

Schmidtchen, Dieter

**Working Paper**

## To Help or not to Help: The Samaritan's Dilemma Revisited

CSLE Discussion Paper, No. 99-09

**Provided in Cooperation with:**

Saarland University, CSLE - Center for the Study of Law and Economics

*Suggested Citation:* Schmidtchen, Dieter (1999) : To Help or not to Help: The Samaritan's Dilemma Revisited, CSLE Discussion Paper, No. 99-09, Universität des Saarlandes, Center for the Study of Law and Economics (CSLE), Saarbrücken

This Version is available at:

<https://hdl.handle.net/10419/23060>

**Standard-Nutzungsbedingungen:**

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

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

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

**Terms of use:**

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

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

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

# To Help or not to Help: The Samaritan's Dilemma Revisited

by Dieter Schmidtchen<sup>\*</sup>

CSLE Discussion Paper 9909

June 1999

## Abstract:

Helping somebody may undermine his incentives to work. What Buchanan identified more than 25 years ago as the Samaritan's dilemma is basically a time-inconsistency problem. The paper discusses possible solutions of the dilemma such as punishment within an iterated game, reshaping the game in the direction of a dynamic one-shot game and the delegation of the power of decision to an agent. The paper shows that only the latter option works.

## ***I. Introduction***

The Samaritan's Dilemma is one of Buchanan's articles written in – what he once called – his pessimistic working period. The essay characterized as an "essay in prescriptive diagnosis" (Buchanan 1977: 169) starts with the observation that modern man has become incapable of making the choices that are required to prevent his exploitation by predators of his own species, whether the predation be conscious or unconscious (see Buchanan 1977: 173). Modern man in twentieth-century Western Society has "gone soft": "His income-wealth position, along with his preference ordering, allows him to secure options that were previously unavailable. What we may call 'strategic courage' may be a markedly inferior economic good, and what we may call 'pragmatic compassion' may be markedly superior." (Buchanan 1977: 173.)

---

<sup>\*</sup> Center for the Study of Law and Economics, Universität des Saarlandes, Bldg. 31, POBox 151 150, D-66041 Saarbrücken, Tel. ++681/302-2132, Fax ++681/302-3591, email csle@rz.uni-sb.de  
I would like to thank Chr. Bier, R. Kirstein and A. Neunzig for helpful comments.

What is the prescription, the direction of reform and improvements? Buchanan's prescription consists of two parts: The first part is "an explicit recognition of the dilemma by those who are caught up in it" (Buchanan 1977: 173). The second part comes in, once this sort of recognition is passed: "(t)he players involved must, individually and collectively, act *strategically* instead of pragmatically" (Buchanan 1977: 173). What that means, is explicitly stated in the paper as "the choice of utility-maximizing rules for personal behavior as opposed to the retention of single period or single-situation choice options. Having once adopted a rule, the Samaritan *should not* be responsive to the particulars of situations that might arise. He should not act pragmatically and on a case-by-case basis" (Buchanan 1977: 177). Both, the choice of a rule and strictly following such a rule requires what Buchanan called "strategic courage". As with strategic moves in general (see Dixit/Nalebuff: 121), strategic courage purposefully limits the freedom of action, thereby altering the beliefs and actions of others in a direction favorable to the decisionmaker.

The final version of Buchanan's paper was prepared for presentation at a conference on Altruism and Economic Theory held in the early seventies. However, it seems to me that the prescriptive diagnosis presented in the paper is as topical today as it was at that time. What Buchanan identified as the Samaritan's dilemma reflects characteristics of the modern Welfare state. The number of those living on transfers has risen dramatically. And there is a widespread belief that the benefits in the modern welfare state are the very reason for undermining the incentives to work (see for a rigorous recent treatment of this subject Lindbeck, Nyberg, Weibull 1999). Thus, we have reason to reflect on what Buchanan wrote many years ago. However, there is another reason for doing so. The parts of the article quoted here extensively make clear that "The Samaritan's Dilemma" can be considered as being an early contribution to the literature on the time-inconsistency problem. Decisionmakers, recognizing that they will be tempted not to do what they announced to do must think of institutional designs making their announcements credible. Otherwise they will get caught in a trap by their own rationality. As Buchanan puts it: "The term dilemma seems appropriate because the problem may not be one that reflects irrational behavior on any of the standard interpretations." (Buchanan 1977: 173.)

The remainder of the paper is organized as follows: Section II presents Buchanan's model. It also deals with the question, in what sense the outcome of the interaction between the Samaritan and those in need of help can be considered being a dilemma. Section III discusses

possible solutions of the dilemma such as punishment within an iterated game, reshaping the game in order to make it a dynamic one-shot game and the delegation of the power of decision to an agent. Section IV concludes the paper.

## ***II. The Samaritan's Dilemma***

Buchanan models the Samaritan's dilemma as a two-by-two matrix game (see Buchanan 1977: 170; see fig. 1).

	B <sub>1</sub>	B <sub>2</sub>
A <sub>1</sub>	I 2,2	II 1,1
A <sub>2</sub>	III 4,3	IV 3,4

Fig. 1: The Active Samaritan's Dilemma

We have two players A and B: A is the potential Samaritan, B its opponent.

Player A has got two strategies:

A<sub>1</sub> =: do not help

A<sub>2</sub> =: help which could mean pay 30\$ to B as a transfer (gift).

Player B has also got two courses of action:

B<sub>1</sub> =: work

B<sub>2</sub> =: do not work.

The pay offs are ordinal utility indicators, with A's pay off ranked first. We assume the matrix as well as rationality of the players being common knowledge. The game has one Nash-equilibrium in pure strategies, namely strategy combination (A<sub>2</sub>, B<sub>2</sub>). A<sub>2</sub> is the dominant

strategy of the Samaritan. Knowing this  $B_2$  is the best reply of its opponent. B gets help and does not work. That is, in a nutshell, the result which many people lament being the typical feature of the modern welfare state.

Note, that the game represented by fig. 1 must be interpreted as implying imperfect information on the side of both players. In other words, the game is a simultaneous game. As a sequential game the player moving second would have to have four strategies.

The Samaritan's dilemma can also occur if the potential Samaritan does not have a dominant strategy. This case is depicted in fig. 2 (see Buchanan 1977: 172).

	$B_1$	$B_2$
$A_1$	I 4,2	II 1,1
$A_2$	III 2,3	IV 3,4

Fig. 2: The Passive Samaritan's Dilemma

This game is set up by simply transposing the pay off numbers for player A as between cells I and III. The game has two equilibria, strategy profiles  $(A_1, B_1)$  and  $(A_2, B_2)$ .

Buchanan calls the games presented in fig. 1 and 2 the active and passive Samaritan's dilemma, respectively. The reason is, that strategic behavior on the part of the potential Samaritan is always dictated in the fig. 1 game whereas in the case of fig. 2 strategic behavior "may be dictated only when a specific gaming situation is forced upon him by his opponent" (Buchanan 1977: 174). That means, if by whatever reason the strategy profile  $(A_2, B_2)$  happens to be the outcome of the game.

From the perspective of game theory strategy profile  $(A_2, B_2)$  is an example of a dilemma on the side of player A because he cannot implement the outcome he prefers most. Thus, both games represent a personal dilemma and not a social dilemma, if the term social dilemma

refers to an equilibrium which is Pareto inefficient. In the game of fig. 1 the equilibrium is Pareto-efficient (as well as cell III). The same holds for the two equilibria in the game of fig. 2.

Of course, one is wondering whether player A with preference ordering as shown in fig. 2 should be considered as being a Samaritan at all, since he prefers an outcome most, in which he does not help. However, his preference ordering might reflect features of a Samaritan because helping is the preferred option in case of B's choice not to work.

The character of player A as revealed by his preferences in the fig. 1 game needs some further comments. As can easily be seen, helping B is A's dominant strategy. Whatever B does to help is preferred by A to not doing so. Since helping has a higher pay off for A whatever B chooses to do, player A can be viewed as an unconditional Samaritan. A's ranking of cell III and IV compared to cell I and II, respectively, mirrors the ranking of these cells by B. That is what we would expect from somebody having altruistic preferences. However, A's character is much more complex as becomes clear by a comparison of cell IV and III. A's utility increases by moving from cell IV to III. With respect to these cells the Samaritan does not care about B's preferences. Although B would suffer an utility loss by the move from IV to III this move increases A's utility. This increase could be explained in several ways: B's work contributes to social product and A can participate in that. Or: From A's point of view being confronted with lazy-bones is a bad, which can be considered as a negative externality lowering his utility. Note, that in the case mentioned first the personal dilemma of the Samaritan would also be a social dilemma.

### ***III. Strategic Courage***

In this section we discuss potential means for the resolution of the Samaritan's dilemma. The focus will be on punishment within an iterated game, the sequentialisation of the one shot game and the delegation of the power to a decision to a third party. Attention will be restricted to the active Samaritan's dilemma.

## 1. Iterated game

The game of fig. 1 can be considered being a stage game that is repeated a number of times. Under the assumption that players can observe the outcomes of previous games before playing a later stage game, players can condition their optimal actions on the other players' behavior in the past. Since this opens up the possibility of punishing and rewarding another player one might ask whether the Samaritan can shape the incentives of his opponent in such a way that he decides to work, i.e. to choose action  $B_1$ . Buchanan already mentioned the punishment strategy opened up by the iteration of the stage game. He argues, that in order to induce B to choose  $B_1$ , i.e. to work, the Samaritan must choose row 1 rather than row 2 when player B is observed or predicted to select column 2 (see Buchanan 1977: 171). But he concludes: "This choice will 'hurt' A. Admittedly, the utility losses may be short-term ones only, and there may be offsetting long-term utility gains in a sequential game, but once the trade-off between short-term utility and long-term utility is acknowledged to be present, we must also acknowledge that A's subjective discount rate will determine his behavior. If this rate is sufficiently high, A may choose nonstrategically, even in the full recognition of the game situation that he confronts." (Buchanan 1977: 171.)

Modern game theory supports the view that the punishment strategy will not work.

- If the game is finitely repeated backwards induction shows that action profile  $(A_2, B_2)$  will be the outcome of each stage game. The reason is, that  $A_2$ , i.e. to help, remains A's dominant strategy in each stage game. The Samaritan's Dilemma still exists.
- In the infinitely repeated game things look better at first glance. One might be reminded of the Folk-Theorem according to which a whole bunch of subgame-perfect Nash equilibria exist if the discount factor is sufficiently close to one and the Nash equilibrium of a finite, static game of complete information is not pareto-efficient (see Gibbons 1992: 88-99). In this case each point in the Pareto-region can be implemented as a subgame perfect equilibrium.

Unfortunately, the Pareto-region of the Samaritan's Dilemma game is empty, because the Nash-equilibrium is pareto-efficient. Thus, there is only one subgame-perfect equilibrium in the infinitely repeated game, namely  $(A_2, B_2)$ . Thus, it is impossible to overcome the dilemma by infinitely repeating the stage game.

If with iteration of the stage game the dilemma cannot be resolved, the solution – if any – must be found in the alteration of the stage game. We will now discuss two alternatives.

Firstly, the design of a dynamic one shot game. Secondly, the involvement of a third player.

## 2. Dynamic one-shot game

In a dynamic one-shot game the interaction of the players is inherently dynamic in the sense, that some players can observe the actions of other players before deciding what to do.

Let us model the interaction between the Samaritan and its opponent as a two-period game, with one player moving in period one and the other in period two under the assumption that the second mover observed the decision of the first mover. Consider fig. 3. The Samaritan A as the first mover has to decide whether to help ( $A_2$ ) or not to help ( $A_1$ ). B having observed A's decision can choose between work ( $B_1$ ) and not to work ( $B_2$ ). The pay offs are from fig. 1. Backwards induction shows, that it is rational for the Samaritan to choose  $A_2$ , i.e. to help, expecting that his opponent plays  $B_2$ . As the equilibrium path represented by doubled edges reveals the Samaritan's Dilemma still exists.

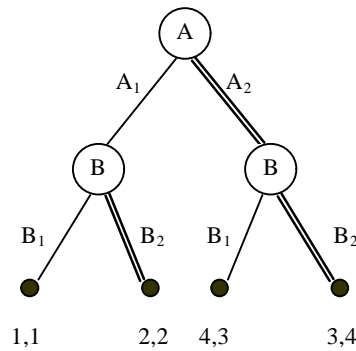


Fig. 3: Dynamic One-shot Game  
(Samaritan as first mover)

If, alternatively, the opponent moves first, the outcome of the game remains the same as an analysis of fig. 4 makes clear.



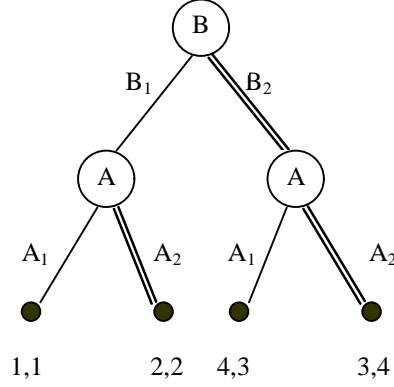


Fig. 4: Dynamic One-shot Game  
(Samaritan as second mover)

Since in the last stage of the game the Samaritan chooses  $A_2$  the opponent chooses  $B_2$ , i.e. not to work. It follows, that sequentialisation of the one-shot game with the preference ordering of fig. 1 does not resolve the Samaritan's dilemma. The reason is, of course, that "help" remains the dominant strategy of the Samaritan.<sup>1</sup>

We can now state a general result: It is impossible to resolve the Samaritan's dilemma given that only the Samaritan and its opponent are the players in the game and given their preference orderings.

### 3. Delegating the power of decision

If the Samaritan would announce in the one-shot game of fig. 1 that he will only help his opponent if the latter chooses  $B_1$ , this announcement would not be credible. As Buchanan already pointed out in his article credibility can be gained if the Samaritan locks himself into a strategic behavior pattern in advance of any observed behavior (see Buchanan 1977: 176).

---

<sup>1</sup> With the preference ordering of fig. 2, however, the Samaritan's dilemma can be overcome, if the Samaritan is the first mover. He will choose  $A_1$  and his opponent will optimally react by choosing  $B_1$ , i.e. to work. It is straightforward that with the opponent being the first mover the equilibrium play will be  $(B_2, A_2)$ .

One way to accomplish this is to "delegate the power of decision in particular choice situations to an agent, one who is instructed to act in accordance with the strategic norms that are selected in advance." (Buchanan 1977: 177.) Buchanan argues that the "agency device" serves two purposes simultaneously: "First, the potential parasite is more likely to believe that the agent will behave in accordance with instructions. Second, by delegating the action to the agent, the Samaritan need not subject himself to the anguish of situational response which may account for a large share of the anticipated utility loss." (Buchanan 1977: 177.)

Whereas the latter purpose does not raise any queries the first one does. Why should the potential parasite believe that the agent will behave in accordance with instructions? Whether the agent will do what the Samaritan expects him to do depends on the incentives generated by the contract governing their relationship. Without knowing the details of the contract as well as the utility function of the agent the potential parasite might not be induced to behave the way the Samaritan would like.

What we have here is a trilateral agency relationship. The Samaritan acts as the principal, the agent is the agent as far as the Samaritan is concerned but he is at the same time the principal in relation to the potential parasite. Let us analyze this three players game with the help of a model.

A (Samaritan) offers D (agent) the following contract: I promise to pay \$  $T > 0$  to you for you watching the behavior of potential parasite B. If you see B working, help; if you see B not working do not help. Help means to give B \$30. I entrust you with \$30 at the moment you accept the contract. It is your duty to pay \$30 back in case of breach of contract. Payment T is due after you have done your job.

We assume for simplicity that the terms of the contract are verifiable in court and that filing suits is costless.

The game tree of the three players Delegation game is depicted in fig. 5.

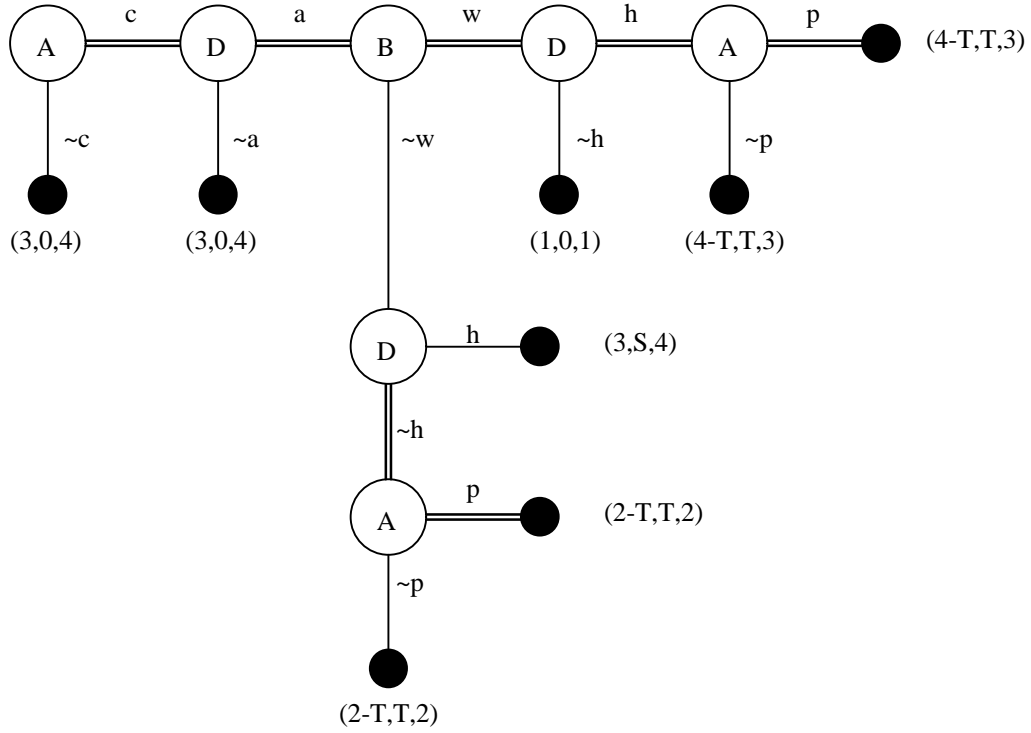


Fig. 5: The Delegation Game

The actions of the players are denoted as follows:

- $c$  =: offer contract
- $\sim c$  =: do not offer contract
- $a$  =: accept contract
- $\sim a$  =: reject contract
- $w$  =: work
- $\sim w$  =: do not work
- $h$  =: help
- $\sim h$  =: do not help
- $p$  =: pay agent
- $\sim p$  =: do not pay agent

The subgames representing court involvement are omitted.

The pay offs are indicated at the endnodes in the order (A, D, B). We assume that the pay offs are the monetary equivalents of utility; the reservation wage of the agent D is normalized to zero, and  $4 - 3 > T > 0$ .

In this game players can either move across or down. The sequence of the moves starts with A's decision to offer a contract or not. D is next to decide whether to accept the contract or not. B as the third mover is confronted with the choice between working and not working. Then it is again D's turn; but now he has to decide whether to help or not to help. The final decision is taken by A, having to choose whether to pay T or not.

Since the court-subgames are omitted some remarks regarding the pay offs of the plays with court involvement seem in order. There are four plays of the game with court involvement.

- (c, a, w, h, ~p)

D did stick to the contract terms but A rejects to pay T. As the pay offs indicate the court perfectly enforces the contract: A pays T to D.

- (c, a, w, ~h)

In this case D breaches the contract. A files suit, in order to get the money back that he entrusted with D.

- (c, a, ~w, h)

As in the previous case D breaches the contract. He decides to help although B does not work. A files suit, in order to get the money back that he entrusted with D.

- (c, a, ~w, ~h, ~p)

This case is similar to the case mentioned first.

The pay offs of the play (c, a, ~w, ~h, p) need some further comments. D sticks to the contract terms and does not help. It is assumed, that he gives A the money back entrusted with him.

This result seems plausible given our assumption of a perfect and costless court system.<sup>2</sup>

As can be seen, A is indifferent regarding p and ~p at both his upper right and lower decision node. To break the tie we assume that he always chooses the options that does not imply court activities.

---

<sup>2</sup> As has been shown elsewhere (Kirstein/Schmidtchen 1997) contractual compliance can be guaranteed in a costly system and with judges committing errors if their detection skills are positive. The latter concept is based on Heiner (1983):

Assuming common knowledge of rationality as well as of the game tree we can derive the subgame perfect equilibrium by backwards induction. This equilibrium is given by the play  $(c, a, w, h, p)$ , leading to the pay off vector  $(A, D, B) = (4 - T, T, 3)$ . The actions chosen by the players are indicated in fig. 5 by doubling the edges.

The backwards induction reasoning is as follows starting with the upper part of the game.

In the last stage of the game A chooses  $p$  given our rule for breaking ties.

Since  $T > 0$  D chooses  $h$ .

Since  $3 > 2$  B decides to work, thus preventing that the play of the game moves downwards.

Since  $T > 0$  D accepts the contract.

Since  $4 - T > 3$  A offers the contract.

As it turns out, Samaritan A can implement his first best option by delegating the decision to an agent. Given the contract and given the workability of the court system the agent acts according to the instructions of the Samaritan.

Of course, there is one question left. If the court system works perfectly and costlessly why did not the Samaritan conclude a contract with B, thereby making B his agent?

By doing this he could save the payment  $T$ , which – in the case of success - would give him higher utility than the outcome of the above mentioned game. However, as it turns out this is an example of fallacious thinking. Whereas in the three players game A has an incentive to sue D in a case of breach A would never bring suit against B. Moreover, we have reason to believe that A will never accept such a contract.

To illustrate:

In the three players game A has an incentive to sue D if D breaches the contract. We consider two cases, action profiles  $(w, \sim h)$  and  $(\sim w, h)$ . If D does not help although B decided to work (action profile  $(w, \sim h)$ ), A brings suit in order to get the money back entrusted with D. The reason is that with a suit he gets pay-off 1 whereas without suit he would get (1 minus utility of the money entrusted with D). He also refuses to pay  $T$ . That is why D helps given B decided to work.

The more interesting case however is given by the action profile  $(\sim w, h)$ . Here A has the right and the incentive to refuse paying  $T$  as well as to get the money back entrusted with D. If D would not give the money back A would receive utility (3 minus utility of the money) instead of 3. D's pay-off  $S$  is negative if he is able to pay back the money. If D's budget constraint

does not allow for that his pay-off would be zero and A's pay-off would go below 3. Since  $T > 0$  D decides not to help, thereby giving B the incentive to work.

What would happen, if A would conclude a contract with B stating that B promises to work and A pays a certain amount of money?

The first question is whether B has an incentive to accept such a contract. The answer is straightforward: B knows that he gets help anyway, therefore he has no incentive to accept.

But assume, for whatever reason, B stepped in and furthermore that the money is paid in advance. Since B decides not to work, we have a breach of the contract. Though being confronted with a breach of contract A would not bring suit against B in order to get the money back, given his preference ordering.

Can we expect a different outcome if the contract states that B promises to work and that the money will be paid after B decided to work? B, knowing that help is A's dominant strategy, will breach the contract, because he can be sure that A will pay the money. The contract does not rule out this possibility.

That a contract between A and B will not be enforced in the case of breach is not surprising. The existence of an option to file suit does not imply the incentive to choose this option. The incentive structure of the contract game between A and B is exactly that of the two versions of the one-shot dynamic game analysed above.

What about A bringing suit in order to force B to work? Here the problem arises, whether the legal order allows for forced labor. If not there is no reason to bring suit. If forced labor is allowed, B would not accept the contract in the first place, knowing that the Samaritan will help anyway.

#### **IV. Conclusion**

It could be shown in this paper that the Samaritan's dilemma cannot be solved if the game remains a game played only by the Samaritan and its opponent. Neither iteration of the simultaneous stage game nor reshaping it in the direction of a one-shot dynamic game nor concluding an enforceable contract work. There is only one way out of the dilemma the delegation of the power to the decision. However, the courts must work in a sufficient manner in order to get the incentives right. With an imperfect court system things might look different.

The analysis has been restricted to interactions with only one Samaritan. However, there may exist more than one potential Samaritan. If all would decide to hire an agent on the terms discussed in the above section the Samaritan's dilemma could be prevented to become a public or social matter. However, if the population of Samaritans is heterogeneous, "enlightened" Samaritans, which have hired an agent, might now have to compete with those having chosen the "soft option". With "potential parasites" preferring to get help from the latter group, something described by Gresham's law result. "Bad" Samaritans drive the "good" ones out of the market. This process might be fuelled by the possibility to give the money to the poor that otherwise would go to the agent. What we have is a kind of a collective prisoner's dilemma, which requires collective action to be overcome.

In modern societies government steps in as the agent. Can we trust government? In a democracy vote shares are decisive. Even if a large number of Samaritans would opt for the strong option, one might ask whether they have sufficient voting power. In any case Buchanan has been sceptical 25 years ago, as can be read out of the following quote: "Government do little more than reflect the desires of their citizens, and the taking of soft options on the part of individuals should be expected to be accompanied by an easing up on legal restrictions on individual behavior." (Buchanan 1977: 185.)

## References

- Buchanan, J.M. (1972): The Samaritan's Dilemma, reprinted in: Buchanan, J.M. (1977): Freedom in Constitutional Contract, Texas A&M University Press: 169-185.
- Dixit, A., B. Nalebuff (1994): Thinking Strategically, W.W. Norton.
- Gibbons, R. (1992): A Primer in Game Theory, Harvester Wheatsheaf, New York.
- Heiner, R. (1983): The Origin of Predictable Behavior, in: American Economic Review 73: 560-595.
- Kirstein, R., Schmidtchen, D. (1997): Judicial Detection Skill and Contractual Compliance, in: International Review of Law and Economics, Vol. 17, No. 4: 509-520.
- Lindbeck, A., Nyberg, St., Weibull, J.W. (1999): Social Norms and Economic Incentives in the Welfare State, in: Quarterly Journal of Economics, Vol. CXIV, Feb.: 1-35.