A comparison of simple action-based and outcome-based policies for emergency-like situations

Tim Friehe [a,b,c], Avraham Tabbach [d,*]

[a] *University of Marburg, Public Economics Group, Am Plan 2, 35037 Marburg, Germany*
[b] *CESifo, Munich, Germany*
[c] *EconomiX, Paris, France*
[d] *Tel Aviv University, Faculty of Law, Tel Aviv University, Tel Aviv 69978, Israel*

HIGHLIGHTS

- We analyze a generalized volunteer's dilemma situation.
- Comparing the efficiency of outcome and action-based social policies to mitigate it.
- Action-based policies enjoy a crucial advantage over outcome-based ones.
- Outcome-based policies always feature an equilibrium with no participation.
- Action-based policies exclude an equilibrium with no participation (a social trap).
- Findings consistent with two common features of the law in emergency like situations.

ARTICLE INFO

ABSTRACT

This paper explores a class of social dilemmas in which the participation of a given number of individuals is required to achieve a social end in the absence of a coordinating authority (e.g., rescuing a person in peril or preventing an imminent crime). After describing the first- and second-best outcomes, we examine whether simple policy instruments such as punishments and rewards can induce the second-best outcome, distinguishing between policies based on an individual's actions (i.e., action-based policies) and policies based on the outcome (i.e., outcome-based policies). For the domain of simple policies considered, we establish that action-based policies enjoy a crucial advantage over outcome-based ones: namely, outcome-based policies always feature an equilibrium with no participation, whereas action-based policies exclude this equilibrium.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Motivation and main results

The well-known volunteer's dilemma (Diekmann, 1985) captures an important social problem. The classical example involves bystanders who witness a person in an emergency situation, such as drowning. The person in peril will suffer severe harm unless one of the bystanders provides assistance.

In the canonical version of the volunteer's dilemma (VD), there is a group of n symmetric bystanders, each of whom can incur a cost to prevent a social harm and will prefer to do so if no other player helps. Because providing assistance is risky or costly, the bystanders may prefer not to act in the hope that a fellow bystander will step up. The game has n asymmetric pure-strategy Nash equilibria in which only one individual volunteers. All pure-strategy equilibria are first-best: harm is prevented at the lowest possible cost. However, these efficient equilibria require coordination. In its absence, the more likely outcome is a symmetric but socially inefficient mixed-strategy equilibrium, in which players randomize between volunteering and not volunteering. In this paper, we are particularly interested in circumstances that exclude meaningful coordination, as is the case in many emergency situations. The time pressure involved and the lack of any authority that could coordinate actions will usually preclude the attainment of the first-best outcome. In addition to the canonical version of the VD, some authors have analyzed a degenerate version of the game in which bystanders strictly prefer not to offer assistance (i.e., independent of what other bystanders are doing) when there is no policy intervention in place (see, for example, Leshem and

* Corresponding author.
E-mail addresses: tim.friehe@uni-marburg.de (T. Friehe),
adtabbac@post.tau.ac.il (A. Tabbach).

Tabbach (2016)). However, the use of policy instruments in the form of punishments or rewards may transform the incentive structure of the degenerate version of the game into its canonical version.

In this paper we analyze a generalized VD – a game with more than one volunteer required to prevent social harm – in which the individual costs of helping are a private-information random variable realized only after the incident has taken place. We thus depart from the classical VD in two important ways. The assumption that the required number of volunteers be greater than one is relevant in many circumstances. The participation of more than one individual is typically required, for example, to help a person trapped inside a crashed car or to resolve a hostage situation. The assumption that the costs of helping are a private-information random variable also seems very apt in practice. Consider the accident example, in which certain individuals may have particularly high costs of helping due to previous traumatic experiences or physical limitations.

In some real-world scenarios, the VD can shed light on bystanders' seeming indifference to the plights of strangers. Two prominent examples, among many others, are the murder of Kitty Genovese (a woman brutally stabbed to death in 1964) and the death of Wang Yue (a 2-year-old girl run over by two vehicles in 2011). In both cases, there was more than one person at the scene who could have helped to prevent the tragic outcome with little effort. Moreover, in both instances, there was no coordinating authority to orchestrate the actions of potential helpers.[1] However, the VD is by no means restricted to emergency-like situations; it is equally relevant in a more general criminal-law context. Consider, for instance, corporate wrongdoing in which multiple employees are privy to information on corporate fraud or corruption. Reporting the wrongdoing will expose the transgressors, but individual employees may have no incentive for whistle-blowing (the degenerate version of the VD) or may prefer that a fellow worker blow the whistle if reporting entails significant personal costs (the canonical version of the VD). As a result, potential whistle-blowers may fail to take action even when each of them individually wants the wrongdoing to be exposed.

From a social perspective a natural question is how the law can shape incentives in order to resolve the VD or mitigate its tragic consequences. In some circumstances, the responsibility to help can be allocated to specific bystanders according to, for example, pre-existing characteristics such as their relationship to the victim (e.g., family members). These targeted individuals should then be rewarded for acting or punished for refraining to act. For the VD in which only one volunteer is required to prevent social harm, the targeted volunteer will therefore have incentives to act in a socially desirable manner. Legislation in the United States follows this path by imposing a duty to protect against unreasonable risk of physical harm on any person who has a specific relationship to the victim, such as an innkeeper or custodian (Restatement (Second) of Torts Section 314A (1965)). Similarly, mandatory reporting laws in most states in the US and Australia break the symmetry among potential volunteers by requiring designated individuals – such as teachers, school administrators, and social workers – to report child abuse and neglect (Mathews and Kenny, 2008, p. 53).

In many instances, however, potential volunteers have no relevant distinguishable characteristics such as a relationship to the victim, a position of authority, relevant training, or knowledge.

Importantly, in scenarios in which the cost of helping is asymmetric, targeting individuals will not necessarily ensure low costs of helping. In addition, highlighting specific bystanders by reference to pre-existing characteristics is likely to be less promising in circumstances in which more than one volunteer is required. Moreover, in many jurisdictions, laws that discriminate among potential volunteers based on "irrelevant" characteristics (e.g., gender, race or ethnicity, age, etc.) run the risk of being declared unconstitutional. As a result, any policy designed to address the generalized volunteer's dilemma must subject all volunteers to the same incentive scheme.

In this paper we analyze how a social planner can structure the law to mitigate the tragic outcomes of the generalized VD in emergency-like situations. In particular, we focus on simple policies with minimal prerequisites, distinguishing between policies that are contingent on the outcome (whether or not harm occurred), which we label outcome-based policies, and policies that depend on the action (whether or not help was provided), which we label action-based policies. These policies are thus comparable to ex-post and ex-ante policies, as action-based policies take effect independently of whether or not an event occurs (and, in that sense, before the uncertainty regarding the outcome that will ultimately result is resolved). Shavell (1993) asserts that the time of intervention is one of three primary dimensions of legal methods of controlling behavior. In addition to the pure cases of intervention either after the action (possibly independent of the occurrence of harm) or after the occurrence of harm, he also highlights mixed scenarios. For example, the use of a weapon against another person – the action – may itself be punishable, but the level of the punishment can depend on the outcome (e.g., injury to or death of the victim). Similarly, tort law and regulations as classical outcome-based and action-based policies apply simultaneously in many circumstances (e.g., De Geest and Dari-Mattiacci, 2007).

For our setup, we start by describing the first-best outcome, in which the required number of players with the lowest realized participation costs act, provided that total costs are lower than the social harm incurred by the adverse event. However, as emphasized above, in practice, there is often no way to coordinate bystanders decisions after an emergency-like situation has occurred. Consequently, we concentrate on a symmetric second-best outcome in which all players are treated equally from an ex-ante point of view. In our setting, attaining the second-best outcome means inducing individuals to participate when their costs incurred by doing so fall short of an optimal cutoff value.

Turning to simple policies, we demonstrate that the second-best outcome can be achieved by resorting to an action-based policy, an outcome-based policy, or a combination thereof. With an outcome-based policy, the sum of reward and punishment that each player should ideally face is equal to the level of the implied externality (which in the degenerate version of the game is the level of social harm). In contrast, with an action-based policy, the sum of reward and punishment is substantially lower than social harm—in fact, in the degenerate version of the game, it should simply be equal to the optimal cutoff cost value, such that players with lower costs than the optimal cutoff level will act and those with higher costs will refrain from acting. More importantly, we find that action-based policies outperform outcome-based ones in one crucial fashion: namely, by ruling out the equilibrium in which all players refrain from helping. The reasoning is simple yet powerful. With outcome-based policies, if no other player renders assistance, there is no point in rendering assistance, as a solitary helper cannot alter the outcome of the game. In contrast, with action-based policies, if no other player renders help, a player may still have an incentive to help if, for example, helping implies avoiding punishment or receiving a reward. Moreover, in the degenerate version of the game, action-based policies can

---