



HAL
open science

Altruism and Strategic Courage. Inside Buchanan's Samaritan's Dilemma

Stefano Dughera, Alain Marciano

► **To cite this version:**

Stefano Dughera, Alain Marciano. Altruism and Strategic Courage. Inside Buchanan's Samaritan's Dilemma. 2022. hal-03684249

HAL Id: hal-03684249

<https://hal.science/hal-03684249>

Preprint submitted on 6 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Altruism and Strategic Courage.

Inside Buchanan's Samaritan's Dilemma

Stefano Dughera^{a,b} and Alain Marciano^c

Abstract: The Samaritan's Dilemma has largely been investigated, frequently by assuming that Samaritans help recipients out of altruism. Yet, Buchanan did not make any behavioral assumption regarding the Samaritan's motives. In this paper, we explicitly introduce this assumption in Buchanan's original model and analyze how this changes the nature of the game. We show that altruism *alone* does not explain the dilemma. A parameter that captures the disutility the Samaritan feels when helping someone who does not reciprocate her benevolence must be introduced to make sense of the different version of Buchanan's Samaritan's Dilemma. We also show that the Samaritan's dilemma is an evolutionary stable outcome, which confirms Buchanan's intuitions. Finally, a third important point put forward in the paper is that the more altruistic are the Samaritans, the less likely it is that they will show the kind of strategic courage envisaged by Buchanan, which is one of the most important traits Samaritans should display to avoid being trapped in a dilemma.

Keywords: Buchanan, Samaritan's Dilemma, Altruism, Strategic Courage

JEL Classification: B15, C73, D64

^a University of Torino, Department of Economics and Statistics, Lungo Dora Siena 100/A – 10153 Torino.

^b LABORatorio R. Revelli Centre for Employment Studies, Piazza Vincenzo Arbarello, 8, 10122 Torino.

^c (corresponding author), e-mail: alain.marciano@umontpellier.fr. MRE and University of Montpellier, University of Montpellier, Faculté d'économie Avenue Raymond Dugrand, CS 79606, F-34960 Montpellier cedex 2 France.

Altruism and Strategic Courage.

Inside Buchanan's Samaritan's Dilemma

Abstract: The Samaritan's Dilemma has largely been investigated, frequently by assuming that Samaritans help recipients out of altruism. Yet, Buchanan did not make any behavioral assumption regarding the Samaritan's motives. In this paper, we explicitly introduce this assumption in Buchanan's original model and analyze how this changes the nature of the game. We show that altruism *alone* does not explain the dilemma. A parameter that captures the disutility the Samaritan feels when helping someone who does not reciprocate her benevolence must be introduced to make sense of the different version of Buchanan's Samaritan's Dilemma. We also show that the Samaritan's dilemma is an evolutionary stable outcome, which confirms Buchanan's intuitions. Finally, a third important point put forward in the paper is that the more altruistic are the Samaritans, the less likely it is that they will show the kind of strategic courage envisaged by Buchanan, which is one of the most important traits Samaritans should display to avoid being trapped in a dilemma.

Keywords: Buchanan, Samaritan's Dilemma, Altruism, Strategic Courage

JEL Classification: B15, C73, D64

1. Introduction

Samaritanism, benevolence, helping people in need, trying to relieve their grief through aid and donations, is considered as one of the main moral duties upon which rest—or should rest—our societies, one duty that we should not try to escape for the good and welfare of all. This seems straightforwardly obvious and yet it is not. To paraphrase the title of a book by Steve Corbett and Brian Fikkert (2014), “helping hurts”. Despite a short-term relief, the recipients of aid or charity, gifts or transfers lose self-reliance, work and save less and adopt riskier behaviors. As a consequence, their wealth decreases, rather than increases; their situation deteriorates rather than improves and they end up increasingly rely on external help. In addition, lured by the short-term positive effects of aid, other persons in need ask for help too. The negative

consequences of aid and charity spread. Indeed, help gives birth to “tragic commons” (Schmidtz, 2000, 686), it creates “charity hazard” (see Raschky and Weck-Hannemann, 2007; Browne and Hoyt, 2000; Smetters, 2003), or, as Buchanan put it (1975a), traps both the helper—the “Samaritan”—and the recipient in a “Samaritan’s Dilemma”.

A rather problematic situation, the Samaritan’s Dilemma also strikes by its “pervasiveness” (Pasour, 1991). It occurs in a large variety of situations: private settings—families, for instance (see Futagami, Kamada and Sato, 2004)—as well as public environments—redistribution and poor relief (Wagner, 2005; Skarbek, 2016), medical care and national health insurance, social programs (Bovard, 1983; Boettke and Martin, 2010; Zelinsky, 2010), international aid, in particular in case of natural disaster such as hurricanes or earthquakes (Boone, 1996; Wilson, Andersson, Ostrom and Shivakumar, 2005; Coyne, 2008, 2013; Williamson, 2010; Stone, 2008; Rajan and Subramanian, 2005; Burnside and Dollar, 2000, among others). Therefore, preventing the dilemma from occurring and finding ways to help Samaritans and recipients to get out of it seems particularly crucial. Studying ways out of the dilemma has thus unsurprisingly been the subject of many works (see for instance, Thompson, 1980; Veall, 1986; Charlton, 1987; Kotlikoff, 1987; Lindbeck and Weibull, 1988; Hansson and Stuart, 1989; Wagner, 1989; Bruce and Waldman, 1990, 1991; Coate, 1995; Easterly, 2003; Lagerlof, 2004; Poulsen and Tinggaard Svendsen, 2005, Blouin and Pallage, 2008; Leeson, 2008; Burns 2009; John and Storr, 2009; Skarbek, 2016).

Buchanan (1975a) himself suggested solutions to prevent the dilemma. In his view, the strategic behaviors of the recipients could be avoided by adopting constitutional rules that would frame the social game. Alternatively, such rules were not necessary if Samaritans could display some strategic courage, or an ethic of responsibility that would stop them from helping persons in need, despite the difficulty and cost such a courage entails. If Samaritans were able

to bear the short-run costs of not helping persons in need, the entire society would end up in a long-run better situation. This is probably one of the main messages Buchanan conveyed in his essay, one of the strongest points of his work, that is very rarely noted—among the exceptions are Schmidtchen (1999); Skarbek (2016); Hertzberg and Goodman (2020).

Yet, in Buchanan's view, the dilemma was likely to persist over time, since the widespread diffusion of charitable and helping behaviors across all fields of human interactions inevitably leads to a pervasive situation where individuals lose their self-reliance, unless willingly devised institutional solutions are put in action to limit its resilience. Because of the Samaritans *soft-heart* and lack of strategic courage, in fact—another strong message of Buchanan's 1975 essay—this unfortunate societal trajectory is unavoidable. Indeed, to the Buchanan of the early 1970s, individuals were no longer able to behave bravely: “[w]hat we may call “strategic courage” may be a markedly inferior economic good, and what we may call “pragmatic compassion” may be markedly superior” (1975a: 75). He insisted, “modern man has “gone soft.”” (ibid.)¹.

From these quotations, one understands that Buchanan's concern went beyond stressing the existence of a specific dilemma affecting welfare programs, redistribution and charity. It was rather to underline the existence of a dynamics that was transforming society as a whole, and shaking its very foundations:

[...] let me emphasize, however, that I am attempting to develop a hypothesis that is generalizable to much of the behavior that we observe in the modern world. The Samaritan example is used for descriptive clarity, in part because I could think of no better one. [...] The hypothesis does apply to certain aspects of the current policy

¹ These ideas are clearly embedded in Buchanan's (1975: 76) own writing, for instance: “A mother may find it too painful to spank a misbehaving child (“This hurts me more than it does you”). Yet spanking may be necessary to instill in the child the fear of punishment that will inhibit future misbehavior” and “the hypothesis is that modern man has become incapable of making the choices that are required to prevent his exploitation by predators of his own species [...] The weakness here may be imbedded in man's utility function” (ivi: 74).

discussion of welfare reform, but this is only one among many applications, and by no means the most important one (Buchanan, 1975a: 74)

This concern is in all likelihood the reason why, in his essay, Buchanan did not actually study the origins of the dilemma but simply postulated its existence. He indeed started with two game matrixes to characterize the two forms—the active and the passive—of the dilemma and *assumed* without explaining that the payoffs of the Samaritan and the beneficiary were ranked in a certain order that would generate the dilemma. Buchanan no more detailed the behavioral assumptions behind such an undesirable situation. He did not explain why and how being charitable could lead a Samaritan to be trapped in such a dilemma. Thus, to put it in other words, Buchanan explained that Samaritans have become too kind, too charitable, too helpful to be able to get out of the dilemma, but he did not wonder if being too charitable could be the cause of the dilemma in the first place.

This is what we do in this paper. We explore some plausible micro-foundations to motivate the payoff ordering chosen by Buchanan. Following a common assumption in the literature about the dilemma, we assume that Samaritans have altruistic concerns towards recipients, although, let us insist, the altruistic motive is not explicitly present in Buchanan's original contribution. We are thus led to ask whether an altruistic Samaritan would face the same dilemma originally postulated by Buchanan back in 1975. Our key finding, in this respect, is quite surprising: an altruistic Samaritan runs no risk of remaining stuck in the dilemma. For, if the situation where helpers assist and recipients remain lazy is still the Nash equilibrium of the game, we show that it is also Pareto efficient for both the Samaritan and the recipient. Although this can be explained rather easily—by fully internalizing the recipient's preference for remaining lazy through her altruistic concerns, the Samaritan feels perfectly fine to help a recipient who free-rides on her assistance, since this is exactly what the recipient wants. Hence, once we explicitly assume that helpers assist out of altruism, the Samaritan dilemma does not

even exist, contrarily to what is usually put forward in the research about the Dilemma that understands the latter as a problem of excessive altruism.

The result does however not take us away from Buchanan. To the contrary, it brings our analysis closer to his, and to the centrality of ethics in both his thought and in his paper on the Samaritan's dilemma. Once we assume that Samaritans are altruistic, in fact, some additional parameters have to be included in the analysis to make sure that the payoff ordering still generates a Samaritan's Dilemma. In this paper, we remain close to Buchanan's thought by adding a parameter that measures the psychological disutility Samaritans feel when they provide help to lazy recipients. We believe that this psychological cost is exactly what Buchanan had in mind when he talked about the ethics of responsibility Samaritans should display to avoid the dilemma. This ethics means that Samaritans should not abandon themselves to the "gratification of instant desires" (1975a: 83) by helping people who make no effort. If they nonetheless follow their soft heart and fail their ethics of responsibility, it should be somehow costly for them. This is what we want to capture by adding a parameter that counterweights the benefits of altruism. By introducing this behavioral motive into the analysis, in fact, our parametrization becomes rich enough to make sense of the Active and Passive versions of Buchanan's original dilemma in which Samaritans are altruist.

We organize this analysis in two parts. In section 2, we go back to Buchanan's original matrixes and formalize the above assumptions on the Samaritans' altruism and ethics of responsibility as plausible behavioral motives generating the dilemma. In section 3, we expand on this work and develop an evolutionary game that, under given parametrizations, results in a Samaritan's dilemma. Using this type of model, we believe, is consistent with Buchanan's societal concerns recalled in the above., as it allows us to capture, thought in a stylized way, the cultural dynamics that Buchanan had in mind when he talked about the diffusion of excessively

charitable behaviors in the society at large. Indeed, our model allows us to characterize a scenario where Samaritanism and opportunism may coevolve as an integrated system of cultural habits.

This analysis offers three types of payoffs. First, by giving microeconomic and behavioral foundations to the matrixes Buchanan used in his essay, we contribute to improve our understanding of Buchanan's work and, at the same time, of the conditions in which a Samaritan's dilemma exists. We show that Buchanan's pessimistic results were not the consequence of an arbitrarily chosen matrix. They apply in settings where more detailed behavioral assumptions are made. At the same time, our analysis also suggests that Buchanan's conclusions cannot be taken for granted when richer constellations of parameters' value (and related behavioral motives) are considered.

Second, by analyzing the existence conditions of Buchanan's original dilemma, we are able to draw conclusions on possible solutions to escape the later. Not only we reiterate the importance of the Samaritan's strategic courage, in line with Buchanan, but we also emphasize an alternative solution that, to the best of our knowledge, receive no previous attention in the literature. What we highlight is that a key but implicit assumption behind Buchanan's work is that recipients have a taste for remaining lazy and that these preferences cannot be engineered. However, when we relax this assumption and contemplate the possibility that recipients can be induced to abandon their "parasitic preferences", we show that Samaritans are no longer forced to use their strategic courage and can behave charitably without fearing the implications of their good heart. This message, we believe, complements Buchanan's on the importance of the Samaritans' strategic courage, suggesting that there may exist situations where a first-best scenario with hard-working recipients and helping Samaritans is achievable. When this is not the case, Buchanan's pessimistic conclusions are the only possible one.

Third, our paper also contributes to the literature on the impact of unilateral altruism. By contrast to what Gary Becker argued in his Rotten Kid theorem (1974), or Robert Axelrod (1981, 1984) with the Tit-for-Tat strategy, we show that unilateral altruism does not always lead egoist to cooperate. Quite the contrary, and well in line with Buchanan's message on the importance of strategic courage, our analysis suggests that the only way in which an altruistic Samaritan can induce an egoist recipient to "cooperate" (that is, to put forward her effort) is by not providing help in the first place, that is, to behave altruistically. In our model, Buchanan's result is closer to the criticisms of Becker's model that one finds in the works of Ted Bergstrom (1989), Jack Hirshleifer, (1977) or Gordon Tullock (1977).

The remainder of the paper is organized as follows. Section 2 discusses Buchanan's original matrix and formalize the assumptions presented above on altruism and responsibly. Section 3 expands on this work by developing an evolutionary model to assess how Samaritanism and opportunisms may coevolve as an integrated system of cultural norms. Section 4 discusses the possible ways out of the dilemma. Section 5 concludes.

2. Buchanan's original work

In his essay, Buchanan described a situation in which a potential helper who clearly has a taste for assisting people in need is stuck in the following dilemma. On the one hand, she may choose to follow her personal inclination and provide assistance to a recipient who, however, will free-ride on her aid by deciding not to work. On the other hand, the helper may decide to show what Buchanan calls "strategic courage" and refrain from providing assistance to the person in need, although this will expose her to the unpleasant situation of watching the potential recipient persisting in her state of distress.

Buchanan presented the dilemma in two different versions. In the first, which he calls the “Active Samaritan Dilemma”, giving aid is the helper’s dominant strategy, regardless of whether the latter is aware that she is in a game at all. In this framework, helping is the Samaritan’s “pragmatic or independent-behavior response to the choice situation that he confronts, whether or not [she] recognizes that [the recipient] exists as a choice-making entity who opposes him in a game-like situation” (Buchanan, 1975a: 72). By looking at the original payoff matrix representing this situation—see Table 1—one can infer three fundamental features of this situation, without finding explicit indications in Buchanan’s discussion on how to interpret the latter. First, Samaritans derive utility from their acts of charity; second, they prefer recipients to put forward their own effort (otherwise their utility would be higher in the Aid/No-Effort situation); and third, recipients have a preference for remaining lazy when receiving assistance from a Samaritan. Given this very precise array of assumptions, the only Nash-equilibrium of the game is the asymmetrical situation where the Samaritan helps and the recipient does not work. In this framework, the helper is stuck in an inferior situation, since, as anticipated, her preferred outcome is that where she helps and the recipient exerts effort.

| | | Recipient | |
|-----------|--------|-----------|-----------|
| | | Effort | No Effort |
| Samaritan | No Aid | 2,2 | 1,1 |
| | Aid | 4,3 | 3,4 |

Table 1: The Active Samaritan’s Dilemma

In the second version of the dilemma, which Buchanan calls the Passive Samaritan’s Dilemma—see Table 2—the helper is well-aware of the strategical implications of her decision,

as can be seen from the fact that she is no longer happy to give unconditionally. This seems to suggest that when the Samaritan becomes aware that her good actions crowd out the recipients' willingness to exert effort, she puts less emphasis on the utility she derives from her acts of charity, and more on the psychological cost of helping a lazy recipient. Even in this strategic situation, however, the Samaritan may end up in an inferior situation. Indeed, the Passive Samaritan's Dilemma has two equilibria in pure strategies, Effort/No-Aid and No-Effort/Aid, and while the former is strictly preferred by the Samaritan, the latter is strictly preferred by the recipient.

| | | Recipient | |
|-----------|--------|-----------|-----------|
| | | Effort | No Effort |
| Samaritan | No Aid | 4,2 | 1,1 |
| | Aid | 2,3 | 3,4 |

Table 2: The Passive Samaritan's Dilemma

The question we ask, at this point, is what kind of preferences may generate these two different scenarios. The remainder of this section is dedicated to put forward some plausible micro-foundations that may explain the Samaritan's behavior in Buchanan's original games. In doing so, we do not aim at modelling—correctly or not—what Buchanan had in mind back in 1975. More modestly, we just wish to test whether and under which conditions his message is valid when the behavioral assumptions behind his original games are specified a little further.

Now, throughout his paper, Buchanan provided little intuition to the assumptions that could explain the preferences guiding the behavior of these two players. In particular, he never said or even hinted that Samaritans could help out of altruism.² Yet, for various reasons—one of them being that the paper was presented in a conference on altruism and published in a book entitled “Altruism, Morality, and Economic Theory” (Phelps, 1975; Fontaine, 2007)—, it is usually believed that this is the case. So, to follow this literature (and because it seems reasonable to claim that someone who enjoys helping others is an altruist), we assume that Samaritans have altruistic concerns towards recipients. As usual, we model altruism through utility interdependence and assume that the Samaritan internalizes the recipient’s payoff through an altruism parameter $\alpha \in (0,1)$. The higher is α , the more altruistic is the Samaritan. In addition, we assume that helping implies a cost measured by $h > 0$. Hence, we rewrite Buchanan’s original matrix as:

| | | Recipient | |
|-----------|--------|------------------|------------------|
| | | Effort | No Effort |
| Samaritan | No Aid | $2\alpha, 2$ | $\alpha, 1$ |
| | Aid | $3\alpha - h, 3$ | $4\alpha - h, 4$ |

Table 3: The Samaritan’s Dilemma with altruism and helping costs

A quick inspection of Table 3 reveals an interesting result: once we allow for altruism alone, the game is a Passive Samaritan’s Dilemma if the helper is not too altruistic—formally, if $\alpha < h/2$ —while it can never be an Active Samaritan’s Dilemma. Indeed, while No-Effort/Aid is

² When he wrote his article, Buchanan was concerned by other issues that had not much to do with altruism (see Fleury and Marciano, 2018).

still the only Nash-equilibrium, the Samaritan is now in her most preferred situation, contrarily to what happens in Buchanan's original game. The reason is as follows: by fully internalizing the recipient's preferences for remaining lazy, a purely altruistic Samaritan is perfectly fine with the fact of providing assistance to someone who exerts no effort. Clearly, this is inconsistent with the game in Table 1, where the helper is stuck in an inferior situation since she would rather be in the Aid/Effort situation. Hence, our first finding is that altruism alone is not enough to explain the emergence of the Active Samaritan's dilemma: some other, compensating mechanism must be at play for the Samaritan to have the kind of preferences suggested by Buchanan. Interestingly, this also suggest that being charitable is not the cause of the Active Samaritan's dilemma. Quite the opposite, our analysis suggests that altruism *per se* would actually annihilate the dilemma.

We believe that this insight conveys an important message, that it is crucial to understand the centrality of ethics in Buchanan's work. To make sense of Table 1, in fact, one must acknowledge that the Samaritan's altruistic concern (and the resulting psychological gain) is mitigated by some other factor arising when she helps a lazy recipient. For instance, some kind of "ethical" preferences measuring the Samaritan's distaste for helping a lazy recipient may lower the helper's utility in the Aid-No-Effort situation. Incidentally, this is consistent with Buchanan's idea that helping is not a matter of altruism (or at least, not exclusively), but rather, that it must be understood as a strategic tool to induce individuals to make efforts and work. Going back to Buchanan's (1975a: 80) own words, these "ethical preference" may capture the "evolutionary acceptance of what we may call "responsible" standards", where "not helping a lazy recipient" may well provide an example of what a "responsible standard" reasonably is. In analogous but more institutional terms, these "ethical preferences" may represent the subjective adherence to a code of behavior (i.e., a social custom or norm) that states that it is

wrong to help those who free-ride on the assistance of others. In this framework, the voluntary acceptance of “responsible” standards requires that Samaritans behave in ways different from those indicated by their direct and apparent self-interest, thus embracing “an individual ethic of responsibility [...] akin to the Kantian generalization principle” (ibid.).

To incorporate this “ethic of responsibility” into the analysis and see how this interact with the introduction of altruism in Buchanan’s original game, we modify Table 3 as in Table 4, where the added parameter $c > 0$ measures the ethical/psychological disutility the Samaritan derives when she fails her ethic of responsibility and help a lazy recipient.

| | | Recipient | |
|-----------|--------|------------------|----------------------|
| | | Effort | No Effort |
| Samaritan | No Aid | $2\alpha, 2$ | $\alpha, 1$ |
| | Aid | $3\alpha - h, 3$ | $4\alpha - h - c, 4$ |

Table 4: The Samaritan’s Dilemma with altruism, helping costs and ethical preferences

The set of behavioral assumptions behind Table 3 is now rich enough to generate parametrizations that are consistent with both versions of the Dilemma. In particular, for the Samaritan’s payoff to be consistent with Buchanan’s Active SD—see Table 1—the newly added parameter must be sufficiently large to offset the altruism-derived utility the Samaritans experience in the No-Effort/Aid equilibrium, but also, sufficiently small to ensure that helping is still the best-response to the recipient’s decision to remain lazy—formally, $\alpha < c < 3\alpha + h$.

In this framework, the dilemma seems to emerge from a profound behavioral contradiction. On the one hand, altruistic Samaritans seem to be too charitable to put a halt to their “pragmatic compassion”—as Buchanan (1975a: 75) calls it—, or, from a different angle, too poorly compliant with the social norm that imposes not to help lazy recipients. As noted by Goodman and Hertzberg (2020:711), in fact,

“[a] Samaritan who believes welfare without work harms the recipient no longer wants to give unconditionally [...] A prospective Samaritan who has fully internalized these beliefs will prefer to only give to those who work, though his underlying altruistic intentions toward the recipient are unchanged”.

On the other hand, they are not charitable enough to fully enjoy the psychological gains of their soft-heart, otherwise they would be perfectly satisfied with the outcome of a game where they help and recipients do not exert any effort, precisely as suggested by Table 3.

Once again, we believe that this insight is key to understand the moral philosophy behind Buchanan’s work. As a matter of history of thought, in fact, it shows that the sizeable stream of research that has been treating the dilemma as a problem of altruism is at least partly mistaken. Indeed, our analysis suggests that the modern man may well have become “too soft”, as argued by Buchanan, but not soft enough to fully enjoy the benefits of her soft-heart. In more institutional terms, this may be interpreted as suggesting that Samaritans have become unable of adhering to a code of behavior that prescribes not to help lazy recipients, but, at the same time, that they are also unable to completely dismiss this norm from their set of behavioral motives.

The next step is to develop an evolutionary version of the game developed so far as to assess the conditions under which Buchanan’s pessimistic conclusion about the dilemma actually apply.

3. Evolutionary Samaritans

3.1. Setup and assumptions

Consider a model society where large populations of potential helpers and recipients (both of mass 1) are randomly paired to play a game that may result in a Samaritan's dilemma. At the beginning of the game, all recipients are in a state of need and receive a normalized utility = 0. When they exit the state of need (which happens with an endogenous probability define later) they receive a normalized utility = 1.

The situation we have in mind is one where players engage across many fields of social interaction. In describing what is willingly a general framework, thus, we do not specify the type of relationship bonding each pair of helpers and recipients. Rather, we describe a dynamics where Samaritanism and self-reliance (intended as the habit of dealing with one's problems without relying on external help) may coevolve as an integrated system of cultural norms under a mechanism of social imitation. As already suggested in the incipit of this paper, we believe that this intention is very close to Buchanan's original purpose—see Buchanan's quotation at p. 3 of the Introduction.

The setup of the game is as follows. At each moment in continuous time there are many random encounters between members of the two population. In each of these, recipients must choose whether to rely on their own forces ($x = 1$) or remain lazy in the hope of receiving assistance from an external helper ($x = 0$). Samaritans, in turn, must decide whether to exert helping effort ($y = 0$) or refrain from doing so ($y = 1$). Given the assumption of random matching and large groups, x turns out to be a measure of the probability that a Samaritan is matched with a self-reliant recipient, and y of the probability that a recipient is matched with a helping Samaritan. Players make choices simultaneously, independently and without information on the others' action. While Samaritans cannot know whether recipients will use

their help as a substitute or a complement to their own effort, recipients cannot anticipate whether they will be relieved by a helping Samaritans when making their effort choice.³

In addition, we assume that all recipients escape this state of need according to a transition probability that depends on the Samaritan's choice as well as on theirs. Formally, we denote such transition probability as $\rho(x, y)$, where the dependencies in the brackets indicate both the recipients' and the Samaritans' decision. Given this, it seems natural to assume that $1 \geq \rho(1,0) \geq \rho(0,0) \stackrel{\geq}{\leq} \rho(1,1) \geq \rho(0,1) = 0$. In words, recipients have a high probability of exiting the state of need when they cope with their own problems and Samaritans additionally help— $x = 1$ and $y = 0$; the probability is intermediate when only one of the two agents exerts effort— $x = 0$ and $y = 0$ or $x = 1$ and $y = 1$; and drops down to zero when no agent exerts any effort— $x = 0$ and $y = 1$. To economize on notation, but without loss of generality, we further assume that $\rho(1,0) = 1$ and $\rho(0,0) = \rho(1,1) = \rho \in (0,1)$.⁴ In addition, we denote the recipient's cost of effort as $e > 0$ and the Samaritans' helping costs as $h > 0$.

As in the analysis developed in Section 2, we further assume that Samaritans care for the well-being of recipients and internalize their payoffs through a $\alpha \in (0,1)$ parameter that captures their degree of altruism. In addition, and always in line with the analysis developed in section 2, we assume that there exists a code of behavior in the Samaritan's community that

³ Of course, one can think of many real-life situations where helpers may condition their assistance on the recipients' future behavior, as well as others where the recipients' intentions can be signaled or inferred during the matching procedure. In addition, allowing for repeated interactions would ultimately change the model's predictions. While each of these scenarios is by no means irrelevant and generate a number of problems that can be addressed through other game-theoretic techniques, our focus here is on the multitude of everyday contingencies that fits with the above assumptions. As an example, consider a driver (he) puncturing a tire under the gaze of a bystander (she). Both players have no connection with one another and most likely, will not interact again in the future. In addition, the bystander has some basic mechanical knowledge, while the driver has no idea on how to handle the flat tire. Clearly, the bystander is facing a Samaritan's dilemma.

⁴ The last assumption, $\rho(0,0) = \rho(1,1) = \rho \in (0,1)$ is made to avoid notational clutter. While all our results would be slightly different in quantitative terms by allowing for the possibility that $\rho(0,0) \neq \rho(1,1)$, none would be altered qualitatively.

states that helping a lazy recipient is wrong. The voluntary adherence to this social custom provides Samaritans with an ethic of responsibility that generates psychological disutility when the norm is disobeyed. Formally, we model this idea by assuming that Samaritans bear a psychological cost, measured by $c > 0$, when they provide help to a lazy recipient.

3.2. Expected payoffs

The assumptions detailed above are resumed in Table 4, where Samaritans are row players and recipients are column-players. Before proceeding, it will be handy to put forward the following Lemma, which will be useful for the analysis that follows.

Lemma 1—*The game in Table 4 is an Active Samaritan's dilemma iff conditions (1) and (2) are simultaneously satisfied; while is a Passive Samaritan's dilemma iff conditions (1) and (3) are simultaneously satisfied:*

$$1 - \rho < e < \rho \tag{1}$$

$$(1 - e)\alpha - h > \rho\alpha - c - h > (\rho - e)\alpha > 0 \tag{2}$$

$$(\rho - e)\alpha > \rho\alpha - c - h > (1 - e)\alpha - h > 0 \tag{3}$$

| | Effort ($x = 1$) | No Effort ($x = 0$) |
|--------------------|-----------------------------------|---------------------------------|
| No Aid ($y = 1$) | $(\rho - e)\alpha;$ $\rho - e$ | $0;$ 0 |
| Aid ($y = 0$) | $(1 - e)\alpha - h;$ $1 - e$ | $\rho\alpha - c - h;$ ρ |

Table 5: Payoff matrix

Given Table 5, we denote the recipients' expected utility as $U(x)$, and the Samaritans' as $V(y)$. The expected payoffs to each strategy are as follows:

$$U(1) = (\rho - e)y + (1 - e)(1 - y)$$

$$U(0) = \rho(1 - y)$$

$$V(1) = (\rho - e)\alpha x$$

$$V(0) = [(1 - e)\alpha - h]x + (\rho\alpha - cy - h)(1 - x)$$

From the above equations, it is straightforward to calculate the following payoff differences,

$$U(1) - U(0) = (2\rho - 1)y + 1 - e - \rho$$

$$V(1) - V(0) = [(2\rho - 1)\alpha - c]x + h + c - \rho\alpha$$

and the curves along which $U(1) - U(0) = 0$ and $V(1) - V(0) = 0$:

$$y = \frac{\rho - (1 - e)}{2\rho - 1} \quad (4)$$

$$x = \frac{\rho\alpha - h - c}{(2\rho - 1)\alpha - c} \quad (5)$$

Observe that from Lemma 1, we already know that $\rho > 1 - e$, that $\rho > 1/2$ and that $\rho\alpha - h - c > 0$. The next step is to describe how the system may evolve under alternative parametrizations.

3.3. Dynamics, equilibria and stability

We model the diffusion of the $x = 1$ and $y = 1$ strategies via the standard replicator-dynamics derived by Jonker and Taylor (1978). The replicator dynamics is a learning-by-imitation model

which postulates that players are boundedly rational, they learn from each other, and they tend to adopt the strategy that performs better than the other, so that relatively successful behaviors are replicated, while unsuccessful behaviors are abandoned. As anticipated, we believe that this type of model is well-suited to describe a process of cultural transmission where opportunism and Samaritanism (or self-reliance and strategic courage) may coevolve as an integrated system of cultural norms. The system's dynamics are given by:

$$\begin{cases} \dot{x} = x(1-x)[U(1) - U(0)] \\ \dot{y} = y(1-y)[V(1) - V(0)] \end{cases} \quad (6)$$

where \dot{x} and \dot{y} are the time derivatives of x and y respectively. Dynamics (6) is defined in the unit square $Q = [0, 1]^2$. As usual with replicator dynamics, all edges of the square are invariant⁵ and the four vertices $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$ where both populations are homogenous—they are both composed of one type only—are always stationary states. In addition, dynamics (6) may admit another stationary states—indicated as (x^*, y^*) , with $0 < x^* < 1$ and $0 < y^* < 1$ —which is located in the interior of Q and corresponds to the intersection, when existing, of the nullclines defined by (4) and (5). In such state, all four types of players coexist. The existence conditions of this stationary point are given in the following Lemma:

Lemma 2— $0 < x^* < 1$ holds iff condition (1) or (7) are satisfied:

$$\rho < e < 1 - \rho \quad (7)$$

⁵ Meaning that all trajectories starting from any initial pair $(x_0, y_0) = (1, \hat{y})$, $(x_0, y_0) = (0, \hat{y})$, $(x_0, y_0) = (\hat{x}, 0)$ and $(x_0, y_0) = (\hat{x}, 1)$ will lie on the side with $x = 1$, $x = 0$, $y = 0$ and $y = 1$ respectively, where $0 \leq \hat{x} \leq 1$ and $0 \leq \hat{y} \leq 1$.

while $0 < y^* < 1$ holds iff condition (8) or (9) are satisfied:

$$\rho\alpha - c - h < 0 \text{ and } (\rho - e)\alpha < (1 - e)\alpha - h \quad (8)$$

$$\rho\alpha - c - h > 0 \text{ and } (\rho - e)\alpha > (1 - e)\alpha - h \quad (9)$$

Confronting Lemma 2 with Lemma 1, we see that (x^*, y^*) always (resp., never) exists when the game is Passive (resp., Active) Samaritan's Dilemma.

Proof: See the Appendix.

Finally, observe that $\dot{x} = 0$ holds along the curve defined by (5) and along the edges of Q where $x = 0$ and $x = 1$, while $\dot{y} = 0$ holds along the curve defined by (4) and along the edges where $y = 0$ and $y = 1$. The next Proposition analyzes the topological properties of system (6):

Proposition 1—*The stationary point (x^*, y^*) , when existing, is either a saddle, a source or a center (Lyapunov-stable). In addition:*

- 1) *The stationary point $(0,0)$ is asymptotically stable iff $1 - e < \rho$ and $\rho\alpha - h - c > 0$.*
- 2) *The stationary point $(1,1)$ is asymptotically stable iff $\rho - e > 0$ and $\alpha(1 - e) - h < \alpha(\rho - e)$.*
- 3) *The stationary point $(1,0)$ is asymptotically stable iff $1 - e > \rho$ and $\alpha(1 - e) - h > \alpha(\rho - e)$.*
- 4) *The stationary point $(0,1)$ is asymptotically stable iff $\rho - e < 0$ and $\rho\alpha - h - c < 0$.*

Proof: see the Appendix.

The stability properties described in Proposition 1 are rich enough to allow for all

dynamic configurations that are compatible with a 2x2 game with no own-population effects⁶, as summarized by the following Proposition:

Proposition 2—*The game admits four monostable regimes, each featuring a corner of Q as the only attractor; two bistable regimes, featuring $(0,0)$ and $(1,1)$ or $(0,1)$ and $(1,0)$ as simultaneous attractors; and two oscillatory regimes where (x^*, y^*) is Lyapunov-stable and all four corners of Q are saddles. In the cyclical regimes, the value of x and y oscillate either clockwise or counter-clockwise around (x^*, y^*) for any initial pair (x_0, y_0) , with $x_0, y_0 \in (0,1)$. The initial distribution (x_0, y_0) is reached again at the end of every cycle.*

Proof: see the Appendix.

4. Discussion

4.1. The existence of the Samaritan's dilemma

From Propositions 1 and 2, the existence of a Samaritan's dilemma, regardless of its form (active or passive), appears to depend on the values of the different coefficients used in the model we propose. This means that many situations are possible, more than the two Buchanan envisaged. Yet, there are values of α, ρ, e and h for which Buchanan's results hold. Indeed, by confronting the results in Proposition 2 with those in Lemma 1—see Appendix A—we see that when the game is an Active Samaritan's Dilemma, the Aid/No-effort equilibrium is the only long-run attractor of the game—for a visualization, see Figure 1. Clearly, this is consistent with Buchanan's analysis reported in Table 1, and therefore, confirms his main finding: when the game has the form of an Active Samaritan's dilemma, Samaritans will always be trapped in a

⁶ When there are no-own population effects, the players' choices do not depend on the behavior of the other members of their population. In our framework, this can be seen from the fact that $\frac{\partial[U(1)-U(0)]}{\partial x} = 0$ and $\frac{\partial[V(1)-V(0)]}{\partial y} = 0$.

welfare depressing situation. What we have added here is “altruism”. Our key result is that altruistic helpers (who adhere to a code of behavior that says that helping lazy recipients is wrong) find themselves trapped in the same dilemma postulated by Buchanan and, just as Buchanan’s “original” Samaritans, have no way to escape the latter. In addition, we also show that the situation is evolutionary stable, which suggests that it should persist in the long-run.

Another conclusion that can be derived from the comparison of Lemma 1 and Proposition 2 is that Aid/No-effort is also a long-run equilibrium of the Passive Samaritan’s Dilemma. But now, we add another result: in a Passive Samaritan’s Dilemma, the No-Aid/Effort equilibrium is also asymptotically stable— for a visualization, see Figure 2. Once again, this is consistent with Buchanan’s analysis of the Passive Samaritan’s Dilemma, which also features two possible equilibria. What we gain from an evolutionary analysis as ours mainly concerns the process of equilibrium selection. When multiple equilibria simultaneously exist, in fact, “history matters”, since the situation the system will eventually reach depends on initial conditions. Taking a look at Figure 2, in fact, we see that for every initial pair (x_0, y_0) lying below—resp., above —the stable manifold of the internal saddle (x^*, y^*) —corresponding to the dotted line in Figure 2—the system will snowball to the Effort/No-Aid equilibrium—resp., to the No-Effort/Aid equilibrium. Hence, the most distinguishing feature of this dynamic regimes is to be found in its coevolutionary and path-dependent nature. In our framework, path-dependency implies that when Samaritanism is initially widespread in the helpers’ population and/or opportunism is diffused in the recipients’, the process of social imitation will lead to an enduring situation where the two cultural habits will stability reinforce each other. Conversely, when few Samaritans are willing to help and self-reliance is widespread in the recipients’ population, the other long-run situation will eventually emerge.

We could possibly connect this analysis with Buchanan's "probabilistic" analysis of free-riding (1965; 1967a: 85-88), according to which the strategy—to cooperate or not—individuals choose depends on how others behave, that is, on the probability each individual has to meet a cooperator or a defector. According to Buchanan, "individuals decide how to behave after having contemplated whether or not their own action will exert some influence on the behavior of others in the group" (1967: 86). The difference between Buchanan's reasoning and that implied by evolutionary game-theory, is that we do not assume that individuals are capable of taking into account the actual number of cooperators and defectors in the society at large: rather, we assume they behave myopically by imitating the strategy that best-performed in the past. Albeit this minor methodological difference, Buchanan's message and ours are fairly consistent. Indeed, we believe that our evolutionary analysis gives a theoretical foundation to Buchanan's claim: in a group in which there is a large number of defectors, the system will evolve towards an equilibrium with no cooperators; in a group in which there is a large number of cooperators, the system will reach an equilibrium with cooperators. Buchanan indeed insisted: free riding, parasitism, defection can be tolerated if they remain under a certain threshold (1967b, 1968).⁷

When multiple attractive equilibria simultaneously exist and the process of equilibrium selection chiefly depends on initial conditions, a key question naturally arises, namely, which of the long-run scenarios is most desirable, at least from a Paretian viewpoint. Even in more straightforward situations featuring a single attractor (as the Active Samaritan's Dilemma), performing a Paretian analysis can be insightful, as it can shed light on the possibility that in

⁷ Moreover, Buchanan makes an additional remark concerning the influence of group size on individual behavior. In small groups, individuals anticipate that their behavior will affect what others do, so they adopt an ethical rule of conduct—they cooperate. In large groups, they believe that what they do will have no consequence and therefore behave egoistically—they do not cooperate.

equilibrium, some player is stuck in an inferior situation because her preferred outcome is unachievable. By simply comparing the agents' payoffs across the five stationary points of dynamics (6) (regardless of their stability properties), we derive an interesting and somewhat puzzling conclusion: in both versions of the dilemma, no equilibrium is ever Pareto-efficient.

In the Active Samaritan's dilemma, in fact, recipients will eventually find themselves in their most preferred situation, while, as amply anticipated, Samaritans will always be stuck in an inferior configuration. However, even if the transition to any other equilibrium were possible, this would always inflict a utility loss upon recipients, who would no longer be in their preferred equilibrium. It is therefore not possible to improve the situation of both categories of players at the same time. Allowing the Samaritan to move out of the trap would plunge the recipient into it. Or, put in different terms, the Samaritan's dilemma would be just transformed into another kind of problematic situation.

The Passive version of the dilemma clarifies this property even further. Indeed, when multiple equilibria simultaneously exist, the Aid/No-Effort equilibrium can be still viewed as a "welfare trap" for Samaritans, who would be better off in the No-Aid/Effort equilibrium (which is now dynamically achievable). However, the latter is clearly not the preferred situation for recipients, who would rather stay in the Aid/No-Effort equilibrium, since their preferences remain as in the Active Samaritan's Dilemma.⁸ Thus, again, moving out of the trap the dilemma represents can only be made at a cost for the recipients.⁹

⁸ The situation is surely not new in game theory, and resembles a classic "Battle of the Sexes" where neither of the two Nash equilibria simultaneously maximizes the payoff of the two players.

⁹ Interestingly, this strategic property of the Samaritan's dilemma is never mentioned by Buchanan in his original contribution, probably, because it sounded like a sort of contradiction to the very point he was trying to make about the need for recipients to make some effort and work.

4.2. Escaping the dilemma

The main issue Buchanan discussed in his 1975 essay relates to how to escape the dilemma. The best solution, however, consisted in avoiding that the dilemma emerges in the first place. From this perspective, Buchanan starts by suggesting that institutional arrangements that prevent Samaritans to get personally involved with recipients should be encouraged. Talking about welfare program, for instance, he maintains that:

“standards for determining welfare eligibility, either for governmental or private programs, should not be left to the discretion of social workers who get personally involved with potential recipients. This institutional arrangement would force social workers into an acutely painful form of the dilemma discussed”.

Another suggested avenue to avoid the dilemma consists in adopting overarching rules whose implementation are not left to the Samaritans’ discretion. When decision-making is delegated to the helpers’ situational ethics, in fact, room is created for interpersonal feelings to distort the Samaritans’ choice, with all the negative repercussions investigated so far.¹⁰

However, when none of these solutions is achievable and helpers do find themselves in Samaritan-like settings, the only way out from the dilemma, according to Buchanan, is to resort to the Samaritans’ strategic courage, that is, to their ability to refuse to help as a mean to induce recipients to put forward their own effort. In this perspective, a courageous Samaritan is one who is aware of the strategic implications of her choice, and having learnt not to accommodate her pragmatic compassion, plays No-Aid regardless of the recipient’s behavior (i.e., No-Aid becomes her dominant strategy). From a strict game-theoretic viewpoint, this can only be achieved if the payoffs she obtains in the No-Aid/Effort equilibrium are strictly larger than those she gets in the Aid/No-Effort equilibrium, a situation which is never contemplated in both

¹⁰ In *The Limits of Liberty* (1975b), written at about the same period “The Samaritan’s Dilemma” was published, Buchanan used the same game to defend a social contract and a contractualist view of institutions.

versions of the dilemma. In our parametrization, this requires that $\alpha < (h + c)/\rho(0,1) \equiv \hat{\alpha}$, which is never satisfied when conditions (2) or (3) are fulfilled. The point is important since it suggests that the more altruistic are the Samaritans, the less likely it is that they will show the kind of strategic courage envisaged by Buchanan.

A closer inspection of the critical threshold $\hat{\alpha}$ allows us to deliver another interesting message. To start, observe that $\hat{\alpha}$ is increasing in the psychological disutility c that Samaritans derive when they fail their ethics of responsibility by assisting a lazy recipient. As previously discussed, this ethics can be understood as the voluntary adherence to a code of behavior or social norm that prescribes that helping free-riders is wrong. Hence, a plausible reasons for the unfortunate societal trajectory described by Buchanan (and allowed for by our model) can be found in the weakening of such social norm, which in turn, can be explained by referring to the declining role of communities as an endogenous enforcement device. In Buchanan's (1975a: 81-82) words:

There may be no escape from the generalized Samaritan's dilemma [...] except through the collective adoption and enforcement of rules that will govern individual situational responses. As they are applied, such rules must be coercive, and they must act to limit individual freedom of action [...] If the collectivity acts to impose uniform behavioral rules on all potential Samaritans, and if these rules are observed to be enforced, the response patterns of potential parasites will be modified

In this perspective, collectivities have the important role of Samaritans not to follow their soft-heart. As it is well-known, the survival of customs can only be ensured if defectors suffer some form of social sanction. Hence the weakening of interpersonal ties that naturally accrues in large social groups can be viewed as a potential explanation of the Samaritan's dilemma. Hence, reinforcing these ties is a possible way of solving it. One way of doing that is to find a social organization based on small groups in which individuals will cooperate, and this is what Buchanan suggested (1978). Alternatively, societies should develop a collective

awareness of how problematic this situation can be. This can however only be achieved through long lasting processes of cultural change that would progressively boost the Samaritans' awareness that helping, though perhaps counterintuitively, should be understood as an act of defection. Individual defection, because it means failing one's ethics of responsibility, and collective defection too, because this personal failure creates a societal dilemma by encouraging recipients to choose not to exert any effort. The "pseudo-cooperation" of helping, thus, becomes a genuine defection that creates more defection, thus blocking the entire society in the vicious circle Buchanan had in mind back in 1975.

So far, the only ways of escaping the dilemma we discussed relate to the Samaritans' behavior and, more specifically, to the role that strategic courage may play in avoiding Samaritan-like settings. Nothing was said about recipients. This is a direct consequence of what Buchanan wrote. Indeed, Buchanan's reasoning implicitly rests on a key assumption, a premise: recipients have preferences that are clearly situation-dependent. Indeed, they choose to free-ride on the Samaritans' help each and every time this option becomes feasible and accept to put forward their own effort only as a second best. The very existence of the dilemma crucially depends on this precise behavioral assumption, that we shall hereafter refer to by saying that recipients have "parasitic preferences". This premise nonetheless deserves further attention. Indeed, as soon as the latter is relaxed, other possible scenarios may emerge, as put forward by our Proposition 2. For instance, when recipients prefer working than remaining lazy upon receiving no assistance, they cannot strategically use their state of need to induce Samaritans to help, a situation which is recurring in both versions of the dilemma. In this framework, recipients have no preferences to behave as parasites, as can be seen from the fact that they have no intention to leverage on the Samaritans' soft-heart. In these cases, the game becomes

something completely different from the Samaritan's dilemma, and it may even result in situations where providing effort is the recipients' dominant strategy.

This suggests that boosting the Samaritans' strategic courage is not the only way out of dilemma, as originally argued by Buchanan. If indeed it is true that this undesirable social situation no longer exists when Samaritans learn to "resist" their pragmatic compassion and stick more strictly to their ethics of responsibility, it is also true that it analogously disappears when recipients do not have the kind of parasitic preferences implicitly assumed by Buchanan. Hence, any behavioral intervention aimed at modifying the recipients' taste for opportunism may well represent another possibility to escape the dilemma. How to proceed? How could that be done? The answer is certainly difficult but we can suggest some important points that must be taken into account. In particular, one of the questions that seem to naturally arise is whether the strategies to boost the Samaritans' courage and to discourage the recipients' parasitism should be used as substitute or complements.

A simple analysis of Buchanan's original matrixes suggests the following: the Samaritans' courage should be boosted only if the recipients' preferences cannot be altered. Intuitively, teaching Samaritans to behave courageously can only makes sense if the act of helping creates perverse incentives for recipients to engage in opportunistic behaviors. Indeed, if one simultaneously modifies the recipient's taste for remaining lazy as well as the Samaritans' pragmatic compassion, one gets at the undesirable equilibrium where "courageous Samaritans" do not help and "non-parasitic recipients" provide effort. This is particularly evident in the extreme case where the players preferences are so altered that "No-Aid" and "Effort" are the Samaritans' and the recipients' dominant strategy. In this case, the game features No-Aid-Effort as the only Nash equilibrium, which however is Pareto-inefficient for recipients, who would be better off in the situation where they provide effort and receive assistance by external helpers.

In this framework, it is advisable to leave the recipients' preferences unchanged, since this would allow the game to reach the Aid-Effort equilibrium, which is Pareto-efficient.

The conclusion to be drawn here is that inciting Samaritans to boost their strategic courage should be used as a measure of last resort when altering the recipients' taste for remaining lazy is impossible. Once again, we believe that this does not take us away from Buchanan. To the contrary, it highlights even further the premises of his thought and, in particular, his pessimistic view about the diffusion of opportunism in our large and modern societies. By saying this, we want to stress that his 1975 essay not only rests on the idea that the modern man has "gone soft", but also, and complementarily, that he has "gone bad": not only he is too weak to refuse to provide aid in a situation where helping encourages opportunistic behaviors, but he is more than ready to free-ride on someone's help each and every time he is given the possibility to do so.

The above discussion highlights another important message that has not been explicitly considered by Buchanan but that follows directly from our approach. Indeed, once we acknowledge that helping recipients to develop a work ethics can be desirable and, when feasible, even preferable than boosting the Samaritans' strategic courage, one may wonder which route to take to achieve this goal. The answer is rather straightforward: to incite recipients to make effort without relying on external help, their cost of effort must be reduced. When individuals find dealing with their own problems too difficult, in fact, they could pursue other strategies to escape their state of need, regardless of whether they may involve free-riding and opportunism. Hence, instead of, or in complement to, donations aimed at temporarily relieving the recipient's distress—which may have the unintended effects of generating incentives for free-riders to engage in opportunistic behaviors—, Samaritans could try to reduce the recipient's costs of efforts and ease their displeasure or difficulty to cope with

their own problems, for instance, by boosting their self-reliance through psychological support, increasing their dexterity through teaching and guidance and so on and so forth.

We thus confirm a finding put forward in another and yet partially communicating literature, namely, that on the Rotten Kid theorem, which suggests that help should be given in nature rather than in cash (Bruce and Waldman, 1990, 1991; Coate, 1995). The difference is that, our explanation—reducing the recipients' costs of efforts—differs from the one usually put forward—transfers in nature cannot be manipulated by recipients. This difference is rather important. Indeed, it implies that in those situations where helping can be supported by other tools that incentivize the recipients' self-reliance, Samaritans can remain just as soft as they wish to be, knowing that their help will not crowd out the recipient's incentives to put forward their own effort. In these frameworks, helping recipients to become autonomous through strategies that are not based on donations and transfers may yield better results than forcing Samaritans to boost their strategic courage. This is a more optimistic result than the one Buchanan put forward.

5. Conclusion

The purpose of this paper was to analyze the scope and extent of the Samaritan's dilemma under the assumption that Samaritans are altruist. What motivated our analysis was that a large literature on the Samaritan's dilemma took that behavioral assumption for granted, as if Buchanan's essay had been about the possible impact of altruism on cooperation. That was not the case: Buchanan made no assumption as to what were the origins of the behaviors of the Samaritan and the recipient. Introducing this assumption allowed us to put forward a first, surprising result: a purely altruistic Samaritan runs no risk of remaining stuck in the dilemma. Indeed, although s/he ends in the same situation originally predicted by Buchanan where s/he

helps and the recipient remains lazy, s/he is now comfortable with this asymmetrical outcome, and this is due to the fact that she internalizes the recipient's preferences not to make any effort through his/her altruistic concerns. Quite interestingly, this suggests that the previous literature which understood the dilemma as a problem of excessive altruism is at least partly mistaken: some other, compensating mechanism must be at play for the game to have the features of Buchanan's dilemma.

The second contribution of the paper is precisely to show that once we allow for altruism, the dilemma can only emerge if Samaritans have complex and partially contradicting preferences. In line with Buchanan, we indeed assume that Samaritans adhere to an ethics of responsibility that prescribes that helping lazy recipients is wrong. By doing so, we show that helpers get caught in the dilemma each and every time they are too altruistic to behave courageously and watch recipients persist in their state of distress, but, at the same time, that they are "moral enough" to know that this act of charity is fundamentally wrong, as it only encourages recipients to free-ride on their assistance. In addition, using an evolutionary approach, we show that the dilemma is stable. This is also important from the perspective of Buchanan's original paper, since Buchanan wanted to insist that the Samaritan's dilemma is problematic because it lasts over time.

Explicitly introducing altruism in Buchanan's original game is also interesting for another reason, that has to do with how to escape from the dilemma. To Buchanan, Samaritans should display strategic courage, an ethics that would lead them to avoid helping individuals unwilling to make some efforts to reciprocate the Samaritans' help. The problem, as we show it in the paper, is that the more altruists Samaritans are, also the less courageous. Put differently, altruism is incompatible with strategic courage. Thus, if altruism is indeed compatible with Buchanan's Samaritan's dilemma, it is not with his purpose to insist on the need for strategic

courage. In that case, alternative solutions must be envisaged. One of these, discussed in the paper and that Buchanan did not study, relates to the behavior of the recipients. Under certain conditions, altering the recipients' behavior can be more efficient than boosting the Samaritan's ethics or courage. This is particularly useful if Samaritans are altruists.

References

- Becker, Gary S. 1981. *A Treatise on the Family*. Cambridge, MA: Harvard University Press.
- Becker, Gary S. 1974. "A theory of social interactions", *Journal of Political Economy*, 82, 1063-93.
- Becker, Gary S. 1976, "Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology", *Journal of Economic Literature*, 4, 817-26.
- Bergstrom, Theodore. 1989. "A fresh look at the rotten kid theorem—and other household mysteries", *Journal of Political Economy*, 97(5), 1138–1159.
- Blouin, Max and Stephane Pallage. 2008. "Humanitarian Relief and Civil Conflict", *Journal of Conflict Resolution*, 52 (4), 548-565.
- Boettke, Peter J. and Adam Martin. 2010. "Exchange, production, and Samaritan dilemmas", mimeo, <http://mpra.ub.uni-muenchen.de/33199/>
- Boone, P. (1996). "Politics and the effectiveness of foreign aid," *European Economic Review*, 40, 289-329.
- Bruce, Neil and Michael Waldman. 1990. "The Rotten-Kid Theorem Meets the Samaritan's Dilemma," *Quarterly Journal of Economics*, 105(1), 155-165.
- Bruce, Neil and Michael Waldman. 1991. "Transfers in Kind: Why They Can Be Efficient and Nonpaternalistic," *American Economic Review*, 81(5), 1345-1351.
- Buchanan, James M. 1975a. "The Samaritan's dilemma", in E.S. Phelps (ed.), *Altruism, morality and economic theory*, New-York, Sage Foundation, 71-85.
- Buchanan, James M. 1975b. *The Limits of Liberty Between Anarchy and Leviathan*, Chicago: University of Chicago
- Buchanan, James M. 1978. "Markets, States, and the Extent of Morals", *American Economic Review*, 68 (2), 364-368

- Burns, Bryan R. 2009. "Solving Samaritan's Dilemmas in Irrigation Investment", mimeo.
- Charlton, Mark W. 1987. "The Food Aid Conundrum and Canadian Aid Strategy: The Issue of Donor Control", *International Journal*, 42 (3), 473-508.
- Coate, Stephe. 1995. "Altruism, the Samaritan's Dilemma, and Government Transfer Policy," *American Economic Review*, 85(1), 46-57.
- Coyne, Christopher J. 2008. *After war: The political economy of exporting democracy*, Stanford: Stanford University Press.
- Easterly, William. 2003. "Can Foreign Aid Buy Growth?" *Journal of Economic Perspectives*, 17 (3), 23-48.
- Edgeworth, Francis Y. 1881. *Mathematical Psychics. An Essay on the Application of Mathematics to the Moral Science*, London, Kegan Paul.
- Faith, Roger L. 1979. "Local Fiscal Crises and Intergovernmental Grants: A Suggested Hypothesis", *Public Choice*, 34 (3/4), 316-331.
- Futagami, Ritsuko, Kimiyoshi Kamada and Takashi Sato. 2004. "Government Transfers and the Samaritan's Dilemma in the Family", *Public Choice*, 118 (1/2), 77-86.
- Gibson, Clark C., Elinor Ostrom, and Sujai Shivakumar. 2005. *The Samaritan's Dilemma. The Political Economy of Development Aid*, Oxford University Press.
- Goodman, Nathan P. and Herzberg, Roberta Q. 2020. "Gifts as governance: Church Welfare and the Samaritan's dilemma". *Journal of Institutional Economics*, 16, 703-714
- Hazlitt, Henry. 1971. "Poor Relief in Ancient Rome", *The Freeman*, 21(4), p. 215-219.
- John, Arielle and Virgil Henry Storr. 2009. "Can the West Help the Rest? A Review Essay of Sachs' The End of Poverty and Easterly's The White Man's Burden", *Journal of Private Enterprise*, 25(1), 125-140.
- Khalil, Elias. 2001. "Adam Smith and Three Theories of Altruism," *Recherches Economiques de Louvain*, 67 (4), pp 421-435.
- Khalil, Elias. 2004a. "What is altruism?", *Journal of Economic Psychology*, 25 (1), 97-123.
- Khalil, Elias. 2004b. "What is altruism? A reply to critics", *Journal of Economic Psychology*, 25 (1), 141-143.

- Khalil, Elias. 2013. "What Determines the Boundary of Civil Society? Hume, Smith and the Justification of European Exploitation of non-Europeans." *Theoria: A Journal of Social and Political Theory*, 59 (135), 26-49.
- Khalil, Elias. 2014. Exploitative Institutions and Efficiency: Is Slavery Inefficient?, mimeo.
- Lee, Dwight R. 1987. "The Tradeoff between Equality and Efficiency: Short-Run Politics and Long-Run Realities", *Public Choice*, 53 (2), 149-165.
- Leeson, Peter T. 2008. "Escaping Poverty: Foreign Aid, Private Property, and Economic Development", *Journal of Private Enterprise*, 23(2), 39-64.
- Lehto, Otto and John Meadowcroft. 2021. "Welfare without rent seeking? Buchanan's demogrant proposal and the possibility of a constitutional welfare state," *Constitutional Political Economy*, 32 (2), 145-164
- Singh, Nirvikar. 1995. "Unilateral altruism may be beneficial: A game-theoretic illustration", *Economics Letters*, 47 (3-4), 275-281.
- Ostrom, Elinor. 2000. "Collective action and the evolution of social norms," *Journal of Economic Perspectives*, 14(3), 137-158.
- Pasour, E. C., Jr. 1991. "The Samaritan's Dilemma and the Welfare State", *Policy*, summer, 55-57.
- Pedersen, Karl R. (2001). "The Samaritan's Dilemma and the Effectiveness of Development Aid," *International Tax and Public Finance*, 8, 693-703.
- Poulsen, Anders U. and Gert Tinggaard Svendsen. 2005. "Social Capital and Endogenous Preferences", *Public Choice*, 123 (1/2), 171-196.
- Schmidtz, David. 2000. "Islands in a Sea of Obligation: Limits of the Duty to Rescue," *Law and Philosophy*, 19: 683-705.
- Skarbek, Emily C. 2016. "Aid, ethics, and the Samaritan's dilemma: strategic courage in constitutional entrepreneurship," *Journal of Institutional Economics*, 12(2), 371-393.
- Stone, Deborah. 2008. *The Samaritan's Dilemma: Should Government Help Your Neighbor?*, Nation Books.
- Tullberg, Jan. 2004. "On Indirect Reciprocity: The Distinction between Reciprocity and Altruism, and a Comment on Suicide Terrorism", *American Journal of Economics and Sociology*, 63 (5), 1193- 1212.

- Veall, Michael R. 1986. "Public pensions as optimal social contracts", *Journal of Public Economics*, 31 (2), 237-251.
- Wagner, Richard E. 1986. "Wealth Transfers in a rent-seeking polity", *Cato Journal*, 6 (1), 155-170.
- Wagner, Richard E. 2005. "Redistribution, Poor Relief, and the Welfare State", in J. Backhaus and R.E. Wagner (eds), *Handbook of Public Finance*, Springer, pp 385-405.
- Williamson, Claudia. 2010. "Exploring the failure of foreign aid: The role of incentives and information," *Review of Austrian Economics*, 23, 17-33.

Appendix A

1. Proof of Lemma 2

The (x^*, y^*) equilibrium exists when x^* and y^* both $\in (0,1)$. This, in turn, requires that two of the following conditions are simultaneously satisfied:

$$2\rho - 1 > \rho - (1 - e) > 0 \quad (4')$$

$$2\rho - 1 < \rho - (1 - e) < 0 \quad (5')$$

$$(2\rho - 1)\alpha - c > \rho\alpha - h - c > 0 \quad (6')$$

$$(2\rho - 1)\alpha - c < \rho\alpha - h - c < 0 \quad (7')$$

Rearranging condition (4') yields condition (1); rearranging conditions (6') and (7') yields, respectively, conditions (8) and (9). The concluding statement in Lemma 1 follows from the fact that conditions (1) and (9) are simultaneously satisfied when the game is a Passive Samaritan's dilemma, since both are implied by condition (3), while the parametrization defined in condition (1) is inconsistent with any of the possible combinations across conditions (4')-(7'), so that (x^*, y^*) never exists when the game is an Active Samaritan's dilemma ■

2. Proof of Proposition 1

The Jacobian matrix of the system is given by:

$$J = \begin{pmatrix} d\dot{x}/dx & d\dot{x}/dy \\ d\dot{y}/dx & d\dot{y}/dy \end{pmatrix} = \begin{pmatrix} (1 - 2x)[U(1) - U(0)] & x(1 - x)(2\rho - 1) \\ y(1 - y)[(2\rho - 1)\alpha - c] & (1 - 2y)[V(1) - V(0)] \end{pmatrix}$$

As it is well-known, a stationary point is attractive if both the eigenvalues of the Jacobian matrix evaluated at that equilibrium have negative real parts. In addition, when either of the partial derivative $d\dot{x}/dy$ and/or $d\dot{y}/dx$ is = 0—which happens whenever $x = 0$ or $x = 1$ and/or $y = 0$ or $y = 1$ —the eigenvalues exactly correspond to the partial derivatives $d\dot{x}/dx$ and $d\dot{y}/dy$. With these facts in mind, the topological properties of the four vertices $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$ can be straightforwardly checked by evaluating the sign of $d\dot{x}/dx$ and $d\dot{y}/dy$ at each of the four vertices, since the partial derivatives $d\dot{x}/dy$ and $d\dot{y}/dx$ are always = 0 at each of these points.

At $(0,0)$ we have that:

$$d\dot{x}/dx = 1 - e - \rho$$

$$d\dot{y}/dy = h + c - \rho\alpha$$

Hence, we have three cases:

- 1) if $1 - e - \rho < 0$ and $h + c - \rho\alpha < 0$, $(0,0)$ is a sink (asymptotically stable).
- 2) If $1 - e - \rho > 0$ and $h + c - \rho\alpha > 0$, $(0,0)$ is source (asymptotically unstable).
- 3) If $1 - e - \rho < 0$ and $h + c - \rho\alpha > 0$, or $1 - e - \rho > 0$ and $h + c - \rho\alpha < 0$, $(0,0)$ is a saddle (asymptotically unstable).

Given Lemma 1, when the game is either an Active or a Passive Samaritan's Dilemma, $(0,0)$ is always a sink.

At $(1,1)$ we have that:

$$d\dot{x}/dx = e - \rho$$

$$d\dot{y}/dy = \alpha(1 - \rho) - h$$

Hence, we have three cases:

- 1) if $e - \rho < 0$ and $\alpha(1 - \rho) - h < 0$, (1,1) is a sink (asymptotically stable).
- 2) If $e - \rho > 0$ and $\alpha(1 - \rho) - h > 0$, (1,1) is source (asymptotically unstable).
- 3) If $e - \rho < 0$ and $\alpha(1 - \rho) - h > 0$ or $e - \rho > 0$ and $\alpha(1 - \rho) - h < 0$, (1,1) is a saddle (asymptotically unstable).

Given Lemma 1, when the game is an Active Samaritan's Dilemma, (1,1) is always a saddle; when the game is a Passive Samaritan's Dilemma, (1,1) is always a sink.

At (1,0) we have that

$$d\dot{x}/dx = \rho + e - 1$$

$$d\dot{y}/dy = h - (1 - \rho)\alpha$$

Hence, we have three cases:

- 1) if $\rho + e - 1 < 0$ and $h - (1 - \rho)\alpha < 0$, (1,0) is a sink (asymptotically stable).
- 2) If $\rho + e - 1 > 0$ and $h - (1 - \rho)\alpha > 0$, (1,0) is source (asymptotically unstable).
- 3) If $\rho + e - 1 < 0$ and $h - (1 - \rho)\alpha > 0$, or $\rho + e - 1 > 0$ and $h - (1 - \rho)\alpha < 0$, (0,1) is a saddle (asymptotically unstable).

Given Lemma 1, when the game is an Active Samaritan's Dilemma, (1,0) is always a saddle; when the game is a Passive Samaritan's Dilemma, (1,0) is always a source.

At (0,1) we have that

$$d\dot{x}/dx = \rho - e$$

$$d\dot{y}/dy = \rho\alpha - h - c$$

Hence, we have three cases:

- 1) if $\rho - e < 0$ and $\rho\alpha - h - c < 0$, (0,1) is a sink (asymptotically stable).
- 2) If $\rho - e > 0$ and $\rho\alpha - h - c > 0$, (0,1) is source (asymptotically unstable).
- 3) If $\rho - e < 0$ and $\rho\alpha - h - c > 0$, or $\rho - e > 0$ and $\rho\alpha - h - c < 0$, (0,1) is a saddle (asymptotically unstable).

Given Lemma 1, when the game is an Active Samaritan's Dilemma, (0,1) is always a saddle; when the game is a Passive Samaritan's Dilemma, (0,1) is always a source.

To prove the stability properties of the (x^*, y^*) equilibrium, we study its Trace and Determinant and see that:

$$\text{TR } J_{(x^*, y^*)} = 0 \text{ and } \text{Det } J_{(x^*, y^*)} = -x^*(1 - x^*)y(1 - y^*)[(2\rho - 1)\alpha - c](2\rho - 1)$$

From Lemma 1, we already know that $(2\rho - 1)\alpha - c > 0$ and $2\rho - 1 > 0$ are simultaneously satisfied when condition (3) is satisfied. In this case, $\text{Det } J_{(x^*, y^*)} < 0$, which proves that (x^*, y^*) is a saddle when the game is a Passive Samaritan's dilemma. In addition, there exists parametrizations for which $\text{Det } J_{(x^*, y^*)} \geq 0$, which completes the Proof of Proposition 1 ■

3. Proof of Proposition 2

From the results in Proposition 1, it is straightforward to derive the following set of results.

First, the existence conditions of the four monostable regimes are as follows:

1. The stationary point $(0,0)$ is the unique attractor iff $e > \max\{\rho, 1 - \rho\}$ and $\alpha > \max\left\{\frac{h}{1-\rho}, \frac{h+c}{e}\right\}$.
2. The stationary point $(1,1)$ is the unique attractor iff $e < \min\{\rho, 1 - \rho\}$ and $\alpha < \min\left\{\frac{h}{1-\rho}, \frac{h+c}{e}\right\}$.
3. The stationary point $(1,0)$ is the unique attractor iff $e < \min\{\rho, 1 - \rho\}$ and $\alpha > \max\left\{\frac{h}{1-\rho}, \frac{h+c}{e}\right\}$.
4. The stationary point $(0,1)$ is the unique attractor iff $e > \max\{\rho, 1 - \rho\}$ and $\alpha < \min\left\{\frac{h}{1-\rho}, \frac{h+c}{e}\right\}$.

Second, the existence conditions of the two bistable regimes are as follows:

1. The stationary points $(0,0)$ and $(1,1)$ simultaneously attract iff conditions (1) and (9) are simultaneously satisfied.
2. The stationary points $(1,0)$ and $(0,1)$ simultaneously attract iff conditions (7) and (8) are simultaneously satisfied.

Third, the existence conditions of the two cyclical regimes are as follows:

1. The system exhibits cyclical behavior with counterclockwise oscillations around (x^*, y^*) iff conditions (7) and (9) are simultaneously satisfied.
2. The system exhibits cyclical behavior with clockwise oscillations around (x^*, y^*) iff

conditions (1) and (8) are simultaneously satisfied.

Observe that when the game is an Active Samaritan's Dilemma, the monostable regime featuring (0,0) as the only attractor results; when the game is a Passive Samaritan's Dilemma, the bistable regime featuring (0,0) and (1,1) as attractors result ■