

Fluet, Claude; Friehe, Tim

Article — Published Version

Optimal law enforcement when individuals are either moral or norm followers

Journal of Public Economic Theory

Provided in Cooperation with:

John Wiley & Sons

Suggested Citation: Fluet, Claude; Friehe, Tim (2023) : Optimal law enforcement when individuals are either moral or norm followers, Journal of Public Economic Theory, ISSN 1467-9779, Wiley, Hoboken, NJ, Vol. 26, Iss. 1, <https://doi.org/10.1111/jpet.12669>

This Version is available at:

<https://hdl.handle.net/10419/288221>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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



<http://creativecommons.org/licenses/by-nc/4.0/>

Optimal law enforcement when individuals are either moral or norm followers

Claude Fluet¹ | Tim Friehe² 

¹CRREP, CRED, Université Laval, Québec City, Canada

²CESifo, EconomiX, University of Marburg, Marburg, Germany

Correspondence

Tim Friehe, CESifo, EconomiX, University of Marburg, Marburg, Germany.

Email: tim.friehe@uni-marburg.de

Abstract

Moral rules and social norms influence whether individuals break the law. We characterize optimal law enforcement when some individuals obey internalized moral rules and implement social norms for other individuals who prefer to comply with them. Moral individuals and norm followers are linked via the endogenous social norm and this induces the social planner to create an expected sanction for norm setters that is higher than without the link. The optimal expected sanction for moral individuals is higher than the one for norm followers if the moral rule is weak but the reverse ranking is true when it is strong.

KEYWORDS

crime, law enforcement, morality, social norm

1 | INTRODUCTION

Behavioral research documents that informal institutions guide decision-making (e.g., Zamir & Medina, 2008). Moral rules are propagated by parents, educators, and others to *influence the behavior* of others who may later suffer guilt after doing the wrong thing and experience virtue after doing the right thing (e.g., Shavell, 2002). In addition, concerns about compliance with social norms provide (sometimes decisive) incentives for individual choice (e.g., Frank, 2020; Krupka & Weber, 2013; Posner & Rasmusen, 1999). Kimbrough and Vostroknutov (2016), for example, explain that people evaluate behavior by comparing it to a socially defined normative standard, and judge their own behavior according to its conformity to the norm. This concern

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© The Authors. *Journal of Public Economic Theory* published by Wiley Periodicals LLC.

about norm compliance may be driven by, for example, a desire to self-signal, be seen as a respected member of society, and avoid stigmatization.

The importance of morality and social norms as determinants of behavior is undisputed. However, they are rarely jointly incorporated into economic analyses of optimal policy (e.g., Bowles & Hwang, 2008; Stringham, 2011). Apart from the exceptions discussed in Section 2, this is also true for an area where the important role of morality and social norms seems particularly striking: the domain of the theory of optimal law enforcement.

We analyze the structure of optimal law enforcement when some individuals (denoted “moral individuals” below) are intrinsically motivated to obey moral rules while other individuals (denoted “followers” below) lack a moral compass but seek to comply with the social norm set by moral individuals’ behavior. In our setting, sanctions and detection probabilities may be type specific; that is, the social planner can condition the expected sanction on the individuals’ type. Moral individuals find an offense justifiable only when the gain exceeds a threshold that may represent the extent to which the moral individual internalizes the crime’s social harm. This approach is comparable to the idea of the fault-based rule in Polinsky and Shavell (2007), ensuring that acts are disapproved of only if the actor failed to obey some standard of behavior made concrete via a gain threshold. Moral individuals’ behavior is supposed to create a social norm and thereby directly influences the decision-making of norm followers. In other words, moral individuals act as role models for followers (e.g., Abel & Brown, 2022; Morgenroth et al., 2015). Portman (2019), for example, argues that celebrities (e.g., actors, athletes, politicians) are increasingly seen as providing a moral compass to people. Followers bear noncompliance costs when their behavior is not aligned with the social norm (e.g., Kimbrough & Vostroknutov, 2016), which creates some deterrence of followers on which the social planner may want to capitalize. Since moral individuals’ behavior changes with law enforcement policy, we analyze an *endogenous* social norm.

Many circumstances exist in which role models and norm-following behavior are relevant. Consider, for example, the case of littering. In 1986, two famous Dallas Cowboys players, Randy White and Ed Jones, informed the public that they strongly disapproved of littering on or along Texas highways in the campaign *Don’t Mess with Texas*, thereby seeking to capitalize on their ability to influence others’ behavior. In fact, in the context of littering, it has been shown that people are susceptible to social influence (e.g., Cialdini et al., 1990). Another example stems from low-income neighborhoods. There, it seems particularly important to guide the behavior of role models using policy because many followers are susceptible to evidence of role-models’ antisocial behavior (e.g., Cheung, 2023; Hurd et al., 2011).

Our setup implies that the deterrence of moral individuals may have a *double dividend* via its direct effect on moral individuals and the indirect effect on the deterrence of followers via the variation in the social norm. However, when moral individuals’ preferences exclude offenses at small criminal gain levels, the social planner may choose to exploit this tendency and to spend exclusively on creating the followers’ expected sanction.

We find that the social planner implements an expected sanction for moral individuals that is higher than the one in the benchmark case without followers if moral rules are too weak. This may seem paradoxical because “better citizens” receive a harsher treatment by law enforcement authorities, but results from the behavioral spillover via the social norm. Otherwise, that is, if moral rules are strong, then the social planner chooses an expected sanction of zero for moral individuals and a positive one for followers. The expected sanctions for moral individuals and norm followers differ in all circumstances. Importantly, we show that the *double-dividend* logic also applies to the circumstances in which the social planner chooses



to implement a positive expected sanction for moral individuals. In other words, for some levels of intrinsic motivation, it is socially optimal to not at all invest in moral individuals' expected sanction when followers are absent but socially optimal to invest at the inflated level when they are present.

Our analysis shows that the structure of law enforcement should take into account prevailing moral rules and social norms to attain maximal welfare. These findings are consistent with, for example, Posner (2000) stating that legal rules must depend on the way that informal rules influence decision-making.

The rest of the paper is structured as follows. We discuss the related literature in the next section. Section 3 presents the model. Section 4 describes privately optimal choices and socially optimal law enforcement policy. Section 5 concludes.

2 | RELATED LITERATURE

Our paper contributes to the literature on the theory of optimal law enforcement which is surveyed by Polinsky and Shavell (2007), for example. In our analysis, we aim at characterizing the optimal level of the sanction and the optimal level of the detection probability conditional on the potential offender's type and the circumstance as described by the moral individuals' behavioral rules.

We are thus concerned with how morality influences the optimal enforcement of law. In a contribution that builds on the discussion in Shavell (2002), Kaplow and Shavell (2007) study characteristics of an optimal moral system that employs feelings of virtue and guilt. An important trade-off considers the incentive effects of guilt on the one hand and the cost of actually feeling guilt on the other. While the moral rules and noncompliance costs in our framework bear some similarity with their guilt, a key distinction lies in the fact that we assume that acts are perceived as justifiable, that is, as implying no moral wrongdoing, when private gains are sufficiently large compared to the external harm; that is, moral noncompliance costs are not incurred when the own gain exceeds a threshold level. Our approach is compatible, for example, with the idea of a deontological threshold function as discussed in Zamir and Medina (2008); see also the notion of a moral due care level in Deffains and Fluet (2013).

Deffains and Fluet (2020) also assume that norms only prescribe that acts should not be carried out when gains are too low and analyze incentives stemming from reputational concerns building on Bénabou and Tirole (2006). Friehe (2020) and Fluet (2020) deal with a setup in which some individuals try to conform with others' choices regarding a criminal opportunity. The conformity concerns are represented by a cost that is incurred independent of the own gain level whenever the offense is undertaken. Their analysis starts from exogenously fixed and symmetric detection probabilities and focuses on the optimal structure of potentially costly sanctions.

In our paper, moral individuals are intrinsically motivated to undertake the act only if their gains exceed the internalized harm from the act. This intrinsic motivation remains intact but becomes irrelevant when moral individuals' expected sanction exceeds the internalized harm from the act. Wenzel (2004), for example, finds supporting evidence for the implication that a higher expected sanction does not influence deterrence when the individual is guided by a strong moral compass, but will be decisive when the moral rules are weak. There is a literature on the possibility that external incentives crowd out intrinsic motivation (surveyed by Bowles & Polania-Reyes, 2012, for example). Lin and Yang (2006), for example, revisit the puzzle posed

by the contribution of Gneezy and Rustichini (2000) about how parents respond to the introduction and removal of a fine for tardiness by exploring the implications of a psychological cost from wrongdoing that is decreasing with the level of the sanction and the proportion of violators. Their result—that a higher sanction may actually produce less deterrence—is not possible within our setup.

In our framework, followers incorporate moral individuals' violation decisions in their choice about what acts to undertake. The idea that violation decisions may (for different reasons) influence others' compliance choices has been articulated by Kahan (1997) and Bargill and Harel (2001), for example. Falk and Fischbacher (2002) present experimental evidence that individuals are indeed more likely to offend if others offend as well. Outside the lab, contributions in the context of tax evasion, for example, find evidence suggesting that individual violation choices are strongly influenced by beliefs about others' violation choices (e.g., Frey & Torgler, 2007). By following the norm, followers may try to avoid being stigmatized.¹ There is a literature about how stigma produces incentives and is itself influenced by the violation decisions of individuals (e.g., Harel & Klement, 2007; Mungan, 2015).

We contribute to the literature by describing how law enforcement should respond to the presence of moral concerns and social norms. There are some contributions discussing the relationship between formal and informal institutions (e.g., Cooter, 2000; Posner & Rasmusen, 1999). McAdams and Rasmusen (2007) survey the literature on norms in law and economics, and Posner (2000) provides an earlier comprehensive treatment. Whereas we follow much of the literature by not modeling why noncompliance with norms creates a cost to individuals, Acemoglu and Jackson (2017) consider other individuals' motivation to blow the whistle on the violator.

With formal and informal institutions operating at the same time, an important question pertains to their interaction. For example, Zasu (2007) analyzes whether norms and formal sanctions function are substitutes or complements. In our analysis, the expected sanction for followers is assessed to complement the deterrence stemming from any norm noncompliance cost, and the policy maker may decide to reduce the detection probability for moral individuals to zero when moral individuals' internalized rule is sufficiently strong, such that the informal institution substitutes for the formal one.

3 | MODEL

The population of risk-neutral potential offenders consists of two types, M and F . The share of individuals of type M is denoted $q \in (0, 1)$. Each potential offender draws the criminal gain g from the uniform distribution on the unit interval (as in, e.g., Bebchuk & Kaplow, 1992; Garoupa, 1999). This means that $g \in [0, 1]$ and that the distribution of gains is independent of offender type. The act imposes a social harm $h < 1$. Thus, for some individuals $g < h$ and for others $g > h$ holds.² Potential offenders' level of wealth represents the upper bound for the monetary sanctions and is normalized at $w = 1$.

¹In general, the individual motivation to align own choices with those of others can be attributed to different rationales (e.g., Cialdini & Goldstein, 2004). For example, a motive to follow others' conduct can also arise from informational aspects (e.g., Bikhchandani et al., 1992).

²The possibility of criminal gains exceeding social harm is a standard assumption in the law-enforcement literature. Starting with Stigler (1970), there is an ongoing controversy in the enforcement literature about the extent to which the criminal gains of the offender represent social value.

Moral individuals (i.e., individuals of type M) consider their internalized moral rule when choosing whether to offend. Their rule says that the offense is morally acceptable when the gains exceed their moral threshold \tilde{h} but that it represents a severe moral infraction otherwise, where $\tilde{h} \in [0, h]$.³ The threshold \tilde{h} can be interpreted as the moral individuals' perception of social harm or of the extent to which it should be internalized. Moral individuals' payoffs from crime can thus be specified as

$$G_M(g; \tilde{h}) = \begin{cases} g - \tilde{h} & \text{if } g < \tilde{h} \\ g & \text{if } g \geq \tilde{h}. \end{cases}$$

In other words, a crime with $g \geq \tilde{h}$ creates the payoff g while a crime with $g < \tilde{h}$ produces $g - \tilde{h} < 0$, which illustrates how our framework differs from the standard one (e.g., Polinsky & Shavell, 2007). As a result, even in the absence of law enforcement, moral individuals abstain from offending when they perceive that the private gain is offset by the internalized harm (i.e., when $g < \tilde{h}$). Note that this payoff function applies to all moral individuals in the same way before the criminal gain g is drawn (since all moral individuals have the same threshold \tilde{h}). However, some moral individuals will eventually draw a criminal gain that is above the threshold \tilde{h} while others draw one below \tilde{h} .

Lacking their own moral compass, individuals of type F are denoted *followers* since they seek to follow a social norm set by moral individuals. The social norm can be represented by a threshold criminal gain level, meaning that undertaking the act complies with the social norm only if the gain level exceeds this threshold.⁴ Thus, followers consider the offense as socially acceptable when their gains are greater than the threshold level g_N prescribed by the norm.⁵ Because the threshold stems from the endogenous choice of moral individuals, we will return to it below. Noncompliance with the social norm creates a personal cost amounting to θg_N , $\theta \in [0, 1)$, where θ represents the importance attached to norm compliance. This personal cost is similar to the disutility of guilt in Kaplow and Shavell (2007) and applies symmetrically for all followers. We assume that the norm g_N results from how moral individuals behave, that is, it is the cutoff gain level of moral individuals to be explained below. Followers' payoffs from crime are then

$$G_F(g; g_N) = \begin{cases} g - \theta g_N & \text{if } g < g_N \\ g & \text{if } g \geq g_N. \end{cases}$$

As illustrated in Figure 1, we obtain $G_F = g_N$ at $g = g_N$, whereas $\lim_{\varepsilon \downarrow 0} G_F(g_N - \varepsilon; g_N) = (1 - \theta)g_N$. Since $(1 - \theta)g_N > 0$, followers may offend even though their gains are less than the norm.

³For example, Deffains and Fluet (2020) similarly consider a behavioral norm as a threshold level of criminal gains. This specification is compatible with the deontological constraints as conceived by Zamir and Medina (2008), for instance. The fault-based sanctioning rules in Polinsky and Shavell (2007) similarly build on the idea of a critical gain level.

⁴This threshold-based formulation is similar to the fault-based sanctioning rule presented in Polinsky and Shavell (2007) where individuals are free from criminal liability as long as their gain exceeds the threshold level.

⁵This understanding is compatible with the idea of *social proof* as explained by Cialdini (2021), for example.

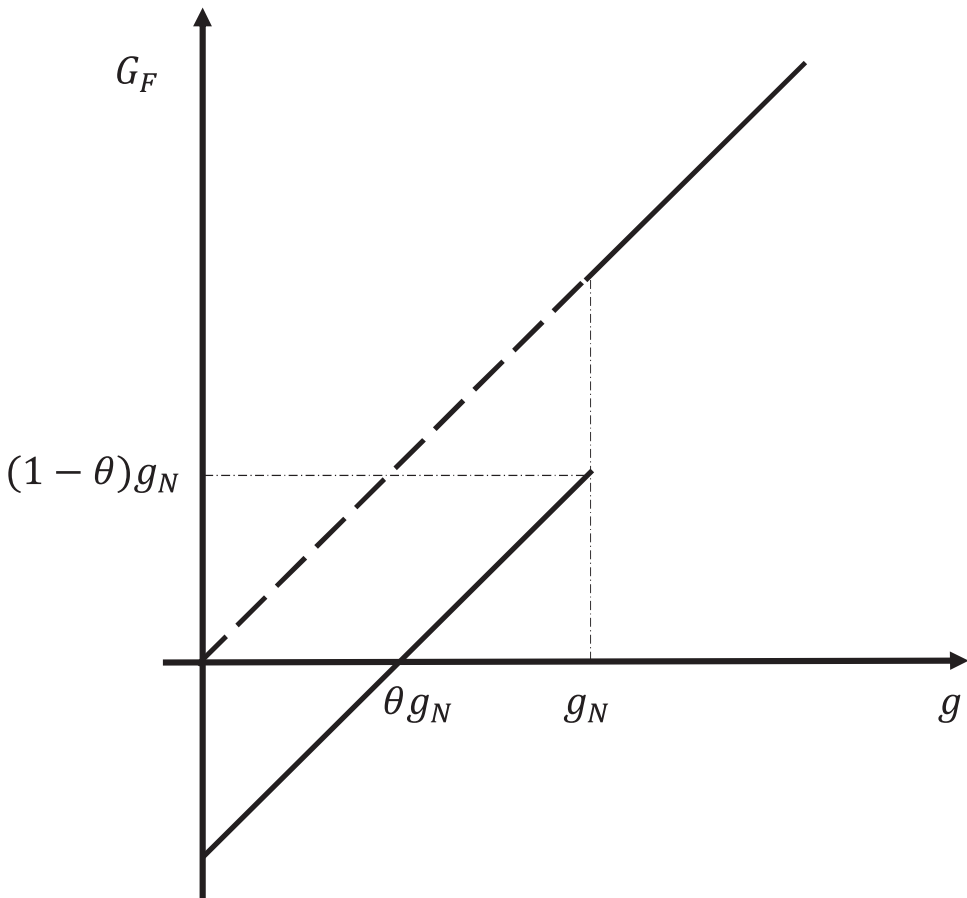


FIGURE 1 Followers' payoffs from crime.

Note the following parallel between G_M and G_F . Moral individuals have a payoff $g - \tilde{h}$ when they offend with a criminal gain below their behavioral standard \tilde{h} . Followers have a payoff $g - \theta g_N$ when they offend with a criminal gain below their behavioral standard g_N . So the specification of payoffs for followers parallels that for moral individuals with the exception that, for followers, the term for the behavioral standard g_N is discounted by $\theta < 1$. This can be explained by the fact that followers' attachment to the norm is assumed to be weaker than the attachment of moral individuals to their internalized moral rules.

The third kind of actor is the social planner who can implement *type-specific* law enforcement. More specifically, we consider a detection probability p_j and the socially costless fine s_j specific to individuals of type j , where $j = F, M$. The overall enforcement cost is

$$C(p_F, p_M; q) = qc(p_M) + (1 - q)c(p_F), \quad (1)$$

where c is the per-capita enforcement cost that is increasing and strictly convex in the level of the detection probability with $c(0) = c'(0) = 0$. Recall that q is the share of moral individuals,

implying that C represents the total enforcement cost for moral individuals who are detected at rate p_M and for followers who are caught at rate p_F .

The timing of the game is as follows: The social planner commits to a law enforcement policy in Stage 1 before potential offenders choose whether or not to offend under the given policy in Stage 2. Detection and sanctioning results in Stage 3 without any more endogenous choices.

4 | ANALYSIS

We first analyze the decision-making of potential offenders for a fixed law enforcement policy. Next, we describe the socially optimal law enforcement policy for our setup.

4.1 | Potential offenders

Moral individuals will undertake the criminal act if

$$g \geq \max\{p_M s_M, \tilde{h}\} = g_N.$$

The level of the moral individuals' threshold g_N that also sets the norm for followers is the criminal gain level at which a moral individual is indifferent between offending and not offending. If $p_M s_M < \tilde{h}$, this critical gain level is defined by the moral individuals' rule. In that case, the moral individual will under no circumstances offend if the criminal gains fall short of \tilde{h} . However, if $p_M s_M > \tilde{h}$, the threshold gain level g_N that applies to moral individuals (and thus the norm that is set for followers) will be equal to the expected sanction for moral individuals. In that scenario, moral individuals are only interested in criminal acts that generate criminal gains strictly exceeding their internalized harm. Also, the level \tilde{h} is in that scenario totally irrelevant for moral individuals' and followers' decision-making.

While followers' crime choice is simultaneous to that of moral individuals, they anticipate the behavioral norm g_N that results from the decision-making of moral individuals.⁶ Followers will offend when

$$g \geq \theta g_N + p_F s_F$$

if their expected sanction is sufficiently small, that is, if $p_F s_F < (1 - \theta)g_N$. In this case, the deterrence level is created by the expected sanction for followers and the cost (e.g., guilt) resulting from noncompliance with the norm. If $p_F s_F$ is at least as large as $(1 - \theta)g_N$, followers undertake the criminal act if

$$g \geq \max\{p_F s_F, g_N\}.$$

⁶This is similar to cases in which potential offenders choose whether to offend anticipating how many other potential offender will violate the law (e.g., Funk, 2005). However, see Section 5.2 for a discussion of the effects of an alternative timing assumption.

Thus, when $(1 - \theta)g_N \leq p_{F,S_F} < g_N$, followers undertake the criminal act only if $g \geq g_N$. When $p_{F,S_F} \geq g_N$, however, the followers' concern for the social norm is irrelevant and their behavior is then governed solely by the expected sanction for followers.

4.2 | Law enforcement policy

The social planner maximizes utilitarian welfare W using the detection probability levels (p_F, p_M) and the sanctions (s_F, s_M) subject to the constraint that sanctions cannot exceed offenders' wealth. Total welfare is

$$W = qW_M + (1 - q)W_F \quad (2)$$

with type-specific welfare defined as

$$W_M = \int_{\max\{p_M s_M, \tilde{h}\}}^1 (g - h)dg - c(p_M), \quad (3)$$

$$W_F = \int_{\max\{p_F s_F, g_N\}}^1 (g - h)dg + \mathbf{1}_{\{p_F s_F < (1-\theta)g_N\}} \int_{\theta g_N + p_F s_F}^{g_N} (g - \theta g_N - h)dg - c(p_F), \quad (4)$$

where $g_N = \max\{p_M s_M, \tilde{h}\}$ is *simultaneously* the deterrence level for moral individuals and the social norm for followers. We use the indicator variable $\mathbf{1}_{\{p_F s_F < (1-\theta)g_N\}}$ equal to 1 if the followers' expected sanction falls short of $(1 - \theta)g_N$ and 0 otherwise.

For any level of expected sanctions, welfare is maximized when sanction levels are set equal to offenders' wealth as this minimizes enforcement costs. With wealth normalized to unity, the socially optimal sanctions are then $s_M^+ = s_F^+ = 1$. Any differentiation of expected sanction levels will result from asymmetric detection probability levels.

To assess socially optimal detection probability levels transparently, we proceed in several steps: First, we explore how the social planner chooses law enforcement policy when all potential offenders are moral. This allows us to characterize the extent to which the social planner will want to rely on the deterrent effect of the moral compass when there are no behavioral spillovers on other types of individuals to consider. Next, we turn to an understanding of optimal policy vis-a-vis norm followers for a given social norm, that is, when we fix the policy toward moral individuals. This part of the analysis identifies to what extent the social planner wants to capitalize on the behavioral spillover between moral individuals and followers. Afterwards, we integrate our insights to arrive at the socially optimal structure of detection probabilities when the social norm is endogenous.

4.2.1 | Moral individuals: When to rely on their moral compass alone

Moral individuals find the offense morally acceptable only if their private gain exceeds their moral threshold. Thus, they are deterred to some extent even without law enforcement. We first

characterize how this level of deterrence compares to the level of deterrence that would be socially optimal for this group of individuals.

If individuals of type M had no moral compass, the social planner would choose p to maximize

$$W_N(p) := \int_p^1 (g - h) dg - c(p), \tag{5}$$

where the subscript N refers to the absence of moral norms. $W_N(p)$ is a strictly concave function and the first-order condition is

$$h - p^* = c'(p^*). \tag{6}$$

The first-order condition shows that $p^* < h$ must hold. The positive marginal enforcement cost makes it optimal that some individuals are underdeterred; that is, some individuals offend when the private gain is less than the social harm. The probability of detection p^* defined in equation (6) will be an important reference level below and induces the welfare level

$$\int_{p^*}^1 (g - h) dg - c(p^*).$$

When individuals are intrinsically motivated not to offend when $g < \tilde{h}$, this creates some deterrence free of enforcement cost. By assumption, intrinsic incentives are insufficient and thus cannot ensure that moral individuals undertake only efficient offenses. Relying on the moral rule yields the welfare level

$$\int_{\tilde{h}}^1 (g - h) dg,$$

which is increasing in \tilde{h} , making clear that the social planner will exclusively rely on the moral compass when \tilde{h} is not too much below p^* . Formal enforcement will be preferred only when the policy with p^* provides sufficient additional deterrence benefits to compensate for the enforcement cost.

Lemma 1. *Welfare W_M is maximized by the probability of detection*

$$\hat{p}_M = \begin{cases} p^* & \text{if } \tilde{h} < h_1 \\ 0 & \text{if } \tilde{h} \geq h_1. \end{cases} \tag{7}$$

where h_1 is the solution to

$$\int_{h_1}^1 (g - h) dg = \int_{p^*}^1 (g - h) dg - c(p^*). \tag{8}$$

Proof. Follows from the above, noting that if formal enforcement does better, then $p > \tilde{h}$ and is therefore set equal to p^* . □

The critical level h_1 must be strictly smaller than the deterrence level resulting with formal enforcement because the indifference formulated in Equation (8) requires that the enforcement cost just offset the improvement regarding deterrence:

$$c(p^*) = \int_{h_1}^{p^*} (h - g) dg.$$

The deterrence level for moral individuals is thus created by formal enforcement and equal to p^* when the moral rule is weak (i.e., when $\tilde{h} < h_1$) and induced by intrinsic incentives alone and equal to \tilde{h} when the moral rule is strong (i.e., when $\tilde{h} \geq h_1$).

The policy maker tolerates a deterrence level that is smaller than the one implemented in the standard framework without moral rule when it comes without enforcement cost, but is willing to spend on enforcement and discard the incentive effects from internalized moral rules when the deviation in terms of deterrence levels becomes too high.

The level of welfare W_M that uses the optimal detection probability \hat{p}_M remains constant when \tilde{h} increases until h_1 is reached and increases afterwards. The welfare from individuals of type M is smooth at h_1 , while a discontinuity in their level of deterrence arises.

4.2.2 | Followers: How to select their detection probability for a given norm

In this section, we explain how the optimal detection probability for norm followers changes with the social norm set by moral individuals. We take the social norm g_N as given and choose p_F to maximize $W_F(p_F, g_N)$ as defined in Equation (4), setting s_F equal to one.

As was true for the enforcement policy regarding moral individuals (Section 4.2.1), the social planner now must decide whether the intrinsic motivation stemming from followers' concern for norm compliance is such that the optimal enforcement policy should capitalize on it or not. When $p_F < (1 - \theta)g_N$, both noncompliance costs θg_N and the expected sanction p_F produce followers' deterrence level. Accordingly, an investment of $p_F = (1 - \theta)g_N$ induces a deterrence level amounting to g_N . In analogy to the argument in Lemma 1, it may be that the norm g_N is too weak and that, as a result, the social planner prefers to implement p^* . To the contrary, an exogenous norm g_N may also be too strong. This results when $p_F = (1 - \theta)g_N > p^*$. In this scenario, the trade-off between deterrence and enforcement cost is again best attained at $p_F = p^*$. In the intermediate cases, the policy maker indeed capitalizes on the deterrence created by the followers' concern about compliance with the norm.

For a given norm g_N stemming from moral individuals' conduct, we can characterize the optimal detection probability for followers as follows (the proof is presented in our appendix):

Lemma 2. *For a given social norm g_N , $W_F(p_F, g_N)$ is maximized (i) by $p_F = p^*$ if $g_N < h_{11}$ for some $h_{11} \in (h_1, p^*)$; (ii) by $p_F = (1 - \theta)g_N$ if $g_N \in [h_{11}, p^*/(1 - \theta)]$; (iii) and by $p_F = p^*$ if $g_N > p^*/(1 - \theta)$.*

In the derivation of the critical norm level h_{11} , we compare the welfare obtained with a deterrence level p^* and enforcement cost $c(p^*)$ to a circumstance with strictly positive but smaller enforcement cost $c((1 - \theta)h_{11})$. This enforcement cost is smaller because the deterrence of followers now includes the influence from the norm set by moral individuals. As a result, the critical level h_{11} must fall inside the interval (h_1, p^*) as

$$c(p^*) = \int_{h_{11}}^{p^*} (h - g)dg + c((1 - \theta)h_{11}),$$

where we know that $c(p^*) = \int_{h_1}^{p^*} (h - g)dg$ from above.

Figure 2 summarizes the optimal policy with respect to norm followers. When the social norm is sufficiently undemanding, that is, $g_N < h_{11}$, it becomes irrelevant in the optimal enforcement policy concerning followers, who are monitored with probability p^* as in the standard model. This is analogous to the irrelevance of moral concerns in the optimal enforcement policy concerning moral individuals when $\tilde{h} < h_1$. When the social norm is sufficiently demanding but not too much so, that is, $g_N \in [h_{11}, p^*/(1 - \theta)]$, the social planner capitalizes on followers' norm noncompliance costs by complementing them with a level of formal incentives $p_F = (1 - \theta)g_N$ just sufficient to induce followers to abide by the norm; that is, the followers' threshold is then $\hat{g}_F = \theta g_N + p_F = g_N$. This means that, in this range, the social planner accepts a deterrence level above or below p^* to benefit from the enforcement cost saving. When the social norm is very demanding (i.e., $g_N > p^*/(1 - \theta)$), inducing followers to comply with the norm is undesirable and followers are again monitored with probability p^* .

4.2.3 | Socially optimal detection probabilities

In Section 4.2.1, we derived \hat{p}_M neglecting spillovers from the deterrence of moral individuals to that of followers. This behavioral spillover can present an additional benefit which makes us conjecture that the optimal detection probability p_M^+ is never smaller than the level \hat{p}_M . Whereas we expect $p_M^+ = \hat{p}_M = 0$ for \tilde{h} sufficiently large, we may have $p_M^+ > \hat{p}_M$ for smaller levels of \tilde{h} .

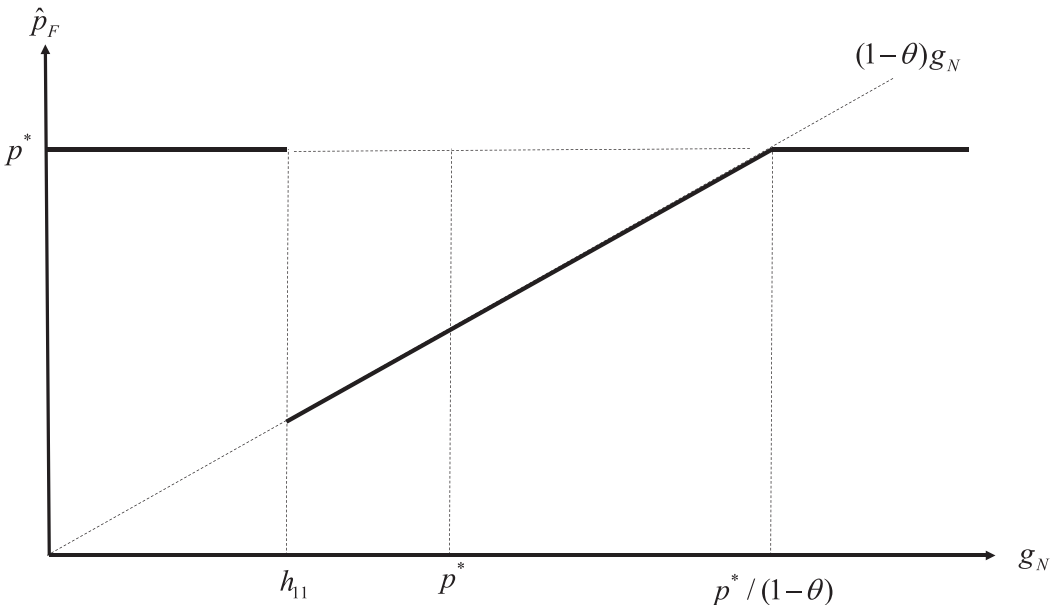


FIGURE 2 Followers' detection probability for varying norms.

To determine the critical moral threshold at which the social planner would switch from p_M^+ to a no enforcement policy with respect to moral individuals, we proceed as in Section 4.2.1 and first consider the case where $\tilde{h} = 0$ so that the social norm is then determined by the detection probability for moral individuals. Using Lemma 2, the socially optimal detection probability for followers as a function of the detection probability for moral individuals is then given by

$$\hat{p}_F(p_M) := \begin{cases} (1 - \theta)p_M & \text{if } h_{11} \leq p_M \leq p^*/(1 - \theta), \\ p^* & \text{if } p_M > p^*/(1 - \theta) \text{ or } p_M < h_{11}. \end{cases} \quad (9)$$

In other words, the detection probability set for moral individuals will influence the detection probability for followers only when the former is set in some intermediate range. With expression (9), we can state welfare as a function of the detection probability for moral individuals as its single argument:

$$\overline{W}(p_M) := qW_N(p_M) + (1 - q)W_F(\hat{p}_F(p_M), p_M), \quad (10)$$

where $W_N(p_M)$ is as defined in (5). Denote the maximand of $\overline{W}(p_M)$ by p^{**} . In Lemma 3, we will show that $p^{**} > p^*$. This is intuitive. If moral individuals influence followers' behavior via the social norm (i.e., their critical gain level), the social planner would want to deter these individuals more than when there is no behavioral spillover on followers.

Apart from the question concerning the detection probability for moral individuals, it is key to assess for what values of \tilde{h} the policy maker prefers a detection probability $p_M = 0$. Similar to our approach in Section 4.2.1 involving the equality of welfare levels in (8), we need to characterize the level of welfare that would be obtained with a no enforcement policy with respect to moral individuals. Welfare then depends on these individuals' moral compass and is equal to

$$Q(\tilde{h}) := q \int_{\tilde{h}}^1 (g - h) dg + (1 - q)W_F(\hat{p}_F(\tilde{h}), \tilde{h}). \quad (11)$$

This expression takes into account the behavioral spillover effect on followers by its use of $g_N = \tilde{h}$. It can be shown that (the proof is presented in our appendix):

Lemma 3. *It holds that the maximand of $\overline{W}(p_M)$ denoted p^{**} lies inside (p^*, h) , $Q(\tilde{h})$ is increasing for $\tilde{h} \leq p^{**}$, and that there is a unique $h_2 \in (h_1, p^{**})$ solving $Q(\tilde{h}) = \overline{W}(p^{**})$.*

The level of Q increases with the internalized harm level because deterrence at p^{**} is desirable even when positive marginal enforcement costs are incurred. Naturally, this will also be true when marginal enforcement costs are incurred only for the deterrence of followers. Regarding the level of the internalized harm level that induces indifference between no formal enforcement for moral individuals and enforcement using detection probability at p^{**} , it is intuitive that the critical level h_2 exceeds the level h_1 derived for moral individuals in isolation. It is now more important to implement a positive detection probability for moral individuals because of the behavioral spillover on followers.

We can now state our main result, which describes the optimal enforcement policy as a function of moral individuals' internalized rule.

Proposition 1. *The optimal law enforcement includes maximal sanctions $s_M^+ = s_F^+ = w$ and the detection probabilities (p_M^+, p_F^+) . If moral individuals' internalized rule is weak (i.e., if $\tilde{h} < h_2$), (ia) $p_M^+ = p^{**} > p_F^+ = (1 - \theta)p^{**}$ if $p^{**} \leq p^*/(1 - \theta)$, and (ib) $p_M^+ = p^{**} > p_F^+ = p^*$ if $p^{**} > p^*/(1 - \theta)$. If the moral rule is strong (i.e., if $\tilde{h} \geq h_2$), $p_M^+ = 0$ and (iia) $p_F^+ = p^*$ if $\max\{\tilde{h}, h_2\} < h_{11}$, (iib) $p_F^+ = (1 - \theta)\tilde{h}$ if $\tilde{h} \in [h_{11}, p^*/(1 - \theta)]$, and (iic) $p_F^+ = p^*$ if $\tilde{h} > p^*/(1 - \theta)$.*

Proof. If $\tilde{h} < h_2$, the policy maker implements a positive detection probability for moral individuals and the results from Lemmas 2 and 3 apply. If $\tilde{h} \geq h_2$, the policy maker chooses a detection probability $p_M^+ = 0$ and the results from Lemma 2 apply. \square

Optimal law enforcement should be tailored to the individual's type via asymmetric detection probabilities. However, whereas the levels of detection probabilities differ in all circumstances, the induced level of deterrence can be symmetric for moral individuals and followers in some circumstances. Symmetric deterrence levels would seem appropriate because the external cost h is independent of offender type and the socially relevant noncompliance costs are fully internalized by the decision-maker. However, the behavioral spillover introduces an important asymmetry between offender types. We highlight some implications from this spillover in:

Corollary 1 (The Spillover Effect). *When we compare the optimal enforcement for moral individuals with and without the spillover to followers, we find that (i) for any $\tilde{h} \in (h_1, h_2)$, the social planner relies on the moral individuals' compass and saves on enforcement costs without spillovers but creates an expected sanction for moral individuals (and thus a higher social norm) with spillovers, and that (ii) the level of the detection probability for moral individuals, if positive, is greater with spillovers than without them.*

It may be intuitive that moral individuals require less attention of law enforcement authorities. This shows in our results only when the moral rule is sufficiently strong. Otherwise, however, moral individuals will actually be confronted with a greater detection probability. This paradoxical result can be rationalized by referring to the social norms that guides followers violation choices and which can be raised by stronger enforcement for moral individuals.

5 | DISCUSSION

In this section, we discuss implications from two key assumptions. First, we assume that the sanction is socially costly. Second, we assume that moral individuals care about deterrence and strategically influence followers.⁷

⁷For example, Esguerra et al. (2023) provide evidence that people are willing to distort own choices to influence others' decision-making.

5.1 | Costly sanctions

Our main analysis is confined to monetary sanctions that represent pure transfers. However, fines may be costly to collect or the sanction may be an imprisonment term. This can be represented by a sanctioning cost parameter k per unit of the sanction actually imposed on offenders.

If we return to the problem of finding the optimal policy for type M individuals in isolation, assuming that these individuals have no moral compass as in (5), we would now first seek to

$$\max_{p,s} W_N(p, s) = \int_{ps}^1 (g - h) dg - kps(1 - ps) - c(p) \quad (12)$$

subject to a constraint $s \leq \tilde{w}$, where \tilde{w} represents either wealth (in the fine interpretation) or years of life remaining (in the prison interpretation). When the sanction s is imposed on violators, the social cost is ks . For instance, if the sanction is a fine, this could represent a collection cost per unit fine of k , where typically $k < 1$. If s is a prison term, then the social cost is the loss suffered by the individual plus the cost to the state, hence $k > 1$.

Despite the social cost of sanctions, it remains optimal to impose maximal sanctions. The argument is that any level of deterrence $\hat{g} = ps$ yields a social sanction cost that depends only on \hat{g} while the detection cost $c(\hat{g}/s)$ decreases with the sanction; see Polinsky and Shavell (2007).⁸

Given the maximum sanction, the marginal change in welfare following a raise in the detection probability is

$$\frac{\partial W_N}{\partial p} = \underbrace{(h - p) - c'(p)}_A + \underbrace{pk - k(1 - p)}_B. \quad (13)$$

In Section 4.1, we have seen the marginal effects from term A. From term B, we find that the marginal deterrence benefit is greater by pk , as deterring additional individuals saves expected punishment cost, and the marginal cost is greater by $k(1 - p)$ as a greater sanction is imposed on undeterred and detected people. The second-order derivative is $\partial^2 W_N / \partial p^2 = 2k - 1 - c''(p)$ and is therefore negative if k is not too large or the marginal cost $c'(p)$ is increasing fast enough, in which case an interior solution is possible. Otherwise, the solution is either no deterrence $p = 0$ or maximal enforcement effort $p = 1$.

When W_N is concave in p , the optimal policy with moral individuals and followers is qualitatively similar to our previous results, except that the relevant thresholds are different. Regarding Lemma 1, we find that the presence of punishment cost induces the social planner to prefer relying on the moral individuals' rules for a wider range of \tilde{h} . Specifically, we obtain a lower threshold for \tilde{h} at which the social planner prefers to switch to the detection probability that solves $\partial W_N / \partial p = 0$.

Regarding Lemma 2, we find that the benefit from the behavioral spillover is similar to the enforcement cost advantage discussed in Section 4.2, except that the benefit now also includes savings in punishment costs. When the deterrence of followers capitalizes on the social norm

⁸When p is exogenously fixed, however, welfare may be nonconcave in s as shown in Kaplow (1990), yielding corner solutions with respect to s . See D'Antoni et al. (2023) for a discussion about which distributions induce extreme costly sanctions.

allowing for a lower p_F , both the detection cost and the punishment cost will be smaller than in a scenario without a behavioral spillover from moral individuals to followers.

In summary, in the well-behaved case, the qualitative pattern of optimal law enforcement in light of moral concerns and norms remains robust to the consideration of costly sanctions. However, when maximizing W_N yields corner solutions, we get corner solutions in the overall problem as well. For instance, when W_N is maximized at $p = 0$, it is then optimal to have no formal sanctions for either moral individuals or followers.

5.2 | Deterrence-Concerned moral individuals

Moral individuals bear internalized cost \tilde{h} when they engage in the crime despite small gains. In this section, we analyze the case in which moral individuals apply that normative standard to followers as well, that is, the case in which moral individuals bear a cost if followers engage in crime despite $g < \tilde{h}$. This can create a rationale for moral individuals' strategic choice of their deterrence threshold.

To make the argument as simple as possible, we assume that there is only one moral individual who moves before many followers, and is aware that the followers incorporate the moral individual's deterrence threshold as a social norm in their decision-making. Consider a policy such that $(p_M, p_F) = (0, p_F^\ell)$ where $\theta\tilde{h} + p_F^\ell < \tilde{h}$ holds true. In this scenario, the moral individual understands that followers will offend although their gains fall short of \tilde{h} . We can imagine that the moral individual seeks to

$$\max_{\hat{g}_N \geq \tilde{h}} \int_{\hat{g}_N}^1 g dg + \gamma \int_{\theta\hat{g}_N + p_F^\ell}^{\tilde{h}} (g - \tilde{h}) dg. \tag{14}$$

The parameter γ captures not only to what extent the moral individual cares about the followers' undertaking of crime with gains $g < \tilde{h}$ but also the number of followers that can be influenced by the moral individual's determination (hence γ could be greater than unity). The marginal payoff change for the moral individual is

$$-\hat{g}_N + \gamma\theta\left(\tilde{h} - \left(\theta\hat{g}_N + p_F^\ell\right)\right). \tag{15}$$

It identifies the marginal cost from a higher threshold (as the moral individual abstains from payoff-increasing criminal opportunities) and the marginal benefit (as the moral individual's higher threshold deters followers from some crimes with gains below the internalized cost \tilde{h}). The marginal payoff change may be negative or zero at $\hat{g}_N = \tilde{h}$, meaning that the moral individual would then not choose to strategically raise the followers' norm. A very small cost for noncompliance with the norm as represented by θ exemplifies a possible circumstance; that is, a corner solution for \hat{g}_N results for sure if $\gamma\theta \leq 1$. Otherwise, the threshold level that sets the marginal payoff change equal to zero is

$$\hat{g}_N^*(p_F^\ell) = \frac{\gamma\theta(\tilde{h} - p_F^\ell)}{1 + \gamma\theta^2}. \quad (16)$$

This means that

$$\frac{d\hat{g}_N}{dp_F} = -\frac{\gamma\theta}{1 + \gamma\theta^2} < 0, \quad (17)$$

such that the *artificial* increase of \hat{g}_N above \tilde{h} by the moral individual is a substitute to the deterrence created by the expected sanction for followers.

It is interesting to inquire about the implications these changes would have in the policy stage. If the policy maker anticipates that the moral individual will move next and offset a small p_F by a strategic increase of \hat{g}_N , the problems reads

$$\max_{p_F} W_F = \int_{\hat{g}_N(p_F)+p_F}^{\hat{g}_N(p_F)} (g - \theta\hat{g}_N(p_F) - h)dg + \int_{\hat{g}_N(p_F)}^1 (g - h)dg - c(p_F). \quad (18)$$

The derivative with respect to p_F gives

$$(h - p_F) - c'(p_F) + (h - \hat{g}_N)\theta \frac{d\hat{g}_N}{dp_F} = 0, \quad (19)$$

which indicates that the social planner has an incentive to somewhat rely on the deterrence created by the moral individual (as the marginal incentives are changed by $(h - \hat{g}_N)\theta \frac{d\hat{g}_N}{dp_F} < 0$ as a result of the strategic influence of the moral individual), implying in particular that $p_F < p^*$. Loosely speaking, the marginal welfare gain from increasing p_F is now smaller because doing so decreases the moral individual's strategic incentive to set an example.

6 | CONCLUSION

Many individuals are guided by moral rules and/or a desire to comply with social norms. Optimal law enforcement policy must acknowledge and take advantage of this fact. Our analysis shows that law enforcement policy should closely follow and respond to prevailing moral rules and social norms to attain maximal welfare. Optimal policy will require differentiation and, perhaps contrary to intuition, higher expected sanctions for moral individuals in some circumstances.

Such a desirable differentiation of law enforcement policy may be constrained by practical aspects such as observability of type and other objectives such as equal treatment before the law. With respect to the observability, it is clearly easier to embed differentiation in later stages (e.g., at the conviction stage) where the individual's type is easier to observe than at earlier ones (e.g., the apprehension stage). Whereas screening of heterogeneous types with private information was considered using menus of fines and imprisonment (e.g., D'Antoni et al., 2023), this seems more difficult to achieve with different levels of the detection probability. When equal treatment is of overriding importance, our results at least help to illustrate its welfare cost in the law enforcement context.

In our analysis with an endogenous social norm, we treated the moral threshold and the relative importance of noncompliance with the norm as exogenous parameters. In a more general framework, these aspects can also be influenced by policy (e.g., Bowles & Hwang, 2008; Cooter, 2000; Kahan, 1997; Kaplow & Shavell, 2007). Another interesting avenue for future research lies in considering the possibility of heterogeneity within type, because we assumed that all moral individuals consider the same internalized threshold and all followers have the same noncompliance cost parameter. This would also introduce the possibility that different followers may consider varying norms depending upon which moral individuals' conduct they incorporate in their decision-making.

ACKNOWLEDGMENTS

We gratefully acknowledge the very helpful suggestions received from an anonymous reviewer and the associate editor. In addition, the authors are grateful for the hospitality of Murat Mungan and the George Mason University Law School which both authors were visiting while most of this work was completed. Open access funding enabled and organized by Projekt DEAL.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Tim Friehe  <http://orcid.org/0000-0002-9390-3537>

REFERENCES

- Abel, M., & Brown, W. (2022). Prosocial behavior in the time of COVID-19: The effect of private and public role models. *Journal of Behavioral and Experimental Economics*, 101, 101942. <https://doi.org/10.1016/j.socec.2022.101942>
- Acemoglu, D., & Jackson, M. O. (2017). Social norms and the enforcement of laws. *Journal of the European Economic Association*, 15, 245–295. <https://doi.org/10.1093/jeea/jvw006>
- Bar-Gill, O., & Harel, A. (2001). Crime rates and expected sanctions: The economics of deterrence revisited. *Journal of Law and Economics*, 30, 485–501. <https://doi.org/10.1086/322055>
- Bebchuk, L., & Kaplow, L. (1992). Optimal sanctions when individuals are imperfectly informed about the probability of apprehension. *Journal of Legal Studies*, 21, 365–370. <https://doi.org/10.1086/467910>
- Bénabou, R., & Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review*, 96, 1652–1678. <https://doi.org/10.1257/aer.96.5.1652>
- Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100, 992–1026. <https://doi.org/10.1086/261849>
- Bowles, S., & Hwang, S. H. (2008). Social preferences and public economics: Mechanism design when social preferences depend on incentives. *Journal of Public Economics*, 92, 1811–1820. <https://doi.org/10.1016/j.jpubeco.2008.03.006>
- Bowles, S., & Polania-Reyes, S. (2012). Economic incentives and social preferences: Substitutes or complements? *Journal of Economic Literature*, 50, 368–425. <https://doi.org/10.1257/jel.50.2.368>
- Cheung, C. (2023). Violent behavior and social network antisocial and prosocial behaviors. *Crime & Delinquency*. <https://doi.org/10.1177/00111287231161526>
- Cialdini, R. (2021). *Influence: The psychology of persuasion*. Harper Business.
- Cialdini, R., & Goldstein, N. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, 55, 591–621. <https://doi.org/10.1146/annurev.psych.55.090902.142015>

- Cialdini, R., Reno, R., & Kallgren, C. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58, 1015–1026. <https://doi.org/10.1037/00223514.58.6.1015>
- Cooter, R. (2000). Three effects of social norms on law: Expression, deterrence, and internalization. *Oregon Law Review*, 79, 1–22.
- D'Antoni, M., Friehe, T., & Tabbach, A. (2022). The optimal use of fines and imprisonment revisited. *American Law and Economics Review*, 24, 2, 449–530. <https://doi.org/10.1093/aler/ahac011>.
- Deffains, B., & Fluet, C. (2013). Legal liability when individuals have moral concerns. *Journal of Law, Economics, & Organization*, 9, 930–955. <https://doi.org/10.1093/jleo/ews002>
- Deffains, B., & Fluet, C. (2020). Social norms and legal design. *Journal of Law, Economics, & Organization*, 36, 139–169. <https://doi.org/10.1093/jleo/ewz016>
- Esguerra, E., Vollmer, L., & Wimmer, J. (2023). Influence motives in social signaling: Evidence from COVID-19 vaccinations in Germany. *American Economic Review: Insights*, 5, 275–291. <https://doi.org/10.1257/aeri.20220163>
- Falk, A., & Fischbacher, U. (2002). ‘Crime’ in the lab—Detecting social interaction. *European Economic Review*, 46, 859–869. [https://doi.org/10.1016/S0014-2921\(01\)00220-3](https://doi.org/10.1016/S0014-2921(01)00220-3)
- Fluet, C. (2020). Comment on “The influence of conformity and moral concerns on the level of optimal sanctions”. *Supreme Court Economic Review*, 28, 141–146. <https://doi.org/10.1086/709737>
- Frank, R. H. (2020). *Under the influence: Putting peer pressure to work*. Princeton University Press.
- Frey, B. S., & Torgler, B. (2007). Tax morale and conditional cooperation. *Journal of Comparative Economics*, 35, 136–159. <https://doi.org/10.1016/j.jce.2006.10.006>
- Friehe, T. (2020). The influence of conformity and moral concerns on the level of optimal sanctions: Some comparative-static results. *Supreme Court Economic Review*, 28, 123–139. <https://doi.org/10.1086/709736>
- Funk, P. (2005). Governmental action, social norms, and criminal behavior. *Journal of Institutional and Theoretical Economics*, 161, 522–535. <https://doi.org/10.1628/093245605774259363>
- Garoupa, N. (1999). Optimal law enforcement with dissemination of information. *European Journal of Law and Economics*, 7, 183–196. <https://doi.org/10.1023/A:1008738128742>
- Gneezy, U., & Rustichini, A. (2000). A fine is a price. *Journal of Legal Studies*, 29, 1–17. <https://doi.org/10.1086/468061>
- Harel, A., & Klement, A. (2007). The economics of stigma: Why more detection of crime may result in less stigmatization. *Journal of Legal Studies*, 36, 355–377. <https://doi.org/10.1086/511893>
- Hurd, N. M., Zimmerman, M. A., & Reischl, T. M. (2011). Role model behavior and youth violence: A study of positive and negative effects. *The Journal of Early Adolescence*, 31(2), 323–354. <https://doi.org/10.1177/0272431610363160>
- Kahan, D. M. (1997). Social influence, social meaning, and deterrence. *Virginia Law Review*, 83, 349–395. <https://doi.org/10.2307/1073780>
- Kaplow, L. (1990). A note on the optimal use of nonmonetary sanctions. *Journal of Public Economics*, 42, 245–247. [https://doi.org/10.1016/0047-2727\(90\)90015-A](https://doi.org/10.1016/0047-2727(90)90015-A)
- Kaplow, L., & Shavell, S. (2007). Moral rules, the moral sentiments, and behavior: Toward a theory of an optimal moral system. *Journal of Political Economy*, 115, 494–514. <https://doi.org/10.1086/519927>
- Kimbrough, E. O., & Vostroknutov, A. (2016). Norms make preferences social. *Journal of the European Economic Association* 14(3), 608–638. <https://doi.org/10.1111/jeea.12152>
- Krupka, E. L., & Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, 11, 495–524. <https://doi.org/10.1111/jeea.12006>
- Lin, C. C., & Yang, C. C. (2006). Fine enough or don't fine at all. *Journal of Economic Behavior and Organization*, 59, 195–213. <https://doi.org/10.1016/j.jebo.2004.04.006>
- McAdams, R. H., & Rasmusen, E. B. (2007). Norms in law and economics. In A. M. Polinsky & S. Shavell (Eds.), *Handbook of law and economics* (Vol. 1). North-Holland.
- Morgenroth, T., Ryan, M. K., & Peters, K. (2015). The motivational theory of role modeling: How role models influence role aspirants' goals. *Review of General Psychology*, 19, 465–483. <https://doi.org/10.1037/gpr0000059>

Mungan, M. (2015). Stigma dilution and over-criminalization. *American Law and Economics Review*, 18, 88–121. <https://doi.org/10.1093/aler/ahv026>

Polinsky, A. M., & Shavell, S. (2007). The theory of public enforcement of law. In A. M. Polinsky & S. Shavell (Eds.), *Handbook of law and economics*, (Vol. 1). North-Holland.

Portman, J. (2019). *Celebrity morals and the loss of religious authority*. Routledge.

Posner, E. (2000). *Law and social norms*. Harvard University Press.

Posner, R. A., & Rasmusen, E. B. (1999). Creating and enforcing norms, with special reference to sanctions. *International Review of Law and Economics*, 193, 369–382. [https://doi.org/10.1016/S0144-8188\(99\)00013-7](https://doi.org/10.1016/S0144-8188(99)00013-7)

Shavell, S. (2002). Law versus morality as regulators of conduct. *American Law and Economics Review*, 4, 227–257.

Stigler, G. J. (1970). The optimum enforcement of laws. *Journal of Political Economy*, 78, 526–536. <https://doi.org/10.1086/259646>

Stringham, E. P. (2011). Embracing morals in economics: The role of internal moral constraints in a market economy. *Journal of Economic Behavior and Organization*, 78, 98–109. <https://doi.org/10.1016/j.jebo.2010.12.011>

Wenzel, M. (2004). The social side of sanctions: Personal and social norms as moderators of deterrence. *Law and Human Behavior*, 28, 547–567. <https://doi.org/10.1023/b:lahu.0000046433.57588.71>

Zamir, E., & Medina, B. (2008). Law, morality, and economics: Integrating moral constraints with economic analysis of law. *California Law Review*, 96, 323–392. <https://doi.org/10.15779/Z38HQ57>

Zasu, Y. (2007). Sanctions by social norms and the law: Substitutes or complements? *Journal of Legal Studies*, 36, 379–396. <https://doi.org/10.1086/511896>

How to cite this article: Fluet, C., & Friehe, T. (2023). Optimal law enforcement when individuals are either moral or norm followers. *Journal of Public Economic Theory*, 26, e12669. <https://doi.org/10.1111/jpet.12669>

APPENDIX A

Proof for Lemma 2

Proof. The optimal p_F is either $p_F \leq (1 - \theta)g_N$ or $p_F > g_N$ because $p_F = g_N$ produces the same deterrence as $p_F = (1 - \theta)g_N$ but at a higher cost. We consider two situations.

Case A: $g_N \leq p^*/(1 - \theta)$

Consider first the scenario in which $p_F \leq (1 - \theta)g_N$. The ranking $p_F \leq (1 - \theta)g_N \leq p^*$ implies that welfare is weakly increasing with the detection probability for followers:

$$\frac{\partial W_F(p_F, g_N)}{\partial p_F} = h - p_F - c'(p_F) \geq 0. \tag{A1}$$

It follows that

$$\max_{p_F \leq (1-\theta)g_N} W_F(p_F, g_N) = W_F((1 - \theta)g_N, g_N).$$

Consider next the scenario in which $p_F > g_N$. For detection probabilities in excess of the norm, the concern about norm compliance is irrelevant. We thus get that

$$\max_{p_F \geq g_N} W_F(p_F, g_N) = \begin{cases} W_F(p^*, g_N) & \text{if } g_N < p^* \\ W_F(g_N, g_N) & \text{if } g_N \geq p^* \end{cases} \quad (\text{A2})$$

Clearly, for a strong norm such that $g_N \in [p^*, p^*/(1 - \theta)]$, the optimum is $p_F = (1 - \theta)g_N$. However, when a weaker norm applies such that $g_N < p^*$, we need to compare

$$W_F((1 - \theta)g_N, g_N) = \int_{g_N}^1 (g - h) dg - c((1 - \theta)g_N)$$

and

$$W_F(p^*, g_N) = \int_{p^*}^1 (g - h) dg - c(p^*)$$

Let $h_{11} \in (h_1, p^*)$ solve

$$\int_{h_{11}}^1 (g - h) dg - c((1 - \theta)h_{11}) = \int_{p^*}^1 (g - h) dg - c(p^*). \quad (\text{A3})$$

Because we are comparing the welfare obtained with a deterrence level p^* and enforcement cost $c(p^*)$ to a circumstance with strictly positive but smaller enforcement cost $c((1 - \theta)h_{11})$, the critical level h_{11} must fall inside the interval (h_1, p^*) . This follows from

$$c(p^*) = \int_{h_{11}}^{p^*} (h - g) dg + c((1 - \theta)h_{11}).$$

Because the left-hand side of (A3) is increasing in h_{11} , the solution to (A3) is unique. This proves the claims (i) and (ii).

Case B: $g_N > p^*/(1 - \theta)$

Now $p^* < (1 - \theta)g_N$. Therefore, over $p_F \leq (1 - \theta)g_N$ and using (A1),

$$\begin{aligned} \max_{p_F \leq (1 - \theta)g_N} W_F(p_F, g_N) &= W_F(p^*, g_N) \\ &= \int_{g_N}^1 (g - h) dg - c(p^*) + \int_{\theta g_N + p^*}^{g_N} (g - \theta g_N - h) dg. \end{aligned} \quad (\text{A4})$$

From (A2) and because $g_N > p^*$,

$$\max_{p_F \geq g_N} W_F(p_F, g_N) = W_F(g_N, g_N) = \int_{g_N}^1 (g - h) dg - c(g_N).$$

The result in (A4) together with the strict inequality $p^* < (1 - \theta)g_N$ implies

$$W_F(p^*, g_N) > W_F((1 - \theta)g_N, g_N) = \int_{g_N}^1 (g - h) dg - c((1 - \theta)g_N).$$

Therefore, $W_F(p^*, g_N) > W_F(g_N, g_N)$, which proves claim (iii). □

Proof for Lemma 3

Proof. To prove that $p^{**} > p^*$, it is sufficient to show that $\bar{W}(p_M)$ is increasing over $[0, p^*]$ and that $\bar{W}'(p^*) > 0$. In (10), $W_N(p_M)$ is increasing as long as $p_M \leq p^*$. Using (9), the second term in (10) is governed by

$$W_F(\hat{p}_F(p_M), p_M) = \begin{cases} \int_{p^*}^1 (g - h) dg - c(p^*) & \text{if } p_M < h_{11} \\ \int_{p_M}^1 (g - h) dg - c((1 - \theta)p_M) & \text{if } p_M \in [h_{11}, p^*/(1 - \theta)] \\ \int_{p_M}^1 (g - h) dg + \int_{\theta p_M + p^*}^{p_M} (g - \theta p_M - h) dg - c(p^*) & \text{if } p_M > p^*/(1 - \theta). \end{cases} \quad (A5)$$

For $p_M \leq p^*$, this expression is either constant in p_M or increasing when $p_M \in [h_{11}, p^*]$, where the latter follows from

$$\frac{dW_F(\hat{p}_F(p_M), p_M)}{dp_M} = h - p_M - (1 - \theta)c'((1 - \theta)p_M) > 0.$$

In particular, because $W_N(p_M)$ is maximized at $p_M = p^*$,

$$\bar{W}'(p^*) = (1 - q)[h - p^* - (1 - \theta)c'((1 - \theta)p^*)] > 0.$$

Next, we show that $\bar{W}(p_M)$ is differentiable and strictly concave over $p_M \geq p^*$. This is clearly so for the first term in (10) which depends on $W_N(p_M)$. For the second term, using (A5),

$$\frac{dW_F}{dp_M} = \begin{cases} h - p_M - (1 - \theta)c'((1 - \theta)p_M) & \text{if } p_M \in (h_{11}, p^*/(1 - \theta)) \\ (h - p_M(2 - \theta))\theta & \text{if } p_M > p^*/(1 - \theta) \end{cases} \quad (A6)$$

and

$$\frac{d^2W_F}{dp_M^2} = \begin{cases} -1 - (1 - \theta)^2c''((1 - \theta)p_M) & \text{if } p_M \in (h_{11}, p^*/(1 - \theta)) \\ -\theta(2 - \theta) & \text{if } p_M > p^*/(1 - \theta). \end{cases} \quad (A7)$$

From (A7), dW_F/dp_M is strictly decreasing in p_M for $p_M \geq p^*$ below and above $p^*/(1-\theta)$. Moreover, dW_F/dp_M exists at $p_M = p^*/(1-\theta)$ because the left and right derivatives at that point satisfy:

$$\begin{aligned} \lim_{p_M \uparrow p^*/(1-\theta)} \frac{dW_F}{dp_M} &= h - \frac{p^*}{1-\theta} - (1-\theta)c'(p^*) \\ &= h - \frac{p^*}{1-\theta} - (1-\theta)(h-p^*) \\ &= \theta \left(h - \frac{p^*}{1-\theta}(2-\theta) \right) \\ &= \lim_{p_M \downarrow p^*/(1-\theta)} \frac{dW_F}{dp_M}. \end{aligned}$$

This completes the proof that $\bar{W}(p_M)$ is differentiable and strictly concave over $p_M \geq p^*$.

It follows that p^{**} is unique and satisfies the first-order condition

$$\bar{W}'(p^{**}) = qW'_N(p^{**}) + \left. \frac{dW_F(\hat{p}_F(p_M), p_M)}{dp_M} \right|_{p_M=p^{**}} = 0$$

Because $p^{**} > p^*$, $W'_N(p^{**}) < 0$ so that we must have $dW_F/dp_M > 0$ at the solution. The statement of (A6) implies $p^{**} < h$.

Finally, we prove that $Q(\tilde{h})$ is increasing for $\tilde{h} \leq p^{**}$ with a unique $h_2 \in (h_1, p^{**})$ solving $Q(\tilde{h}) = \bar{W}(p^{**})$. Combining the previous results, $\bar{W}(p_M)$ is increasing so long as $p_M \leq p^{**}$. Because $Q(\tilde{h}) = \bar{W}(\tilde{h}) + qc(\tilde{h})$, $Q(\tilde{h})$ is increasing as well. Hence, $h_2 < p^{**}$ and is unique. We now show that $h_2 > h_1$. Because $h_1 < h_{11}$ and by definition of h_1 ,

$$\begin{aligned} Q(h_1) &= q \int_{h_1}^1 (g-h) dg + (1-q)W_F(p^*, h_1) \\ &= q \left\{ \int_{p^*}^1 (g-h) dg - c(p^*) \right\} + (1-q)W_F(p^*, h_1) \\ &< q \left\{ \int_{p^*}^1 (g-h) dg - c(p^*) \right\} + (1-q)W_F(\hat{p}_F(p^*), p^*) \\ &< \bar{W}_M(p^{**}). \end{aligned} \tag{A8}$$

The inequality before last follows from

$$\begin{aligned} W_F(p^*, h_1) &= \int_{p^*}^1 (g-h) dg - c(p^*) < \int_{p^*}^1 (g-h) dg - c((1-\theta)p^*) \\ &= W_F(\hat{p}_F(p^*), p^*) \end{aligned}$$

thus concluding the proof. □