

Adriani, Fabrizio; Sonderegger, Silvia

Article

The signaling value of punishing norm-breakers and rewarding norm-followers

Games

Provided in Cooperation with:

MDPI – Multidisciplinary Digital Publishing Institute, Basel

Suggested Citation: Adriani, Fabrizio; Sonderegger, Silvia (2018) : The signaling value of punishing norm-breakers and rewarding norm-followers, Games, ISSN 2073-4336, MDPI, Basel, Vol. 9, Iss. 4, pp. 1-32,
<https://doi.org/10.3390/g9040102>

This Version is available at:

<https://hdl.handle.net/10419/219223>

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.



<https://creativecommons.org/licenses/by/4.0/>

Article

The Signaling Value of Punishing Norm-Breakers and Rewarding Norm-Followers

Fabrizio Adriani ¹ and Silvia Sonderegger ^{2,*}

¹ Department of Economics, University of Leicester, University Road, Leicester LE1 7RH, UK; fa148@le.ac.uk

² School of Economics, University of Nottingham, CEDEX, University Park, Nottingham NG7 2RD, UK

* Correspondence: silvia.sonderegger@nottingham.ac.uk

Received: 31 August 2018; Accepted: 15 November 2018; Published: 13 December 2018



Abstract: We formally explore the idea that punishment of norm-breakers may be a vehicle for the older generation to teach youngsters about social norms. We show that this signaling role provides sufficient incentives to sustain costly punishing behavior. People punish norm-breakers to pass information about past history to the younger generation. This creates a link between past, present, and future punishment. Information about the past is important for youngsters, because the past shapes the future. Reward-based mechanisms may also work and are welfare superior to punishment-based ones. However, reward-based mechanisms are fragile, since punishment is a more compelling signaling device (in a sense that we make precise).

Keywords: norms; punishment; reward; signaling

JEL Classification: D82, D83, C72

1. Introduction

It is well known that people are willing to punish norm-breakers, even in situations where punishment is costly and does not provide any material benefit to the punisher ¹ and literature thereafter. This raises the question: Why should people be willing to incur personal costs in order to punish norm deviators? This paper focuses on an information-based rationale for punishment. The existence of an informational content of punishment is the basis of the so-called denunciation (or expressive) theory in legal philosophy. ² This theory emphasizes the role of punishment in teaching people what are the social norms and what is not acceptable. Much anecdotal evidence such as the widespread tradition of bringing children to see public executions also points to a potential information role of punishment. ³ Our analysis builds on this idea. We focus on information transmission from parents to children, although it should be quite clear that the analysis applies more broadly to environments in which an “experienced” individual may convey information to a “naive” individual, towards whom he feels altruistic—examples include teacher and pupil, senior worker and junior worker, etc. We argue that the desire to transmit the “correct” information may underpin punishment or ostracism of norm breakers. For instance, a Muslim mother who wears a veil and wishes to raise her daughter along similar lines may be reluctant to engage in friendly relations with unveiled women.

¹ See e.g., Fehr and Gächter (2000) [1].

² Advocates of this theory include Feinberg (1970), Kahan (1998), Cooter (1998) and Sunstein (1996) [2–5].

³ In 18th century Rome, for instance, fathers would bring their sons to see executions and would slap them at the moment the executioner would strike his blow, to “keep that vision vivid in their memory.” This tradition was documented for instance in a sonnet by Roman poet G. G. Belli “Er Ricordo” in 1830. Similar practices existed also in other countries such as Great Britain.

Similarly, to signal to his child that pre-marital sex is reprehensible, a parent might demand the removal of an unmarried pregnant teacher working in the child's school.⁴

We build a model of overlapping generations where parental behavior informs children about norms in society. Parents are paternalistic: they evaluate the child's cost of being punished in the future using the lens of their own (higher) discount factor. This generates a conflict of interests between parents and children that justifies the use of costly signaling for information transmission. We concentrate on norms that are individually costly. Our running example focuses on unproductive behaviors, such as dressing codes, lengthy rituals, or elaborate etiquette, although the analysis can equally apply to behaviors (such as public good provision) that generate externalities.⁵ Norm-following is thus purely sustained by fear of punishment. By contrast, the metanorm of sanctioning norm violators is sustained by signaling motives.

The game has an equilibrium where nobody punishes, and nobody follows the norm. However, it also has a punishment equilibrium where norm-following and punishment occur. A central role is played by the notion that parents have better information than children about the history of play. Children know that society may be in one of different regimes, which differ in the extent to which punishment and norm-following are carried out, but they do not know which one. They use their parents' behavior to update their beliefs about the prevailing regime—and, thus, about the likelihood of being met with punishment if they break the norm. Parents punish past norm-breakers since by doing so they send the correct signal to their child.

An important contribution of the analysis is to highlight the role of history in sustaining norms. In equilibrium, people punish when they observe punishment featuring in recent shared history. Importantly, people (correctly) expect that recent history will shape individual behavior in the future. We show that this effect alone can provide sufficient incentives to sustain an otherwise dominated behavior such as costly punishment.

More precisely, adults active at t know that other t adults will punish past norm breakers, since they have been exposed to the same shared history as they have. At $t + 1$, recent history (capturing events at t) will thus contain punishment. In turn, this will induce people in $t + 1$ to punish as well. The connection between past, present and future punishment means that information about the past is valuable for youngsters, since the past molds the future. It is worth highlighting that in our setup history matters not only because it anchors beliefs but also, more concretely, because the total cost of punishing past norm breakers depends on how many people broke the norm in the previous period (since this determines how many people are to be punished).

Our results underscore the role of *continuity*, namely the link between past, present and future, to sustain norms (and the metanorm of punishment). The analysis suggests that, if continuity breaks down (as e.g., in the case of an external shock such as a war) this would trigger a process of unraveling that would affect behavior. This regime shift could happen very quickly—in contrast with the predictions of preference-based theories of punishment, where a shock to preferences would translate only very slowly onto behavior (since the older generations would have to die off).⁶

We show that the analysis produces non-obvious comparative statics. For instance, we find that the norm being sustained cannot be too cheap. Intuitively, if it was, the conflict of interests between parents and children would vanish, and this would eliminate the need for costly signaling.

⁴ This is presumably the rationale why a school may end up dismissing a female employee who becomes pregnant while unwed. Recent high-profile (and controversial) instances of this practice can be found at <http://www.jewishpress.com/blogs/muqata/school-fires-unmarried-pregnant-teacher-the-whole-story/2013/03/05/> and at http://www.huffingtonpost.com/2013/03/01/teri-james-pregnant-woman-fired-premarital-sex-christian-school_n_2790085.html.

⁵ Further discussion of the types of norms that fit our analysis can be found in Section 7.1.

⁶ In their entry on social norms in the Stanford Encyclopedia of Philosophy, Bicchieri and Muldoon (2014) [6] report that one of the defining features of norms is that they can change rather abruptly, with a sudden and unexpected demise of old patterns of behavior.

While our signaling equilibrium generates interesting insights, any explanation for behavior X that relies on the existence of an equilibrium where X is a signaling device should be taken with caution. Critics might argue that almost *any* behavior might be rationalized by constructing a setup where failure to adopt that behavior would be interpreted in a sufficiently negative light by the receiver. A convincing account should therefore provide a *more compelling* rationale why X (rather than some other behavior) should act as a signaling device. To this aim, we construct a setup where signaling operates through an alternative mechanism—namely, costly rewards.⁷ Sure enough, we show that the same mechanism at work in a punishment equilibrium may also sustain a reward-based equilibrium. However, a reward equilibrium may emerge only under some conditions. We find that a scenario where the (marginal) cost of rewarding equals the (marginal) benefit from being rewarded, as in the case of monetary transfers, would not work. This is consistent with the observation that, in social interactions, cash rewards are seldom used—in fact, there is almost a taboo against them.

The next step in the analysis is to compare punishments and rewards by constructing a measure of robustness—would a punishment equilibrium be immune to people switching to costly rewards for signaling, and would a reward equilibrium be immune to people starting to use punishment instead of reward? We show that a reward-based equilibrium can generally be “invaded” by punishment, but not vice-versa. Although the reward equilibrium generates more welfare than the punishment equilibrium, it is also more fragile, since punishment is a more compelling signaling method than reward. Intuitively, this is because the cost of punishing *decreases* in the share of norm compliers—since there are less people to punish. Hence, by using punishment instead of, say, reward, a parent can credibly signal that the norm is sufficiently widespread, while at the same time incurring low signaling costs. By contrast, the cost of rewarding *increases* in the share of compliers—since there are more people to reward. This makes rewarding norm compliers a less compelling signaling device of the widespread norm than punishing norm breakers.

1.1. Related Literature

One possible explanation for costly punishment that has been offered by the literature is that people may have direct in-built preferences for punishing norm-deviators, which may outweigh the material costs involved in punishing.⁸ This explanation is especially plausible when the norms in question are the result of a joint process of evolution and acculturation, such as may for instance be the case for norms against incest. When this does not apply, however, the direct preference approach leaves some important questions unanswered, since it does not explain why people who exhibit a preference for engaging in costly punishment (and who thus suffer a material disadvantage compared to those who do not exhibit this preference) are not wiped out by evolutionary forces.

Another possible rationale that has been offered is that those who do not punish norm-breakers may themselves be seen as deviators and thus be punished—such as for instance in Akerlof (1976) [8].⁹ Using Axelrod’s (1986) [10] terminology, a norm for X typically requires the existence of a *metanorm* to sanction people who fail to do X , one to sanction people who fail to sanction those who fail to do X , and so on. The applicability of this explanation depends on the environment. Although norm breaking can often be observed directly (e.g., because an individual fails to dress/ behave/ speak in a manner that complies with the norm), the failure to punish norm-breakers—and to punish those who fail to punish norm breakers, etc.—are often much harder to identify.¹⁰ It is thus plausible that

⁷ Clearly enough, the absence of punishment may be seen as a reward, and, thus, the punishment equilibrium could be rephrased in those terms. However, in that case the activity of rewarding involves no cost (it is the absence of rewarding that is costly). By contrast, in Section 5 we consider rewards that are costly to those who implement them.

⁸ See e.g., Rotemberg (2008) [7] for a formal model.

⁹ See also Peski and Szentes (2013) [9].

¹⁰ The notion that, in some instances, information flows may be too restricted for reputation-based mechanisms to work effectively is well recognized, see e.g., Ghosh and Ray (1996) [11].

these metanorms may be sustainable in small communities, where monitoring is easy, but not in larger environments, where it is harder.

The central theme of our analysis is that norms (or, more precisely, metanorms) may be the object of signaling. The notion that signaling-based explanations may help clarify apparent “behavioral puzzles” has been presented elsewhere, such as in Glazer and Konrad (1996), Ellingsen and Johannesson (2008) and Hopkins (2014). Sliwka (2006) and Gneezy and Rustichini (2000) [12–16] explicitly consider norm-signaling. Differently from those works, in our setup parental behavior informs children about the behavior they are likely to encounter in the future in the future. This shares similarities with Adriani and Sonderegger (2018), Adriani et al. (2018) and Kotsidis (2018) [17–19]. However, this paper studies the mechanics and the signaling value of punishment versus reward, and is thus very different in focus.

Our work also adds to the literature on intergenerational transmission of information and (self-)signaling. Bénabou and Tirole (2011) [20] consider an anticipatory-utility setup where punishment of deviant behavior helps the individual shield himself from negative anticipatory feelings. Other relevant works include Bénabou and Tirole (2002, 2004 and 2006) [21–23] and Dessí (2009) [24]. These papers emphasize how anticipatory feelings or altruism towards future generations may generate selective memory. Bisin and Verdier (2000) [25], Corneo and Jeanne (2009, 2010) [26,27], Cervellati and Vanin (2013) [28], Carvalho (2013) [29] and Verdier and Zenou (2018) [30] are also related to the present paper. However, these works focus on intergenerational transmission of *values* or *cultural traits* (rather than information). Also related is Van der Weele (2012) [31], who studies the signaling role of sanctions when the authorities have private information about the fraction of egoists in society. An important difference with our analysis is that, in our setup, strategic complementarities among parents play a central role, while in his model these complementarities are entirely absent since punishment is fully centralized.

The effectiveness of “stick” versus “carrot” as incentive devices and their interaction are clearly important for understanding norm-following, and have been extensively studied by the experimental literature.¹¹ However, the theoretical literature has largely remained silent on the subject. The exception is Herold (2012) [33], who proposes a model of evolution of preferences that explicitly compares preferences for punishing with preferences for rewarding.¹² The rationale why preferences for punishing eventually crowd out preferences for rewarding is that, when cooperation is sufficiently widespread, punishing non-cooperators becomes cheaper than rewarding cooperators. Our analysis supplements this intuition with a new, information-based effect. We show that, to successfully “invade”, a signal (punish/reward) should not only be cheap, but it should also be *credible*. Punishment is a credible signal that the norm is sufficiently widespread because it naturally satisfies incentive compatibility: It is cheaper when the norm is widespread than when it is rare. By contrast, reward follows the opposite path.¹³

Finally, our work is related to the literature that studies the interplay between history and norms—the most relevant recent examples for our purposes include Acemoglu and Jackson (2014), Rohner et al., (2013) and Bidner and Francois (2013) [34–36].¹⁴ A central theme of this literature is that

¹¹ See e.g., Andreoni et al., (2003) [32] and literature thereafter. Interestingly, one of the findings of Andreoni et al. is that people tend to reward less when punishment is also available (but not vice-versa). This is consistent with our results in Section 6, where we show that punishment typically crowds out reward (but not vice-versa).

¹² This paper belongs to a wider literature on endogenously derived preferences, discussed in Section 7.

¹³ It may be worth speculating about the possible implications of our analysis from an evolutionary viewpoint. We have shown that punishing behavior may emerge as the equilibrium of a game involving players with “standard” preferences. This may be interpreted as providing a possible rationale why direct preferences for punishing behavior may emerge—since they generate optimal behavior. In fact, our story predicts the possible coevolution of two sets of preferences: Preferences for punishing behavior (in parents), and preferences for following the diktat of parents, provided that it is backed up by parental behavior that is consistent with it (in children). These preferences, it may be argued, are “mutual best responses” under some conditions.

¹⁴ See also Tirole (1996) [37] for a classic account of how past history may affect current incentives.

the history of past play leaves a lasting legacy, by affecting behavior for a long time afterwards. This is also the case in our model. However, rather than it being a case of past history revealing information about the players' types, in our model *the history of play itself* is the object of signaling. By perpetuating the past, people can send information about past history (and, thus, indirectly, about the future, since in equilibrium the past shapes the future).

Similar to us, Acemoglu and Jackson (2014) [34] also assume that players cannot fully observe the history of past play. However, their paper has a rather different focus, since it does not study punishment.

The remainder of the paper is organized as follows. Section 2 introduces our key idea within a very simple setup. Section 3 describes the model, while Section 4 characterizes the punishment equilibrium. Section 5 considers the case where individuals use costly rewards to signal the dominant norm to their child. Section 6 investigates the robustness of the punishment equilibrium to the use of rewards, and the robustness of the reward equilibrium to punishment. Finally, Section 7 provides discussions and offers some final thoughts.

2. A Very Simple Example

Consider the following sequential game between a parent and his child. The parent moves first, and has two actions: Punish past norm-breakers, and Not Punish. Punishing norm-breakers is costly for the parent, while not punishing is not. The child moves second, after having observed the parent's move. He also has two actions: Follow (some unspecified costly social norm) and Not Follow. The parent possesses private information about the consequences that the child will face if he does not follow the norm. For a start, we take these consequences as being exogenously determined. There are two states of the world: (i) *punishment state*: if the child breaks the norm, he is punished by a third party and incurs a punishment cost; (ii) *no punishment state*: if the child breaks the norm, he incurs no punishment. The cost for the child from following the norm is positive but smaller than the cost he incurs if he is punished. Hence, if the child knows that norm-breaking will be punished by the third party, he finds it optimal to follow the norm. However, the child does not know the state of the world, and looks at the parent's behavior to gain information about it. The parent is altruistic towards the child. We now argue that the following may emerge in equilibrium:

- The child follows the norm if he observes the parent punishing, breaks the norm otherwise.
- The parent punishes in state (i) (when the third party punishes), does not punish in state (ii) (when the third party does not punish).

It is clear that, given the parent's strategy, the child's strategy prescribes optimal behavior. Consider now the parent. Taking the child's strategy as given, if the parent does not punish in state (i), then the child will break the norm and be punished. By contrast, if the parent punishes past norm breakers the child will infer that the state is (i) and will therefore follow the norm (and thus avoid being punished). For parameter values, it is clear that the parent may prefer the latter to the former, and would thus find it optimal to punish in state (i), even though punishing is a costly endeavor.¹⁵

In the equilibrium described above, the parent punishes past norm-breakers to signal to his child that, if he breaks the norm, he will be punished. Hence, punishment is motivated by signaling concerns. A limiting feature of this example is that the consequences the child will face if he breaks the norm are exogenous, rather than being derived as an equilibrium feature. However, it is possible to augment the model to allow for that. Suppose that, in each period t , a different parent-child pair plays the game described above. If the child breaks the norm, he may or may not be punished in period $t + 1$ (by the parent of the $t + 1$ parent-child pair). Payoffs are the same as before, except that the action of the "third party" is no longer exogenous. The third party is in fact the $t + 1$ parent,

¹⁵ It is straightforward to see that punishing is always dominated in state (ii).

who faces a problem analogous to that of the t -period parent. Following the reasoning above, under some parameter restrictions the t -period parent will then choose to punish if he believes that the $t + 1$ parent is sufficiently likely to punish. Similarly, the $t + 1$ parent will choose to punish if he believes that the $t + 2$ parent is sufficiently likely to punish, and so on.

This simple example describes in a very crude way the type of story we have in mind. However, there are several issues that the example does not address.

- (1) What is the exact nature of the information asymmetry between parents and children?
- (2) Why should the parent use punishment as a way to communicate to his child? Would norm-following by the parent not be enough? Even more fundamentally, why can the parent not just “tell” the child what to do?

Our fully blown model addresses these issues. As we will see, a specific feature of the setup will allow us to address concern (2), namely, the presence of a paternalistic element in the parent’s motives. Furthermore, it is quite clear that the logic we have highlighted would equally work if instead of focusing on punishment, we were to consider a reward-based system. The question then is: are there any differences between these two mechanisms? Our full-blown setup will allow us to meaningfully compare the signaling value of punishments and rewards for information-transmission across generations.

3. Model

We start off by constructing a model aimed at formalizing the intuition sketched in Section 2. As we will argue in Section 5, this setup can also be employed to study the case where signaling operates through rewards, rather than punishment.

Overlapping Generations We consider an environment populated by overlapping generations. Each period t is characterized by a continuum of active parent-child pairs. Period t parents select to follow or not the norm, and to punish or not those individuals who broke the norm in the previous period. Period t children observe their parent’s behavior and decide whether to follow the norm. At $t + 1$, t parents and t children become purely passive and may be punished if they broke the norm in the previous period. Just before $t + 2$, t parents die. The t children observe the state of the world and recent history (more on this below) and become adults, i.e., the $t + 2$ parents, with one offspring each.

The following table summarizes our notation for the players’ actions.

$n = 0$: Youngster breaks the norm;	$n = 1$: Youngster follows the norm.
$N = 0$: Adult breaks the norm;	$N = 1$: Adult follows the norm.
$M = 0$: Adult fails to punish;	$M = 1$: Adult punishes,

where the letter “ M ” is a mnemonic for “metanorm”.

Timing Timing is illustrated in the diagram below (Figure 1).

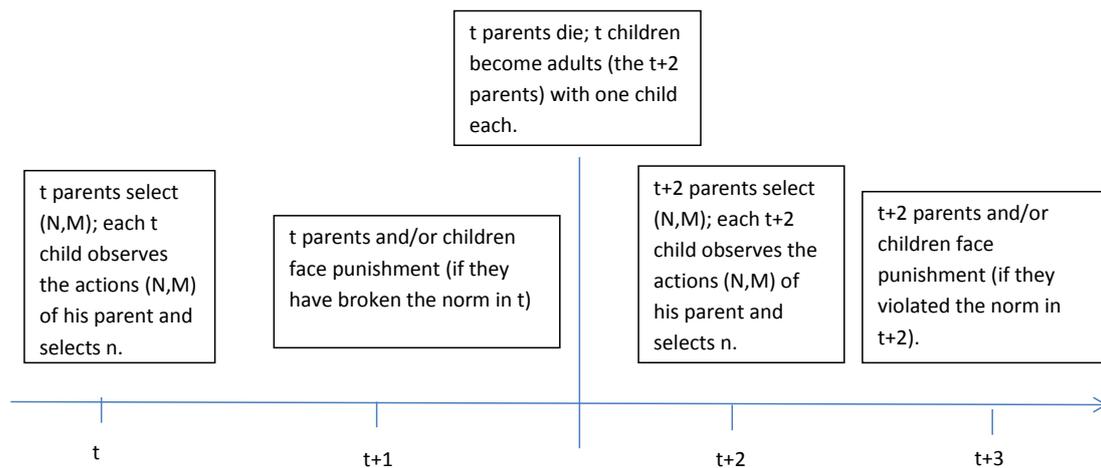


Figure 1. Timing 1.

Note that if an individual breaks the norm at t he faces punishment at time $t + 1$. Intuitively, this reflects the notion that punishment often takes the form of being confronted with ostracism or bias in future interactions. For instance, someone who gets drunk and behaves obnoxiously at a party will probably find his party invitations greatly reduced in the future. Similarly, someone who fails to respect religious practices may find himself socially shunned in future exchanges with members of his religious community, and so on. This lag in punishment implies that period t adults punish people who broke the norm at time $t - 1$. In turn, the children who followed/violated the norm at $t - 1$ (and who are punished at t) become adults (and potential punishers) at $t + 1$.

Cost of punishing An individual who punishes norm-breakers incurs a cost. A natural benchmark is the case where punishment takes the form of ostracism, in which case the cost incurred arises from the missed opportunity of forming an economically beneficial link. Moreover, as argued by Elster (1989) [38], expressing disapproval is always costly—at the very least, it requires energy and attention that could have been used for other purposes. One might also alienate or provoke the target individual, at some cost or risk to oneself. Letting the mass of norm-breakers be denoted as $b \geq 0$, the total cost of punishing is given by $b\theta + \varepsilon$, where $\theta > 0$. Punishment is more expensive the greater the mass of people who must be punished, but also involves a fixed cost, ε which we assume is strictly positive and possibly very small. One way to think about this small fixed cost is as a shortcut to reflect the presence of a small share of “behavioral” agents who always violate the norm.¹⁶ As a result, a parent who chooses to punish will always incur a positive cost, even if $b = 0$. This makes it clear that our results do not rely on the cost of punishing being exactly zero in any state of the world.

Following the norm is costly and involves a positive cost. In what follows, when making welfare statements, we will focus on the benchmark case in which the behavior prescribed by the norm generates no externalities, although we will discuss how the presence of positive/negative externalities may affect our results when appropriate.

Availability of the punishment option We allow for the possibility that, for some parents, punishment may actually be unfeasible. Accordingly, we assume that there are two states of the world.

- with probability γ , the state of the world is l : the punishment option is available for all parents.
- with probability $1 - \gamma$ the state of the world is h : the punishment option is unavailable to a share $q \in (0, 1)$ of parents and is available for the remaining $1 - q$ share.

¹⁶ We could have modelled this explicitly, by letting the mass of norm violators be $b + \varepsilon$, for some $\varepsilon > 0$. This would not affect the results in any meaningful way, but would make the algebra less transparent. This, we feel, would detract attention from the main messages we wish to convey.

For simplicity we assume that the state of the world is the same in all periods. Parents observe the state of the world, while children do not. A natural benchmark is the case where γ is arbitrarily close (but not equal to) to 1, implying that state l is “almost certain”.

Adults We introduce the following additional notation:

m_t is the mass of punishers in any given period t , and
 Θ is the (marginal) cost of being punished.

The material payoff of a t -period adult is summarized below.

	Follow the norm ($N_i = 1$)	Do not follow the norm ($N_i = 0$)
Punish ($M_i = 1$)	$-c - \theta b_{t-1} - \varepsilon$	$-\theta b_{t-1} - \Theta m_{t+1} - \varepsilon$
Do not punish ($M_i = 0$)	$-c$	$-\Theta m_{t+1}$

The discount factor is implicitly set equal to one. Throughout the analysis, we assume that $\Theta - c > \varepsilon$. Note that $\Theta > c$ is actually a necessary requirement for punishment to be an effective incentive device for norm-following. The adult’s direct material payoff can equivalently be expressed as

$$u_i(b_{t-1}, m_{t+1}, N_i, M_i) = -cN_i - (\theta b_{t-1} + \varepsilon) M_i - \Theta m_{t+1}(1 - N_i). \tag{1}$$

Note that the direct payoff (1) provides an incomplete description of individual welfare, since it omits the utility that an individual derives from altruistic concerns towards his child. A full description of individual welfare is provided below, in (3).

Information Before selecting whether to follow the norm and punish, adults observe m_{t-1} and b_{t-1} , namely the mass of adults (from the previous generation) who punished norm deviators in the previous period, and the mass of norm-breakers in the previous period. They also observe the state of the world (l or h) and, in state h , they observe whether, for them, the punishing option is available or unavailable. The realizations of m and b in periods before the last cannot be observed.

Children Each adult i is endowed with one child, denoted by \bar{i} . Children do not observe (m_{t-1}, b_{t-1}) , the state of the world (l or h), or whether the punishment option is/is not available to their parents. However, each youngster perfectly observes his parent’s actions (N_i, M_i) . After observing his parent’s actions, he chooses $n_{\bar{i}} \in \{0, 1\}$ (violate the norm/follow the norm). The cost of following the norm for children is c , the same as an adult’s.

If a youngster does not follow the norm, he becomes the object of social punishment.¹⁷ We assume that the stigma associated with norm-breaking does not carry through adulthood. Hence, as adults, all individuals start off with a clean slate, independently of how they behaved as youngsters.¹⁸

The payoff of a t -period child is as follows:

Follow the norm ($n_{\bar{i}} = 1$): $-c$
 Do not follow the norm ($n_{\bar{i}} = 0$): $-\delta\Theta m_{t+1}$

where $\delta < 1$ indicates the child’s discount factor. This can equivalently be expressed as

$$\bar{u}_{\bar{i}}(m_{t+1}, n_{\bar{i}}) = -n_{\bar{i}}c - \delta\Theta m_{t+1}(1 - n_{\bar{i}}). \tag{2}$$

¹⁷ Another possibility, mentioned for instance in Bidner and Eswaran (2014) [39], is that the parents of violators may be punished, rather than the violators themselves. We do not explore this here.

¹⁸ This ensures stationarity of the problem at hand, since the problem an individual faces as an adult is independent of the action he took as a youngster.

Note that we assume that children discount the future more heavily than adults.¹⁹ This reflects the idea that age affects preferences and choice. Robson and Samuelson (2007, 2009) [40,41] study the evolution of discount rates and find that, under some natural conditions, these should fall with age.

Adults' Utility Adults are altruistic towards their children, but this altruism is “impure.” More specifically, we assume that parents evaluate the child's payoff using their own (higher) discount factor—although they are aware that this differs from the child's discount factor. This generates what Doepke and Zilibotti (2014) [42] call a *paternalistic element* in the parents' motives, which is standard in models of parent-child interactions.²⁰ The total utility of an adult i is thus given by

$$U_i(b_{t-1}, m_{t+1}, N_i, M_i, n_{\bar{i}}) \equiv u_i + \bar{u}_{\bar{i}} - \Theta m_{t+1}(1 - n_{\bar{i}})(1 - \delta). \quad (3)$$

As will become clear below, parental paternalism introduces a conflict of interests between parents and children, which motivates the use of costly signaling for information transmission. Intuitively, if the interests of parents and children were perfectly aligned then parents could just “tell” children what to do (and children would always find it optimal to follow what their parents tell them). The existence of a conflict of interests eliminates this possibility, since parents have an incentive misrepresent their information. The use of costly signaling is then necessary to address this credibility problem.

4. Punishment Equilibrium

We concentrate on Perfect Bayesian Equilibria. The equilibrium concept in our setup thus satisfies the following requirements.

- (1) Children update their beliefs from parent's action using Bayes rule whenever possible.
- (2) At each information set, each player's strategy specifies optimal actions, given his beliefs and the strategies of the other players

In what follows, we concentrate on *pure strategy symmetric* equilibria. We also restrict attention to equilibria that are *stationary*, in the sense that all individuals follow strategies that independent of calendar time. Finally, to simplify the analysis, we will restrict attention to equilibria in which:

Refinement 1. *Children who observe their parent violating the norm never follow the norm.*

Imposing this refinement allows to rule out counterintuitive scenarios at the outset. Note however that the equilibrium we characterize would continue to hold even if Refinement 1 was lifted. Moreover, its key property, namely that those who punish do so to signal that norm-breakers are likely to be punished, would arise more generally also in other possible equilibria. However, a full characterization would be lengthy and, we believe, not very illuminating.

The crucial feature of our setup is that past history of play cannot be perfectly observed. Agents active at t have a common prior over history up to that period. We impose a “grain of truth” restriction on the prior. In particular, we assume that prior beliefs at t assign a strictly positive probability to the *actual* (real) realization of history up to that period. Adults also obtain some direct evidence of past history, since they observe (m_{t-1}, b_{t-1}) , namely how much norm breaking and how much punishing there was in the previous period. A strategy for parent i maps (m_{t-1}, b_{t-1}) and the state of the world (l or h) into (N_i, M_i) . In contrast, children have no direct information about past history, except through their parents' actions. A strategy for child \bar{i} maps (N_i, M_i) into $n_{\bar{i}}$. From (2), the child's optimal action

¹⁹ This assumption is made for concreteness, but our key results would still go through if we assumed that the parent's discount factor *when evaluating his own payoff* equals the child's, namely δ . Rather, what is really crucial is the presence of paternalistic motives (see below), namely that, *when evaluating the child's payoff*, the parent should discount the future less heavily than the child.

²⁰ This paternalistic element shares similarity with the concept of imperfect empathy introduced by Bisin and Verdier (2001) [43].

at t depends on his expectation of m_{t+1} , namely the share of punishers at $t + 1$. The child will decide to follow the norm if and only if ²¹

$$\delta \Theta E(m_{t+1} | N_i, M_i) > c. \quad (4)$$

Our first lemma establishes a benchmark result.

Lemma 1. *The game has a history-independent equilibrium where nobody follows the norm, and nobody punishes.*

Proof. In Appendix A.

Clearly enough, if nobody punishes norm-followers, then norm-following and punishing are dominated actions. We now consider a more interesting scenario, where norm-following and punishing emerge in equilibrium. By Refinement 1, we are restricting attention to environments where the posterior beliefs about m_{t+1} of a child who has observed his parent breaking the norm induce him to break the norm too. It remains to address the child's posterior beliefs when he observes his parent following the norm. \square

Lemma 2. *In an equilibrium with punishment, the following must hold. A child's expectation of m_{t+1} when he has observed his parent both following the norm and punishing at t must exceed that of a child who has observed his parent following the norm but not punishing.*

Proof. In Appendix A.

For punishment to emerge it must be that, by observing his parent punishing, the child gains useful information about the likelihood of being met with future punishment if he breaks the norm. However, note that punishment conveys information about (m_{t-1}, b_{t-1}) , the parent's private information. This raises the question: how can information about the *past* be useful to predict the *future*? For this to occur, there must be a link between the past and the future, so that, by gaining information about the past from the actions of their parents, children are able to make inferences about the future. In other words, past history must have a bearing in shaping the future.

The next result says that the information conveyed by parental punishment must actually be determinant for the child's choice. \square

Lemma 3. (Strategy A.) *In an equilibrium with punishment, the following must hold. A child who has observed his parent both following the norm and punishing will follow the norm. A child who has observed his parent following the norm but not punishing will break the norm.*

Proof. In Appendix B.

Intuitively, if a parent could induce his child to follow the norm by simply following the norm himself, he would never select to punish, as this involves unnecessary punishing costs. Since all parents would follow this reasoning, punishment would disappear. This rules out that, in a punishment equilibrium, children may decide to follow the norm whenever they observe their parent following the norm, with no concern for punishing behavior.

Suppose that Refinement 1 holds, and that children follow strategy A (described in Lemma 3). We now compute the payoffs that a parent active at time t may obtain from each of the available action-pairs. First, the payoff from following the norm and punishing—namely, $(N_i, M_i) = (1, 1)$ —is given by

$$U^t(1, 1 | m_{t-1}, b_{t-1}) = -2c - \theta b_{t-1} - \varepsilon. \quad (5)$$

²¹ We adopt the convention that, if indifferent, the child breaks the norm and, if indifferent, the parent does not punish. This is immaterial for the results.

This follows since a parent selecting (1, 1) will induce his child to follow the norm. Note that history has a direct bearing on payoffs, and is thus more than a simple correlation device. This is because the cost of punishing is higher the higher the share of past norm breakers.

Second, the payoff from $(N_i, M_i) = (1, 0)$ —namely, following the norm and failing to punish—is

$$U^t(1, 0 \mid m_{t-1}, b_{t-1}) = -c - \Theta m_{t+1}(m_{t-1}, b_{t-1}) \quad (6)$$

since in that case the child will not follow the norm and will thus be punished by all those adults at $t + 1$ who choose to punish. The expression $m_{t+1}(m_{t-1}, b_{t-1})$ gives the value of m_{t+1} conditional on the information at the parents' disposal, namely m_{t-1} and b_{t-1} .²²

Third, the payoff from (0, 0)—namely, breaking the norm and failing to punish—is

$$U^t(0, 0 \mid m_{t-1}, b_{t-1}) = -2\Theta m_{t+1}(m_{t-1}, b_{t-1}) \quad (7)$$

since in that case both the parent and the child will be punished at $t + 1$.

Finally, the payoff from $(N_i, M_i) = (0, 1)$ —namely, breaking the norm and punishing—is

$$U^t(0, 1 \mid m_{t-1}, b_{t-1}) = -\theta b_{t-1} - \varepsilon - 2\Theta m_{t+1}(m_{t-1}, b_{t-1}). \quad (8)$$

Children We are now able to describe more precisely how children process the information conveyed by parental actions. Children start with a prior about the history of past play. Upon observing their parent's behavior (N_i, M_i) , they combine their prior information with the information conveyed by parental behavior to form posterior beliefs about the pair (m_{t-1}, b_{t-1}) . Since parental strategies depend on recent history, this posterior induces a probability distribution over (m_t, b_t) , which in turn determines a probability distribution over m_{t+1} , which is the variable of interests to children—since it determines the payoff from norm-breaking. Intuitively, children look at parental behavior to gain a clue about recent history. This matters, since recent history affects adult behavior in the following periods. \square

4.1. Steady States

We concentrate on equilibria where economy is in a *steady state*: The optimal reply by parents to $m_{t-1} = m^*$, $b_{t-1} = b^*$ generates $m_t = m^*$, $b_t = b^*$. To characterize the possible steady states, we need to check for incentive compatibility. For instance, for $m^* = 1$, $b^* = 0$ to be a possible steady state, we need (5) to exceed (6)–(8) whenever we impose the steady state conditions $b_{t-1} = b^* = 0$ and $m_{t+1} = m^* = 1$. The following lemma characterizes the possible steady states that may emerge.

Lemma 4. *Suppose that children follow strategy A described in Lemma 3. The possible steady states are:*

- (i) *High-punishment: All parents follow the norm and (if they can) punish. As a result, in state l: $m^* = 1$, $b^* = 0$, while in state h: $m^* = 1 - q$, $b^* = q$.*
- (ii) *No-punishment: $m^* = 0$, $b^* = 2$. Nobody punishes, nobody follows the norm.*

Proof. In Appendix A.

Lemma 4 shows that there are two possible steady states: a no-punishment, essentially equivalent to the history-independent equilibrium, and a high-punishment, where all parents follow the norm and punish norm-breakers (if they can). This multiplicity arises because parent-child signaling generates strong complementarities among parents: People punish norm-breakers because this is what others

²² This could in principle be an expected value, although as we will see that will not be the case in equilibrium.

in society do, and vice-versa. [Further discussion of the mechanism that underpins punishment is provided below, when discussing Proposition 1.]

It is important to note that, as mentioned in the proposition, the actual share of punishers and norm-breakers in steady state may actually take *three* different values. More specifically:

- If the steady state is high-punishment and the state of the world is l : $m^* = 1, b^* = 0$. We denote this scenario as regime H.
- If the steady state is high-punishment and the state of the world is h : $m^* = 1 - q, b^* = q$. We denote this scenario as regime I (for intermediate).
- If the steady state is no-punishment: $m^* = 0, b^* = 2$. We denote this scenario as regime L.

We now turn to children's expectations. As a first approximation, it is useful to assume that children have rational expectation, and thus know that the economy is in a steady state (although they do not know which one). Prior beliefs thus assign positive probability to at most three possible regimes. We let the children's prior assigns a probability p_H to regime H, probability p_I to regime I and probability and $p_L = 1 - p_I - p_H$ to regime L. We assume that both p_H and p_I are strictly greater than zero. The children's posterior beliefs are then well defined and can be computed from Bayesian updating:

$$\pi_H(1,1) = \frac{p_H}{p_H + (1-q)p_I}; \pi_I(1,1) = \frac{(1-q)p_I}{p_H + (1-q)p_I}; \pi_I(1,0) = 1; \pi_H(1,0) = 0. \quad (9)$$

To refine the child's out-of-equilibrium beliefs, we apply Cho and Kreps (1987) [44] Intuitive Criterion: Following an out-of-equilibrium move by the parent, a child assigns zero probability to the deviation emanating from a regime in which the deviation is equilibrium dominated. \square

Proposition 1. *Suppose that both p_H and p_I are strictly greater than zero and that*

$$\min\{\Theta(1-q) - \varepsilon - \theta q, \frac{p_H + (1-q)^2 p_I}{p_H + (1-q)p_I} \delta \Theta\} > c \geq \delta \Theta(1-q). \quad (a)$$

Then punishment may emerge in equilibrium. In the refined steady state equilibrium with punishment, the following holds.

- (i) *Children follow the norm if and only if they observe their parent punishing (Strategy A).*
- (ii) *Adults follow the norm and punish (if they can).*

Proof. The remainder of the proof can be found in Appendix A.

In what follows, we will use the term "punishment equilibrium" to indicate the steady state equilibrium that features punishment. Proposition 1 highlights the role of history in sustaining norms and punishment meta-norms. People look at recent shared history to form their expectations of the future. Adults active at t who have been exposed to a recent history containing punishment have an incentive to punish because they know that other t adults will punish, since they have observed the same shared history as they have. The behavior of other t adults matters because, in the aggregate, it shapes history at $t + 1$. If punishment occurs at t , recent history in $t + 1$ will feature punishment. In turn, this will induce people in $t + 1$ to punish as well (since they correctly expect that their shared past will be projected into the future). In sum, people punish norm-breakers to pass information about the past to the younger generation. This creates a link between the past and the future which ensures that information about the past is important (since the past shapes the future).

In our analysis, the link between past, present and future is crucial to sustain norms. If this link were to break up (as e.g., in the case of a major disruption such as a war) this might start a process of unraveling that may work very quickly. This remark is also consistent with the observation

that, sometimes, seemingly irrelevant events may have rapid and lasting consequences for norms (Bicchieri and Muldoon 2014 [6]). Note that this role of history and continuity is entirely absent from other accounts, such as e.g., preference-based theories of punishment. Indeed, in their survey article, Bicchieri and Muldoon [6] conclude that ²³

Studies as disparate as the analysis of Prohibition support, racial integration [and] the sexual revolution in the 1960s (...) all lend credibility to a model of norms grounded on individuals' (...) expectations of what others will do (...).

Finally, note that, although we do not explore this explicitly, norm-following and punishment may survive only in sufficiently homogeneous environments, where parents share a similar understanding of the prevailing norm and metanorm. Intuitively, that is required to ensure that parental behavior is informative about the behavior of other people in society. This shares similarities with Adriani and Sonderegger (2009) [45] and suggests a possible rationale for favoring homogeneous societies, namely that they may make punishment-sustained pro-social norms easier to maintain. ²⁴

In the benchmark case where γ , the probability that the state of the world is l , is arbitrarily close to (although strictly smaller than) one, p_I is arbitrarily small and, thus:

$$\pi_H(1,1) \rightarrow 1 \text{ and } \pi_I(1,1) \rightarrow 0 \quad (10)$$

As a result, condition (a) simplifies to

$$\min\{\Theta(1-q) - \varepsilon - \theta q, \delta\Theta\} > c \geq \delta\Theta(1-q). \quad (11)$$

This has the advantage of being independent of prior beliefs (which are very difficult to measure). Note that, although children know that state h is extremely unlikely, upon observing $(1,0)$ they will necessarily conclude that society finds itself in regime I, since this is the only scenario where $(1,0)$ may emerge as an equilibrium move.

We now turn to welfare. From Lemma 1, the game supports a history-independent equilibrium where nobody follows the norm and nobody punishes in any period, and where all players obtain a payoff of zero. By contrast, the punishment equilibrium supports payoffs that are *below* that amount. In fact, if we consider the benchmark case where norm-following generates no direct externalities (and is thus purely wasteful), we can show that, \square

Proposition 2. *In any punishment equilibrium material welfare is lower than in the history-independent equilibrium.*

Proof. In Appendix A.

Intuitively, this follows because punishment is a very wasteful activity, since it imposes costs on both the punisher and the individual being punished. Clearly enough, however, this observation may be reversed if we were to consider norms that generate positive externalities (such as norms of public good provision). Proposition 2 highlights that positive externalities are a necessary (although not sufficient) condition for a regime featuring punishment to generate higher welfare than the history independent equilibrium. As we will see below in Proposition 4, this result stands in contrast with what happens when reward rather than punishment is used as a signaling device.

We now turn to the conditions that underpin the punishment equilibrium. First, the condition $\frac{p_H + (1-q)^2 p_I}{p_H + (1-q)p_I} \delta\Theta > c$ implies that, upon observing his parent following the norm and punishing, a child believes that regime H is sufficiently likely to induce him to follow the norm. Although children

²³ References to the studies in question can be found in Bicchieri and Muldoon (2014) [6].

²⁴ See Grout et al., (2015) [46] and Kets and Sandroni (2018) [47] for possible norm-related accounts of why mixed societies may instead be preferable.

fear punishment less than parents, they nonetheless find it optimal to follow the norm if they believe that punishment is sufficiently widespread in society. Consider now a child who observes his parent following the norm but not punishing—in which case the child will infer that society is in regime I. Condition $c \geq \delta\Theta(1 - q)$ ensures that, in that case, the child will *not* find it worthwhile to follow the norm. Third, the condition $\Theta(1 - q) - \varepsilon - \theta q > c$ ensures that parents who can punish actually choose to punish in regime I (as well as regime H), to induce their children to follow the norm. This creates a wedge between the child's and the parents' views. Intuitively, $\Theta(1 - q) - \varepsilon - \theta q > c$ implies $\Theta(1 - q) > c$. Hence, in the I regime, parents would like their children to follow the norm, since the (parental perception of the) benefit this involves in terms of spared punishment outweighs the cost of norm-following. However, as we have seen, if the child knew that the regime is I he would not follow the norm. The following corollary summarizes the nature of the conflict of interests between parents and children in the punishment equilibrium. \square

Corollary 1. *In the punishment equilibrium, the conflict of interests between parents and children takes the following form: Parents would like their children to follow the norm in both the I and the H regimes. By contrast, children (if they had perfect information) would like to follow the norm only in the H regime.*

Note that for the conflict of interests to arise, the child's discount factor (i.e., δ) should be sufficiently smaller than the parent's discount factor (i.e., one). If $\delta = 1$, condition (a) fails and punishment does not emerge.

Robustness to direct communication Corollary 1 helps us address one of the questions we presented in Section 2, namely, why can the parent not just “tell” the child what to do? Suppose that, instead of punishing norm deviators, a parent may just send a message (involving a positive but arbitrarily small cost) to his child, urging him to follow the norm.²⁵ Formally, let $M \in \{0, 1, 2\}$, where $M = 0$ corresponds to “not punish and not send the message”, $M = 1$ corresponds to “not punish but send the message.”, and $M = 2$ corresponds to “punish and do not send the message.” Starting from an equilibrium where communication is not used (i.e., nobody selects $M = 1$), would a parent have an incentive to deviate and use direct communication instead of punishment?²⁶ In the L regime, sending the message is equilibrium-dominated for the parent. Hence, if we refine out-of-equilibrium beliefs using a standard refinement such as Cho and Kreps's Intuitive Criterion, the child should rule this regime out. By the same token, it is clear that the parent might have an incentive to send the message *both* in the H and the I regimes, if he thought the message sufficiently likely to succeed. Note however that the parent's *net* return from the deviation in regime I exceeds that in regime H.²⁷ This follows since the gross payoff from the deviation is the same in both states, while the parent's equilibrium payoff is actually higher in regime H, implying that the deviation is less appealing. There are thus three cases that may arise. The deviation is either (a) equilibrium-dominated in both regimes, or (b) profitable in both regimes, or (c) it is profitable in the I regime and it is equilibrium-dominated in the H regime. In case (a) the Intuitive Criterion does not pin down the child's posterior beliefs following the deviation. In case (b), the Intuitive Criterion establishes that, following the deviation, the child may assign positive probability to I and/or H. In case (c), the Intuitive Criterion establishes that, following the deviation, the child should assign probability one to the I

²⁵ If the message's cost is exactly zero then the argument is strengthened, since the child cannot rule out the message emanating from the L regime.

²⁶ Note that the parent has no incentive to deviate by *both* punishing *and* sending the message since this would involve unnecessary costs (recall that in the equilibrium we are considering the act of punishing is sufficient to induce the child to follow the norm).

²⁷ This stands in contrast with the net returns from punishing, which are actually greater in regime H: the net gain from choosing (1,1) over (1,0) in regime H is $\Theta - \varepsilon - c$, while in regime I it is $\Theta(1 - q) - \theta q - \varepsilon - c < \Theta - \varepsilon - c$. Another difference is that, while the child's beliefs following the deviation are refined through the Intuitive Criterion, his beliefs following the equilibrium move (1,1) are given by standard Bayesian updating, and assign probability $\pi_I(1,1) < 1$ to regime I, as shown in (9).

regime. Overall, we conclude that, if the child believes that the out-of-equilibrium move originated from regime I with probability one, this is consistent with the Intuitive Criterion. The child's best reply is then $n_{(1,2)} = 0$, which of course makes the deviation suboptimal for the parent. The punishment equilibrium is thus robust to direct communication. Note that this argument can be seen as a special case of the analysis provided in Section 6.1 (this is formalized in the Proof of Proposition 6).

Comparative statics Proposition 1 highlights interesting comparative statics. The second inequality in condition (a) requires

$$c \geq \delta\Theta(1 - q). \quad (12)$$

In other words, the cost of following the norm cannot be too low. This may at first appear strange. If a norm is cheaper than another, how could it ever be harder to sustain? The intuition is subtle. If the cost of following the norm is very low, this would eliminate the conflict of interest between adults and youngsters and would thus render signaling redundant. More specifically, if the cost of norm following is very low, then the child would find it optimal to follow the norm even society is in the I regime. By simply following the norm, a parent could then signal that the regime is I, and thus induce the child to follow the norm with no need for additional signaling. This would trigger a process of unraveling which would destroy the use of punishment (and norm-following) in equilibrium.

Note that the RHS of (12) is an increasing function of Θ , the cost of being punished. This implies that a higher cost of being punished raises the lower bound on c that needs to be met for punishment to emerge in equilibrium. The reason is again linked with the necessity of a conflict of interests between parents and children.

These comparative statics set our model apart from other possible explanations for punishment, such as the existence of direct preferences for punishing (which actually delivers no comparative statics at all), or the idea that people may punish to avoid being punished themselves.

5. Reward Equilibrium

In this section, we construct a setup where signaling operates through an alternative mechanism that has been presented by the literature as relevant to sustain norms (e.g., Herold 2012 [33]), namely rewards. At first glance, a full-blown analysis of this case may appear redundant. After all, one may see the failure to punish someone as a "reward", and the failure to reward someone as a "punishment". This intuition suggests that there may be a one-to-one correspondence between the two cases. However, this analogy is misleading. To see this, suppose that rewards are costly to give—if they were not, their use to incentivize norm-following would be trivially sustained. In this case, then the act of punishing someone by failing to reward them would actually make the punisher better off, since it would save him the cost of rewarding.²⁸ This shows that the analogy between punishing and rewarding fails if we are interested in comparing the use of costly rewards and costly punishments as possible incentive devices for norm-following.

5.1. Setup

The setup mirrors that of Section 3 in most respects. Let $M_i \in \{0, 1\}$ indicate whether an individual follows the metanorm of rewarding ($M_i = 1$) or failing to reward ($M_i = 0$) norm-followers. Similar to existing literature (Herold 2012) we model the act of rewarding as zero-one. The payoff from (N_i, M_i) for a t parent is equal to

$$u_i(b_{t-1}, m_{t+1}, N_i, M_i) = (Rm_{t+1} - c) N_i - rM_i(2 - b_{t-1}) \quad (13)$$

²⁸ In an equilibrium of that nature, the act of rewarding would actually be part of the norm itself. An example may be a norm which consists of "sharing resources with all those who also share their resources", and where punishment for failing to do so takes the form of being excluded from sharing. In this case, punishment is not a costly act to perform. As a result, its use to sustain norm-following is not surprising from a theoretical viewpoint.

where $2 - b_{t-1}$ is the mass of $t - 1$ parents and children who followed the norm at $t - 1$ (and who are being rewarded at t) and m_{t+1} is the mass of $t + 1$ parents who reward norm followers. Note that if an individual follows the norm at t , he is rewarded at time $t + 1$. Intuitively, the type of rewards we have in mind take the form of receiving costly favors at some point in the future. Examples include preferential treatment when looking for a job or asking for a loan, or more generally receiving help when the need arises. This lag implies that time t adults reward people who followed the norm at time $t - 1$. In turn, the children who followed/violated the norm at $t - 1$ (and who are rewarded at t) become adults (and potential rewarders) at $t + 1$. The marginal benefit from being rewarded for following the norm is R . We assume that $R > c$ to ensure that rewards can be used as an effective incentive for norm-following. The parameter $r > 0$ captures the marginal cost of rewarding an additional individual.

Availability of the rewarding option Similar to the case of punishing, we allow for the possibility that, for some parents, rewarding may be unfeasible or prohibitively costly. There are two states of the world.

- with probability γ , the state of the world is l : the rewarding option is available for all parents.
- with probability $1 - \gamma$ the state of the world is h : the rewarding option is unavailable to a share $q \in (0, 1)$ of parents.

Children For children, the payoff from $n_{\bar{i}}$ is given by

$$\bar{u}_{\bar{i}}(m_{t+1}, n_{\bar{i}}) = (\delta R m_{t+1} - c) n_{\bar{i}} \tag{14}$$

where $\delta < 1$ is the child’s discount factor (as in Section 3). We thus maintain the assumption that children discount the future more heavily than parents. Parents are altruistic but exhibit imperfect empathy. The total utility of an adult i is thus given by

$$U_i(b_{t-1}, m_{t+1}, N_i, M_i, n_{\bar{i}}) \equiv u_i + \bar{u}_{\bar{i}} + n_{\bar{i}} m_{t+1} R (1 - \delta). \tag{15}$$

Timing The timing of events is as follows (Figure 2)

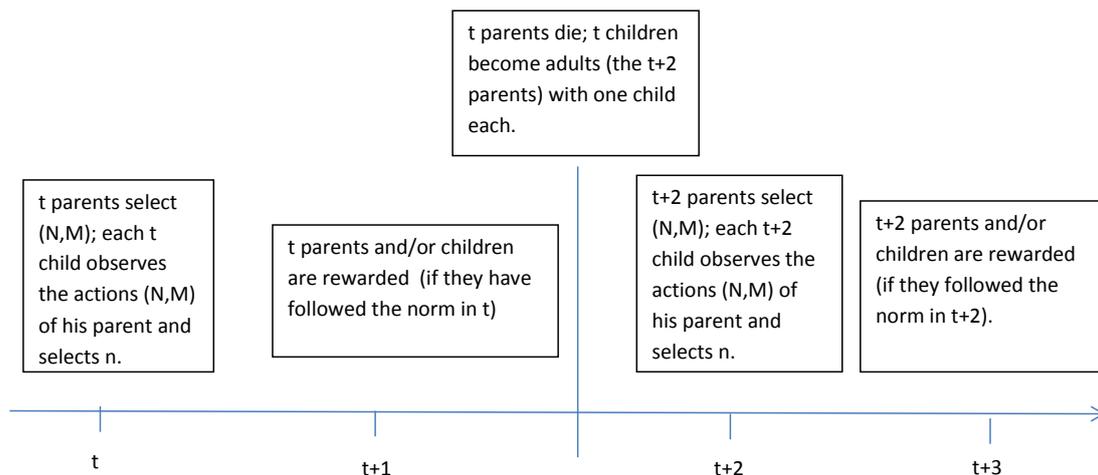


Figure 2. Timing 2.

5.2. Signaling

Similar to Section 4, we restrict attention to environments where Refinement 1 holds. Similar to the case of punishment, it is easy to show that, in steady state, there are three possible regimes:

- H regime: all parents follow the norm and reward norm-followers.

- I regime: all parents follow the norm but only a share $1 - q$ rewards since the remaining share q faces prohibitive costs of rewarding
- L regime: nobody rewards and nobody follows the norm.

The following proposition identifies the conditions for reward to feature in a refined steady state equilibrium.

Proposition 3. *Suppose that p_H and p_I are strictly greater than zero and*

$$\min\{R - 2r, R(1 - q) - r(2 - q), \frac{p_H + (1 - q)^2 p_I}{p_H + (1 - q)p_I} \delta R\} > c > \delta R(1 - q) \quad (b)$$

Then rewarding behavior may emerge in equilibrium. In the refined steady state equilibrium with reward, the following holds.

- (i) *Children follow the norm if and only if they observe their parent rewarding.*
- (ii) *Adults follow the norm and reward (if they can). If condition (b) fails, then the unique refined steady state equilibrium is one where nobody follows the norm and nobody rewards.*

Proof. See Appendix B.

In what follows, we will use the term “reward equilibrium” to indicate the steady state equilibrium that features reward.²⁹

Condition (b) shares many similarities with condition (a). For instance, it shows that, for a reward equilibrium to emerge, the cost of following the norm cannot be too low, similar to what we found in the case of punishment. Another noteworthy feature of condition (b) is the requirement $R - 2r > c$. Since $c > 0$, this requirement implies $R > r$, i.e., the value of the reward for the person who receives it should exceed the cost for the rewarder. This arises from the parent’s incentive compatibility constraint in the H regime. If it did not hold, then the parent would always be better off by selecting (1,0), i.e., follow the norm but not reward, instead of (1,1), i.e., follow the norm and reward. Rewarding would therefore be a dominated strategy.³⁰ Intuitively, this is because in that case the costs of rewarding norm-followers would be too high to be worth incurring.

This has implications for the *type* of reward that may emerge. An arrangement where people reward by donating resources they have been abundantly endowed with and are rewarded by receiving goods they have been scarcely endowed with may fit this description.³¹ By contrast, rewards that are costly for those bestowing them but have little value for the receivers would not work. The same would be true of an environment where the cost of rewarding equals the benefit from being rewarded, so that $r = R$, as for instance may be the case if rewards take the form of monetary transfers. This may help explain why, in social exchanges, people tend to reward others in *kind* rather than through money.³² Differently from the punishment equilibrium, in the case of rewards. □

²⁹ Similar to the case of punishment, it is instructive to look at the benchmark case where γ , the probability that the state of the world is l , is arbitrarily close to (although strictly smaller than) one (so p_I is arbitrarily small). In that case, condition (b) simplifies to

$$\min\{R(1 - q) - r(2 - q), \delta R\} > c > \delta R(1 - q). \quad (16)$$

independent of prior beliefs.

³⁰ In the H regime, the payoff from (1,1) is $-2c - 2r + 2R$ while that from (1,0) is $-c + R$. If $R < c + 2r$, then the latter clearly exceeds the former.

³¹ This argument assumes that utility from consumption is concave, and that there are non-negligible transaction costs for converting goods to cash.

³² In a different setup, Ellingsen and Johannesson (2011) [48] argue that a signaling explanation may lie at the core of the taboo against rewarding people using cash which we observe in many social occasions. Our story can be seen as complementary to theirs.

Lemma 5. *In any reward equilibrium, the adult's material payoff (i.e., calculated by ignoring the child's welfare) is always larger than zero.*

Proof. In Appendix A.

The direct implication is that, even in the benchmark case where norm-following generates no direct externality (and is thus purely wasteful from a welfare viewpoint), the following holds. \square

Proposition 4. *In any reward equilibrium aggregate material welfare is higher than in the history-independent equilibrium.*

Proof. In Appendix A.

Hence, in stark contrast with the punishment case, here any regime that features norm-following is necessarily welfare-superior to the regime where the norm is ignored (and no rewarding takes place). This happens despite the fact that the norm is actually socially wasteful. The intuition is that, as discussed, the net benefits created by the act of rewarding compensate for the welfare loss generated by costly norm-following. Clearly enough, Proposition 4 continues to hold if we allow for direct positive externalities of norm-following, but may cease to hold if we consider a norm which generates sufficiently strong *negative* direct externalities. \square

6. Robustness

This section studies the robustness of the punishment and reward equilibria characterized in Propositions 1 and 3. More specifically, we wish to test whether punishment equilibria are robust to people starting to use rewards instead, and whether reward equilibria are robust to people starting to use punishment. We also wish to establish whether there are asymmetries, i.e., whether one signaling mechanism is somehow more robust than the other.

6.1. Robustness of Reward Equilibrium to Punishment

We start off by considering the robustness of the reward equilibrium to deviations involving punishment. Suppose that, instead of rewarding norm followers, a parent may select to punish norm deviators. Formally, let $M \in \{0, 1, 2\}$, where $M = 0$ corresponds to “not reward, not punish”, $M = 1$ corresponds to “reward, not punish”, and $M = 2$ corresponds to “not reward, punish.” We wish to ask the following question. Starting from a reward equilibrium (i.e., where no parent selects $M = 2$), and keeping the behavior of all other parents fixed, when would an individual find it optimal to (unilaterally) use punishment rather than reward as a signaling device? As in the previous sections, we apply the Intuitive Criterion. To evaluate whether a deviation may possibly be profitable for a parent, we consider the standard case of unilateral deviations. As in Section 3, we let the cost incurred by punishing be given by $\theta b + \varepsilon$.

The first observation is that, in the L reward regime, deviating cannot possibly bring any benefits, since the parent does not actually want his child to follow the norm. By contrast, deviating may possibly be beneficial in the H and I reward regimes, since in those cases the parent *does* want the child to follow the norm (and ideally would like to achieve that by incurring the smallest possible signaling cost). Consider now a unilateral deviation that consists in following the norm and punishing norm-breakers. Suppose that the deviation induces the child to follow the norm. The key remark is that this deviation would deliver a higher payoff in the H reward regime than the I reward regime (since, in the latter, more people break the norm and should therefore be punished). Moreover, the difference between the two payoffs is increasing in θ , the marginal cost incurred by punishing. Compare now the payoff obtained from deviating with equilibrium payoffs. Clearly enough, if ε is small, a successful deviation would always be profitable in the H reward regime. Moreover, if θ is sufficiently high, we can ensure that the deviation in the I reward regime is equilibrium-dominated. In other words, we can always find suitable parameter values that would induce parents to deviate only in the high-reward

state. The deviation is thus a *credible* signal that the regime is high-reward. Figure 3 illustrates the logic of the argument.

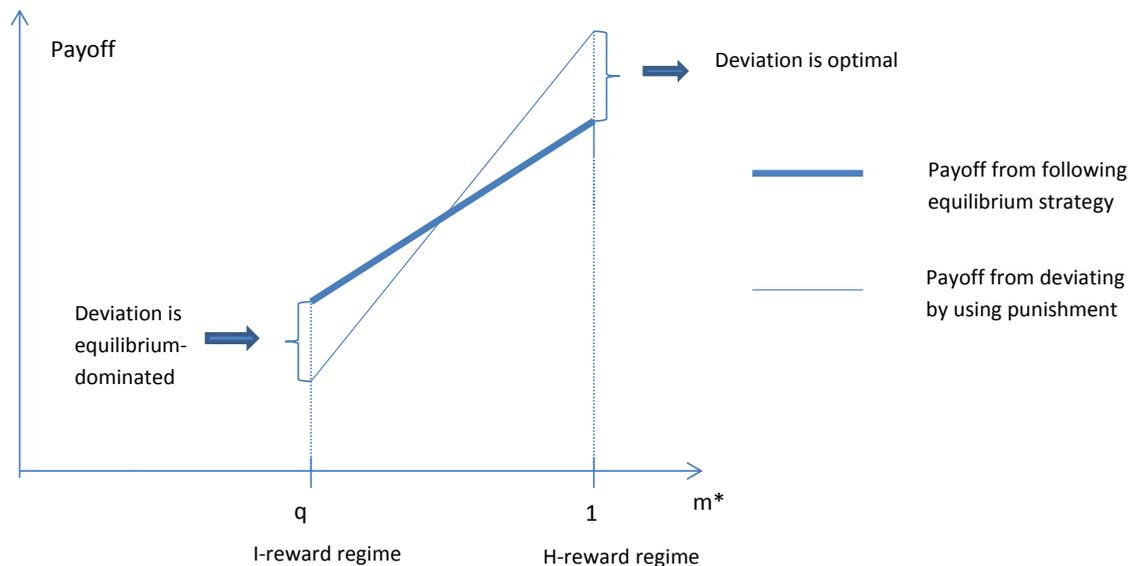


Figure 3. Robustness 1.

Proposition 5. For each $r > 0$ there exist values θ^* and $\epsilon^* > 0$ such that, if $\theta > \theta^*$ and $\epsilon < \epsilon^*$, the reward equilibrium of the augmented game fails the Intuitive Criterion.

Proof. See Appendix A. □

6.2. Robustness of Punishment Equilibrium to Costly Rewards

We now consider the robustness of the punishment equilibrium to deviations involving costly rewards. Let $M \in \{0, 1, 2\}$, where $M = 0$ corresponds to “not punish, not reward”, $M = 1$ corresponds to “punish, not reward”, and $M = 2$ corresponds to “not punish, reward norm followers.” As in Section 5, we let the marginal cost incurred by rewarding norm followers be given by r .

Clearly enough, in the no-punishment regime, deviating is always dominated. By contrast, deviating may possibly be beneficial in the H and I punishment regimes (since in those regimes the parent actually wants the child to follow the norm). The second observation is that the payoff a parent may obtain by deviating and using rewards is always lower in the H punishment regime than in the I punishment regime. Intuitively, this is because in the H regime more people follow the norm and must therefore be rewarded.³³ Consider now the parent’s equilibrium payoff. It is straightforward to show that the equilibrium payoff in the H punishment regime is always higher than in the I punishment regime (since in the latter there are more people who violate the norm and must therefore be punished). This implies that, if a parent deviates in the H punishment regime, he will also find it optimal to deviate in the I punishment regime. Figure 4 illustrates graphically this logic.

³³ Note that this stands in contrast with the *equilibrium payoffs* obtained in the reward equilibrium—these are higher in the H reward regime than the I reward regime. Intuitively, this is because the equilibrium payoff includes the rewards people enjoy when they follow the norm.

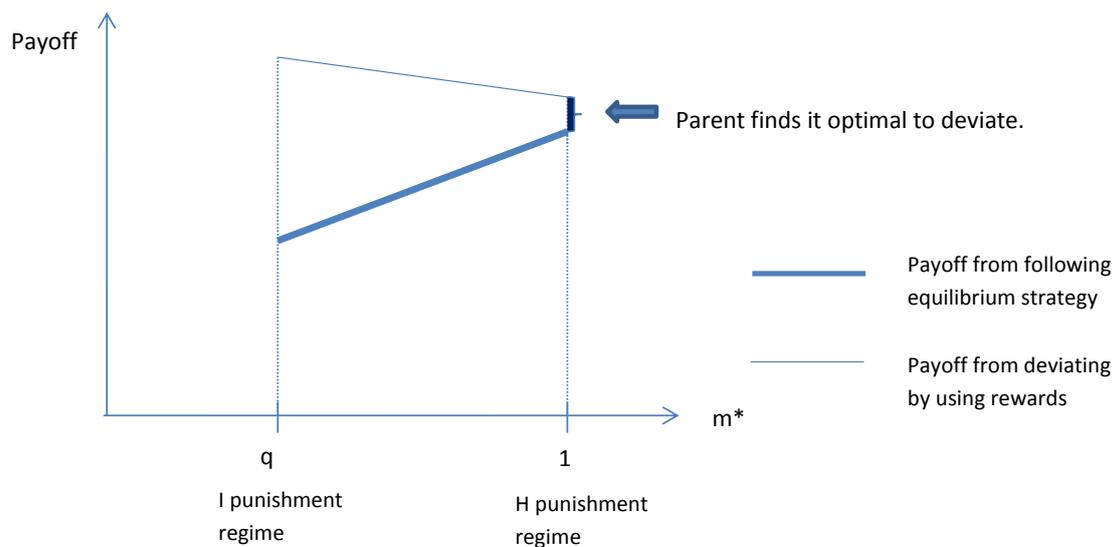


Figure 4. Robustness 2.

The Intuitive Criterion is thus consistent with the belief that the deviation originated in the I punishment regime. As a result, the child will find it optimal to break the norm upon observing the deviation, thus rendering the deviation suboptimal. The punishment equilibrium thus satisfies the Intuitive Criterion. This is summarized in the next proposition.

Proposition 6. *In the augmented game, the punishment equilibrium satisfies the Intuitive Criterion for all $r > 0$.*

Proof. See Appendix A. \square

6.3. Discussion

The intuition for the results in this section can be summarized as follows. Reward does not destabilize the punishment equilibrium because the payoff obtained when using reward out-of-equilibrium (i.e., when everyone else uses punishment) is *decreasing* in the share of norm-followers. Out-of-equilibrium reward is thus ill-suited to effectively signal that norm compliance is high. By contrast, the payoff obtained when using punishment is always *increasing* in the share of norm-followers. Punishment is thus a compelling tool for signaling high compliance.

It is important to highlight that the punishment equilibrium is robust to reward *no matter* how low reward costs may be. To see this, consider an extreme scenario, where reward costs are arbitrarily close to zero. In that case, a successful deviation (i.e., that induces the child to follow the norm) involving rewards *always* yields a higher payoff than the equilibrium move. However, the parent has an incentive to deviate *both* in the H- and the I-punishment regimes. In other words, a deviation involving rewards cannot credibly signal that the regime is high punishment. This is the rationale for the punishment equilibrium's robustness. It is straightforward to see that a similar rationale also guarantees robustness in the case where the parent may deviate by selecting an alternative signal whose cost is independent of the share of norm-followers—e.g., because it involves a fixed cost ϕ , for any $\phi \geq 0$.

Consider now a deviation from the reward equilibrium that involves punishment. If the fixed cost of punishing (ϵ) is sufficiently small, the deviation payoff exceeds the equilibrium payoff in the H-reward regime. However, as we have argued, this is only part of the argument. The additional crucial ingredient is credibility. The payoff from out-of-equilibrium punishment is higher in the H-reward regime than the I regime. For θ high enough, this payoff increases so steeply in the share of norm-followers that it cuts the equilibrium payoff from below, as in Figure 2, thus credibly signaling

that the regime is H. In turn, this ensures that, upon observing this deviation, the child finds it optimal to follow the norm, thus rendering the deviation optimal.³⁴

7. Further Discussion and Concluding Remarks

In this section, we briefly provide further discussions of our results and their implications.

7.1. Which Norms?

Our mechanism continues to work even if we consider norms that do not impose a net individual cost to all, at least as far as adults are concerned. Consider for instance “good lifestyle” behaviors, in matters such as drinking, personal hygiene, sexual promiscuity, eating habits and so on. These behaviors often involve a trade-off, since they require effort, but they also generate benefits, both in the present and in the future (for instance, in the form of improved health). If youngsters discount the future more heavily than adults, they will value the delayed benefits less. Alternatively, adopting the good behavior may require greater effort for youngsters than for adults. It is then possible that, while for parents the net cost of following the good lifestyle norm may be negative (i.e., the benefits more than offset the costs), the opposite may hold for youngsters. Suppose then that we modify our setup as follows. The net cost for parents of following the norm is $c^{parents} < 0$, while the net cost for youngsters is $c^{children} > 0$.³⁵ Clearly enough, our results still apply. In this modified setup parents always want the child to follow the norm, even in the no punishment regime (since, in their eyes, the health benefits from following the norm more than justify its costs). The conflict of interests between parents and children is thus even stronger.

7.2. Infinitely Repeated Games

Many “folk-theorem related” models consider environments where deviations trigger a punishment phase. In the most commonly used framework, punishment consists in a permanent reversion to the static Nash equilibrium (see e.g., Friedman 1971 [49]). Clearly enough, in these models, the problem of sustaining punishment does not emerge (since the punishment phase constitutes a Nash equilibrium). Our work concentrates on environments where the existence of punishment cannot be justified on these grounds. Moreover, in models that use Nash reversion as punishment, equilibrium payoffs are always weakly higher than those of the static Nash equilibrium. This stands in contrast with our results where, as we have seen, in the punishment equilibrium people may end up worse off than they would be in the history-independent equilibrium. The history-independent equilibrium is the unique equilibrium that would emerge if the game were finite (and adults knew their position with respect to the last period).³⁶ The literature on repeated games has, to be sure, identified scenarios where the payoffs obtained in the repeated game may fall below those of the static Nash equilibrium. However, these typically involve so-called “carrot and stick” finite punishment cycles, where the punishment phase is followed by a reward phase (see, e.g., Abreu 1986 [50]). Even leaving all other differences aside, it is clear that our story is fundamentally different.

³⁴ Note that this argument does not necessarily imply that rewards will *never* be observed. In particular, one can envisage that, starting from an environment the norm is rare, norm-following may actually spread thanks to the use of rewards. What we argue, however, is that, if it is successful (in the sense of inducing generalized norm-following), the reward mechanism is liable to be supplanted by punishment.

³⁵ Note that imperfect empathy would then take a dual form: (i) Parents evaluate the child’s cost of following the norm using the lens of their own cost; and (ii) parents evaluate the child’s penalty from being punished in case of norm-breaking using their own discount factor—this is the same as in Section 3.

³⁶ Note however that, if we were to consider the setup discussed in Section 7.1, then this would not be the case. More precisely, in that setup, punishment could emerge even if the game were finite (provided that, in the last period, children are unaware that the game is in its last period).

7.3. More General Parental Utility

In our setup, the parent's utility is given by the sum of his own payoff and the child's payoff (evaluated using the parent's discount factor). This is however a special case of a more general setup where the child's payoff is assigned a weight $\beta \in \mathbb{R}_{>0}$ in the parent's utility, so that (7) becomes

$$U_i(b_{t-1}, m_{t+1}, N_i, M_i, n_i) \equiv u_i + \beta \bar{u}_i - \beta \Theta m_{t+1} (1 - n_i) (1 - \delta). \quad (17)$$

It is easy to verify that this would not affect our findings. In particular, a punishment equilibrium continues to exist, but the necessary and sufficient condition for this is slightly more involved. Condition (a) becomes

$$\min\left\{\Theta - \frac{\varepsilon}{\beta'}, \frac{\beta \Theta (1 - q) - \varepsilon - \theta q}{1 + \beta} \frac{p_H + (1 - q)^2 p_I}{p_H + (1 - q) p_I} \delta \Theta\right\} > c \geq \delta \Theta (1 - q). \quad (a')$$

which of course collapses to (a) when $\beta = 1$.³⁷ A similar observation applies to the reward equilibrium. Condition (b) becomes

$$\min\left\{R - \frac{2r}{\beta}, R(1 - q) - \frac{r(2 - q)}{\beta}, \frac{p_H + (1 - q)^2 p_I}{p_H + (1 - q) p_I} \delta R\right\} > c > \delta R(1 - q) \quad (b')$$

In fact, all our results (including those on robustness outlined in Section 6) would carry through for all values of $\beta \in \mathbb{R}_{>0}$, as well as the limiting case $\beta \rightarrow \infty$ —implying that the parent's concern for the child's payoff is infinitely higher than his concern for his own payoff.³⁸

7.4. Concluding Remarks

Recently, a literature on endogenously derived preferences has sprouted within economics.³⁹ This literature derives preferences from first principles or *ultimate* causes, and thus complements the behavioral literature, which focuses on *proximate* causes.⁴⁰ We believe this is an important agenda. One of the advantages of building models of ultimate causes of behavior is that it generates sharper predictions and comparative statics. We hope that this paper may contribute to this literature by showing that, in an environment with asymmetric information, costly punishment of norm-breakers may emerge optimally as a mean to transmit information to the next generation.

Finally, while the model assumes that the marginal cost of punishing is fixed, it is possible to think of environments where this does not hold. For instance, if punishment takes the form of social boycott, then the opportunity cost incurred (in terms of lost networking opportunities or sources of information) when socially excluding a norm-breaker may be *lower* if others also exclude him. By contrast, the opportunity cost incurred by refusing to, say, offer employment to a past norm-breaker will be *higher* if others practice the same policy (since in that case he would have lower outside options and would therefore be willing to work for less). This suggests that some forms of punishment (e.g., social boycott) may be more likely to be effective signaling devices than others (e.g., refusing employment). Future work might be devoted to refining this intuition.

³⁷ The requirement $\Theta - \varepsilon > c$ is not explicitly included in condition (a) since it is assumed at the outset, in the model section.

³⁸ However, if the parent's utility assigned a weight of *exactly zero* to his own payoff, then Proposition 5 would no longer hold since the parent would be unconcerned with choosing a deviation that allows to send the signal more cheaply.

³⁹ The literature actually dates to Güth and Yaari (1992) [51], but has recently experienced renewed impetus—examples include Robson (2001) [52], Samuelson (2004) [53], Samuelson and Swinkels (2006) [54], Rayo and Becker (2007) [55], Netzer (2009) [56], Herold (2012) [33], Adriani and Sonderegger (2015) [57]. Robson and Samuelson (2010) [58] provide a comprehensive survey.

⁴⁰ See e.g., Binmore (2005) [59] for a discussion of ultimate causes and proximate psychological mechanisms.

Author Contributions: Formal analysis, F.A., S.S.; Writing—Original draft preparation, F.A., S.S.; Writing—Review and editing, F.A., S.S.

Funding: This research received no external funding.

Acknowledgments: We thank two anonymous referees as well as Dough Bernheim, Eddie Dekel, Eugenio Proto, Simon Gächter, Florian Herold, Nick Netzer, Massimo Morelli, Pietro Ortoleva, Luis Rayo, Larry Samuelson, Daniel Seidmann, Joel Sobel, Jean Tirole and seminar audiences at various institutions for comments and discussions. All errors are our own.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Proof of Lemma 1. Let $n_{(N,M)} \in \{0, 1\}$ denote the child’s equilibrium action upon observing (N, M) . We now prove that the following strategies constitute an equilibrium: (i) Children select $n_{(0,0)} = 0$, and $n_{(N,M)} \in \{0, 1\}$ for $(1, 0)$, $(0, 1)$ and $(1, 1)$. [Refinement 1 imposes $n_{(0,1)} = 0$ but since this is not a necessary condition for the equilibrium to hold, we provide a more general proof here.] (ii) Parents always select $(0, 0)$. To see why this represents an equilibrium, consider first parents. In the proposed equilibrium, parents expect that $m_{t+1} = 0$. The expected payoffs from different action-pairs are: $U^t(0, 0) = 0$, $U^t(1, 0) = -c \left(1 + n_{(1,0)}\right)$, $U^t(0, 1) = -cn_{(0,1)} - \theta b_{t-1} - \varepsilon$, $U^t(1, 1) = -c \left(1 + n_{(1,1)}\right) - \theta b_{t-1} - \varepsilon$. Clearly, $(1, 0)$, $(0, 1)$ and $(1, 1)$ are strictly dominated by $(0, 0)$, and this is true independently of the precise values taken by b_{t-1} , $n_{(1,0)}$, $n_{(0,1)}$ or $n_{(1,1)}$. Consider now children. Given the parents’ strategy, the children’s best reply along the equilibrium path is clearly to select $n_{(0,0)} = 0$. The action-pairs $(1, 0)$, $(0, 1)$ and $(1, 1)$ are out-of-equilibrium moves, and, thus, the child’s beliefs upon observing them are not well defined. However, as we have seen, the precise values taken by $n_{(1,0)}$, $n_{(0,1)}$ or $n_{(1,1)}$ are unimportant for the equilibrium to hold. \square

Proof of Lemma 2. Let $n_{(N,M)} \in \{0, 1\}$ be the child’s action upon observing (N, M) . In Lemma 3, below, we prove that, in any equilibrium where punishment occurs, $n_{(1,1)} = 1$ and $n_{(1,0)} = 0$. From (2), it is clear that the child will follow the norm iff $c < \delta \Theta E(m_{t+1} | N, M)$, and will break the norm otherwise. Hence, $E(m_{t+1} | 1, 1) > c / \delta \Theta \geq E(m_{t+1} | 1, 0)$. \square

Proof of Lemma 3. We divide the proof in parts. Part (i): In any equilibrium with punishment, youngsters whose parent follows the norm but does not punish break the norm. Let $n_{(N,M)} \in \{0, 1\}$ be the child’s action upon observing (N, M) . Consider an equilibrium where $n_{(1,0)} = 1$. The parent’s expected payoff from $(1, 0)$ is then $-2c$. Consider now the parent’s payoff from $(1, 1)$. This is

$$-c \left(1 + n_{(1,1)}\right) - \theta b_{t-1} - \varepsilon - \Theta \left(1 - n_{(1,1)}\right) m_{t+1}. \tag{A1}$$

It is straightforward to see that if $n_{(1,1)} = 1$ then (A1) $< -2c$. If $n_{(1,1)} = 0$ then (A1) becomes

$$-c - \theta b_{t-1} - \varepsilon - \Theta m_{t+1}. \tag{A2}$$

Expression (A2) weakly exceeds $-2c$, namely the payoff from $(1, 0)$, iff

$$c \geq \theta b_{t-1} + \Theta m_{t+1} + \varepsilon \tag{A3}$$

If (A3) does not hold, then $(1, 1)$ is dominated by $(1, 0)$. Consider now the parent’s payoff from $(0, 0)$. From refinement 1, $n_{(0,0)} = 0$, and, hence, the payoff from $(0, 0)$ is

$$-2\Theta m_{t+1}. \tag{A4}$$

It is straightforward to see that if (A3) holds, then (A4) is strictly higher than (A2). To recap, when $n_{(1,0)} = 1$ we have two possibilities. If $n_{(1,1)} = 1$ then $(1, 1)$ is surely dominated by $(1, 0)$. If $n_{(1,1)} = 0$, then $(1, 1)$ is either dominated by $(1, 0)$ or it is dominated by $(0, 0)$. Hence, $(1, 1)$ cannot be selected

in equilibrium. Consider now (0,1). By refinement 1, $n_{(0,1)} = 0$. This implies that (0,1) is strictly dominated by (0,0). This proves that, if $n_{(1,0)} = 1$, then punishment cannot emerge in equilibrium. *Part (ii).* In any equilibrium with punishment, youngsters whose parent both follows the norm and punishes follow the norm. The necessary condition for punishment to emerge is that (1,1) should not be a dominated action. From (i), we know that in any equilibrium with punishment it is necessary that $n_{(1,0)} = 0$. The parent’s expected payoff from (1,0) is then $-c - \Theta m_{t+1}$. Consider now the parent’s payoff from (1,1), namely (A1). It is straightforward to see that, if $n_{(1,1)} = 0$, then (1,1) is strictly dominated by (1,0). Hence, in any equilibrium with punishment, we need $n_{(1,1)} = 1$. \square

Proof of Lemma 4. We first establish some general properties. Inspection of (5)–(7) reveals a number of things. First, as already stated, (0,1) is clearly dominated, and, thus, cannot be an equilibrium strategy. In what follows, we will thus ignore the existence of this strategy. Second, we cannot have $U^t(1,1) = U^t(1,0) = U^t(0,0)$. Third, for punishment to emerge at t , it is necessary that $U^t(1,1) \geq U^t(1,0)$ —or else punishing would always be dominated. In turn, it is straightforward to see that this requires that $U^t(1,0) > U^t(0,0)$ must also hold. Hence, if punishment occurs at t , then (0,0) is dominated by (1,0), which implies that all adults must choose to follow the norm. Conversely, for (0,0) not to be dominated, we require $U^t(0,0) \geq U^t(1,0)$. In turn, this requires that $U^t(1,0) > U^t(1,1)$ must hold, i.e., (1,1) is dominated by (1,0). We now explore the implications of these properties for steady state. Above we have concluded that, for (0,0) not to be dominated at t , (1,1) must be dominated by (1,0) and, thus, $m_t = 0$ (recall that (0,1) is always dominated). In steady state, this then implies $m^* = 0$, i.e., nobody punishes. Since $m^* = 0$, it is clear that following the norm cannot possibly be optimal. This proves that (0,0) may never coexist with either (1,1) or (1,0). Hence, (0,0) may emerge in a steady-state equilibrium only if all parents select this action pair (and, consequently, all children break the norm), i.e., $m^* = 0$ and $b^* = 2$. From Lemma 1, we know that, in that case, the optimal action-pair is (0,0). Hence, $m^* = 0$ and $b^* = 2$ is a possible steady state regime. The other possible steady states are (i) all adults who can punish select (1,1), while those who cannot punish select (1,0); (ii) all adults select (1,0) (so that by Lemma 3 all children break the norm).

In case (i) we have two possibilities: (1) the state of the world is l . In that case, $m^* = 1$ and $b^* = 0$. Payoffs are: $U(1,1) = -2c - \varepsilon$, $U(1,0) = -c - \Theta$, $U(0,0) = -2\Theta$. It is straightforward to show that, provided that

$$\Theta - \varepsilon > c \tag{A5}$$

holds, (1,0) and (0,0) are dominated by (1,1). Hence, $m^* = 1$ and $b^* = 0$ is a possible steady state regime. The second possibility for case (i) is (2) the state of the world is h . In that case, $m^* = 1 - q$ and $b^* = q$. Payoffs are: $U(1,1) = -2c - \varepsilon - \theta q$, $U(1,0) = -c - \Theta(1 - q)$, $U(0,0) = -2\Theta(1 - q)$. Consider first parents who cannot punish. For them, (1,1) is obviously not an option. Selecting (1,0) yields a higher payoff than (0,0) if

$$\Theta(1 - q) > c \tag{A6}$$

Consider now those parents who can punish. The necessary and sufficient condition for option (1,1) to dominate both (1,0) and (0,0) is

$$\Theta(1 - q) - \varepsilon - \theta q > c. \tag{A7}$$

Note that, if (A7) holds, (A6) and (A5) automatically hold too (but not vice-versa).

In case (ii), $m^* = 0$ and $b^* = 1$. Payoffs are: $U(1,1) = -2c - \theta - \varepsilon$, $U(1,0) = -c$, $U(0,0) = 0$. Hence, (1,0) is strictly dominated by (0,0). This implies that case (ii) cannot be a steady-state equilibrium. \square

Proof of Proposition 1. (i) We prove the necessity of condition (a). Suppose that parents follow the strategy described in the proposition, and let prior beliefs be p_H and p_L , both strictly positive.

From Lemma 3, we know that in any equilibrium with punishment, $n_{(1,0)} = 0$ and $n_{(1,1)} = 1$ must hold. Upon observing (1,1), the child's posterior expectation of m_{t+1} is equal to $\pi_H(1,1) + (1 - q)\pi_I(1,1)$. To obtain $n_{(1,1)} = 1$ in equilibrium, we thus require $-\left[\pi_H(1,1) + (1 - q)\pi_I(1,1)\right]\delta\Theta < -c$. Substituting for $\pi_H(1,1) = p_H / (p_H + (1 - q)p_I)$ and $\pi_I(1,1) = (1 - q)p_I / (p_H + (1 - q)p_I)$ and rearranging, we obtain

$$\frac{p_H + (1 - q)^2 p_I}{p_H + (1 - q)p_I} \delta\Theta > c. \tag{A8}$$

Moreover, we know from the proof of Lemma above that, for parents who can punish, the necessary and sufficient condition for option (1,1) to dominate (1,0) and (0,0) in both regimes H and I is that

$$\Theta(1 - q) - \varepsilon - \theta q > c. \tag{A9}$$

This explains the LHS of condition (a). Consider now the RHS of (a). Upon observing (1,0), the child's posterior expectation of m_{t+1} is $(1 - q)$. To obtain $n_{(1,0)} = 0$ in equilibrium, we thus require $-(1 - q)\delta\Theta \geq -c$. Rearranging, this gives the RHS of (a). This proves that (a) is necessary for a punishment equilibrium. (ii) *We prove the necessity of the requirement $p_H > 0$ and $p_I > 0$.* First, note that, if $p_H = 0$, then an equilibrium with punishment cannot exist since in that case $n_{(1,1)} = 1$ and $n_{(1,0)} = 0$ cannot hold simultaneously. Suppose now that $p_I = 0$, so that (1,0) is an out-of-equilibrium move by the parent. Let $n_{(1,0)} \in \{0, 1\}$ be the child's action following this out-of-equilibrium move. The parent's payoff from selecting (1,0) in the H regime is

$$-c \left(1 + n_{(1,0)}\right) - \Theta \left(1 - n_{(1,0)}\right). \tag{A10}$$

The parent's payoff from selecting (1,0) in the L regime is

$$-c \left(1 + n_{(1,0)}\right). \tag{A11}$$

It is straightforward to see that (1,0) is necessarily equilibrium-dominated in the L regime (since (A11) < 0). However, this is not the case in the H regime: If $n_{(1,0)} = 1$, then (A10) becomes $-2c$ and is therefore higher than the equilibrium payoff in the H regime, namely $-2c - \varepsilon$. Under the Intuitive Criterion, upon observing the out-of-equilibrium move (1,0) the child should therefore conclude that the regime is H. The child's optimal reply is thus $n_{(1,0)} = 1$. In turn, this makes punishing dominated for parents (which in turn renders norm-following suboptimal as well). This implies that, if $p_I = 0$, the H regime would unfold. The only possible refined steady state equilibrium would then be one where nobody follows the norm and nobody punishes. We conclude that $p_H > 0$ and $p_I > 0$ are necessary for punishment to emerge in a refined steady state equilibrium. (iii) *We show that the punishment equilibrium cannot be upset by out-of-equilibrium moves.* Since both p_H and p_I are strictly positive, the only possible cases of zero-probability events that may arise are $(N_i, M_i) = (0, 1)$ or $(N_i, M_i) = (0, 0)$ and $p_N = 0$. It is straightforward to show that selecting (0,1) is strictly dominated by (1,1) in both the H and the I regimes. Hence, the parent could never gain from selecting this out-of-equilibrium move in either the H or the I regime. Suppose now that $p_L = 0$ and consider now a parent who selects the move (0,0). In the H regime, this is strictly dominated by (1,1) independently of the precise value of $n_{(0,0)}$. Hence, under the Intuitive Criterion, upon observing (0,0) the child must rule the H regime out. The child's optimal reply is then $n_{(0,0)} = 0$. Given this, it is straightforward to see that a parent could also never gain from selecting the out-of-equilibrium move (0,0) in the I regime. [Note that by refinement 1 we could have imposed $n_{(0,0)} = 0$ at the outset. However, as we have shown, this is actually not necessary for the argument.] (iv) *Finally, we argue that, given p_H and p_I are strictly positive, (a) is a sufficient condition for the punishment equilibrium to exist.* To see this, note that, under (a), the child's best replies are $n_{(1,0)} = 0$ and $n_{(1,1)} = 1$. Given (A9), the parent's best reply is (1,1) if he can punish and (1,0) if he cannot. □

Proof of Proposition 2. Total material welfare in the history-independent equilibrium is zero. Consider now the H regime. In any given period t , the material payoff obtained by each adult is $-c - \varepsilon$, while the material payoff obtained by each child is $-c$. Since each of these constituents is <0 , it is clear that material welfare here is strictly negative. Consider now the I regime. In any given period t , the material payoff obtained by a share $1 - q$ of adults is $-c - \theta(1 - q) - \varepsilon$, while that of the remaining q share is $-c$. The material payoff obtained by a share $1 - q$ of children is $-c$, while that obtained by the remaining q share is $-\Theta(1 - q)$.⁴¹ Since each of these constituents is <0 , it is clear that material welfare in the I regime is also strictly negative. \square

Proof of Lemma 5. In the H regime, the parent’s direct payoff is $-c - 2r + R$ (since he rewards a mass 2 of people, namely all adults and all children active in the previous generation, and he is rewarded by a mass one of people, namely all adults active next period). By condition (b), this is greater than zero. In the I regime, the parent’s direct payoff is $-c - r(2 - q) + R(1 - q)$ if he can punish, and $-c + R(1 - q)$ otherwise. Under condition (b), both are strictly greater than 0. \square

Proof of Proposition 4. In Lemma 5 we have shown that in any reward equilibrium the material payoff of parents is strictly positive. Consider now children. By breaking the norm, a child knows that he can obtain 0. Hence, in any equilibrium with reward, children must obtain at least 0. [Note that this argument holds independently of whether we use the child’s or the parent’s discount factor to evaluate the child’s material welfare.] Summing up, this implies that in any reward equilibrium welfare is higher than in the history-independent equilibrium. \square

Proof of Proposition 5. We formalize the graphical intuition provided in the text. Recall that equilibrium payoffs for the parent in the different reward regimes are as follows. H reward regime:

$$-2c - 2r + 2R \tag{A12}$$

I reward regime (if he can reward):

$$-2c - r(2 - q) + 2R(1 - q) \tag{A13}$$

if he can not reward $-c + R(1 - q)$

Consider the out-of-equilibrium move (1, 2), in which the parent follows the norm and punishes norm-breakers (rather than rewarding norm followers). Suppose that $n_{(1,2)} = 1$. In the H reward regime, the payoff from the deviation is

$$-2c + 2R - \varepsilon, \tag{A14}$$

in the I reward regime it is

$$-2c + 2R(1 - q) - \theta q - \varepsilon \tag{A15}$$

while in the L reward regime it is $-2c - 2\theta - \varepsilon$. Clearly, the deviation is equilibrium-dominated in the L regime, since there the parent obtains 0 in equilibrium. If

$$\frac{\theta q + \varepsilon}{2 - q} > r > \frac{\varepsilon}{2} \tag{A16}$$

then the deviation payoff in the I regime, (A15), is smaller than the equilibrium payoff (A13), while the deviation payoff in the H regime, (A14), exceeds the equilibrium payoff (A12). When (A16) holds, the Intuitive Criterion requires that, upon observing (1, 2), the child must rule the L and I reward regimes out, and must assign probability 1 to the H reward regime. The child’s optimal reply is thus

⁴¹ We are using the parent’s discount factor, namely 1, to evaluate the child’s material welfare. This is immaterial.

$n_{(1,2)} = 1$, which renders the deviation (1,2) optimal in the H reward regime. The sufficient conditions for (A16) are: (i) $\theta > r(2 - q)/q$ and (ii) $\varepsilon < 2r$. We conclude that, for each $r > 0$ there exist values θ^* and $\varepsilon^* > 0$ such that, if $\theta > \theta^*$ and $\varepsilon < \varepsilon^*$, the reward equilibrium of the augmented game fails the intuitive criterion. \square

Proof of Proposition 6. Recall the equilibrium payoffs for the parent in the different punishment regimes. H punishment regime:

$$- 2c - \varepsilon \tag{A17}$$

I punishment regime:

$$- 2c - \theta q - \varepsilon \tag{A18}$$

$-c - \Theta(1 - q)$ if they can not punish.

(i) Consider first the out-of-equilibrium move (1,2), in which the parent follows the norm and rewards norm-followers. It is clear that if $n_{(1,2)} = 0$ the deviation cannot be profitable. Suppose now that $n_{(1,2)} = 1$. In the H punishment regime, the payoff from the deviation is ⁴²

$$- 2c - 2r \tag{A19}$$

in the I punishment regime it is

$$- 2c - r(2 - q). \tag{A20}$$

while in the L punishment regime it is $-2c$. Clearly, the deviation is equilibrium-dominated in the L regime, since there the parent obtains 0 in equilibrium. Consider now the H punishment regime. For the deviation to be profitable, we require (A19) to exceed (A17), the equilibrium payoff. This happens if

$$\varepsilon > 2r. \tag{A21}$$

Note that if (A21) holds then (A20) necessarily exceeds the equilibrium payoff in the I punishment regime when the parent can punish. ⁴³ This implies that, under the Intuitive Criterion, the child's posterior beliefs following (1,2) cannot *simultaneously* rule out the I punishment regime *and* assign positive probability to the H punishment regime. The deviation is either (a) equilibrium-dominated in both regimes, or (b) profitable in both regimes, or (c) it is profitable in the I punishment regime and it is equilibrium-dominated in the H punishment regime. In case (a) the Intuitive Criterion does not pin down the child's posterior beliefs following the deviation; In case (b), the Intuitive Criterion establishes that following the deviation the child should rule out regime L (but may assign positive probability to I and/or H); In case (c), the Intuitive Criterion establishes that following the deviation the child should assign probability 1 to the I punishment regime. All three cases (a)–(c) are consistent with the child's reply following the out-of-equilibrium move being $n_{(1,2)} = 0$. If we set $n_{(1,2)} = 0$ (which we can do, since as we have just argued this does not contradict the Intuitive Criterion), then the out-of-equilibrium move (1,2) is unprofitable in all regimes. (ii) Consider now the out-of-equilibrium move (0,2), in which the parent breaks the norm and rewards norm-followers. It is clear that if $n_{(0,2)} = 0$ the deviation cannot be profitable. Suppose then that $n_{(0,2)} = 1$. In the H punishment regime, the payoff from the deviation is

$$- c - 2r - \Theta \tag{A22}$$

in the I punishment regime it is

$$- c - r(2 - q) - \Theta(1 - q). \tag{A23}$$

⁴² We restrict attention to $\tau = r$ since, clearly, the deviation is feasible only if the parent faces non-prohibitive rewarding costs.

⁴³ This follows since $r(2 - q) - \theta q < 2r$.

while in the L punishment regime it is $-2c$. Clearly, the deviation is equilibrium-dominated in the L punishment regime. Consider now the H punishment regime. For the deviation to be profitable, we require (A22) to exceed the equilibrium payoff, namely (A17). This happens if

$$\varepsilon > 2r - c + \Theta. \quad (\text{A24})$$

Note that if (A24) holds then (A23) necessarily exceeds the equilibrium payoff in the I punishment regime when the parent can punish.⁴⁴ This implies that, under the Intuitive Criterion, the child's posterior beliefs following (0, 2) cannot *simultaneously* rule out the I punishment *and* assign positive probability to the H punishment regime. The remainder of the argument is then the same as in part (i).

Robustness to direct communication. Consider the out-of-equilibrium move in which the parent follows the norm and sends the message (rather than punish). Let $t \geq 0$ be the cost of sending the message. If upon receiving the message the child follows the norm, the payoff from deviating for the parent is $-2c - t$ both in the H and the I punishment regimes. Since equilibrium payoff in the H punishment regime exceeds that in I regime, it follows that, under the Intuitive Criterion, the child's posterior beliefs following the deviation cannot simultaneously rule out the I punishment regime and assign positive probability to the H punishment regime. The case where the parent's out-of-equilibrium move consists of breaking the norm and sending the message is analogous. \square

Appendix B. Reward Equilibrium

To characterize the reward equilibrium, we retrace the steps we took in Section 4.

Lemma A1. *(Strategy B.) In an equilibrium with reward, the following must hold. A child who has observed his parent both following the norm and rewarding will follow the norm. A child who has observed his parent following the norm but not rewarding will break the norm.*

Proof. The proof of Lemma A1 is a straightforward adaptation of the proof of Lemma 3 and is therefore omitted.

Suppose then that Refinement 1 holds, and that children follow strategy B. Consider a parent active at time t . We now derive his payoff in each of the available action-pairs. First, the payoff from following the norm and rewarding—namely, $(N_i, M_i) = (1, 1)$ —is given by

$$U^t(1, 1) = -2c - \tau(2 - b_{t-1}) + 2Rm_{t+1}. \quad (\text{A25})$$

The payoff from $(N_i, M_i) = (1, 0)$ —namely, following the norm and failing to reward—is

$$U^t(1, 0) = -c + Rm_{t+1}. \quad (\text{A26})$$

Third, the payoff from $(0, 0)$ —namely, breaking the norm and failing to reward—is

$$U^t(0, 0) = 0. \quad (\text{A27})$$

Finally, the payoff from $(N_i, M_i) = (0, 1)$ —namely, breaking the norm and rewarding—is

$$U^t(0, 1) = -\tau(2 - b_{t-1}). \quad (\text{A28})$$

Several remarks are in order. First, (0, 1) is clearly a dominated strategy. Second, we cannot have $U^t(1, 1) = U^t(1, 0) = U^t(0, 0)$. Third, for reward to emerge at t , it is necessary that $U^t(1, 1) \geq U^t(1, 0)$. In turn, it is straightforward to see that this requires that $U^t(1, 0) > U^t(0, 0)$ must also hold. Hence,

⁴⁴ This follows from $r(2 - q) - \theta q + \Theta(1 - q) < 2r + \Theta$.

if reward occurs at t , then $(0,0)$ is dominated, so that all adults must choose to follow the norm. Conversely, for $(0,0)$ not to be dominated, we require $U^t(0,0) \geq U^t(1,0)$. In turn, this requires that $U^t(1,0) > U^t(1,1)$ must hold, i.e., $(1,1)$ is dominated. Having established these general properties, we can now characterize possible steady states. \square

Lemma A2. *Suppose that (1) children follow strategy B and (2) refinement 1 holds. The possible steady states are:*

- (i) *High-reward: All parents follow the norm and (if they can) reward. As a result, in state l : $m^* = 1, b^* = 0$, while in state h : $m^* = 1 - q, b^* = q$.*
- (ii) *No-reward: $m^* = 0, b^* = 2$. Nobody rewards, nobody follows the norm.*

Proof. First, the action-pair $(0,1)$ is always strictly dominated and therefore cannot be an equilibrium move. In what follows, we will therefore ignore strategy $(0,1)$. Second, the existence of a no-reward steady state follows from lemma 1. Moreover, from the argument provided in the text above, we know that $(0,0)$ may never coexist with either $(1,1)$ or $(1,0)$. Hence, the possible steady state regimes are: (i) All adults who can select $(1,1)$, while those with prohibitively costly rewarding costs (if any) select $(1,0)$; (ii) All adults select $(1,0)$.

In case (i), we have two possibilities: (1) The state of the world is l . In that case, $m^* = 1, b^* = 0$. Payoffs are $U(1,1) = -2c - 2r + 2R, U(1,0) = -c + R$ and $U(0,0) = 0$. Provided that

$$R - 2r > c \tag{A29}$$

$(1,0)$ and $(0,0)$ are dominated by $(1,1)$. This proves that $m^* = 1, b^* = 0$ is a possible steady state. The second possibility that may arise in case (i) is: (2) the state of the world is h . In that case, $m^* = 1 - q, b^* = q$. Payoffs are: $U(1,1) = -2c - \tau(2 - q) + 2R(1 - q), U(1,0) = -c + R(1 - q)$ and $U(0,0) = 0$. Consider first parents who can reward. Provided that

$$R(1 - q) - r(2 - q) > c \tag{A30}$$

$(1,0)$ and $(0,0)$ are dominated by $(1,1)$. Note that when (A30) hold, (A29) automatically holds, but not vice-versa. Consider now parents who cannot reward. It is straightforward to show that, under (A30), they strictly prefer $(1,0)$ to $(0,0)$. In case (ii), $m^* = 0, b^* = 1$. As a result, $U(1,1) = -2c - \tau, U(1,0) = -c$ and $U(0,0) = 0$. Clearly, $(1,0)$ is strictly dominated by $(0,0)$. This implies that $m^* = 0$ and $b^* = 1$ is not a possible steady state. \square

We are now able to prove Proposition 3.

Proof of Proposition 3. (i) *We first prove that condition (b) is necessary.* The necessity of (A30) is proved above in the proof of Lemma 4. We now prove the necessity of the condition

$$\frac{p_H + (1 - q)^2 p_I}{p_H + (1 - q) p_I} \delta R > c. \tag{A31}$$

Suppose that parents follow the strategy described in the proposition, and let prior beliefs be p_H and p_I , both strictly positive. From Lemma 3, we know that in any equilibrium with punishment, $n_{(1,0)} = 0$ and $n_{(1,1)} = 1$ must hold. Upon observing $(1,1)$, the child's posterior expectation of m_{t+1} is equal to $\pi_H(1,1) + (1 - q)\pi_I(1,1)$. To obtain $n_{(1,1)} = 1$ in equilibrium, we thus require $-c + [\pi_H(1,1) + (1 - q)\pi_I(1,1)] \delta R > 0$. Substituting for $\pi_H(1,1) = p_H / (p_H + (1 - q)p_I)$ and $\pi_I(1,1) = (1 - q)p_I / (p_H + (1 - q)p_I)$ and rearranging, we obtain (A31). We now prove the necessity of the RHS of condition (b). Upon observing $(1,0)$, the child's posterior expectation of m_{t+1} is $1 - q$. To obtain $n_{(1,0)} = 0$ in equilibrium, we thus require $-c + (1 - q)\delta R \leq 0$. Rearranging, this gives the RHS of (b). This proves that (b) is necessary for a reward equilibrium. (ii) *We prove that $p_H > 0$ and*

$p_I > 0$ are necessary. First, note that, if $p_H = 0$, the requirements $n_{(1,1)} = 1$ and $n_{(1,0)} = 0$ cannot hold simultaneously. Suppose now that $p_I = 0$, so that $(1, 0)$ is an out-of-equilibrium move. Let $n_{(1,0)} \in \{0, 1\}$ be the child's action following this out-of-equilibrium move by the parent. The parent's payoff in the H regime is

$$(R - c) \left(1 + n_{(1,0)} \right). \quad (\text{A32})$$

The parent's payoff in the L regime is $-c$. It is straightforward to see that $(1, 0)$ is necessarily equilibrium-dominated in the L regime (since $-c < 0$). However, this is not the case in the H regime: If $n_{(1,0)} = 1$, then (A10) becomes $2(R - c)$ and is therefore higher than the equilibrium payoff in that regime, namely $2(R - c - \tau)$. Under the Intuitive Criterion, upon observing the out-of-equilibrium move $(1, 0)$ the child should therefore conclude that the regime is H. The child's optimal reply is thus $n_{(1,0)} = 1$. In turn, this makes rewarding (and, thus, norm-following) dominated for parents. This implies that, if $p_I = 0$, the H regime would unfold. The only possible refined steady state equilibrium would then be one where nobody follows the norm and nobody rewards. We conclude that $p_H > 0$ and $p_I > 0$ are necessary for reward to emerge in a refined equilibrium. (iii) *We show that the reward equilibrium cannot be upset by out-of-equilibrium moves.* It is straightforward to show that selecting $(0, 1)$ is strictly dominated by $(1, 1)$ in both the H and the I regimes. Hence, the parent could never gain from selecting this out-of-equilibrium move in either the H or the I regime. Consider now $(0, 0)$. In the H regime, this is strictly dominated by $(1, 1)$, independently of the precise value of $n_{(0,0)}$. Hence, under the Intuitive Criterion, upon observing $(0, 0)$ the child must rule the H regime out. The child's optimal reply is then $n_{(0,0)} = 0$. Given this, it is straightforward to see that $(0, 0)$ is equilibrium-dominated also in the I regime. The parent can therefore not gain from selecting this out-of-equilibrium move. (iv) *We prove that, given p_H and p_I are strictly positive, condition (b) is sufficient for a reward equilibrium to exist.* Under (b), the child's best replies are $n_{(1,0)} = 0$ and $n_{(1,1)} = 1$. Given (A30), the parent's best reply is $(1, 1)$ if he can reward and $(1, 0)$ otherwise. \square

References

1. Fehr, E.; Gächter, S. Cooperation and punishment in public goods experiments. *Am. Econ. Rev.* **2000**, *90*, 980–994. [[CrossRef](#)]
2. Feinberg, J. *Doing and Deserving: Essays in the Theory of Responsibility*; Princeton, N.J., Ed.; Princeton University Press: Princeton, NJ, USA, 1970.
3. Kahan, D.M. Social meaning and the economic analysis of crime. *J. Legal Stud.* **1998**, *27*, 661–672. [[CrossRef](#)]
4. Cooter, R.D. Expressive law and economics. *J. Legal Stud.* **1988**, *27*, 585–608. [[CrossRef](#)]
5. Sunstein, C.R. On the expressive function of law. *Univ. Law Rev.* **1996**, *144*, 2021–2031. [[CrossRef](#)]
6. Bicchieri, C.; Muldoon, R. Social Norms. In *The Stanford Encyclopedia of Philosophy*; Zalta Edward, N., Ed.; Spring: Palo Alto, CA, USA, 2014.
7. Rotemberg, J.J. Minimally acceptable altruism and the ultimatum game. *J. Econ. Behav. Organ.* **2008**, *66*, 457–476. [[CrossRef](#)]
8. Akerlof, G. The Economics of Caste and of the Rat Race and Other Woeful Tales. *Q. J. Econ.* **1976**, *90*, 599–617. [[CrossRef](#)]
9. Peski, M.; Szentes, B. Spontaneous discrimination. *Am. Econ. Rev.* **2013**, *6*, 2412–2436. [[CrossRef](#)]
10. Axelrod, R. An evolutionary approach to norms. *Am. Political Sci. Rev.* **1986**, *80*, 1095–1111. [[CrossRef](#)]
11. Ghosh, P.; Ray, D. Cooperation in community interaction without information flows. *Rev. Econ. Stud.* **1996**, *63*, 491–519. [[CrossRef](#)]
12. Glazer, A.; Konrad, K. A signaling explanation for private charity. *Am. Econ. Rev.* **1996**, *86*, 1019–1028.
13. Ellingsen, T.; Johannesson, M. Pride and prejudice: The human side of incentive theory. *Am. Econ. Rev.* **2008**, *98*, 990–1008. [[CrossRef](#)]
14. Hopkins, E. Competitive altruism, mentalizing and signalling. *Am. Econ. J.* **2014**, *6*, 272–292.
15. Sliwka, D. Trust as a signal of a social norm and the hidden costs of incentive schemes. *Am. Econ. Rev.* **2006**, *97*, 999–1012. [[CrossRef](#)]
16. Gneezy, U.; Rustichini, A. A fine is a price. *J. Legal Stud.* **2000**, *29*, 1–18. [[CrossRef](#)]

17. Adriani, F.; Sonderegger, S. Signaling About Norms: Socialization Under Strategic Uncertainty. *Scand. J. Econ.* **2018**, *120*, 685–716. [[CrossRef](#)]
18. Adriani, F.; Matheson, J.; Sonderegger, S. Teaching by Example and Induced Beliefs in a Model of Cultural Transmission. *J. Econ. Behav. Organ.* **2018**, *145*, 511–529. [[CrossRef](#)]
19. Kotsidis, V. *Call to Action: Intrinsic Motives and Material Interest*; Mimeo: New York, NY, USA, 2018.
20. Bénabou, R.; Tirole, J. Identity, morals and taboos: Beliefs as assets. *Q. J. Econ.* **2011**, *126*, 805–855. [[CrossRef](#)]
21. Bénabou, R.; Tirole, J. Self-confidence and personal motivation. *Q. J. Econ.* **2002**, *117*, 871–915. [[CrossRef](#)]
22. Bénabou, R.; Tirole, J. Willpower and personal rules. *J. Political Econ.* **2004**, *112*, 848–887. [[CrossRef](#)]
23. Bénabou, R.; Tirole, J. Belief in a just world and redistributive politics. *Q. J. Econ.* **2006**, *121*, 699–746. [[CrossRef](#)]
24. Dessì, R. Collective memory, cultural transmission, and investments. *Am. Econ. Rev.* **2009**, *98*, 534–560. [[CrossRef](#)]
25. Bisin, A.; Verdier, T. Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits. *Q. J. Econ.* **2000**, *115*, 955–988. [[CrossRef](#)]
26. Corneo, G.; Jeanne, O. A theory of tolerance. *J. Public Econ.* **2009**, *93*, 691–702. [[CrossRef](#)]
27. Corneo, G.; Jeanne, O. Symbolic values, occupational choice and economic development. *Eur. Econ. Rev.* **2010**, *54*, 241–255. [[CrossRef](#)]
28. Cervellati, M.; Vanin, P. Thou shalt not covet: Prohibitions, temptation and moral values. *J. Public Econ.* **2013**, *103*, 15–28. [[CrossRef](#)]
29. Carvalho, J.-P. Veiling. *Q. J. Econ.* **2013**, *128*, 337–370. [[CrossRef](#)]
30. Verdier, T.; Zenou, Y. Cultural leaders and the dynamics of assimilation. *J. Econ. Theory* **2018**, *175*, 374–414. [[CrossRef](#)]
31. Van der Weele, J. The signaling power of sanctions in social dilemmas. *J. Law Econ. Organ.* **2012**, *28*, 103–25. [[CrossRef](#)]
32. Andreoni, J.; Harbaugh, W.; Vesterlund, L. The carrot or the stick: Rewards, punishments, and cooperation. *Am. Econ. Rev.* **2003**, *93*, 893–902. [[CrossRef](#)]
33. Herold, F. Carrot or Stick: The evolution of reciprocal preferences in a haystack model. *Am. Econ. Rev.* **2012**, *102*, 914–940. [[CrossRef](#)]
34. Acemoglu, D.; Jackson, M.O. History, expectations, and leadership in the evolution of social norms. *Rev. Econ. Stud.* **2014**, *82*, 423–456. [[CrossRef](#)]
35. Rohner, D.; Thoenig, M.; Zilibotti, F. War signals: A theory of trade, trust, and conflict. *Rev. Econ. Stud.* **2013**, *80*, 1114–1147. [[CrossRef](#)]
36. Bidner, C.; Francois, P. The emergence of political accountability. *Q. J. Econ.* **2013**, *128*, 1397–1448. [[CrossRef](#)]
37. Tirole, J. A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *Rev. Econ. Stud.* **1996**, *63*, 1–22. [[CrossRef](#)]
38. Elster, J. Social norms and economic theory. *J. Econ. Perspect.* **1989**, *3*, 99–117. [[CrossRef](#)]
39. Bidner, C.; Eswaran, M. A Gender-Based Theory On the Origin of the Caste System in India. Unpublished manuscript, 2014.
40. Robson, A.J.; Samuelson, L. The evolution of intertemporal preferences. *Am. Econ. Rev.* **2007**, *97*, 496–500. [[CrossRef](#)]
41. Robson, A.J.; Samuelson, L. The evolution of time preference with aggregate uncertainty. *Am. Econ. Rev.* **2009**, *99*, 1925–1953. [[CrossRef](#)]
42. Doepke, M.; Zilibotti, F. Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission. Unpublished manuscript, 2014.
43. Bisin, A.; Verdier, T. The economics of cultural transmission and the dynamics of preferences. *J. Econ. Theory* **2001**, *97*, 298–319. [[CrossRef](#)]
44. Cho, I.-K.; Kreps, D.M. Signaling games and stable equilibria. *Q. J. Econ.* **1987**, *102*, 177–222. [[CrossRef](#)]
45. Adriani, F.; Sonderegger, S. Why do parents socialize their children to behave pro-socially? An information-based theory. *J. Public Econ.* **2009**, *93*, 1119–1124. [[CrossRef](#)]
46. Grout, P.; Mitraille, S.; Sonderegger, S. The Costs and Benefits of Coordinating with a Different Group. *J. Econ. Theory* **2015**, *160*, 536–556. [[CrossRef](#)]
47. Kets, W.; Sandroni, A. *A Theory of Strategic Uncertainty and Cultural Diversity*; Mimeo: New York, NY, USA, 2018.
48. Ellingsen, T.; Johannesson, M. Conspicuous generosity. *J. Public Econ.* **2011**, *95*, 1131–1143. [[CrossRef](#)]

49. Friedman, J. A noncooperative equilibrium for supergames. *Rev. Econ. Stud.* **1971**, *38*, 1–12.
50. Abreu, D. Extremal equilibria of oligopolistic supergames. *J. Econ. Theory* **1986**, *39*, 191–225. [[CrossRef](#)]
51. Güth, W.; Yaari, M. An evolutionary approach to explain reciprocal behavior in a simple strategic game. In *Explaining Process and Change: Approaches in Evolutionary Economics*; Witt, U., Ed.; The University of Michigan Press: Ann Arbor, MI, USA, 1992; pp. 23–34.
52. Robson, A.J. The biological basis of economic behavior. *J. Econ. Lit.* **2001**, *39*, 11–33. [[CrossRef](#)]
53. Samuelson, L. Information-based relative consumption effects. *Econometrica* **2004**, *72*, 93–118. [[CrossRef](#)]
54. Samuelson, L.; Swinkels, J. Information, evolution and utility. *Theory Econ.* **2006**, *1*, 119–142.
55. Rayo, L.; Becker, G. Evolutionary efficiency and happiness. *J. Political Econ.* **2007**, *115*, 302–337. [[CrossRef](#)]
56. Netzer, N. Evolution of time preferences and attitudes toward risk. *Am. Econ. Rev.* **2009**, *99*, 937–955. [[CrossRef](#)]
57. Adriani, F.; Sonderegger, S. Trust, Trustworthiness and the Consensus Effect: An Evolutionary Approach. *Eur. Econ. Rev.* **2015**, *77*, 102–116 [[CrossRef](#)]
58. Robson, A.J.; Samuelson, L. The evolutionary foundations of preferences. In *Handbook of Social Economics*; Alberto, B., Matt, J., Eds.; Elsevier: Amsterdam, The Netherlands, 2010; pp. 221–310.
59. Binmore, K.G. *Natural Justice*; Oxford University Press: Oxford, UK, 2005.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).