NBER WORKING PAPER SERIES

THE IMPACT OF COUNTERFEIT VICTIMS IN THE VICTIM MARKETPLACE

Mingyuan Ban
Yi Qian
Qiang Gong
Karl Aquino

Working Paper 33327
http://www.nber.org/papers/w33327

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2025

The authors would like to thank Wenjia Ba, Tony Ke, Michelle Lu, Mahesh Nagarajan, Eddie Ning, Hao Zhang, and participants at the BIGs Conference and the VSE/Sauder IO Workshop for helpful comments. Karl and Yi would like to acknowledge funding support from SSHRC IDG 430-2023-01156. Yi would also like to acknowledge funding support from SSHRC IG 435-2018-0519 and 435-2023-0306. Qiang would like to acknowledge funding support from the National Social Science Fund of China (No.23AZD029) and National Natural Science Foundation of China (No.72073146). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Impact of Counterfeit Victims in the Victim Marketplace
Mingyuan Ban, Yi Qian, Qiang Gong, and Karl Aquino
NBER Working Paper No. 33327
January 2025
JEL No. C7

ABSTRACT

Just and efficient allocations of charity have attracted much academic and media attention. The sources of inefficiency and unjust are important to understand yet understudied. Our study aims to fill this void by directly modelling the victims' market in a collective reputation framework. By analyzing three types of individuals who signal their victim status with different personalities and incentives, we derive the honest, dishonest and unfunded equilibria as well as the mixed equilibrium where both types of these equilibria could coexist. Our analyses of the social welfare under each equilibrium shed light on key parameters that could potentially serve as policy instruments for improving social welfare. We also reveal that the mechanisms analogous to bank run and lemons market could take place in the victims' market as much as in other markets. In particular, when charity resources are scarce, more strategic signallers could rush to emit false victim signals and drive the market to the dishonest equilibrium with lower social welfare. The need for screening signallers could drive up the psychological costs of authentic victims to the extent that they voluntarily drop out of the market and suffer alone, resulting in misplaced charity funds and severe deadweight losses. When there is psychological utility associated with cheating for the hedonic signallers, the social welfare is even worse off.

Mingyuan Ban
Peking University
banmingyuan2008@126.com

Qiang Gong
Zhongnan University of Economics and Law
China
qgongpku@gmail.com

Yi Qian
Sauder School of Business
University of British Columbia
2053 Main Mall
Vancouver, BC V6T 1Z2
and NBER
yi.qian@sauder.ubc.ca

Karl Aquino
University of British Columbia
karl.aquino@sauder.ubc.ca

1. Introduction

Humans routinely emit distress signals communicating that they are experiencing pain, fear, anxiety, mistreatment, or any of the countless other unhappy states and conditions that blight their lives. When communicated to others, we consider these descriptions to be complex signals of distress that can exert a powerful and sometimes irresistible force upon the minds and hearts of observers to provision the signaler with material (e.g., money, food), socio-emotional (e.g., status, attention), or symbolic (e.g., legitimacy, recognition) resources. Determining whether a distress signal is an accurate indicator of a person's adverse condition is an important concern for recipients of the signal because we assume that most would prefer to avoid being exploited by someone who presents an exaggerated or false plea for help. Human beings possess many mental capacities that allow them to detect when they are being misinformed (Sperber, et al., 2010), but these capacities are not infallible. In this paper, we analyze by way of a theoretical model how the presence of people who emit false distress signals impact the victims and charity market and we unravel the subsequent social welfare consequences of this assessment.

There are good reasons to believe that inaccurate distress signals are emitted with some regularity in the general population because lying is among the most common of all social behaviors. On average, people lie at least twice a day (DePaulo, et al., 1996). Although many of these lies are so-called "white lies" that are inconsequential, people are willing to be dishonest about their victimization in ways that can be personally beneficial. A study by Ok and colleagues (Ok, et al., 2020) surveyed 1,509 Canadians who are recruited through a national consumer survey panel and found that 24% of them reported that at some point in their lives they had pretended to be hurt or harmed physically or psychologically to get something they wanted. Seventy-two percent said they personally knew someone who had done so. In the United States, false victim claims are estimated to cost insurance companies, governmental aid agencies, and charities billions of dollars per year (e.g., insurance fraud; Coalition Against Insurance Fraud, 2017). False claims of child abuse are common in child custody disputes (Avieli, 2022). In short, there is ample evidence that a fair number of people lie about their victimization.

A perennial challenge faced by all would-be benefactors[1] of those who claim to be in distress is that there will inevitably be actors who intentionally emit false signals so they can benefit from the misplaced generosity or concern of others. The emission of false signals can therefore have a negative effect on the recipient of the signal. However, this consideration, at the same time, reveals diverse motivations of individual donors (Macquarrie 1972, Exley 2020, Schmitz 2021, Chao and Fisher 2022, Kumar and Chakrabarti 2023) and doubtful calculations could compromise the sincerity and subsequent effectiveness of charity. We show that it can also impose psychological costs upon honest signalers, leading to a deterioration of social welfare.

Hereafter we will refer to actors who emit inaccurate distress signals as *false victim signalers (FVS)*. We use the term "victim" to highlight our interests in particular kind of signal emitted by people who want to portray themselves to an audience as fitting the definition of a victim as someone "who is acted on and usually adversely affected by a force or agent" and "subjected to oppression, hardship, or mistreatment." (Oxford English Dictionary). The agent alleged to have caused suffering or hardship can be human or non-human (e.g., an organization or the state) and the requests the victim signaler makes to their audience can vary. We also define *false victim signaler* (FVS) as one who purposefully generates indicators of distress that might also be observed in an authentic signaler such as shedding crocodile tears, altering the pitch of one's voice so they sound despondent or desperate, making facial contortions to convey sadness or fear[2]. These examples of counterfeit victims recognize that humans' capabilities for detecting duplicity are matched or exceeded by their ability to use dissimulation and artifice to achieve their goals. What matters for our purposes is that the FVS publicly communicates an aversive state or condition with the intention of influencing charitable organizations or individuals to provide them with a resource that will alleviate it. However, they know the information they are communicating is distorted, incomplete, exaggerated, or false.

[1] We use the term "benefactor" to refer to any human or non-human entity who can provide a person who emits a distress signal with resources that can benefit the signaler and potentially alleviate their distress.
[2] See for example the recent high-profile case of actor Jussie Smollet: https://www.cnn.com/2022/03/10/us/jussie-smollett-sentencing-trial/index.html

To answer the question of what the impact on social welfare might be if the number of false victim signalers in a population rises relative to the number of authentic victim signalers, we draw from the literature of counterfeiting and collective reputation as novel and useful micro-foundations to analyze the consequences of victim signaling. We extend models from this literature to show the social welfare consequences of false victim signaling or solicitation as it spreads throughout a population. When using the term "signal", "signaling", or "signaler" in our paper, we want to clarify that we adopt a definition proposed in social psychological research (Ok et. al. 2020) and that our theoretical framework is a collective-reputation model not one based on traditional signaling theory.

We develop a game-theory model for victims' marketplace dynamics to investigate how the presence of false victim signaler might influence potential benefactors' willingness to transfer resources to anyone who asks for them. Within the category of false victim signalers, we further distinguish between *strategic* and *hedonic* types. We contrast them with *authentic victim signalers (AVS)* defined as actors who publicly communicate an aversive state or condition hoping to persuade others to help alleviate it, but who do so without any intention to distort, embellish, or fabricate facts to create an inaccurate impression to an audience. In other words, they try to the best of their abilities to avoid misinforming would-be benefactors into providing them with a desired resource. In contrast, we propose that *hedonic signalers* always emit false victim signals because they find it personally satisfying or rewarding to manipulate others into giving them what they want. *Strategic signalers* sometimes falsely signal based on the cost-benefit tradeoffs.

One way a potential benefactor can defend themselves from being duped by a false distress signal is to rely on signaler's reputation as an input for judging its accuracy. However, when individual signaler's reputation cannot be accurately observed, then reputation of the whole group of signalers is relied upon to infer individual reputation. The benefactor will be more skeptical if the signaler comes from a group that has a reputation for making dishonest or exaggerated claims. The colloquial expression describing this kind of signaler is that they "cry wolf". We utilize collective reputation as a theoretical tool to sketch the negative externalities of FVS. Treating the victim as a brand identity that people are expected to respond to with sympathy, compassion, or concern makes it amenable to a theoretical analysis based on collective

reputation. Our model characterizes the collective reputation of authenticity and honesty of a group (i.e., victims) that can benefit those who are categorized as being part of this group. In this way, false victim signalers attempt to free ride off of the victim brand. We show that FVS not only corrupts an individual's reputation, but may trigger ripple effects that strategic signalers imitate and that eventually contaminate the collective reputation of all people who signal victimhood. A collective reputation for duplicity is both a cause and consequence of individual acts of duplicity and can cause the collapse of the victims' marketplace.

We innovatively link the false victim signal to the introduction of a "counterfeit persona" into a marketplace of other victim signalers. The idea that there is a marketplace of victim signalers is a fundamental premise of research on competitive victimhood (e.g., Young & Sullivan, 2016; Sullivan, et al., 2012), an interdisciplinary area of study that investigates how groups compete for attention and resources based on which one has been subjected to greater hardship, oppression, or disadvantage. We extend competitive victimhood research by modeling victim competition at the individual rather than group level and introducing the concept of counterfeit personas into our analysis. We treat the counterfeit persona like a counterfeit product by assuming that the counterfeit persona presents the appearance of being a genuine victim without actually being one. We assume that counterfeit personas compete with authentic personas of people whose presentation of their victimhood is not intended to deceive an audience. The potential audience for victim signalers can be a charitable organization, government agency, an individual, or any entity that has resources the victim signaler might want to obtain. These targets are analogous to consumers in a market who can be persuaded to exchange resources with a seller (i.e., a victim signaler) in exchange for some benefit.

A critical difference between a consumer who responds to the plea of counterfeit victims by transferring resources to them and the consumer who buys a counterfeit product is that the latter often knows that the counterfeit product is not genuine. However, we presume that the would-be benefactor of counterfeit victims believes they are helping someone who is acting in good faith by accurately depicting their suffering and the circumstances that produced it. We do not expect most people to help a victim who they knew was feigning their victimhood. Distinguishing false from authentic victim signalers is difficult

in practice. The difficulty is particularly acute if signalers make claims about subjective states like their mental well-being, emotions, or physical sensations that are inaccessible to observers. A case of counterfeiting that is a closer analogy to the situation to which our model applies would be if the counterfeiter attempts to fool a consumer into buying a fake product.

Our model is perhaps the first to disentangle how counterfeit victims give rise to social costs from the perspectives of collective reputation. Our analysis points out three sources of social costs. First, from the perspective of authentic victim signalers, counterfeit victim signalers of both the hedonic and strategic types not only obtain resources that should be provided to those in genuine need, but also may aggravate the psychological burden of worrying that a person who victim signals might be misjudged as counterfeit victims. The latter may discourage the authentic victims to stay in marketplace while deteriorating the collective reputation of all victim signalers. We call these two effects the *crowd-out effect* and *lemon market effect,* respectively. From the perspective of strategic victim signalers, the observed success of hedonic victim signalers at obtaining what they want may tempt them to also emit false victim signals to extract benefits from the environment thus intensifying the competition for funding. We term this effect as *run effect*. Last but not the least, from the perspective of potential benefactors, counterfeit victims shrink incentives to input resources for fear that they will be misdirected thereby reducing the overall pool available for public assistance. A potential benefactor who is concerned about social welfare may not be willing to contribute resources to the overall pool if they anticipate a dishonest equilibrium incurred by the prevalence of counterfeit victims.

Our research provides valuable managerial insights for charity and resource allocators in a broader sense. The natural human tendency to provide for victims might be exploited by hedonic and utilitarian individuals. The presence of such individuals undermines the collective reputation of victims, leading to trust crises within organizations. Our research emphasizes the importance of virtue, social norms, and resource planning in organizational governance. Scarce resources may intensify the competition for support and thus deteriorate collective reputation and overall welfare. The good news is that the improvement of tracking technology in the current digitalization wave (e.g., the application of blockchain) may potentially

help to develop an efficient resource allocation within organization. At the same time, we emphasize the harm when organizations impose excessive screening on the victims, as this may cause those who truly need help to bear an overwhelming psychological burden, making them hesitate and even ashamed to speak about their distresses.

2. Literature

2.1 What Are the Characteristics of False Victim Signalers?

We treat FVS as synonymous with other terms used by economists and psychologists to describe people who intend to mislead others through lying, deception, or cheating. The traditional economic view on deceptive behavior is that people trade-off maintaining personal integrity for wealth-maximizing purpose (Becker, 1968). Economic explanations have highlighted how incentives can encourage deception while the potential costs of being caught discourage it (Abeler et al., 2014). On the seller's side, the incentive to deceive a consumer that a counterfeit is authentic is charging a higher price than they otherwise would for a known counterfeit product (Qian, 2014). Attempting to deceive a consumer can have uncertain consequences for the counterfeiter because there is always a risk of being caught and not obtaining the expected benefits (Celse et al., 2019). This may deter some counterfeiters from being deceptive, yet others may believe the potential benefits in a given environment are greater than the risk of being detected (Qian, 2014). These counterfeiters are likely to use deception strategically depending on external conditions.

Individual differences among victim signalers are accounted for in our modeling by the *authentic, strategic,* and *hedonic* victim-signaling types. Supporting our view that different character types will vary in their willingness to emit false victim signals, decades of research on lying and deception have shown that individual differences in personality, dispositional tendencies, goals, or values predict dishonesty. For example, altruistic individuals (Alan et al., 2020) and those who are highly concerned about inequality (Amato et al., 2020) cheat and deceive less. In contrast, opportunistic individuals (Pate, 2018) and individuals with Dark Tetrad personalities (Pfattheicher et al., 2019) are more likely to deceive. The counterfeit self, implicit self-theory, ethical mindset, and moral emotion are among the pivotal factors that can lead to corrupt behavior. Abraham et. al. (2022) describe contexts that consist of groups of people with

particular (un)ethical mindsets, a proneness to experience certain moral emotions, and to hold certain theories about the self. Rosenbaum et al. (2014) reviewed the economic and psychological experiments on dishonesty and find that in addition to strategic cheaters, a subset of *unconditional cheaters* regularly emerges in market contexts. These types of cheaters will not be as concerned about monitoring or the intrinsic psychological costs of deception like feeling guilty and may even derive psychological utilities from them. These observations provide a micro-foundation for the rationales encapsulated by our theoretical framework.

We distinguish three social types of victims by their willingness to emit false victim signals to achieve this goal and the conditions under which they are likely to be deceptive. These could have implications for workplace victimization too (Aquino and Bradfield, 2000). Recall that an AVS accurately portrays their state of deprivation, need, or suffering. In contrast, the strategic signaler may emit a victim signal even when they are not currently in a state of need, or they may exaggerate the severity of their situation and distort the events that led to it, if they perceive that the environment provides opportunities to obtain resources they desire by portraying themselves as a victim. To the strategic signaler, the instrumental calculation of the benefits and costs of victim signaling will be relatively more important for guiding their behavior than adhering to a personal moral code. The strategic victim signaler is analogous to someone with a Machiavellian personality, a trait that reflects an amoral and primarily instrumental orientation that has been shown to be a reliable predictor of lying in many situations (Christie and Geis, 1970).

The hedonic signaler is similar to the strategic signaler because they also assess environment to determine whether victim signaling can be an effective means for getting what they want. They differ from the authentic signaler, and perhaps even some strategic signalers, because they are completely uninhibited by conscience, the preservation of their integrity, or a commitment to a moral code. What distinguishes the hedonic signaler from the strategic signaler is that the former derives hedonic satisfaction from being able to dupe or manipulate an audience of potential benefactors and *not solely* from getting what they want through dishonest means. The hedonic signaler is what in everyday language would be called a pathological liar who enjoys lying for its own sake and not only for its instrumental benefits. The clinical psychology

literature has long recognized this type of person as someone with a psychopathic or sociopathic personality (Hare, 2006; Glenn & Raine, 2014). For psychopaths, false victim signaling is a way of life and one indicator of a general antisocial orientation. Their ability to emit false victim signals with impunity is facilitated by the absence of guilt or remorse that most people feel when violating social norms. This character trait gives the hedonic signaler an advantage over other kinds of signalers because it is easier for them to construct elaborate and believable, but ultimately untruthful, stories about their victimhood.

Having described the character types of three victim signalers, we now turn to how the literature on counterfeit products can be applied to explain how these types can influence would-be benefactors, the victim market, and social welfare.

2.2 Counterfeit Products and Counterfeit People

The extant marketing literature on non-deceptive counterfeiting is a starting point for our theory. Counterfeiters usually produce products inferior to the authentic goods but that resemble them enough to deceive an undiscriminating consumer (Grossman and Shapiro, 1988). Most of the findings in this literature identify status signaling or emotional benefits as motivating the consumption of non-deceptive counterfeits like those sold in shoddy or concealed retail environments and that openly acknowledge their inauthenticity (Kalyoncuoglu et al., 2017; Yoganarasimhan, 2012; Kuksov and Wang, 2013; Wilcox et al., 2009; Mbura, et al., 2020; Fink et. al. 2016 provides a detailed review). Buyers of counterfeits attempt to elevate their status by displaying a counterfeit brand that people perceive to be genuine (Qian, 2014). However, we depart from the assumptions of non-deceptive counterfeiting research by modeling a market where the buyers, represented in our model by charitable organizations, are motivated to avoid allocating resources to FVS because they want to be responsible stewards of their resources and deliver them only to those who are in genuine need. In other words, they do not want to be deceived by counterfeit victims.

A deceptive counterfeiter claims trademarks they do not own to pass off their product as legitimate. Rational consumers may be aware that such counterfeits are present in a market, though they are not always able to distinguish well-made counterfeits from the authentic product. Similarly, we assume most charities recognize that there will be some proportion of people in the victim marketplace who want to receive

assistance without actually being in the state of need that they allege. Analogous to products of different quality tiers, we treat victim signalers (people who apply for charity support) to have different *qualities of authenticity*. The highest tier of quality is the AVS. Their presence in the victim marketplace is analogous to having genuine products in markets. An FVS has a lower quality of authenticity and their presence in the victim marketplace resembles the introduction of counterfeits into a product market. One consequence of lower priced counterfeit products in the marketplace is that consumers may be less willing to pay the full price of a high-quality good leading producers of high-quality goods to exit the market (Akerlof, 1970). We model similar effects in the victim marketplace where the presence of FVS could crowd out some AVS.

There is an extant literature that links consumer personality, attitudes, and purchases of counterfeits (Swami et al., 2009; Hidayat et al., 2013; Harun et al., 2012; Liang et al., 2012; Mohammad et al., 2020). We integrate this literature with the growing body of work on moral psychology to understand the complex dynamics of psychology that can lead to corruption and other counterfeit behaviors. Corruption is an act of abusing power for personal gain (Zaloznaya, 2014). While most studies treat it as a phenomenon or a product, Abraham et. al. (2022) is among the few to discuss the psychological evolvement that leads to it. Moore (2009) hypothesized that everyday unethical behaviors by individuals can eventually lead to corruption in wider society. Similar moral slippery slope and degradation are reported in Gino et al. (2010), who show that using counterfeit products (e.g., wearing counterfeit sunglasses) could dull moral senses and broaden tolerance for unethical behaviors, giving rise to more cheating. The influence and reinforcement of counterfeiting attitudes and behaviors have resemblance to the dishonest equilibria we derive in our model settings for counterfeit victims' market.

2.3 What are the Social Consequences of Deceptive Counterfeiting

Darby and Karni (1973) were the earliest to model the deceptive nature of counterfeiting and its impact on markets by describing how counterfeit producers can free ride on a brand's trademark and reputation. Analogously, we propose that FVS could free-ride on the reputation of AVS. Qian (2008, 2015) studied the marketing impacts of entry by counterfeiters and the heterogeneous sales impacts in the Chinese footwear industry. For authentic producers to counteract the free-riding of counterfeiters, they need to differentiate

their products from the counterfeits through quality upgrades, price and non-price signals, and other self-enforcement strategies. These results are corroborated in later studies (Gao et al., 2017, 2018; Grolleau et. al. 2023). The heterogeneous impacts identified by Qian (2015) further explain authentic firms' incentives for moving up the quality ladder, where they might leverage the advertising effects of counterfeits while escaping the business-stealing effects, as Lu and Bendle (2020) also documents for piracy. Similar substitution effects might occur in the victim marketplace when different types of victim signalers vie for charity resources.

Qian et. al. (2015) show theoretically that searchable quality can serve as a signal for the degree of experiential quality. Many of these theoretical predictions and findings have been corroborated by other work (e.g., Berger et al., 2012; Li and Yi, 2017). In this paper, we analyze a victim market where victim signals are driven by a combination of the social (i.e., market) context and the victim signaler's character type, the latter mapping directly onto qualities of authenticity. Honesty is then a differentiating attribute analogous to the experiential quality of a product. The accuracy of past victim signaling behaviors provides information to an audience about quality of signalers' collective integrity or reputation.

Our model accounts for how the behavior of other victim signalers can influence whether individuals emit false victim signals. Villeval (2024) provides a comprehensive review of experimental studies on the social determinants of unethical behavior. Consider corruption, a dishonest behavior frequently discussed in the economics literature as a collective (i.e., group level) activity. Dong et al. (2012) showed that a person's willingness to engage in corruption is influenced by the perceived activities of peers and other individuals who are also corrupt. They call this corruption contagion. In addition to corruption, other structural conditions can motivate dishonesty in the domain of victim signaling. For example, when resources are allocated unequally, the relatively disadvantaged often act more dishonestly (Birkelund and Cherry, 2020). One explanation for their behavior is competition. As competitors cheat to acquire advantage, individuals would also cheat in order to "level the playing field" (Atanasov and Dana, 2011). These conditions are among those that can produce the competitive victimhood dynamics. We model such intra- and inter-group dynamics, especially with respect to the three groups of signalers we mentioned, and solve

conditions for game-theoretic equilibria where we demonstrate the intricacies of these mechanisms and the resulting welfares.

Finally, our model considers an attribute of the victim signaler that could be used by potential benefactors (e.g. charity organizations) to avoid being misled: the signaler's reputation for (dis)honesty. We use personal reputation as a proxy for the accuracy of victim signal, and we consider the reputation of a group of signalers as the sum of the individual reputations. Our theoretical framework for incorporating a victim signaler's reputation draws from the collective reputation literature and especially Tirole's (1996) pioneering model. A typical application of collective reputation is on collective brands in marketing. Collective brands benefit firms by decreasing the cost of maintaining customers' trust (Saak, 2012) and enlarging market share to provide customer with better access to firms (Fishman et al., 2018). However, a collective brand can also generate incentives to free-ride on the other firm's investment (Neeman et al., 2019) by providing lower quality that is sub-optimal for the group (Winfree and McCluskey, 2005). This is analogous to individuals falsely signaling as victims in our model, which can tarnish the reputation of the victims' group and social welfare. By introducing individual personality as a potential explanation for how collective reputation can be compromised, our theory contributes to the emerging literature of economic modeling of personality traits and virtues (Heckman et. al. 2023) and illustrates the effects of individual virtues or vices.

3. Model Setup

We consider a victims' marketplace in an infinite-period game ($t = 0,1,2,\dots\infty$). The market is composed of two sets of players. One is a set of homogenous charitable organizations, and the other is a continuum of individuals (w.l.o.g., we use she and her as pronouns throughout) with a measure of one. All individuals and the charity are risk-neutral, randomly distributed in the market. In each period, a continuum of individuals is matched with a new charity, i.e., no chance to meet the same charity twice. As such, individual reputation would affect individual incentives only through feeding into the collective reputation of the respective group. This is important to consider in the real world, as many fraudulent behaviors occur in society with the fraudsters plan on one-time interaction. A close analysis of the interplay between

individual and collective reputations and the subsequent impacts on the market could help charity and other social planners design effective mechanisms in resource allocations. Informed education of this issue could potentially improve individual's willingness to maintain collective reputation and the wellbeing of a community.

In each period, there is a proportion of λ individuals remaining in the market for the next period, and a proportion of $1-\lambda$ individuals exiting the market, as will be offset by the same number of newcomers in the next period. The relevant discount factor is δ_0, and the effective discount factor becomes $\delta \equiv \delta_0\lambda < 1$. There is a probability of τ that an individual is assigned in the victim state, suffering hardship, and a probability of $1-\tau$ that an individual is assigned in the normal state, delivering normal output from production. An individual who is in the victim state applies for 1 unit of charity fund to maintain a living. Every charity is endowed with a total amount of funds denoted as *m* for each period, which is exogenously collected from donors. If there is any surplus of funds, the charity uses for other charitable purposes, and it will not be retained to the next period.

Individual types: Individuals could be classified into three types according to heterogeneous personality. We refer to these types as *honest*, *strategic*, and *hedonic individuals.* The presence of these three types of individuals have been previously discussed in a review of experimental literature in Rosenbaum et. al. (2014). We highlight the use of signalers here does not imply a signaling game, but for consistent denotation of these individuals in prior psychology literature (Ok et. al. 2020). We separate the nature of individual types (being) and their actions (doing) in our model as a truthful depiction of the world. All types of individuals apply for charity funding when they are in the victim state, yet they differ in whether they would feign victimhood to deceive funding in the normal state.

The first type of honest individual, with a proportion of α, are people who will never fabricate victimhood; that is, they always accurately portray their needs and the conditions that led to them to the best of their knowledge. The second type is a strategic individual, with a proportion of β, who will rationally compare costs and benefits in order to strategically to be authentic or emit false victimhood solicitations.

The third type is a hedonic individual, with a proportion of γ, who expects to win sympathy through false solicitation and therefore will readily apply for charity funding. These three types of individuals with different personalities would have different incentives to behave.

In the baseline model here, we consider the case where all individuals will request help in times of genuine victimhood. We consider complications where the authentic victims may shy away from that in Section 6. Both the hedonic and strategic individuals could be counterfeit victims. We model their strategic decisions and analyze the unraveling of the market equilibria and the subsequent social welfare. In some ways, our model elements of psychological utility and costs depict the frictions when individuals' doings (actions/strategies) do not match their being. We further consider psychological utilities of the hedonic individuals in the Appendix, which would yield more complex equilibria outcomes.

Action: The action for individuals and charity are chosen in each time period (instead of a one-shot decision). In each period, each individual has two actions to choose from: Authentic supplication (AS) or False solicitation (FS). A strategy describes a contingent plan of action under different states. If an individual chooses authentic supplication, she only applies for funding in the victim state; and if an individual chooses false solicitation, she applies for funding in both victim and normal states.

In Ok et al. (2020), victim signaling (either authentic supplication or false solicitation) takes the forms of victimhood expressions and communications directed to an audience. Based on the characterization of individual's personality, honest and hedonic individuals always (unconditionally) choose AS and FS, respectively. The strategic individuals rationally choose one to maximize their long-term expected utility. These preferences are consistent with Tirole (1996).[3] Unlike the traditional collective reputation or adverse selection settings, where individuals signal ability through efforts or performance, the individuals in victims

[3] Tirole (1996) mentioned that the presences of unconditionally honest and dishonest agents are standard in reputation models. Such assumption could help refine spurious multiplicity of equilibria. Apart from technical convenience, Rosenbaum et. al. (2014) also provide evidence for the persistent existence of unconditional cheaters and honest individuals in the real world. In the victim marketplace, there are many scams by counterfeit victims that motivate this assumption (e.g., Gofundme scams, etc.).

market signal their inability. This is more analogous to the counterfeit products market where the counterfeiters opt for low-cost strategies of free-riding.

In each period, the matched charity decides whether to enter the victim market or not. If not enter, no individual obtains funding in this period. Upon entry, the charity screens on the funding applicants and distributes funding among approved applicant. The approval rule will be introduced after the assumption on information.

Payoff: Individual's expected payoff in each period depends on the state and whether she obtains funding from the charity. An individual in the normal state gains from production with payoff $r \in (0,1)$, while in the victim state suffers from no production with payoff $-R < 0$. If an individual in victim state (authentic victim) obtains one unit of funds from charity, her payoff is recovered to 0. If an individual in normal state but deceive funds by false solicitation (counterfeit victim), she yields 1 unit of payoff and no longer has any incentive to produce. We suppose $R > 1$. This implies that funding an authentic victim is more socially beneficial compared to funding a counterfeit one, as the gain from avoiding suffering loss exceeds one unit of resource input. Charity acts like a social planner, whose payoff is the social welfare in the victim market, that is, the aggregation of individuals' utility deducted by the supply of funding. In the baseline model, individual's payoff is independent of type. We further incorporate psychological utility and costs in model extensions, to sketch how personality-based utility affects actions and yields more complex equilibria outcomes.

	If not being funded	If being funded
Normal state	r	1
Victim state	$-R$	0

Table 1. Individual's payoff

Information: We use the extensive form of the game (Figure 1) to show how events and information flow in a single period. In Step 0, a charity matches with individuals. The type of personality is privately observed by an individual. The charity cannot identify the exact type of each individual, yet knows the distributions of individual types. In Step 1, the charity decides whether to enter the victim marketplace. In

Step 2, each individual chooses a strategy simultaneously. After strategies are chosen by individuals, and before the timeline evolves into Step 3, nature randomly assigns a state for each individual. After observing the realized state, each individual executes the strategy she chose.[4]

In Step 3, the charity conducts screening over victims. Through the screening process, the charity observes an imperfect record $\xi_t \in \{G, B\}$ (good, bad) of each victim. The imperfect record implies that the charity detects the past deception (if any) with probability x.[5] We define deception as any incidence of successfully being funded through false solicitation. Should the individual false solicitated yet being unfunded, then it will not be considered as a deception.[6] Based on this imperfect record, the charity decides whether to reject or approve each victim.

In Step 4, the charity distributes one unit of fund to each victim who has been approved in Step 3. The funding service is on a first-come-first-serve basis (sequential-service), until the funds have been exhausted. Note that there is a possibility that some victims are approved yet remain unfunded due to the scarcity.

Discussion on assumptions: The assumption that individual's past deception is imperfectly observed is important to motivate the role of collective reputation (Tirole, 1996; Levine, 2009). If past deception is fully unobserved, the strategic individual would have no incentive to sustain their own reputation. And if past deception is fully observed, the collective reputation would play no role because individual reputation is identifiable. The heterogenous individual types and imperfect record for individual's past behavior are both necessary conditions to drive collective reputation in group. It is also worth emphasizing that the

4 Note that the sequence in which nature chooses the state and individuals choose their strategies does not affect the solution. This is because the difference between the AS and FS strategies only arises in the normal state, as all individuals inherently adopt AS strategy in the victim state. If nature determines the state before individuals choose their strategies, the individual's value function for long-run payoffs could become contingent on the state, unnecessarily complicating the model's solution. This inconvenience can be easily resolved by assuming that individuals choose a plan of actions contingent on anticipated states. In this case, the individual's value function is based on the expected payoff (accounting for state uncertainty), which does not alter the conditions under which they choose the AS or FS strategy.

[5]A more general assumption is that the charity detects the past deception with probability $x_k$if the victim has k times of deceptions, and $x_{k+2} - x_{k+1} < x_{k+1} - x_k, \forall k \in N^+$ (Tirole, 1996).

[6] In practice, charity organization's screening could have implications on the cause and we model potential psychological costs of authentic victims in Section 6. We thank participants of the BIGs conference for discussions on this point.

charity's discovery of deception each period is time-independent, because the match between individuals and charity is non-repeated.

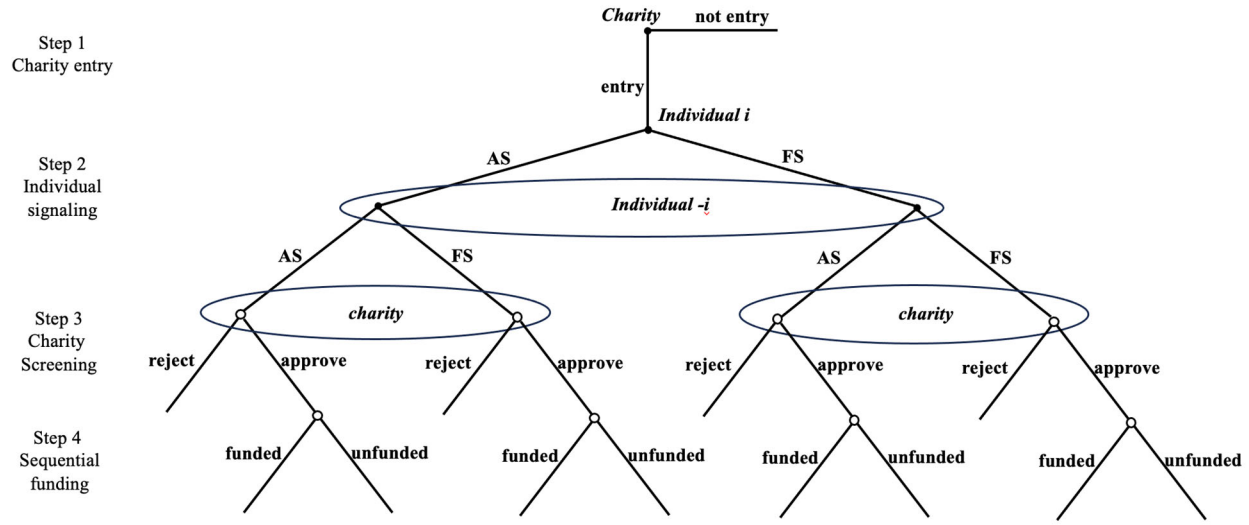

Figure 1. Extensive form of game in a single period

Note: "Individual" in Figure 1 refers to strategic individuals. Considering that extensive forms are commonly used to depict the game with discrete decision nodes, we illustratively simplify the simultaneous-move of a continuum of individuals into a simultaneous-move of individual *i* and individual -*i*.

The difference between our model and existing collective reputation literature is as follows: We extend the matching between one principal and one agent to the matching between one principal (charity) and multiple agents (victims) in each period. This makes a substantial difference in Step 4, because the approved agents are competing for the opportunities of being funded, and each victim's payoff is affected by other individuals' strategies. Such interdependent payoffs among agents do not exist in a one-to-one matching game, since the agents make decisions independently. The reason we adopt the one-to-many matching setup is that it is more relevant to the victim marketplace in practice. More importantly, this framework helps us analyze the social costs brought about by counterfeit victims, because funding a counterfeit victim comes at the expense of depriving another authentic victim of the opportunity to be funded.

4. Individual and collective reputation

In this section, we show how individual reputation and collective reputation are derived and how they evolve in a dynamic setting.

4.1 Reputation system

History: We define history $H_{i,t} = \{a_{i,1}, a_{i,2}, \dots, a_{i,t}\}$ as the set of individual i's past actions $a_{i,t}$ up to period t. According to whether the individual chose FS in the past, the history are categorized into two cases $H_{i,t} \in \{H^{AS}, H^{FS}\}$, where the authentic history $H^{AS} = \{a_{i,\tau} = AS\ for\ all\ \tau \leq t\}$ denotes the case when the individual has never acted false solicitation so far; and the counterfeit history $H^{FS} = \{a_{i,\tau} = FS\ for\ any\ \tau \leq t\}$ denotes the case when the individual has at least once acted false solicitation so far.

Individual reputation: Individual reputation is shaped by individual's history. We introduce $\phi_{i,t} \in \{\phi_A, \phi_C\}$ to represent individual i's reputation in period t, indicating whether the individual has ever successfully deceived funding up to period t. The authentic reputation ϕ_A denotes never successfully deceived funding, while the counterfeit reputation ϕ_C denotes at least once successfully deceived funding.

It is important to point out that false solicitation does not necessarily lead to counterfeit reputation. Because false solicitation may not ultimately "succeed" in deceiving funding. The "unsuccessful" case may happen when victims (both authentic or counterfeit) are too much or the funding resources are too little, false solicitation ends up with not being funded. This definition of individual reputation is reasonable in the victim marketplace, as we have stated that counterfeit behavior is too difficult to distinguish and instead, the funding outcome is more observable and reliable to develop the individual reputation. Sometimes the temptation for false solicitation arises because doing so is 'almost lossless' for the individual, as not being funded means no risk of downturn in reputation. Our analysis will demonstrate that, strategic individuals may become more opportunistic under this setting of individual reputation.

It is straightforward that an individual who never acted false solicitation in the past keeps an authentic reputation; i.e., $prob\{\phi_t = \phi_A | H_t = H^{AS}\} = 1$. However, an individual who has ever acted false solicitation in the past may probabilistically lead to counterfeit reputation ϕ_C or authentic reputation ϕ_A. θ_t is defined as the probability of owning counterfeit reputation for a victim with counterfeit history H_t^{FS}. Later on, we will derive the steady-state closed-form solution of θ_t. Without loss of generality, individual reputation is updated at the end of each period, i.e., after payoff is realized.

$$\theta_t \equiv prob\{\phi_t = \phi_C | H_t = H^{AS}\}$$

Tracking record: Given each individual's reputation ϕ_{t-1}latest updated at period $t-1$, the tracking record ξ_t is generated for each individual and observed by charity at period t.

$$prob\{\xi_t = B | \phi_{t-1} = \phi_A\} = 0$$

$$prob\{\xi_t = B | \phi_{t-1} = \phi_C\} = x$$

A bad record always suggests the individual is counterfeit, whereas a good record does not necessarily suggest the individual is authentic. The tractability parameter x plays a crucial role in the model. A higher tractability means individual reputation is more observable, and collective reputation is leveraged less in the victim marketplace.

By the tracking record, a counterfeit victim (who owns counterfeit history H^{FS}) is approved with probability $Y(\theta_{t-1})$. This could happen either when the individual owns counterfeit reputation yet being un-detected with probability $(1-x)\theta_{t-1}$, or when the individual remain keeps an authentic reputation with probability $1-\theta_{t-1}$.[7]

$$Y(\theta_{t-1}) = prob\{\xi_t = G | H_{t-1} = H^{FS}\} = \underbrace{(1-x)\theta_{t-1}}_{\substack{counterfeit\ reputation \\ yet\ un-detected}} + \underbrace{(1-\theta_{t-1})}_{\substack{authentic \\ reputation}}$$

Collective reputation: The tracking record updates charity's belief on the authenticity among *approved* victims, which is the measure of collective reputation in the victim marketplace. By Bayes rule, the collective reputation is given by

$$\Phi_t \equiv prob\{H_{t-1} = H^{AS} | \xi_t = G\} = \frac{\mu(AS)\tau}{\mu(AS)\tau + \big(1-\mu(AS)\big)Y(\theta_{t-1})}$$

To illustrate, $\mu(AS)\tau$ is the proportion of authentic victim (who owns authentic history H^{AS}), $\big(1-\mu(AS)\big)Y(\theta_{t-1})$ is the proportion of counterfeit victim (who owns counterfeit history H^{FS}).

Individual reputation contributes to collective reputation. This is shown as collective reputation Φ_t is one-to-one correlated with θ_{t-1}, the distribution of individual reputation among counterfeit victims. Since

[7] Note that our derivation of Y_t is technically equivalent to Tirole (1996)'s concept of the average probability that past corruption activities go unnoticed.

counterfeit victims and authentic victims are mixed together in marketplace, the charity uses collective reputation to infer the authenticity of each victim, and this belief is shaped by all victims' past behaviors. Considering the monotone relationship between Φ_t and θ_{t-1}, we use θ_{t-1} as the proxy for collective reputation in the following context.

4.2 Dynamics and steady state of reputation system

The reputation system evolves over time in the victim marketplace. On the one hand, the dynamics is driven by the stochastic entry and exit. In each period, some new victims enter market and replenish the proportion of authentic victims, while some old victims exit the market. On the other hand, the dynamics is driven by the victim's behaviors. We define the switching function $\phi_t = g(\phi_{t-1}, s_t)$ to illustrate the dynamics of individual reputation. Contingent upon the individual not quitting in market from period $t-1$ to t, the switching process for the individual reputation is governed by:

$$g(\phi_{t-1} = \phi_A, a_t = AS) = \phi_A$$
$$g(\phi_{t-1} = \phi_C, a_t = AS) = \phi_C$$
$$g(\phi_{t-1} = \phi_C, a_t = FS) = \phi_C$$
$$g(\phi_{t-1} = \phi_A, a_t = FS) = \begin{cases} \phi_C, & prob = (1-\tau)p(\theta_{t-1}) \\ \phi_A, & prob = 1-(1-\tau)p(\theta_{t-1}) \end{cases}$$

It is obvious that authentic supplication does not switch individual reputation. The reputation switching happens only when an individual maintains authentic reputation so far, while acts false solicitation and successfully deceives funding in the current period. The probability of being funded among approved victim is denoted as $p(\theta_{t-1})$, which is endogenously determined by the reputation system. We will soon revisit the derivation of $p(\theta_{t-1})$. Figure 2 displays the dynamics of individual reputation for counterfeit victims.

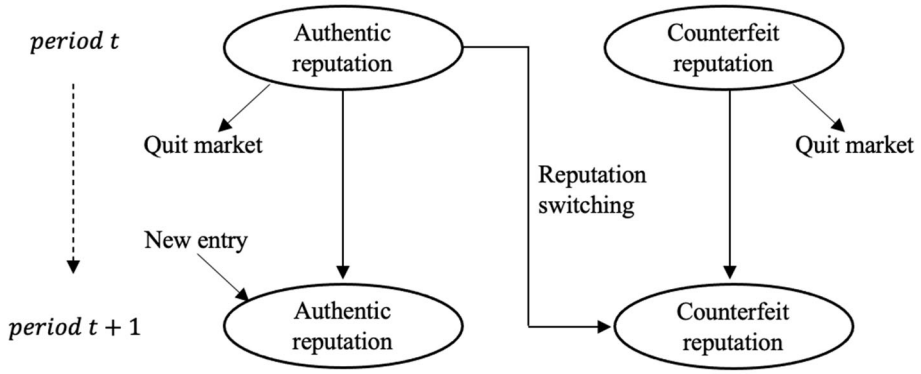

Figure 2. Dynamics of individual reputation for counterfeit victim

The distribution of individual reputation for counterfeit victim evolves into the steady state when $\theta_t = \theta_{t-1}$. The steady-state condition is sketched in Equations (1) and (2), which are mathematically equivalent.

$$\underbrace{1-\lambda}_{newcomers} + \underbrace{(1-\theta_{t-1})\lambda - (1-\theta_{t-1})(1-\tau)p(\theta_{t-1})\lambda}_{status-unswitched\ survivors} = 1-\theta_t \quad (1)$$

$$\underbrace{\theta_{t-1}\lambda}_{counterfeit\ survivors} + \underbrace{(1-\theta_{t-1})(1-\tau)p(\theta_{t-1})\lambda}_{status-switching\ survivors} = \theta_t \quad (2)$$

The economic implications of Equations (1) and (2) are as follows. Equation (1) is from the perspective of individuals with authentic reputation; the steady state comes when the counterfeit-victim quitters are offset by the new entrants. Equation (2) is from the perspective of individuals with counterfeit reputation; the steady state comes when the new status-switching survivors are replenished by the new entrants. Without loss of generality, the following analysis on equilibrium revolves around the steady state of individual reputation, and thus the time subscript t is omitted.

We now turn to the derivation of the probability of being funded $p(\theta)$. After the charity filters some counterfeit victim by the tracking record, the funding is randomly attributed to the approved victim. When the funding resources is plenty enough, each approved victim is funded.

$$p(\theta) = min\left\{\frac{m}{\mu(AS)\tau + (1-\mu(AS))Y(\theta)}, 1\right\}$$

The proportion of authentic and counterfeit victims determines how likely any application is to be funded. As the preference of authentic individual and hedonic individual are given, the choice of strategic individual is the key determinant for the probability of being funded. We focus on two possible symmetric pure strategies for strategic individuals, and accordingly two possible outcomes of $(\theta, p(\theta))$.

AS strategy: Given charity enters the market, the strategic individual acts authentic supplication in each period. Under AS strategy, $\mu(AS) = 1-\gamma$. The collective reputation and the probability of being funded are both valued high at θ^H and $p(\theta^H)$ respectively. θ^H and $p(\theta^H)$ are endogenously determined by the following set of equations:

$$\begin{cases} \theta^H \lambda + (1-\theta^H)(1-\tau)p(\theta^H)\lambda = \theta^H \\ p(\theta^H) = min\{\dfrac{m}{(1-\gamma)\tau + \gamma Y(\theta^H)}, 1\} \end{cases}$$

FS strategy: Given charity enters the market, the strategic individual acts false solicitation in each period. Under FS strategy, $\mu(AS) = 1-\beta-\gamma$. The collective reputation and the probability of being funded are both valued low at θ^L and $p(\theta^L)$ respectively. θ^L and $p(\theta^L)$ are endogenously determined by the following set of equations:

$$\begin{cases} \theta^L \lambda + (1-\theta^L)(1-\tau)p(\theta^L)\lambda = \theta^L \\ p(\theta^L) = min\{\dfrac{m}{(1-\beta-\gamma)\tau + (\beta+\gamma)Y(\theta^L)}, 1\} \end{cases}$$

To end this section, we plot Figure 3 below to summarize the interrelations among individual's type, history, individual reputation, tracking record, and the funding outcome.

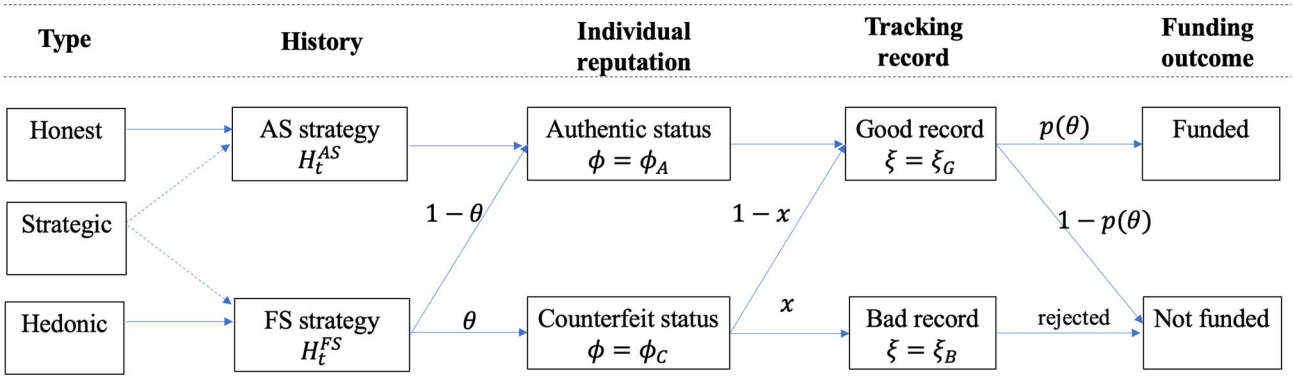

Figure 3. Interrelation among main variables

5. Steady-state equilibria

We solve for the pure-strategy Perfect Bayesian Equilibrium with backward induction. First, consider the strategic individual's action given that the funding is available in the market. The strategic individual is aimed to maximize long-run expected utility, with the tradeoff between current utility gain and future reputation loss of false solicitation. The maximum long-run expected utility is defined as $V(\phi_t; \theta_t)$ for $t = 0,1,2, \ldots \infty$, given both individual reputation ϕ_{t-1} and collective reputation θ_{t-1}. According to Bellman's principle of optimality, the value function satisfies:

$$V(\phi_{t-1};\theta_{t-1}) = \max_{s_t \in \{AS,FS\}} \mathbb{E}_t[u(a_t;\phi_{t-1},\theta_{t-1}) + \delta V(\phi_t;\theta_t)]$$

$$s.t.\, \phi_t = g(\phi_{t-1}, a_t)$$

where $\mathbb{E}_t u(a_t;\phi_{t-1},\theta_{t-1})$ is the expected utility in the current period and $g(\phi_{t-1}, a_t)$ is the switching function for individual reputation as introduced in section 4.2. As the dynamic programming problem is discussed under steady-state, time subscript is omitted in the following context.

To solve individual's dynamic programming problem, we first consider individual with counterfeit reputation ϕ_C. The current expected utility for the individual with counterfeit reputation is given as follows.

$$\mathbb{E}_t u(FS;\phi_C,\theta) = (1-\tau) - \kappa\big(p(\theta)\big)[(1-\tau)(1-r) + \tau R]$$

$$\mathbb{E}_t u(AS;\phi_C,\theta) = (1-\tau)r - \tau R\kappa\big(p(\theta)\big)$$

where $\kappa(p(\theta)) \equiv (1-x)(1-p(\theta)) + x$ is rejection rate for the individual with counterfeit reputation.

Solve the dynamic program for the individual with counterfeit reputation and yield Lemma 1.

Lemma 1. The individual with counterfeit reputation acts false solicitation as the optimal choice moving forward.

Proof: see Appendix.

Lemma 1 indicates that individuals with counterfeit reputation will be locked into false solicitation. Since the likelihood of rejection is irrelevant with the number of times of false solicitation, and the individual who has deceived funding successfully will always keep a counterfeit reputation going forward, the optimum choice for the individual is to repeat false solicitation. Lemma 1 helps us simplify the analysis by ruling out the time-varying strategies in the strategy space. Technically, we can obtain similar results to Lemma 1 under a more general assumption that the probability of past deception being rejected, increases at a decreasing rate as the number of false solicitations increase (Tirole, 1996).

Next consider the individual with authentic reputation ϕ_A. The current expected utility for the individual with authentic reputation is given as follows.

$$\mathbb{E}_t u(AS;\phi_A,\theta) = p(\theta)\tau R + (1-\tau)r - \tau R$$

$$\mathbb{E}_t u(FS;\phi_A,\theta) = p(\theta)[(1-\tau)(1-r) + \tau R] + (1-\tau)r - \tau R$$

It is obvious that $\mathbb{E}_t u(FS; \phi_A, \theta) > \mathbb{E}_t u(AS; \phi_A, \theta)$, which implies the individual with authentic reputation can gain a higher expected utility in the current period if she acts false solicitation. However, the dark side of false solicitation is that the individual may face with the reputational loss in the future. The continuation value can be decomposed into reputation switching and reputation unswitched two cases.

$$V(g(\phi_A, FS); \theta) = \underbrace{(1-\tau)p(\theta)V(\phi_C; \theta)}_{reputation\ switching} + \underbrace{\big(1-(1-\tau)p(\theta)\big)V(\phi_A; \theta)}_{reputation\ unswitched}$$

The strategic individual trades off the current gain from false solicitation and future loss from possibly being rejected with counterfeit reputation. Note that collective reputation affects individual's tradeoff in two ways. On one hand, it affects how costless the false solicitation is in terms of reputational concern. On the other hand, it influences how attractive the authentic supplication is in terms of the probability of being funded. If the collective reputation is poor, an authentic individual is more likely to not receive funding when she is really in need someday. In other words, individual reputation is costly kept but mostly useless.

There are two possible pure-strategy equilibria in the steady state. Suppose all agents hold the rational expectation of high level of collective reputation as θ^H. Every strategic individual expects that other strategic individuals choose AS as their best responses. Given the high level of collective reputation, the individual's incentive compatibility (IC) condition for choosing AS is[8]:

$$\mathbb{E}_t[u(AS; \phi_A, \theta^H) + \delta V(g(\phi_A, AS); \theta^H)] \geq \mathbb{E}_t[u(FS; \phi_A, \theta^H) + \delta V(g(\phi_A, FS); \theta^H)]$$

Suppose all agents hold the rational expectation of low level of collective reputation as θ^L. Accordingly, every strategic individual expects that other strategic individuals choose FS as their best responses. Given the low level of collective reputation, the individual's incentive compatibility (IC) condition for choosing FS is:

$$\mathbb{E}_t[u(AS; \phi_A, \theta^L) + \delta V(g(\phi_A, AS); \theta^L)] \leq \mathbb{E}_t[u(FS; \phi_A, \theta^L) + \delta V(g(\phi_A, FS); \theta^L)]$$

We next revisit the first step of charity's decision on whether to enter the market and provide funding. A welfare-concerned charity is willing to provide funding only if the expected social welfare with funding is higher than the case when funding is withdrawn. The expected social welfare is the aggregation of

[8] One-shot deviation principal is satisfied for pure-strategy equilibrium. Because if no deviation is better than any possible deviations henceforth, no deviation must be better than one-shot deviation.

individuals' expected utilities deducted by funding distributed in the victim market. The charity estimates social welfare based on the rational expectation of collective reputation of the victim group. Suppose the charity enters the market with the rational expectation of collective reputation $\theta \in \{\theta^H, \theta^L\}$, the expected social welfare is given by:

$$\mathbb{E}_t SW(\theta) = \underbrace{\mu(AS)\mathbb{E}_t u(AS; \phi_A, \theta)}_{authentic\ victims} + \underbrace{\big(1-\mu(AS)\big)[\theta \mathbb{E}_t u(FS; \phi_C, \theta) + (1-\theta)\mathbb{E}_t u(FS; \phi_A, \theta)]}_{false\ victims}$$

$$- \underbrace{[\mu(AS)\tau + (1-\mu(AS))(\theta(1-x) + (1-\theta))]}_{funding\ distributed}$$

If charity does not enter the market and provide funding, fraction $1-\tau$ of individuals produce and fraction τ of individuals suffer loss. The expected social welfare is given by

$$\mathbb{E}_t \underline{SW} = (1-\tau)r - \tau R.$$

The charity enters the market if and only if doing such is beneficial for social welfare. Note that providing funding does not necessarily increase social welfare, because the counterfeit victim no longer has incentive to produce once successfully deceives funding. Given the collective reputation $\theta \in \{\theta^H, \theta^L\}$, charity's IR condition is given by: $\mathbb{E}_t SW(\theta) \geq \mathbb{E}_t \underline{SW}$.

To solve strategic individual's IC condition and charity's IR condition, we introduce three technical assumptions to streamline analyses.

Assumption 1: $R > \frac{1-\delta+\delta(1-x)(1-\tau)}{\delta x \tau}(1-r)$.

Assumption 1 is the necessary condition for strategic individuals to possibly choose the AS strategy. If Assumption 1 is violated, the victimhood loss becomes incredibly low, making the gain from false solicitation overwhelming. In such cases, there is never any incentive for strategic individuals to choose the AS strategy.

Assumption 2: $R < 1 + \frac{1-\tau}{\tau} r$.

Assumption 2 is the necessary condition for the charity to possibly not enter the victim market. If Assumption 2 is violated, the victimhood loss is so substantial that charities are always willing to provide resources to improve social welfare regardless of the rational expectation of collective reputation.

Assumption 3: $\lambda(x+\tau) \leq 1$.

Assumption 3 is a technical assumption to ensure that our definition of the probability of being funded in steady state is well-defined within [0,1].

5.1 Equilibrium analysis

Given Assumptions 1-3, the steady-state equilibria are sketched in Proposition 1.

Proposition 1: The pure-strategy Perfect Bayesian Equilibria in steady state are characterized as follows.

(1) If $m \geq max\left\{\widetilde{m_c^H}, \widetilde{m_v^H}\right\}$, the market has the Honest Equilibrium. The charity enters the market and approves funding victims with good tracking records, whereas rejects victims with bad tracking records. All strategic individuals choose authentic supplication in every period. The belief on collective reputation is $\theta = \theta^H$.

(2) If $m < min\left\{\widetilde{m_c^D}, \widetilde{m_v^D}\right\}$, the market has the Dishonest Equilibrium. The charity enters the market and approves funding victims with good tracking records, whereas rejects victims with bad tracking records. All strategic individuals choose false solicitation in every period. The belief on collective reputation is $\theta = \theta^L$.

(3) If $\widetilde{m_c^D} \leq m < \widetilde{m_v^D}$, the market has the Type-F Unfunded Equilibrium. The charity does not enter the market. On the off-equilibrium path, all strategic individuals choose false solicitation in every period.

(4) If $\widetilde{m_c^H} \leq m < \widetilde{m_v^H}$, the market has the Type-A Unfunded Equilibrium. The charity does not enter the market. On the off-equilibrium path, all strategic individuals choose authentic supplication in every period. Note:

$$\widetilde{m_c^H} = \frac{\tau(1-\gamma)(1-\lambda)(r+R-1)\big((R-1)\tau-r\gamma(1-\tau)\big)}{\lambda(r-(r+R-1)\tau)\big(\gamma(1-x)(1-\tau)r-\tau(R-1)(1-x\gamma)\big)}$$

$$\widetilde{m_c^D} = \frac{\tau(1-\beta-\gamma)(1-\lambda)(r+R-1)\big((R-1)\tau-r(\beta+\gamma)(1-\tau)\big)}{\lambda(r-(r+R-1)\tau)\big((\beta+\gamma)(1-x)(1-\tau)r-\tau(R-1)(1-x(\beta+\gamma))\big)}$$

$$\widetilde{m_v^H} = \frac{(1-r)(1-\delta)\Big[(1-\tau)(1-r)(\lambda-\delta)\big(\tau+\gamma(1-\tau-x)\big)+x\delta\tau(1-\lambda)\Big(1-r+(R-1+r)\big(\tau+\gamma(1-\tau)\big)\Big)\Big]}{\delta\big(Rx\tau-(1-r)(1-x)(1-\tau)\big)\big[\delta x(1-\lambda)\big(1-r+\tau(R-1+r)\big)+(1-r)(1-\tau)(\lambda-\delta)\big]}$$

$$\widetilde{m_v^D} = \frac{(1-r)(1-\delta)\left[(1-\tau)(1-r)(\lambda-\delta)\left(\tau+(\beta+\gamma)(1-\tau-x)\right)+x\delta\tau(1-\lambda)\left(1-r+(R-1+r)\left(\tau+(\beta+\gamma)(1-\tau)\right)\right)\right]}{\delta\left(Rx\tau-(1-r)(1-x)(1-\tau)\right)\left[\delta x(1-\lambda)\left(1-r+\tau(R-1+r)\right)+(1-r)(1-\tau)(\lambda-\delta)\right]}$$

Proof: see Appendix.

Proposition 1 indicates that the equilibrium is determined by both the level of funding supply and the other market characteristics. To illustrate the insights of Proposition 1, we use a set of numerical parameters to demonstrate each equilibrium in the space of (γ, m) in Figure 3. When the proportion of hedonic individual γ is low and the funding supply m is sufficiently abundant, the market has the Honest equilibrium (see regions I, II, IV, and V in Figure 3). In our context, and henceforth, we refer to the funding as "abundant" when all approved applicants can be funded; otherwise, the funding is "scarce". Holding a low level of proportion of hedonic individual, if the funding supply m is scarce, the market has the Dishonest equilibrium (see regions IV and VII in Figure 3). In this case, although strategic individuals falsely solicit, the charity is still willing to provide funding because counterfeit victims are few overall.

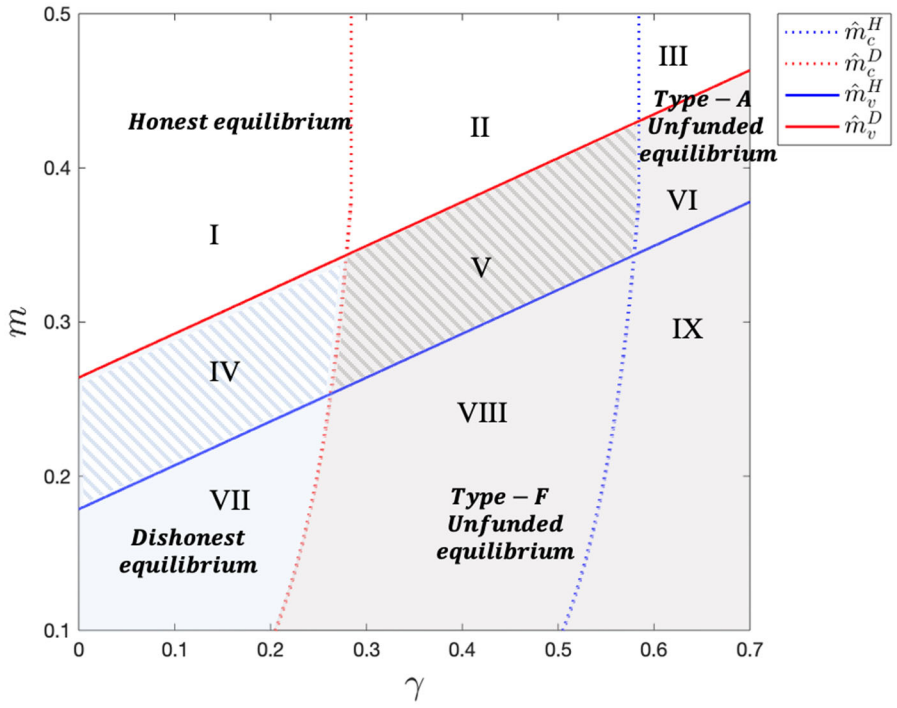

Figure 3. Equilibria in the steady state

Note: The red solid line and dotted lines are $\widetilde{m_v^H}$ and $\widetilde{m_c^H}$ solved from individual's IC and charity's IR condition of the honest equilibrium. The blue solid line and dotted lines are $\widetilde{m_v^D}$ and $\widetilde{m_c^D}$ solved from individual's IC and charity's IR condition of the dishonest equilibrium. The union of regions I, II, IV and V represents the Honest Equilibrium. The union of regions IV and VII (blue- shadowed) represents the Dishonest Equilibrium. The union of regions V, VI, VIII

and IX (grey-shadowed) represents the Type-F Unfunded Equilibrium. The union of regions III, VI and IX represents the Type-A Unfunded Equilibrium. The blue-line shadowed area shows coexistence of the Honest Equilibrium and Dishonest Equilibrium, and the grey-line shadowed area shows coexistence of the Honest Equilibrium and Type-F Unfunded Equilibrium. Baseline parameters are set as follows: $\lambda = 0.85, \delta = 0.8, \beta = 0.3, x = 0.6, \tau = 0.2, r = 0.6, R = 2$.

The third equilibrium comes when the proportion of hedonic individual γ is moderate and the funding supply m is scarce, for which we call Type-F unfunded equilibrium (see regions V, VI, VIII, and IX in Figure 3). "F" stands for strategic individuals choose FS strategy in the off-equilibrium path. The charity forgives to provide funding because the limited funding may induce strategic individuals to false solicit, and this delivers a worse welfare outcome compared to the scenario without funding. The other unfunded equilibrium is called Type-A unfunded equilibrium when the proportion of hedonic individual γ is very large (see regions III, VI, and IX in Figure 3). Type-A unfunded equilibrium is worse than Type-F unfunded equilibrium, because the charity gives up funding supply regardless of how abundant the funding resources are. The Type-A unfunded equilibrium implies that too many hedonic individuals irreversibly hurt the victim marketplace.

One interesting result is the existence of multiple equilibria under specific market conditions. For example, the honest equilibrium and the dishonest equilibrium co-exist when the proportion of hedonic individual is low, and the dishonest equilibrium and the Type-F unfunded equilibrium co-exist when the proportion of hedonic individual is moderate. It can be proven that multiple equilibria hold as long as Assumption 3 is guaranteed. The multiple equilibria arise from the heterogenous belief on collective reputation. Victims are competing for chances of being funded. If a strategic individual believes that other strategic individuals are counterfeit victims so that the collective reputation is bad, her best response is to be a counterfeit victim as well. This demonstrates the strategic complementarity between victims. Echoing our finding, Jiang and Villeval (2024) propose that groups can trap into "tragedy of dishonesty", because the group members pursue selfish interests and there is a lack of coordination within group members.

The feature of multiple equilibria can be viewed as a kind of run behavior. The "run" refers to situations where market conditions are feasible for the honest equilibrium, while the shift in belief on collective reputation leads to deviation, consequently altering the honest equilibrium into the dishonest

equilibrium or unfunded equilibrium. The behavior of run-on funding is also similar to bank run (Diamond and Dybvig, 1983). The difference between bank run and run behavior we propose here is that the latter one relates to reputation concern; i.e., the strategic individual may suffer individual reputational loss in future if she joins the run-on funding. Compared to bank run, the reputation concern increases the costs to join the run and reduces the likelihood of a run.

Corollary 1: (1) When the proportion of hedonic individuals is sufficiently large, $\gamma > \hat{\gamma} \equiv \hat{\gamma}(x,\tau,\lambda,r,R) = \frac{(R-1)(1-\lambda\tau)\tau}{(1-\tau)(r(1-\lambda(x+\tau))+\lambda\tau x(r+R-1))}$, the market never supports the honest equilibrium in the steady state, no matter how abundant the funding is. (2) When the supply of funding is absolutely scarce, $m < \hat{m} \equiv \hat{m}(\lambda,\tau,x,r,R) = \frac{(1-r)(1-\delta_0\lambda)\tau}{\delta_0\lambda(Rx\tau-(1-r)(1-x)(1-\tau))}$, the market never supports the honest equilibrium in the steady state, no matter how few the proportion of hedonic individuals is.

Proof: see Appendix.

Corollary 1 shows the maximum quantity of hedonic individual and the minimum funding supply to support the Honest equilibrium. As victims are competing for funding, the quantity and hedonic individual and funding supply jointly determine how much the victim is affected in terms of probability of being funded by the peers' false solicitation actions. This is somewhat like peer pressure effect. Having more hedonic individuals in victim marketplace or less funding supply intensifies the peer pressure among victims, which in turn motivates strategic individuals to abandon authentic reputation and further deteriorates the expectation of collective reputation. The importance of resource abundance to maintain social conduct is supported by a lot of evidence. There's one Chinese idiom saying people know etiquette only when the stock is plenty and understand honor and humility when there are sufficient clothes and food. The general insight also aligns with the wisdom of asking for just enough (Proverbs 30:8).

The collective reputation plays an important role from both charity and victim sides. For the charity side, the expectation of collective reputation determines willingness of funding supply. For the victim side, the expectation of collective reputation determines the intensity of peer pressure effect. Most of classic collective reputation models (e.g., Tirole 1996 and Levin 2009) typically only outline the former effect.

The latter effect, which is how our work stands out, characterizes how victims' incentives are interdependent through collective reputation and, in turn, endogenously shape the collective reputation itself. Compared to the classic models, the mechanism of peer pressure we introduce makes collective reputation more vulnerable to group member's deviation. This intuition is similar to the run behavior we have discussed. The deviation of peers adversely impacts the individual's expected utility, thereby reducing the individual's incentive to maintain a good individual reputation and collective reputation.

5.2 The impact of improvement in tracking technology

With the development of information technology, tracking technology has been widely used in the charity and donation sector. For example, blockchain technology is widely used to address issues of transparency, fraud, and inefficiency. One of the most significant advantages of blockchain in charitable organizations is transparency. Donors often question how their contributions are used, and blockchain offers an open, accessible record of all transactions. Donation platform leveraging blockchain technology ensure that every donation can be tracked, from the donor to the final beneficiary. Blockchain makes it easier to detect fraud, as the history of all transactions and donations is visible and cannot be altered. In our model, the increment of x can be interpreted as the improvement in tracking technology. Our model sheds light on how digital technology benefits the victim marketplace.

Corollary 2: $\frac{\partial \hat{\gamma}}{\partial x} > 0$, and with the support of $\gamma \leq \hat{\gamma}$, $\frac{\partial \widetilde{m_v^H}}{\partial x} < 0$. The higher the accuracy of tracking technology is, the more likely the victim marketplace supports the honest equilibrium.

The accuracy of tracking record x plays an important role in the victim marketplace. The associated numerical results of Corollary 2 are plotted in Figure 4. A special case is when funding supply is abundant. In this case, the improvement in tracking technology increases the likelihood of being detected and deteriorates expected utility for counterfeit victim, weakening the strategic individual's incentive to falsely solicit. From the view of charity, enhanced ability to detect counterfeit victims means a more efficient funding allocation, strengthening the charity's incentive to provide funding. A more interesting case comes when the funding supply is non-abundant. Apart from the above effect, a novel effect of advanced tracking

technology is that it increases the incentive for individuals to make authentic supplications. As the traceability enhances, more counterfeit victims are being detected, which leaves the remaining applicants more likely to be funded. By considering both the decreasing incentive for false solicitation and the increasing incentive for authentic supplication, improvements in tracking technology make the market more likely to support an honest equilibrium.

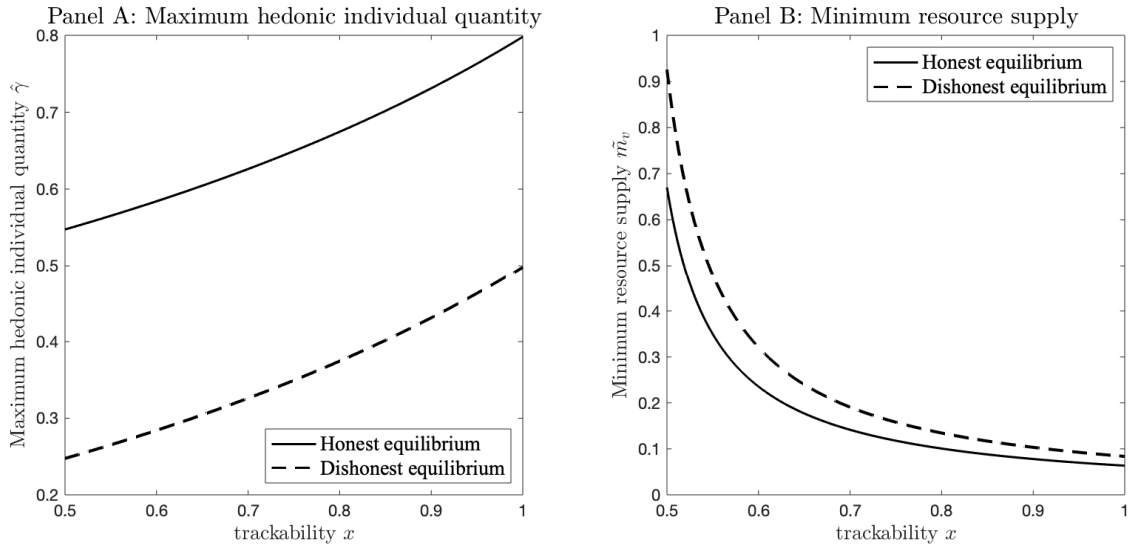

Figure 4. The impact of improvement of tracking technology

Note: Panel A represents the maximum level of hedonic individuals allowed for the honest equilibrium (solid line) and dishonest equilibrium (dashed line). Panel B represents the minimum resource supply for the honest equilibrium (solid line) and dishonest equilibrium (dashed line). Parameters are set as same as Figure 3 except that $\gamma = 0.2$.

5.3 Other derived Results

In this section, we show some comparative statics results on our equilibrium analysis. Corollary 2 demonstrates the sensitivity of the necessary conditions that support the honest equilibrium to the underlying market characteristics. The associated numerical results are plotted in Figure 5.

Corollary 3: $\frac{\partial \hat{\gamma}}{\partial \tau} > 0, \frac{\partial \hat{\gamma}}{\partial \lambda} > 0$, and with the support of $\gamma \leq \hat{\gamma}$, $\frac{\partial \widetilde{m_v^H}}{\partial \tau} < 0, \frac{\partial \widetilde{m_v^H}}{\partial \lambda} < 0$. The higher the victimhood risk and stagnant rate, the more likely the victim marketplace supports the honest equilibrium.

Proof: see Appendix.

The victimhood risk τ has a profound impact on the equilibrium outcome. A high risk of victimhood encourages individuals to behave authentically and strengthens charitable efforts through funding. On the

one hand, an increase in victimhood risk means that false solicitation behavior is more likely to be recorded and possibly corrupt individual reputation. On the other hand, it intensifies the punishment once individuals are locked into counterfeit reputation by depriving them of the right to be aided in the victim state. The heightened victimhood risk thereby poses a credible threat to individuals, thereby weakening the incentive to engage in false solicitation. A high victimhood risk also levels up charity's tolerance on the prevalence of counterfeit victims, as the welfare loss will be substantial if funding is unavailable. Overall, the degree of victimhood risk suggests the extent to which victims rely on social assistance.

Stagnant rate λ indicates the time horizon that the individual participates in the victim marketplace. A higher stagnant rate implies that individual has a longer horizon and values more on reputation and future payoff, so the individual has weaker incentive to FS. The stagnant rate also determines the renewability of collective reputation in steady state. A higher stagnant rate means that the collective reputation in steady state is more contributed by the incumbent individual and less contributed by new comer. Our finding that short horizon may induce more cheating is supported by Effron et al. (2015), who find that the odds of cheating are almost three times higher than usual when the individual is facing with the last round of game.

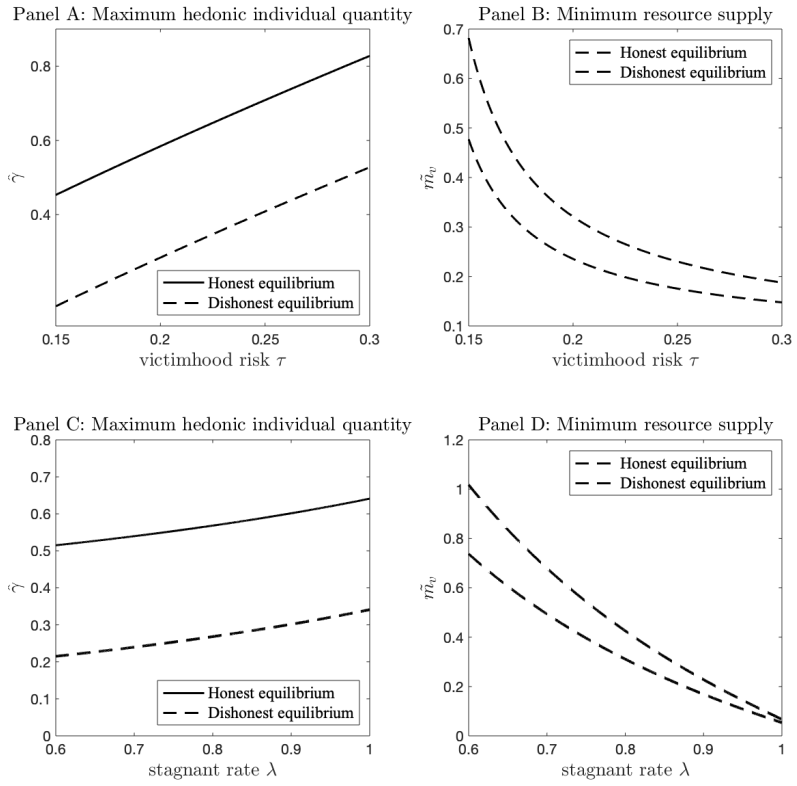

Figure 5. Sensitivity analysis of equilibrium outcome to victimhood risk and stagnant rate

Note: Panel A and C represent the maximum level of hedonic individuals for the honest equilibrium (solid line) and dishonest equilibrium (dashed line) along with victimhood risk and stagnant rate, respectively. Panel B and D represent the minimum resource supply for the honest equilibrium (solid line) and dishonest equilibrium (dashed line) along with victimhood risk and stagnant rate, respectively. Parameters are set as same as Figure 3 except that $\gamma = 0.2$.

We also conduct the large-scale comparative statics analysis for the necessary condition of the honest equilibrium in Figure 6, by conditioning on a large range of proportions of hedonic individuals γ and funding supply m. Figure 6 shows that the conclusion shown in Figure 4 are robust for any specific level of proportion of hedonic individual γ. We also conduct the large-scale comparative statics analysis for the necessary condition of the dishonest equilibrium. The results are consistent with the honest equilibrium, and we relegate these to the Appendix B.

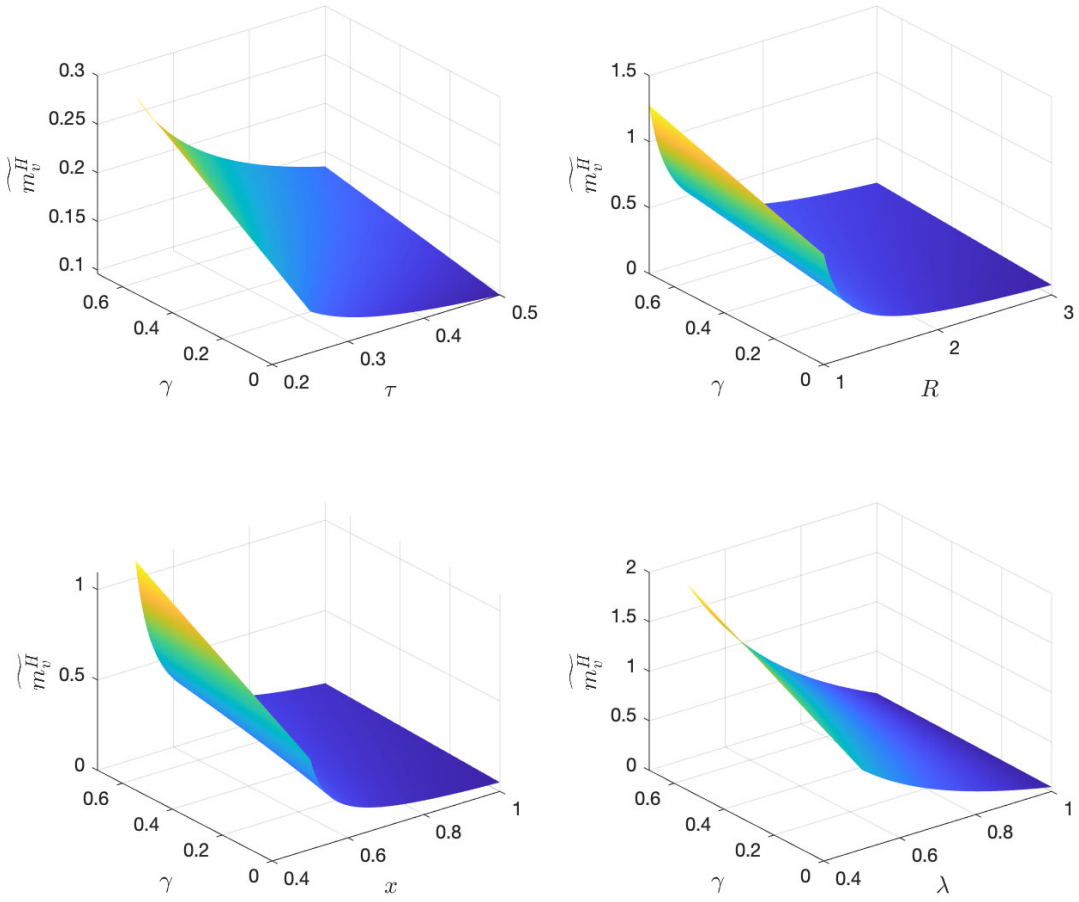

Figure 6 Large-scale comparative statics analysis

Note: The 3D surfaces plot the relationship between the individual IC condition for the honest equilibrium under different combinations of m, γ and the respective exogenous variables as labeled on the main horizontal axis of each panel. All the unspecified parameters are set at their baseline levels.

5. Welfare analysis

This section aims to analyze the welfare impact of counterfeit victims and uncover the mechanisms leading to social welfare changes across different equilibrium outcomes. To compare the social welfare under different equilibria, we first suppose the first-best allocation for comparison. Suppose there exists a social planner who perfectly observes each individual's type and state and precisely allocate the funding to authentic victims. The social welfare under the first-best allocation is given by:

$$\mathbb{E}_t SW^F = (1-\tau)r - max\{\tau - m, 0\}\, R - \tau.$$

Compare the social welfare under different equilibria and the first-best allocation, we obtain Proposition 2.

Proposition 2: The social welfare under the honest equilibrium is strictly higher than that under the dishonest equilibrium, i.e., $\mathbb{E}_t SW(\theta^H) > \mathbb{E}_t SW(\theta^L)$. The social welfare is non-decreasing with the supply of funding, i.e., $\frac{\partial \mathbb{E}_t SW(\theta)}{\partial m} \geq 0$ and $\frac{\partial^2 \mathbb{E}_t SW(\theta)}{\partial m^2} \geq 0$ for $\theta \in \{\theta^H, \theta^L\}$.

Proof: see Appendix.

Figure 7 depicts the comparison of social welfare under different equilibria. The social welfare under the first-best allocation is the highest, followed by those of the honest equilibrium and the dishonest equilibrium, and the unfunded equilibrium has the lowest welfare. The social welfare increases with the supply of resources m until the resources are abundant to serve all approved applicants. The minimum resources necessary to achieve the maximum social welfare in the dishonest equilibrium are higher than those in the honest equilibrium, followed by those in the first-best allocation.

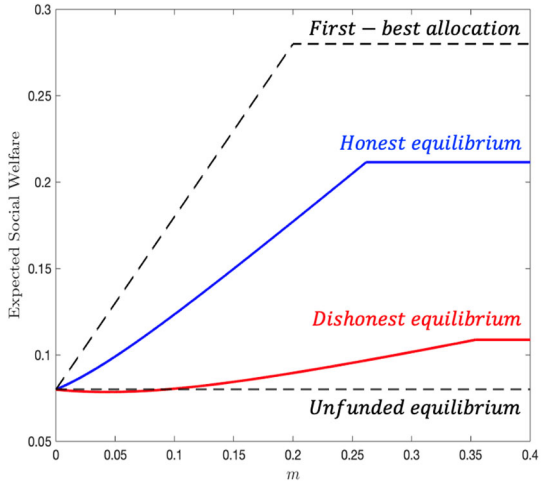

Figure 7. Social welfare in the steady state

Note: The dashed (top) line, blue solid line, red solid line, and dashed (bottom) line, respectively, plot the social welfare under the first-best allocation, the honest equilibrium, the dishonest equilibrium, and the unfunded equilibria (either type-A or type-F) with $\gamma = 0.2$ and other baseline parameters.

Proposition 2 shows that the increment of one unit of funding on social welfare under the honest equilibrium is higher compared to that under the dishonest equilibrium. This is because funding are utilized more effectively under the honest equilibrium, as the funding are more likely to be given to the authentic victim. This result may help explain why donors have higher willingness to donate in a trustworthy market. Consider a mass of donors who care about social welfare (the charity in our model, to some extent, acts like a delegator of a mass of welfare-concerned donors). Each donor is more willing to donate if she anticipates the market is feasible for the honest equilibrium, because the donation is more valuable for welfare in the honest equilibrium compared to the dishonest equilibrium. The charity IR condition characterizes under what condition the donor is willing to donate. Our theoretical prediction echoes the effective-altruism movement, which advocates that donors allocate their limited resources to charity that maximize welfare (Berman et. al., 2018). Information disclosure for the charity is highlighted as important.

Another novel finding shown in Proposition 2 is that the increment of social welfare is marginally increasing in the resource supply. This is distinct from the traditional diminishing marginal utility theory. The reason for this surprising result is that the expansion of resource supply not only increases the quantity of resources but, more importantly, improves the efficiency of resource allocation. To be specific, the increment in the funding supply upward-shifts the distribution of individual reputation in steady state (i.e., $\frac{\partial \theta}{\partial m} > 0$), which means false solicitation is more likely to corrupt individual reputation. As a consequence, more counterfeit victims could be rejected by the spotted tracking record, and correspondingly, resources are more efficiently allocated to authentic victims. The quantity and efficiency effects jointly explain the second-order increment of funding supply on social welfare. We argue that the presence of counterfeit victims may reverse the common belief that more resources are accompanied by less efficiency. In contrast, limited resources distort individual incentives through peer pressure, leading to lower efficiency.

Corollary 4: $m(\theta)^* \equiv argmax\ \mathbb{E}_t SW(\theta)$, $\frac{\partial m(\theta)^*}{\partial \gamma} > 0, \frac{\partial \mathbb{E}_t SW(m(\theta)^*;\theta)}{\partial \gamma} < 0$ for $\theta \in \{\theta^H, \theta^L\}$.

Proof: see Appendix.

Corollary 4 shows that the resource supply that maximizes social welfare is increasing in the proportion of hedonic individual, while the maximized social welfare is decreasing in the proportion of hedonic individual. More counterfeit victims in the marketplace implies more severe funding misallocation, resulting in more deadweight losses. To offset the adverse effects of counterfeit victims on social welfare, charity has to raise more money, which could be considered as the monetary costs of the counterfeit victims. We again highlight the inefficient resource allocation at the presence of counterfeit victims. That is, more resource input is needed, while it unfortunately ends with a lower welfare outcome.

Why does the presence of counterfeit victims deteriorate social welfare? We further investigate how social welfare varies along with the quantity of hedonic individuals (see Figure 8). The results show that social welfare is decreasing with the quantity of hedonic individuals for both the honest equilibrium and the dishonest equilibrium. The charity gives up funding as long as social welfare is lower than the unfunded equilibrium. The presence of counterfeit victims incurs deadweight loss that accounts for the social welfare gap between the equilibria outcome and the first-best allocation. The social welfare of the honest equilibrium is lower than that of the first-best allocation, because counterfeit victims (hedonic individuals) crowd out the resources that could be granted to authentic victims. We refer to the first component of deadweight loss as *crowd out effect.* Resource misallocation harms welfare for two reasons. On one hand, resources that should be allocated to authentic victims are instead consumed by counterfeit victims, leaving the authentic victims unaided and suffering welfare losses. On the other hand, the resources misallocated to counterfeit victims cause these recipients to shirk and cease production, which decreases overall output.

The second component of deadweight loss is *run effect*, which explains the social welfare gap between the honest equilibrium and the dishonest equilibrium. The strategy switching by strategic individuals cause the run on resources, which deteriorate welfare for two reasons. On one hand, the change in the ratio of counterfeit victims to authentic victims re-adjusts the payoff structure of social welfare (i.e., counterfeit victims account for a larger share of welfare). On the other hand, if the payoff structure of social welfare remains unchanged, strategic individuals intensify the competition for funding when resources are limited,

which results in a lower probability of funding for all victims and marginally decreases social welfare. The latter reason can be treated as *secondary crowd out effect*, as authentic victims are faced with the joint crowd out effect from both the hedonic and strategic individuals in the dishonest equilibrium.

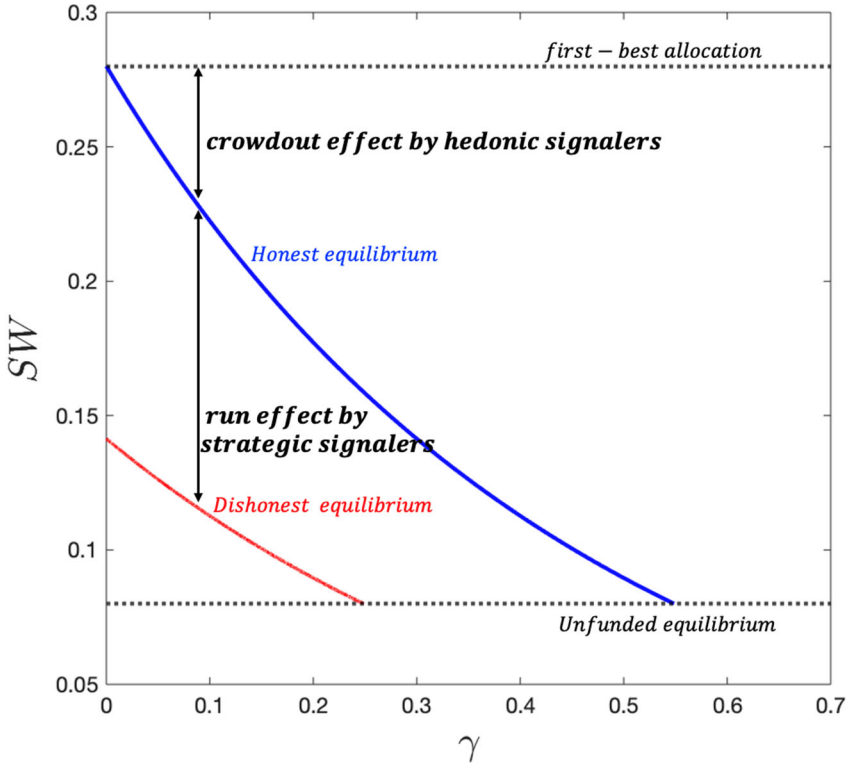

Figure 8. Decomposition of deadweight loss in social welfare

Note: The blue solid line is the social welfare under the honest equilibrium in the steady state, and the red solid line is that under the dishonest equilibrium. The two dotted lines from top to bottom, respectively, are social welfare under first-best allocation and the unfunded equilibrium (both of type-A and type-F). Except for $m = 0.2$, all parameters are set to their baseline levels.

Given that overall social welfare decreases with the increase in counterfeit victims, an interesting question arises: does anyone benefit at the expense of others' utility deteriorating? To answer this question, we compare the individual-level utility of different types of individuals under different equilibria (see Figure 9). The result shows that the utilities of all types of individuals are lower in the dishonest equilibrium than in the honest equilibrium. In other words, the dishonest equilibrium is Pareto inferior. The point we address here is that strategic individuals also suffer in the dishonest equilibrium. As mentioned above, acting false solicitation simultaneously changes the payoff structure and the peer pressure for strategic

individuals. When resource supply is limited, the latter adverse effect on peer pressure exceeds the former positive effect, resulting the decrement in overall expected utility for strategic individuals. No one nor the organization gain in the dishonest equilibrium compared to the honest equilibrium. We highlight how inadequate resources may lead to the collapse of collective reputation and further make every group member worse off. This theoretical prediction echoes Jiang and Villeval (2024)'s experiment finding. They find that individuals who pursue selfish interests may cause the collective failure of members of their communities, including themselves. We also interpret this finding as a link between competition intensity and organizational efficiency, which closely relates to the rat-race phenomenon. Competition on limited resources or opportunities drives aggressive (and possibly unethical) behaviors without necessarily leading to more productive outcome (Akerlof, 1976).

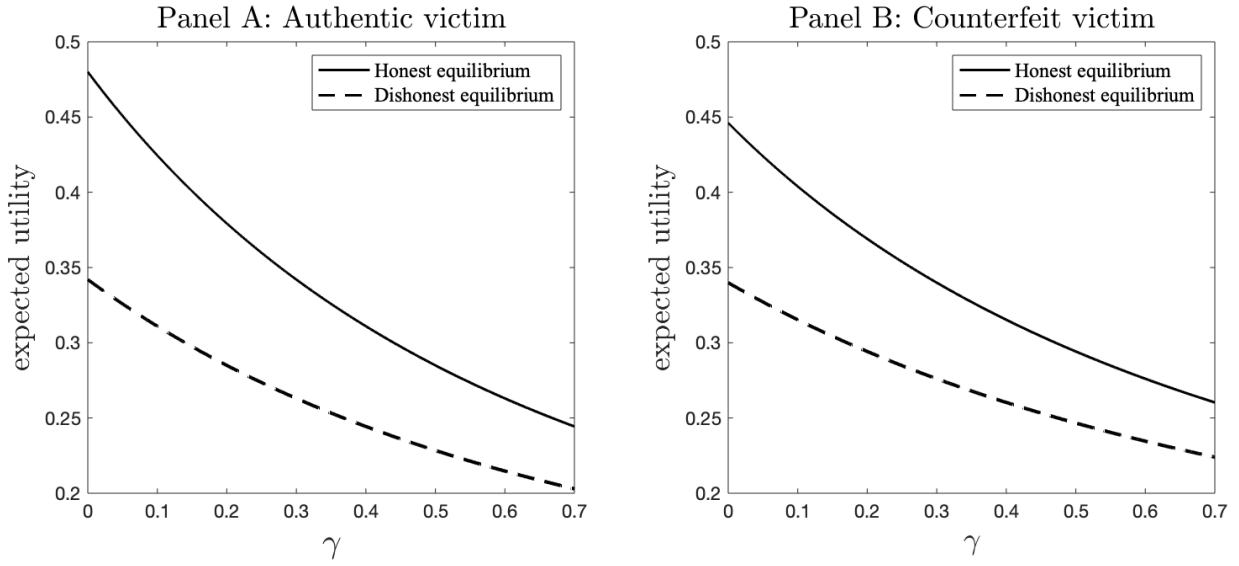

Figure 9. Victim-level expected utilities

Note: Panel A represents the expected utility for authentic victim under the honest equilibrium (solid line) and the dishonest equilibrium (dashed line), respectively. Panel B represents the expected utility for counterfeit victim under the honest equilibrium (solid line) and the dishonest equilibrium (dashed line), respectively. Parameters are set as same as Figure 3 except that $m = 0.2$.

6. Model Extension: Psychological Uncertainties and Costs

In this section, we consider the extension to the baseline model where authentic individuals could suffer psychological costs due to a few factors: 1. the inaccuracy of charity organization's screening,

resulting in uncertainty of the application payoffs; 2. uncertainty about own victim status, e.g., in the diagnosis phase of the status; 3. stigma associated with victimhood status that can result in social rejection, devaluation, and discrimination (Lasky-Fink and Linos 2022); and 4. "add insult to injury" in the most traditional sense. A recent working paper attempts to analyze the implications of the risk and the hurt that results from rejection (Bénabou et. al. 2022). Past studies also show that 20 to over 50 percent of households do not utilize government poverty-aid programs for which they are eligible (Bhargava & Manoli, 2015; Blumenthal, Erard, & Ho, 2005; FNS, 2020; Giannarelli, 2019). Our model incorporates these costs, as well as other consequences like shame, anger, and false guilt from being second-guessed or misunderstood.

We assume that the charity organization may misjudge any authentic individual in the screening process and reject her application for funding, with probability $z < 1$. If the authentic individual is misjudged, she not only incurs a monetary loss that could help relieve her deprivation, but also the psychological cost of suspecting that her plea for help was ignored, not believed, or discounted. We denote such psychological costs as PC. For an authentic individual, $PC = c$, and for the strategic and hedonic individual, $PC = 0$. Due to psychological costs, authentic individuals may voluntarily exit the victims market to avoid being misjudged. When all authentic individuals exit the market and all individuals left in the market falsely solicit, the market falls into the *lemon equilibrium*. The concept "lemon" is borrowed from the well-documented phenomenon of lemon cars driving out high-quality cars (Akerlof, 1970). The Lemon equilibrium is an unfortunate market equilibrium that echoes lab experimental results on lying in that strategic individuals emit more lies to counteract screening, and the ones how are more adept in lying remain in the market (Aquino and Becker, 2005).

The introduction of misjudgment and the associated psychological cost leads the authentic individuals to tradeoff between exiting and staying in the market. If she exits, the outside expected utility is $u^q = (1-\tau)r - \tau R$. Otherwise, if she stays in the victim market, her expected utility is:

$$\mathbb{E}_t u(\phi_0, AS; \theta) = (1-\tau)r + \tau(\underbrace{(1-z)\big(1-p(\theta)\big)(-R)}_{being\ crowd-out} + \underbrace{z(-R-c)}_{being\ misjudged}).$$

The incentive compatibility condition for authentic individual to stay in the market is given by

$$\mathbb{E}_t u(\phi_0, AS; \theta) \geq u^q.$$

The Lemon Equilibrium (LE) could be considered as the worst case of collective reputation, as all remaining individuals falsely solicit and only the proportion τ among them are authentic victims. Given Assumption 2, the charity chooses to close funding when he anticipates that all authentic individuals quit the market. In this case, the LE would be an unfunded equilibrium.

Proposition 4: When the misjudgment and the associated psychological cost is considered, the honest individuals exit the market when $m < m^q(\gamma, x, \tau, \lambda, r, R) \equiv \frac{cz(\beta+\gamma)(R(1-\lambda)(1-z)+cz\lambda(1-\tau)(1-x))}{R(1-z)(R(1-\lambda)+cz\lambda(1-\tau))}$.

Prove: see Appendix.

$\frac{\partial m^q}{\partial \gamma} > 0$ implies that the authentic individuals are more likely to quit the market if the proportion of hedonic individuals increases. Due to the crowd-out effect, the increment of hedonic individuals may lower the probability of being funded and the expected utility to apply for funding for the authentic individuals. Moreover, we can prove that $\frac{\partial m^q}{\partial z} > 0$ and $\frac{\partial m^q}{\partial c} > 0$. The higher the probability of being misjudged and the higher the psychological cost is, the more likely the authentic individuals will quit the market. As C.S. Lewis once said, "I have learned now that while those who speak about one's miseries usually hurt, those who keep silence hurt more."

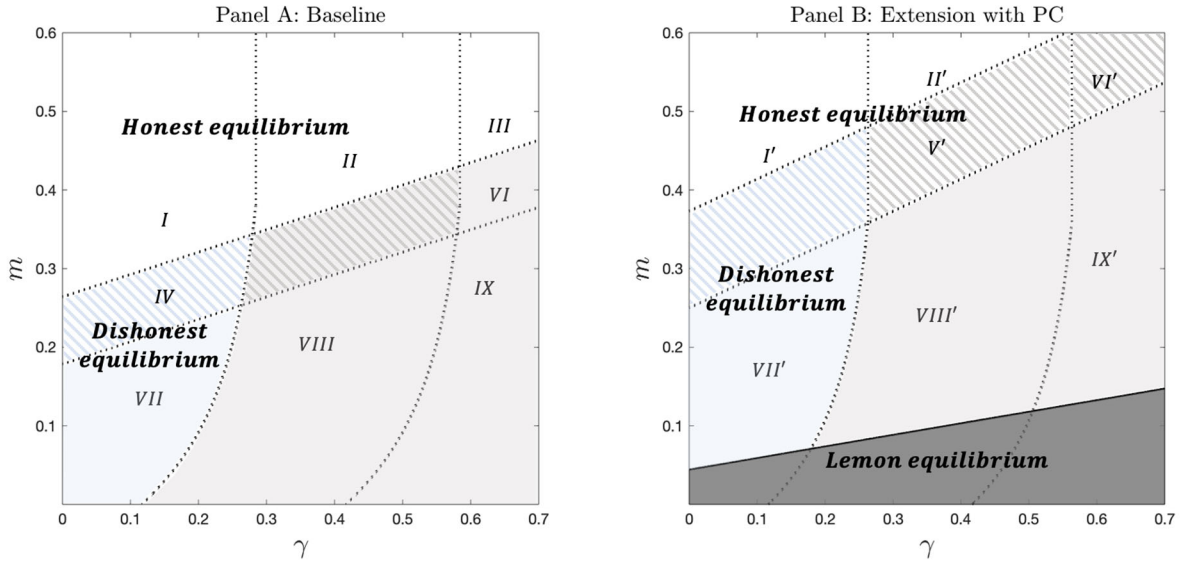

Figure 10. Comparison on Equilibria with and without psychological costs

Note: Panel A is the replication of Figure 2 for comparison. The regions I' to IX' in Panel B are the counterparts of I to IX in Panel A. The shadow area under the solid line that is solved from honest individual's IC condition represents the lemon equilibrium. To save space, the unfunded equilibria is not displayed. Except for $c = 4$ and $z = 0.1$, other parameters are chosen at their baseline levels.

The misjudgment not only motivates the honest individuals to quit the market, but also enhance the incentive for the strategic individuals to falsely solicit. Figure 9 displays the equilibria with (baseline) and without the misjudgment. It shows that the market is more likely to unravel into the dishonest equilibrium given the same condition as that in the baseline model. Because the misjudgment lowers the value of keeping an authentic reputation, it reduces the incentives for the strategic individuals to authentically behave. As a consequence, the charity need to supply more resources to prevent the strategic individuals from false solicitation.

Another finding is that, when misjudgment is introduced, multiple equilibria prevail more likely with a larger parameter space under the same market conditions. For illustration, compare the size of region IV+V in Panel A and B Figure 10. Intuitively, misjudgment reduces the difference in approval probabilities between authentic supplication and false solicitation, thereby dispersing the beliefs on collective reputation and paving the way for the coexistence of multiple equilibria due to uncoordinated perceptions. From another perspective, the imprecise screening makes the honest equilibrium more vulnerable to the run behavior caused by strategic individuals' deviations.

We also find the social welfare of the honest and the dishonest equilibria, when misjudgment is introduced, are lower than their counterparts in the baseline model under the same market conditions. Even though the psychological costs are not included in social welfare, the misjudgment of screening leads more authentic victims suffer victimhood loss (due to being mistakenly rejected), thus deteriorating overall welfare. The social welfare of the lemon equilibrium is exactly the same as that in the unfunded equilibrium. If the psychological costs are considered as part of social welfare, then the lemon equilibrium would result in even lower social welfare. Should we further consider the rippling effects of lying in the FVS practices, the misallocated charity resources could even be used for other rent-seeking activities, leading to further

social welfare deteriorations. We do not model these illegal behaviors in the current model, yet one could easily build extensions to capture such chain reactions.

In Appendices C, we explore and discuss the consequence hedonic individual's personality-based utility. Hedonic individuals differ from strategic individuals in that they enjoy extra psychological utility from false solicitation. This insight has practical implications especially in an age where people have at their disposal more sophisticated tactics of deception that reward the deceiver not only physically (such are charity fund), but also psychologically from being able to successfully exploit a trusting benefactor. The results in the baseline model still hold when psychological utility is introduced, except that the equilibria concepts are more complicated. Overall, the baseline is a special case when psychological utility is sufficiently large such that hedonic individuals are unconditional false victims, and the market is more likely to unravel into the dishonest and unfunded equilibria when the psychological utility is larger.

7. Discussions and Conclusions

Counterfeiting has received increasing attention in the recent years. We offer a novel application of models on counterfeits and collective reputation to predict what might happen when people in a marketplace of other people-seeking relief or deliverance from alleged suffering emit authentic or false victim signals. We argue that the FVS is conceptually similar to the introduction of a counterfeit product into a commodity market. We offer a novel application and extension of the collective reputation model to predict the potential social welfare costs when an increasing number of people in a society emit false victim signals. By sketching ingredients of victims' mindsets and subsequent economic ramifications, we show the mechanism and relevant environmental factors that would moderate and determine the subsequent equilibria in which the market performs.

We show that the market is more likely to gravitate to the dishonest equilibrium when the proportion of hedonic signalers increases, resulting a reduction in social welfare. An improved ability to screen and detect false signalers would direct the market toward an honest equilibrium and protect social welfare. The abundance of charity resources also has implications, including a phenomenon that is analogous to bank

run in case of scarcity where more strategic signalers decide to emit false victim signals to capture a portion of the diminishing resources.

Further consideration of the potential psychological cost of authentic supplication and being screened by the charity reveals an interesting lemon equilibrium in which the authentic individuals who suffer from victimhood shy away from the victims' market for fear of being misjudged as strategic or hedonic signalers (FVS). This is analogous to the well-documented phenomenon of lemon cars driving out high-quality cars (Akerlof, 1970).

We explore and discuss a plethora of other settings, including multiple equilibria, and welfare implications. Our model extension incorporating an extra psychological utility of cheating for the hedonic signaler results in more dishonest equilibrium when the psychological utility of deception increases. This insight has practical implications especially in an age where people might use more sophisticated tactics of deception that reward the deceiver and impose costs on others. When the deceiver not only receives material rewards, such as charity money, but also gains psychological utility from being able to successfully exploit a trusting benefactor, the society is more likely to unravel into an equilibrium where more of the strategic and hedonic signalers would choose to false signal at the expense of the honest signalers. Our model framework has broader applications (e.g., grant and job screening, etc.) and implications.

It is worth noting that we elect not to refine the multiple equilibria in our model. While many articles employ various sophisticated refinement approaches to obtain a unique equilibrium, we believe that doing so is not in line with our research interests. Instead, our focus is on providing economic mechanisms for how counterfeit victims worsen social welfare. We agree that the honest equilibrium is a Pareto-dominant equilibrium as it enhances utilities across all agent types, it is nonnegligible that the dishonest equilibrium always could be a case with positive possibility, especially when the proportion of counterfeit victims increases. The existence of the dishonest equilibrium is nontrivial, as it is commonly observed in reality (see examples in Introduction as evidence). Therefore, for economic relevance and significance, we do not refine the dishonest equilibrium. This is analogous to placing limited emphasis on refining the undesirable (run) equilibrium in bank run models, which share some similarities with the nature of our multiple

equilibria. One strong support comes from the Nobel prized work by Dybvig (2023). Referring to their seminal work (Diamond and Dybvig, 1983), Dybvig (2023) commented that, "*Multiple equilibria are still shunned to some extent, and there are still plenty of papers that make strong and sometimes elaborate assumptions to get rid of multiple equilibria. We should remember that the comparative statics in these models come from the strong assumptions, not from economics*" (p.19). We highly agree that in our setting, "*having multiple equilibria is part of the economic content of the paper, not a defect in the model*" (p.19).

That said, a potential direction for future research in need of deriving a unique equilibrium is to employ global game techniques, also known as regime-change game, typically by allowing the agent to observe a private noisy signal that can coordinate actions among agents (see as an example, Goldstein and Pauzner, 2005). This could enable further exploration of how charitable organizations choose the optimal contract (payment structure) to prevent the marketplace from being stuck in the dishonest equilibrium and, consequently, to maximize social welfare. Furthermore, other interesting results could surface should one allow the funding and proportions of different types of individuals to differ from period to period. Last but not the least, while our research makes one of the first strides to explore and explain the victims' market in a collective reputation framework, it does not encapsulate all potential complexities of the players and market designs yet. Modeling the motivations of the donors, the principal-agent dichotomy between the donors and charity organizations, different arrival rates of victim state, and the sponsor-victim dynamics overtime could all be fruitful directions for future research.

References

Abeler J, Becker A, Falk A (2014) Representative Evidence on Lying Costs. *J. Public Econ*. 113: 96–104.

Abraham J, Prayoga T, Murti K, Azizah A, Krishti N S, Fajrianti, S. P., ... & Manurung R H (2022) How Everyday Counterfeit Behavior That Disrupts Self Authenticity Might Lead to Corruption Tendencies. *Psychol Res. Behav. Management*. 15: 637-663.

Akerlof G (1970) The Market for Lemons: Quality Uncertainty and the Market Mechanism. *Quart. J. Econom.* 84(3): 488-500.

Akerlof G (1976) The Economics of Caste and of the Rat Race and Other Woeful Tales. *Quart. J. Econom*. 90(4): 599-617.

Alan S, Ertac S, Gumren M (2020) Cheating and Incentives in a Performance Context: Evidence from a Field Experiment on Children. *J. Econom. Behav. Organ.* 179: 681-701.

Amato C, Gino F, Montinari N, Sacco P (2020) Cheating, Inequality Aversion, and Appealing to Social Norms. *J. Econom. Behav. Organ.* 179: 767-776.
Aquino, K., & Becker, T. E. (2005). Lying in negotiations: How individual and situational factors influence the use of neutralization strategies. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, *26*(6), 661-679.
Aquino, K., & Bradfield, M. (2000). Perceived victimization in the workplace: The role of situational factors and victim characteristics. *Organization Science*, *11*(5), 525-537.
Atanasov P, Dana J (2011) Leveling the Playing Field: Dishonesty in the Face of Threat. *J. Econom. Psychol.* 32(5): 809-817.
Avieli, H. False Allegations of Domestic Violence: A Qualitative Analysis of Ex-Partners' Narratives. *J Fam Viol* **37**, 1391–1403 (2022). https://doi.org/10.1007/
Bénabou R, Jaroszewics A, Loewenstein G (2022) It Hurts to Ask. NBER Working Paper (w.30486).
Berger F, Blind K, Cuntz A (2012) Risk Factors and Mechanisms of Technology and Insignia Copying – A First Empirical Approach. *Res. Policy*. 41(2), 376-390.
Berman, J. Z., Barasch, A., Levine, E. E., & Small, D. A. (2018). Impediments to effective altruism: The role of subjective preferences in charitable giving. *Psychological science*, *29*(5), 834-844.
Bhargava S, Manoli D (2015) Psychological frictions and the incomplete take-up of social benefits: Evidence from an irs field experiment. *Amer. Econ. Rev.* 105(11): 3489–3529.
Birkelund J, Cherry TL (2020) Institutional Inequality and Individual Preferences for Honesty and Generosity. *J. Econom. Behav. Organ.* 170: 365-361.
Blumenthal M, Erard B, Ho C (2005) Participation and compliance with the Earned Income Tax Credit. *National Tax Journal*. 58(2): 189–213.
Celse J, Max S, Steinel W, Soraperra I, Shalvi S (2019) Uncertain Lies: How Payoff Uncertainty Affects Dishonesty. *J. Econom. Psychol.* 71: 117-125.
Chao M, Fisher G (2022). Self-interested giving: the relationship between conditional gifts, charitable donations, and donor self-interestedness. *Management Sci*. 68(6): 4537-4567.
Christie, R., Geis, F. L. (1970). Chapter I-Why Machiavelli. *Studies in machiavellianism*, 1-9.
Coalition Against Insurance Fraud. (2017). By the numbers: Fraud statistics. Retrieved from http://www.insurancefraud.org/statistics.htm#1 on Nov. 13, 2024.
DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., Epstein, J. A. (1996). Lying in Everyday Life. *Journal of Personality and Social Psychology*, 70(5), 979-995.
Darby M, Karni E (1973) Free Competition and the Optimal Amount of Fraud. *J. Law Econom*. 16(1): 67–88.
Diamond DW, Dybvig PH (1983) Bank Runs, Deposit Insurance, and Liquidity. *J. Polit. Econom*. 91(3): 401–419.
Dong B, Dulleck U, Torgler B (2012) Conditional corruption. *J. Econom. Psychol.* 33(3): 609-627.
Dybvig PH (2023) Nobel Lecture: Multiple Equilibria. *J. Polit. Econom*.131(10):2597-2622.
Effron DA, Bryan CJ, Murnighan JK (2015) Cheating at the End to Avoid Regret. *J. Pers. Soc. Psychol*. 109(3): 395-414.
Exley CL (2020) Using charity performance metrics as an excuse not to give. *Management Sci*. 66(2): 553-563.
Fink C, Maskus K, Qian Y (2016) The Economic Effects of Counterfeiting and Piracy: A Review and Implications for Developing Countries. *World Bank Research Observers.* 31(1): 1-28.
Fishman A, Finkelstein I, Simhon A, Yacouel N (2018) Collective Brands. *Int. J. Ind. Organ*. 59: 316-339.
Food and Nutrition Service (FNS). (2020). Estimates of state Supplemental Nutrition Assistance Program participation rate in 2017. U.S. Department of Agriculture. https://www.fns.usda.gov/SNAP/estimates-state-snap-participation-rates-2017
Gao SY, Lim WS, Tang C (2017) The impact of the potential entry of copycats: Entry conditions, consumer welfare, and social welfare. *Decision Sci*. *48*(4): 594-624.

Gao Y (2018) On the use of overt anti-counterfeiting technologies. *Marketing Sci. 37*(3): 403-424.
Gao Y, Wu Y (2023) Regulating probabilistic selling of counterfeits. *Management. Sci*. 69(8):4363-4971
Goldstein I, Pauzner A (2005) Demand-Deposit Contracts and the Probability of Bank Runs. *J. Finance*. 60(3): 1293-1327.
Giannarelli L (2019) What was the TANF participation rate in 2016? (pp. 1–2). Urban Institute.
Gino F, Norton MI, Ariely D (2010). The counterfeit Self: The Deceptive Costs of Faking It. *Psychol. Sci*. 21(5):712–720.
Glenn, A. L., Raine, A. (2014). Psychopathy: An Introduction to Biological Findings and Their Implications (Vol. 1). NYU Press.
Grolleau G, Evon J, Qian Y (2022) How fine wine producers can make the best of counterfeiting. *Strategic Change*. *31*(5), 515-522.
Grossman G, Shapiro C (1988) Counterfeit Product Trade. *Amer. Econom. Rev*. 78(1): 59-75.
Hare, R. D. (2006). Psychopathy: A Clinical and Forensic Overview. *Psychiatric Clinics*, 29(3), 709-724.
Harun A, Beldram NAAR, Suki NM, Hussein Z (2012) Why Customers Do Not Buy Counterfeit Luxury Brands? Understanding the Effects of Personality, Perceived Quality and Attitude on Unwillingness to Purchase. *Labuan e-Journal of Muamalat and Society (LJMS)*, 6:14-29.
Heckman JJ, Galaty B, Tian H (2023). The Economic Approach to Personality, Character and Virtue. National Bureau of Economic Research (No. w31258).
Hidayat A, Diwasasri AHA (2013) Factors Influencing Attitudes and Intention to Purchase Counterfeit Luxury Brands Among Indonesian Consumers. *Int. J. Marketing Stud*. *5*(4): 143-151.
Jiang S, Villeval MC (2024) Dishonesty as a collective-risk social dilemma. *Economic Inquiry*. 62(1): 223-241
Kalyoncuoglu S, Sahin B (2017) Moderating Role of Materialism in the Effect of Perceived Value on Purchase Intention of Counterfeits of Luxury Brands. *Int. J. Marketing Stud*. 9(4): 76-88.
Kuksov D, Wang K (2013) A model of the "it" products in fashion. *Marketing Sci.* 32(1): 51-69.
Kumar, A., & Chakrabarti, S. (2023). Charity donor behavior: A systematic literature review and research agenda. *Journal of Nonprofit & Public Sector Marketing*, *35*(1), 1-46.
Levin J (2009) The Dynamics of collective reputation. *The B.E. J. Theor. Econom*. 9(1): Article 27.
Li F, Yi Z (2017) Counterfeiting and Piracy in Supply Chain Management: Theoretical Studies. *J. Bus. Ind Marketing*. 32(1): 98-108.
Liang SJS, Qiong W, Jian W (2012) The Relationship Between the Willingness of Buying Counterfeit Goods and Consumer Personality Traits. *ICPM. Crisis Management in the Time of Changing World.*
Lu S, Wang X, Bendle N (2020) Does piracy create online word of mouth? An empirical analysis in the movie industry. *Management. Sci*, 66(5): 2140-2162.
Macquarrie, John (1972). *Paths in Spirituality*.
Mbura S, Uiso H, Kapaya S (2020) Assessment of Level of Formal Education on Perceived Behavioural Control and Consumers' Purchase Intention of Counterfeit Apparel Products in Tanzania. *Int. J.Bus. Manag. Econom. Rev*. 3(6): 01–17.
Mohammad B, Heidarimoghadam R, Saeidnia H, Zemestani M (2020) Relationship Between Personality Characteristics and Attitude Toward Purchase of Counterfeit goods in the Iranian Population. *Cogent Psychol*. 7:1.
Moore C (2009) Psychological Processes in Organizational Corruption. *Psychological Perspectives on Ethical Behavior and Decision Making*. 35–71.
Neeman Z, Öry A, Yu J (2019) The Benefit of Collective Reputation. *RAND J. Econom*. 50 (4): 787-821.
Ok E, Qian Y, Strejcek B, Aquino K (2020) Signaling Virtuous Victimhood as Indicators of Dark Triad Personalities. *J. Pers Soc Psychol*. 120(6):1634-1661.
Pate J (2018) Temptation and Cheating Behavior: Experimental Evidence. *J. Econom. Psychol.* 67: 135-148.
Pfattheicher S, Schindler S, Nockur L (2019) On the Impact of Honesty-Humility and a Cue of Being Watched on Cheating Behavior. *J. Econom. Psychol*. 71: 159-174.
Qian Y (2008) Impacts of Entry by Counterfeiters. *Quart. J. Econom.* 123(4), 1577–1609.

_____ (2014) Brand Management and Strategies against Counterfeits. *J. Econom. Management Strategy.* 23(2): 317-343.

_____ (2015) Counterfeiters: Foes or Friends? How Counterfeits Affect Sales by Product Quality Tier. *Management. Sci*. 60(10): 2381-2400.

Qian Y, Gong Q, Chen Y (2015) Untangling Searchable and Experiential Quality Responses to Counterfeiting. *Marketing Sci.* 34(4): 522-538.

Rosenbaum SM, Billinger S, Stieglitz N (2014) Let's Be Honest: A Review of Experimental Evidence of honesty and Truth-telling. *J. Econom. Psychol*. 45: 181-196.

Saak AE (2012) Collective Reputation, Social Norms, and Participation. *Amer. J. Agric. Econom*. 94(3): 763-785.

Schmitz J (2021). Is charitable giving a zero-sum game? The effect of competition between charities on giving behavior. *Management. Sci*. 67(10): 6333-6349.

Sullivan, D., Landau, M. J., Branscombe, N. R., Rothschild, Z. K. (2012). Competitive Victimhood As a Response to Accusations of Ingroup Harm Doing. *Journal of Personality and Social Psychology*, 102(4), 778-795.

Swami V, Chamorro-Premuzic T, Furnham A. (2009) Faking it: Personality and Individual Difference Predictors of Willingness to Buy Counterfeit Goods. *Journal of Socio-Economics*. *38*(5): 820-825.

Tirole J (1996) Theory of Collective Reputations (with Applications to the Persistence of Corruption and to Firm Quality). *Rev. Econom. Studies*. 63(1): 1-22.

Villeval, M.C. (2024). The social determinants of unethical behavior. In A. Bucciol and S. Quercia (Eds.) Research Handbook on Unethical Behavior. Edward Elgar Publishing Ltd.

Wilcox K, Kim H, Sen S (2009) Why do Consumers Buy Counterfeit Luxury Brands? *J. Marketing Res*. 46(2): 247-259.

Winfree J A, McCluskey J J (2005) Collective Reputation and Quality. *American Journal of Agricultural Economics*. 87 (1): 206-213.

Yoganarasimhan H (2012) Cloak or flaunt? The fashion dilemma. *Marketing Sci*. 31(1): 74-95.

Young, I. F., Sullivan, D. (2016). Competitive Victimhood: A Review of the Theoretical and Empirical Literature. *Current Opinion in Psychology*, 11, 30-34.

Zaloznaya M (2014) The Social Psychology of Corruption: Why It Does Not Exist and Why It Should. *Sociology Compass*. 8(2):187–202.