient, avoided inequality	54 % (N=26)	57 % (N=23)	7 % (N=28)
State 2	Socially inefficient, created inequality	38 % (N=26)	46 % (N=24)	14 % (N=28)
State 3	Socially efficient, avoided inequality	100 % (N=26)	96 % (N=24)	86 % (N=29)
State 4	Socially efficient, created inequality	92 % (N=26)	86 % (N=21)	82 % (N=28)
Total (without state 0)		71 % (N=104)	71 % (N=92)	48 % (N=113)

Observed frequencies of breach reveal that retaliation induced a change in behavior of the seller in all treatments in states 1 and 2. The possibility of being punished reduced the frequency of breach by approximately half, providing evidence for hypothesis 5 and for the fact that sellers did change their behavior when faced with the possibility of irrational punishment by the buyer.

In treatment trade, sellers that would virtually never trade in the absence of retaliation (and that rather breached in around 90% of cases) changed their behavior in order to avoid punishment by the buyer, and refused to perform much less often when they could suffer losses from retaliation. Similarly, in treatment promises, promisors kept their promises around half of the times when the promisee could retaliate to breach, remarkably more than in the absence of retaliation, when sellers almost never kept their promises. Retaliation induced a change in behavior of sellers, but the effect of retaliation to breach of promise is not different from the effect of retaliation against a mere refusal to trade, as sellers did not anticipate higher retaliation to breach of promise.

In treatment contract, retaliation further contributed to induce performance by the seller beyond expectation damages. In states 1 and 2, expectation damages already provided incentives for sellers to perform, and should have lead to full compliance. They however fell short in achieving that goal by 28%, corresponding to observed rates of breach in those states under expectation damages. Retaliation contributed to induce efficient performance beyond that level, decreasing rates of breach from 28% to 5% (state 1) or 14% (state 2), thereby contributing to social welfare in inducing efficient behavior, and in deterring inefficient breaches.

Results concerning the buyer's decision to retaliate were, in contrast to those concerning the behavior of the seller (at least in the absence of the possibility of retaliation) much less in line with the strict rational choice hypotheses. They reveal that costly punishment, even if irrational and predicted not to emerge, is a strong behavioral tendency. Table 10 below presents the frequency of retaliation by disappointed buyers in each treatment and contingency.

Average rates of retaliation should, under strict rational choice hypotheses, be equal to zero in all treatments, and in all possible contingencies. Observed rates of retaliation reveal, however, that retaliation is present in all treatments. They are, moreover, much higher when the seller had promised to perform, in some specific contingencies (states 2 and 4). Lastly, they are substantially lower when the buyer is entitled to claim damages for breach.

In effect, individuals engage in costly punishment of sellers that did not implement the exchange in all contingencies already in treatment trade, differently from what strict rational choice models predict. There are different possible reasons for this behavior even in single interactions, and in the absence of any form of agreement and commitment. Inequality-averse individuals have a reason to retaliate in states 2 and 4, since selling the good to the third party in those states realizes an unequal outcome where the seller earns something while the buyer does not. Individuals that care for social welfare have a reason to retaliate in states 1 and 2, since not implementing the exchange in those states did not maximize welfare. The implemented control group captures those effects, as this was its function by design.

Table 10. Frequency of retaliation to breach [43]

	Characteristics of breach	Trade	Promises	Contract
State 0		. (N=0)	. (N=0)	. (N=0)
State 1	Socially inefficient, avoided inequality	36 % (N=33)	50 % (N=24)	0 % (N=3)
State 2	Socially inefficient, created inequality	31 % (N=26)	58 % (N=24)	12 % (N=8)
State 3	Socially efficient, avoided inequality	21 % (N=52)	17 % (N=46)	20 % (N=51)
State 4	Socially efficient, created inequality	29 % (N=47)	49 % (N=41)	23% (N=48)
Total (*excluding status quo*)		28 % (N=158)	40 % (N=135)	20 % (N=110)

Firstly, individuals retaliate more often to breach of promise than to the seller's mere refusal to trade. The total rate of retaliation, reported in the last row of table IV.10, is substantially higher in treatment promises (40%) than in treatment trade (28%). Statistical tests for the equality of proportions, reported in table 11 below, reveal that this difference is significant (p=0.047). It delivers *supporting evidence for hypothesis 2 that breach of promise induced retaliation by victims.*

[43] The reader concerned that ordering effects may have driven these results can consult the following table, which excludes all data obtained in part 3 of the experiment (as described in table 6 *supra*):

Table 10.b. Frequency of Retaliation to Breach (excluding data from part 3)

	Characteristics of breach	Trade	Promises	Contract
State 0		. (N=0)	. (N=0)	. (N=0)
State 1	Socially inefficient, avoided inequality	43 % (N=14)	38 % (N=13)	0 % (N=2)
State 2	Socially inefficient, created inequality	20 % (N=10)	73 % (N=11)	0 % (N=4)
State 3	Socially efficient, avoided inequality	15 % (N=26)	13 % (N=23)	24 % (N=25)
State 4	Socially efficient, created inequality	25 % (N=24)	56 % (N=18)	39 % (N=23)
Total (without state 0)		24 % (N=74)	40 % (N=65)	28 % (N=54)

However, higher rates of retaliation to breach of promise are restricted to some specific contingencies, revealing that promisees did not punish all breaches equally, and did not punish all violations of the moral norm of keeping promises indiscriminately. Punishment of breach of promise depended on the consequences of breach, and the planned design of the experiment allows for that identification. In fact, in state 3, where breach of promise was efficient and fair, buyers did not retaliate more often than in treatment trade. When breach joined both desirable consequences of avoiding inefficiency and avoiding inequality, breach of promise was not punished.

In contrast, in states 2 and 4, where breach of promise created an unfair result, breach of promise was punished significantly more (the results of the statistical tests are presented in table 11 below). These breaches conjoin two different negative elements, the wrong in breaking promises and the unfairness of the resulting outcome. Results reveal that retaliation to breach of promise committed to achieve a higher profit from an outside transaction is more pervasive than sole retaliation to breach or sole retaliation to unfairness.

Secondly, in the presence of damages for breach, observed rates of retaliation to breach of promise were substantially lower (in total, $p=0.001$). This is observed in all states except for state 3, where retaliation to breach of promise in fact did not emerge, since breach there was not unfair, and did not create any inequality in the final payoffs. These differences rely on fewer observations but are still highly significant ($p=0.041$ in state 2, and $p=0.014$ in state 4). There is hence *supporting evidence for hypothesis 4 that compensatory remedies in effect crowd out retaliation to breach.*

Table 11 presents results of Fisher's exact test between observed rates of retaliation. The first test, reported in the fifth column, concerns the hypothesis that differences in rates of retaliation between treatments trade and promises is due to the presence of promissory commitment. It provides evidence that individuals retaliate to breach of promise more often than they retaliate to a mere refusal to trade, and that this effect occurs only when there was breach of promise *and* the creation of inequality (in the overbidder paradigm).

The second test, presented in the last column, delivers results of that same test of hypothesis between treatments promises and contract. It reveals that observed differences are statistically significant, and hence due to the effect of expectation damages, thereby providing evidence that damages in fact crowd out retaliation to breach. This effect is, as expected, observed only in those contingencies where breach of promise induced retaliation, revealing how compensation crowds it out in those states.

Compensation seems to be necessary when the promisor breaches in order to profit from an outside transaction, but not when the promisor breaches in order to avoid incurring losses because of a spike in the costs of production. In this case, breach avoided the creation of inequality, and can arguably be perceived as the fair thing to do. It did not induce retaliation, and promisees possibly did not perceive it as a wrong in need of redress. It is in the case of outside offers, when the promisor profits from her own wrong, that compensation is necessary to crowd out the observed tendency of disappointed promisees to reciprocate to perceived wrongful behavior.

Table 11. Frequency of retaliation to breach (two-sided tests)[44]

	Characteristics of breach	Trade	Promises	Fisher's exact	Promises	Contract	Fisher's exact
State 0		. (N=0)	. (N=0)		. (N=0)	. (N=0)	
State 1	socially inefficient, avoided inequality	36 % (N=33)	50 % (N=24)	p=0.416	50 % (N=24)	0 % (N=3)	p=0.230
State 2	socially inefficient, created inequality	31 % (N=26)	58 % (N=24)	p=0.086	58 % (N=24)	12 % (N=8)	p=0.041
State 3	socially efficient, avoided inequality	21 % (N=52)	17 % (N=46)	p=0.800	17 % (N=46)	20 % (N=51)	p=0.800
State 4	socially efficient, created inequality	29 % (N=47)	49 % (N=41)	p=0.082	49 % (N=41)	23 % (N=48)	p=0.014
Total		28 % (N=158)	40 % (N=135)	p=0.047	40 % (N=135)	20 % (N=110)	p=0.001

[44] Excluding data from part 3 of the experiment, results are consistent with the exception of the crowding out effect in state 4, which is not significant when excluding half of the observations. The crowding out effect of retaliation in state 2 is still significant, with p=0.0260.

Table IV.11.b. Frequency of Retaliation to Breach (excluding data from part 3)

	Characteristics of breach	Trade	Promises	Fisher's exact	Promises	Contract	Fisher's exact
State 0		. (N=0)	. (N=0)		. (N=0)	. (N=0)	
State 1	Socially inefficient, avoided inequality	43 % (N=14)	38 % (N=13)	p=1.000	38 % (N=13)	0 % (N=2)	p=0.520
State 2	Socially inefficient, created inequality	20 % (N=10)	73 % (N=11)	p=0.030	73 % (N=11)	0 % (N=4)	p=0.026
State 3	Socially efficient, avoided inequality	15 % (N=26)	13 % (N=23)	p=1.000	13 % (N=23)	24 % (N=25)	p=0.470
State 4	Socially efficient, created inequality	25 % (N=24)	56 % (N=18)	p=0.059	56 % (N=18)	39 % (N=23)	p=0.360
Total		24 % (N=74)	40 % (N=65)	p=0.067	40 % (N=65)	27 % (N=54)	p=0.180

B. Regression Results

The first regression analyzes the effect of promissory commitment, expectation damages, and the possibility of retaliation on the behavior of sellers. That is, interest lies in the determinants of performance, and on whether promises, damages, and retaliation can induce sellers to perform. The analysis distinguishes these effects according to the consequences of the seller's decision in terms of inefficiency and inequality.

Since subjects made the same decision multiple times in each session, standard errors are clustered by subject. Moreover, different controls for possible ordering effects due to the treatment having been implemented in later parts of each session are also included in the model. The seller's decision was, in the game, dichotomous: if the seller decided to "produce the good and deliver it to the buyer," then $perform = 1$, otherwise $perform = 0$. Accordingly, the parametric model has the following form:

$$Prob(perform|promises, contract, retaliation, inef, ineq) = \beta_0 + \beta_1 inef + \beta_2 ineq + \beta_3 promises + \beta_4 promises \cdot inef + \beta_5 promises \cdot ineq + \beta_6 contract + \beta_7 contract \cdot inef + \beta_8 contract \cdot ineq + \beta_9 retaliation + \beta_{10} retaliation \cdot inef + \beta_{11} retaliation \cdot ineq + u$$

where

$inef$	is an indicator for the creation of losses for both parties (inefficiency) as a consequence of the seller's decision, equal to unity only in states 1 and 2;
$ineq$	is an indicator for the creation of inequality in the parties' final earnings as a consequence of the seller's decision, equal to unity only in states 2 and 4;
$promises$	is an indicator for promissory commitment, equal to unity in case both parties agreed to the exchange of promises (present in treatments promises and contract);
$contract$	is an indicator for the possibility of buyers to claim expectation damages to breach promise, equal to unity whenever parties agreed to the enforceable exchange of promises (present only in treatment contract).
$retaliation$	is an indicator for the buyer's possibility to retaliate, equal to unity if the observation was collected in a treatment with retaliation.

Regressors for the different treatments are resumed in table IV.12:

Table 12. Regressors

Main Treatment	Without retaliation	With retaliation
Trade	Promises = 0 Contract = 0 Retaliation = 0	Promises = 0 Contract = 0 Retaliation = 1
Promises	Promises = 1 Contract = 0 Retaliation = 0	Promises = 1 Contract = 0 Retaliation = 1
Contract	Promises = 1 Contract = 1 Retaliation = 0	Promises = 1 Contract = 1 Retaliation = 1

Different regressions with observations collected only in part 1, in part 1 and part 2, and in all three parts of each session are all reported in the three columns of table 13 below.

Consider first the isolated impact of inefficiency and inequality. The fact that the seller's decision would lead to an overall loss of resources (being thus inefficient) or to an unequal distribution (being thus "unfair") induced some sellers to trade even in the absence of any form of commitment or enforcement. This reflects the same findings reported in tables 8 and 9 above, where in state 3 (the contingency in which deciding *not* to perform created neither inefficiency nor inequality), sellers in fact never decided to perform (frequency of breach was 100% in the absence of retaliation, and 96% in its presence).

In all other states, sellers at times decided to trade even when doing so was individually unprofitable, thereby avoiding the realization of those negative consequences. The magnitude of the estimated impact of inefficiency and inequality on that decision is quite comparable, and the probability of trade was, on average, around 20 percentage points higher when trade was socially efficient or fair.

Promises, on the other hand, never had any effect upon the seller's decision, as reported in tables 8 and 9. Rates of trade in the absence and presence of promises were virtually identical, and regression results confirm that there is *no evidence supporting hypothesis 1* (promises induce trade in the absence of any type of enforcement mechanism). In comparison to other empirical studies that attest the existence of that effect when breach of promise caused a material loss of resources for the promisee, the experiment provides no evidence for the existence of that same effect when breach of promise caused only loss of expectancy.

Retaliation further induced performance by the promisor, although to a lower extent than damages. There is a general effect of retaliation that does not depend on efficiency or equality, and a specific one, through which retaliation induced higher rates of performance when performance was socially efficient (except when considering only data from part 1). There is hence *support for hypothesis 5* (fear of irrational retaliation induces a change in behavior of sellers).

Table 13. Regression results on seller's decision to perform

DEP. VARIABLE:	SPECIFICATION (average marginal effects reported)		
Seller's decision to perform	Probit Part 1	Probit Parts 1,2	Probit Parts 1,2,3
Inefficiency	0.25*	0.21**	0.18*
	(0.14)	(0.09)	(0.10)
Inequality	0.17**	0.15**	0.18***
	(0.07)	(0.06)	(0.06)
Promises	0.06	0.04	0.05
	(0.09)	(0.07)	(0.06)
Promises · inefficiency	-0.11	-0.04	0.02
	(0.08)	(0.07)	(0.07)
Promises · inequality	0.02	-0.01	-0.05
	(0.09)	(0.07)	(0.06)
Contract	0.11	0.14**	0.12*
	(0.08)	(0.07)	(0.06)
Contract · inefficiency	0.32**	0.39***	0.34***
	(0.15)	(0.12)	(0.10)
Contract · inequality	-0.08	-0.09*	-0.05
	(0.08)	(0.05)	(0.05)
Retaliation	0.15**	0.15***	0.14***
	(0.07)	(0.05)	(0.05)
Retaliation · inefficiency	0.16	0.15**	0.16**
	(0.12)	(0.07)	(0.08)
Retaliation · inequality	-0.10	-0.05	-0.08
	(0.06)	(0.04)	(0.05)
Controls	none	for part 2	for part 2 and part 3
Observations	306	623	939
Pseudo R^2	0.382	0.433	0.425
χ^2 (degrees of freedom)	120.3 (11)	264.2 (23)	773.7 (41)
p	0	0	0
Number of clusters	80	80	80

NOTES: *Controls for part 2* include an indicator variable for that part and its interaction with all regressors, controlling for possible general ordering effects as well as for specific ordering effects. *Controls for part 3* are equally defined and further include an indicator variable for part 3 implemented immediately after another part with retaliation, as well as its interaction with all regressors, controlling for possible effects of implementing retaliation immediately after retaliation (vis-à-vis implemented after no retaliation). Stars notation: *** p<0.01, ** p<0.05, * p<0.1 (two-sided tests).

All standard errors corrected for clustering by Subject.

The second set of regressions analyzes the determinants of retaliatory behavior in contractual relationships. The effect of breach of promise and of expectation damages is again distinguished according to the motives and consequences of breach. The buyer's decision was also dichotomous, and the model used to estimate those effects, separately with data from part 1, parts 1 and 2, and all three parts is the following:

$$Prob(retaliate|promises, contract, inef, ineq) = \beta_0 + \beta_1 inef + \beta_2 ineq + \beta_3 promises + \beta_4 promises \cdot inef + \beta_5 promises \cdot ineq + \beta_6 contract + \beta_7 contract \cdot inef + \beta_8 contract \cdot ineq + u$$

where regressors are defined as in table 12 above.

Table 14 below reports the results. Results are consistent and do not depend, for the hypotheses under consideration, on whether only data obtained in part 1, parts 1 and 2, or all parts of each session is considered. There is evidence that buyers retaliate more often against sellers that did not take the socially optimal course of conduct (the coefficient of inefficiency, considering only data from parts 1 and 2, has a p-value of 0.1). There is no evidence that inequality-aversion, by itself, leads to retaliation in the designed trade game.

The main factor driving retaliatory behavior is unfairness in breach of promise, or breach that created an unequal distribution of gains between the parties. Promisees retaliated substantially more against sellers that profited from their own wrong. There is therefore *supporting evidence for hypothesis 2* (breach of promise induce retaliation by the promisee), restricted however to breaches responsible for an unequal distribution.

Lastly, there is considerable evidence for the effect of compensatory damages to crowd out retaliation to breach by promisees. This effect is found exactly under the same circumstances where promises induced retaliation, namely when breach created inequality. This presents *supporting evidence for hypothesis 4* (damages for breach crowd out retaliation by promisees), although the scope of this effect seems to be restricted to breaches to profit from a higher outside option.

Regression results are hence all consistent with results of the different test of hypotheses implemented in the previous sections. Promises did not induce a change in behavior of sellers, but both expectation damages and the possibility of retaliation by the buyer did so. While the latter lead to rates of breach that were roughly half of those observed in the absence of any enforcement mechanism, expectation damages proved to be much more effective in deterring breaches, and deterred not half of all inefficient breaches, but around three quarters of them.

Promises, on the other hand, induced a substantial change in the behavior of buyers, and breach of promise lead to rates of retaliation that were around twice as high as the ones observed in the absence of commitment. Additionally, this effect was restricted to those contingencies where the seller breaches to realize higher profits from an outside transaction, profiting from the act of breach. Compensatory remedies such as expectation damages were then capable of substantially reducing those rates, as both tests of hypotheses and regression analyses confirm.

Table 14. Regression results on buyer's decision to retaliate

DEP. VARIABLE:	SPECIFICATION (average marginal effects reported)			
Buyer's decision to retaliate	Probit Part 1	Probit Parts 1,2	Probit [45] Parts 1,2,3	Probit Parts 1,2,3
Inefficiency	0.28**	0.15	0.10**	0.10**
	(0.14)	(0.09)	(0.05)	(0.05)
Inequality	0.14	0.02	0.04	0.04
	(0.11)	(0.09)	(0.06)	(0.06)
Promises	0.02	-0.07	-0.03	-0.03
	(0.13)	(0.10)	(0.07)	(0.07)
Promises · inefficiency	-0.16	0.13	0.13**	0.14**
	(0.11)	(0.09)	(0.06)	(0.06)
Promises · inequality	0.35**	0.40***	0.20**	0.20**
	(0.17)	(0.12)	(0.10)	(0.09)
Contract (promises · damages)	0.07	0.13	0.00	-0.00
	(0.17)	(0.13)	(0.09)	(0.09)
Contract · inefficiency	Not estimable because of lack of variance [46]			-0.24***
				(0.08)
Contract · inequality	-0.19**	-0.20**	-0.20***	-0.18**
	(0.09)	(0.09)	(0.07)	(0.07)
Controls		for part 2	for part 2 and part 3	for part 2 and part 3
Observations	105	187	392	403
Pseudo R^2	0.159	0.147	0.171	0.178
χ^2 (degrees of freedom)	29.5 (7)	81.2 (15)	3175 (31)	3311 (32)
p	0.0001	0	0	0
Number of clusters	43	80	80	80

NOTES: *Controls for part 2* include an indicator variable for that part and its interaction with all regressors, controlling for possible general ordering effects as well as for specific ordering effects. *Controls for part 3* are equally defined and further include an indicator variable for part 3 implemented immediately after another part with retaliation, as well as its interaction with all regressors, controlling for possible effects of implementing retaliation immediately after retaliation (vis-à-vis implemented after no retaliation). Stars notation: *** p<0.01, ** p<0.05, * p<0.1 (two-sided tests).

All standard errors corrected for clustering by Subject.

[45] As seen in the descriptive results, there was one instance where the buyer decided to retaliate to breach in state 2 with data from all three parts. Therefore, it is possible to estimate the effect of damages to crowd out retaliation with all data, but given the few observations for that, the table above distinguishes both types of regressions.

[46] As reported in the descriptive results, there was no observation of retaliation to breach, under expectation damages, in states 1 and 2 when not considering data from part 3.

C. Analysis of Social Welfare

Once seen that damages for breach not only induce performance by the promisor, but also reduce retaliation by victims of breach – at least when breach created inequality inside the parties' promissory relationship –, the question that arises concerns the gains of social welfare provided by each of these functions of legal relief. Overall social welfare was given by

$$SW(q_s, r_b, \theta) = q_s\big(V - c(\theta)\big) + (1 - q_s) \cdot (z(\theta) - r_b R - r_b K)$$

In case of trade, realized social welfare was equal to the value of performance to the buyer minus the costs of production of the seller. In case of breach, it was given by the net gains of trade with the outside party, minus the loss that seller and buyer incur in case of retaliation to breach. For the chosen parameters, it corresponds to

$$SW(q_s, r_b, \theta) = q_s\big(30 - c(\theta)\big) + (1 - q_s) \cdot (z(\theta) - 12r_b)$$

Trade was always expected to take place in state 0, independent of promissory commitment or of any type of enforcement mechanism, and is therefore excluded from the analysis. In all other states, sellers were, in the absence of damages, always predicted not to trade with the buyer and to break their promises. This, however, does not mean that the predicted gains from trade were zero, since in states 2 and 4 the seller earns all the gains from trade by selling the good to the outside party, and breach to profit from an outside option does contribute to social welfare.

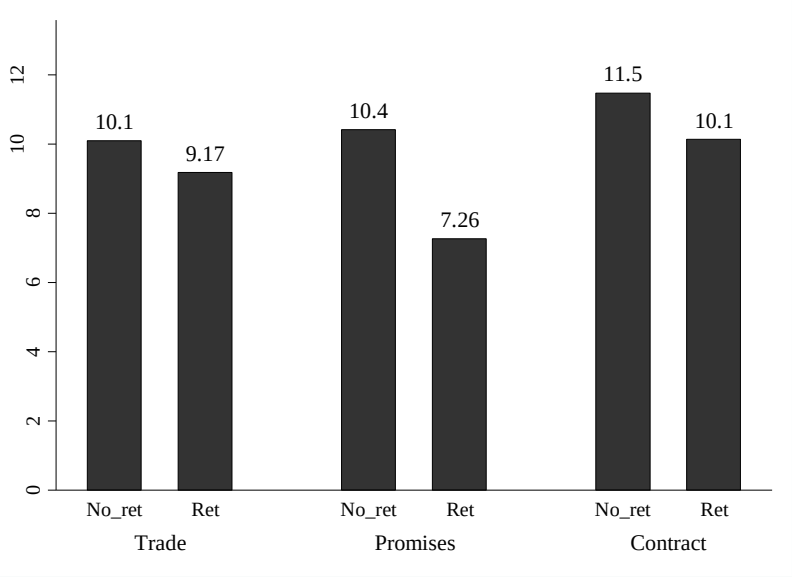
Expectation damages are predicted to create gains of welfare by inducing performance in states 1 and 2. In those, gains from trade are expected to be foregone in the absence of damages (in treatments trade and promises), and under expectation damages are predicted to be exhausted. This is the predicted contribution of damages to welfare in inducing performance of contracts.

Table 15. Expected social welfare in each treatment

Contingency \ Treatment	Trade and promises		Contract	
	Predicted choice	Resulting outcome	Predicted choice	Resulting outcome
State 1 (efficient trade)	Breach	(0,0)	Performance	$(-5,10)$
State 2 (efficient trade)	Breach	(15,0)	Performance	(10,10)
State 3 (inefficient trade)	Breach	(0,0)	Breach	$(-10,10)$
State 4 (inefficient trade)	Breach	(25,0)	Breach	(15,10)
Total expected SW	10		12.5	

Graph 1 presents the realized social welfare depending on the type of commitment, and on damages and retaliation.

Graph 1. Social welfare by treatment



The graph reveals that retaliation caused heavy losses of welfare. In the absence of promises, it created a loss of almost 0.9 in welfare. In the presence of promises, that loss was more than three times higher, of more than 3. In the presence of damages, it was minimized, and losses from retaliation to *breach* were almost all eliminated, being brought back to levels similar to those observed in treatment trade. For the investigation of the social welfare function of legal redress, there is the need to consider the possible gains created by retaliation in inducing performance and not only its social costs. Table 16 below presents those in detail, and separately.

Retaliation did deliver some gains of social welfare by inducing performance by the seller, and thus often leading to trade by itself, without the need of legal enforcement. Sellers traded more often with retaliation than without retaliation, as reported in the descriptive and regression results above. In treatment trade, accordingly, retaliation generated gains from trade of 1.15, but at a social cost of 2.08, thus delivering an overall loss of social welfare of 0.83.

In treatment promises, retaliation was, as described before, more pervasive but did not induce higher rates of performance by the seller than it did in the absence of promises. Retaliation to breach of promise delivered therefore minor gains of only 0.25 in inducing performance by the seller, and only at a very high social cost of 3.39. It delivered a major net loss of welfare of 3.14.

In treatment contract, retaliation did induce higher rates of performance of contracts than in its absence (and sole presence of damages), but provided by itself a minor gain of social welfare of 0.23. Since the buyer was entitled to compensatory damages, retaliation was lower in treatment contract than in treatment promises, and was responsible for social losses of 1.59, considerably lower than the 3.39 observed in treatment promises. Retaliation still caused a loss of welfare, in the presence of compensation, of 1.36, which is, because of the crowding out function of legal

redress, much lower than the observed 3.14 in treatment promises.

Table 16. Social Welfare generated by promises, damages and retaliation

Treatment	Without Retaliation	With Retaliation
Trade	SW = 10.1	SW = 9.17 Total gains from trade induced by ret. = 11.25 (net 1.15) Losses from ret. = 2.08 Net welfare effect of retaliation = - 0.83
Promises	SW = 10.4	SW = 7.26 Gains from trade induced by ret. = 10.65 (net 0.25) Losses from ret. = 3.39 Net welfare effect of retaliation = - 3.14
Contract	SW = 11.5	SW = 11.5 Gains from trade induced by ret. = 11.73 (net 0.23) Losses from ret. = 1.59 Net welfare effect of retaliation = - 1.36
Welfare created by damages in	inducing performance = 1.1	Crowding out retaliation to breach of promise by disappointed promisees = 1.78

In conclusion, promises alone provided for almost no gain of welfare. Expectation damages, in inducing socially efficient performance, delivered a net gain of 1.1. In crowding out retaliation, expectation damages avoided a net loss of 1.78. At a minimum, both functions were, in the parameters of the implemented trade game, equally important for the maximization of social welfare. There is evidence that the function of damages to substitute for private redress can be, at times, as under the parameters of the game, even more important than the function of damages to induce socially efficient performance.

IV. Conclusion

The experimental study provided evidence that contractual parties do not behave exclusively based on their own material self-interest, and that deviating behavior is attributable to the behavior of promisees in the presence of the possibility to retaliate. Promisors in fact behaved, in the absence of retaliation, much in line with predictions from rational choice, and did not keep their promises whenever doing so was individually unprofitable. They however clearly anticipated the payment of damages for breach when deciding to perform or breach.

Accordingly, obtained results do not provide support for hypothesis 1 (promises induce performance), and there is no evidence for the acceptance of that hypothesis when breach caused

no more than loss of expectancy. This result differs from those obtained in previous studies on promise-keeping.[47] Although it is difficult to pinpoint the reason for this difference, previous experiments investigated whether individuals keep promises when the promisee suffered a real monetary loss from breach, and not only loss of expectancy. Additionally, they most often did not reward subjects for each decision, but rather randomly selected only one decision for that. Lastly, the incomplete contract and the different types of contingencies rendered breach unprofitable for the promisor only in some interactions, and, because of that, the experiment differs substantially from the previous ones where promisors faced a situation where breach was unprofitable in every round. Still, in the present experiment, where only the expectation interest is at stake, the moral force of keeping promises was not enough to induce promisors to keep their deals.

In contrast, results deliver strong evidence for hypothesis 3 (expectation damages induce performance if and only if performance is socially efficient). The implemented tests of hypotheses and the obtained regression results favor its acceptance. It is worth noting that expectation damages did not induce all sellers to perform whenever socially efficient, and an estimated 28% behaved irrationally and still breached the contract in those circumstances. The reason underlying that decision lies possibly in their belief that buyers would not claim damages, somehow condoning breach, perhaps because sellers knew buyers did not really lost their own money because of breach.

Buyers, in turn, very often behaved irrationally with respect to the predictions derived from strict rational-choice theory, and did not abstain from costly punishment even in one-shot, single interactions with anonymous counter-parties. Obtained results show that retaliation, in treatment trade, and thus in the absence of promissory commitment and damages for breach, was present when the seller's decision was inefficient.

Sellers anticipated irrational retaliation by the buyer (yet again, differently from the strict rational choice prediction) and rates of trade were substantially higher in the presence of the possibility of ex post retaliation. Sellers anticipated that reaction mostly in contingencies where trade was socially efficient. There is some support for hypothesis 5 (the possibility of irrational retaliation induces a change in behavior of sellers, and leads sellers to trade even when doing so is individually unprofitable).

Results present supporting evidence for hypothesis 2 (breach of promises induces higher rates of retaliation), even when controlling for the unfairness and inefficiency of the result, present in treatment trade. Higher rates of retaliation in treatment promises were not driven by inequality-aversion or by a desire to punish behavior that did not maximize social welfare, since these were captured in the control group, and do not explain differences between treatments.

Retaliation to breach of promise was concentrated in situations in which breach created inequality in the final distribution, and in which the promisor profited from it. There is no evidence that promisees retaliate against any type of breach of promise, as average retaliation to breach of promise when breach was socially efficient and avoided the creation of inequality (in state 3) was

---

47 *See supra* note 22.

virtually equal to average rates of retaliation in the absence of promise. On the other hand, whenever the promise-breaker profited from her wrong, rates of retaliation were twice as high as in the absence of promissory commitment.

Lastly, the experiment delivers evidence that supports hypothesis 4 (damages for breach crowd out retaliation), and damages for breach do fulfill the function to substitute for private redress and to crowd out retaliation by victims of breach. This effect in concentrated in those situations where breach in fact induced retaliation, in states 2 and 4. As the estimated impact upon social welfare reveals, this function is, at times, neither secondary nor marginal, for its contribution to overall social welfare was more important, in the experiment, than the well-known contribution of damages to induce efficient performance.

Of course, these experimental results do not allow for a generalized conclusion that one function is, in reality, more important than the other. They do allow, however, for the claim that the function of legal relief to substitute for private redress has its own contribution to social welfare, and is therefore justified even under strict economic terms. Compensation for victims of breach shall be taken seriously into consideration among the functions of remedies for breach, and, even if only for reasons of social welfare, is in fact fundamental in the law of remedies for breach.