The waning and restoration of social norms

A formal model of the dynamics of norm compliance and norm violation

Paul T. de Beer* and Robert H.J. Mosch**

This version: 5 February 2007

Abstract

Recent debates about crime on the one hand and the purported deterioration of values and norms on the other hand implicitly refer to two kinds of coordination mechanisms. Crime is supposed to be a consequence of deficient material incentives in the form of detection and formal sanctions. Values and norms are related to immaterial incentives, such as feelings of shame or loss of reputation, originating from informal social supervision, and by feelings of guilt or repentance, originating from the personal conviction that one ought to behave otherwise. In order to investigate the relationship between rising crime and other breaches of norms on the one hand and the deterioration of both material and immaterial incentives on the other hand, a simple rational choice model is developed. This model proves to be a useful framework for analyzing behavior in a context in which different incentives for individual behavior, i.e. individual self-interest, formal sanctions, internalized social norms and social norms that are enforced by external sanctions, play a role simultaneously. It is shown that the waning and restoration of norms are asymmetric processes in which the restoration of eroded norms involves a much larger effort than was needed to maintain the norms in the original situation.

* Henri Polak professor of industrial relations at the University of Amsterdam; affiliated to the Amsterdam Institute of Advanced Labour Studies (AIAS) and De Burcht (Dutch Center for Industrial Relations). Address: AIAS, Plantage Muidergracht 4, NL-1018 TV Amsterdam, The Netherlands, tel. +31 20 525 7128, email: p.t.debeer@uva.nl.

** Affiliated to the De Nederlandsche Bank, Supervisory Strategy, PO Box 98, NL-1000 AB, Amsterdam, The Netherlands, tel. +31 20 524 5817, email: r.h.j.mosch@dnb.nl.
The waning and restoration of social norms
A formal model of the dynamics of norm compliance and norm violation

1 INTRODUCTION

Recent debates about crime on the one hand and the purported deterioration of values and norms on the other hand implicitly refer to two kinds of coordination mechanisms. Crime is supposed to be a consequence of deficient material incentives in the form of detection and formal sanctions. Values and norms are related to immaterial incentives, such as feelings of shame or loss of reputation, originating from informal social supervision, and by feelings of guilt or repentance, originating from the personal conviction that one ought to behave otherwise (e.g. Etzioni 2000, 2001). This paper investigates the relationship between rising crime and other breaches of norms on the one hand and the deterioration of both material and immaterial incentives on the other hand by way of a simple rational choice model. In this model ‘formal control and enforcement’ and ‘informal norms’ are regarded as two types of coordination mechanisms that help societies in functioning more efficiently (Mosch 2004). Since these formal and informal coordination mechanisms can influence each other, they have to be taken into account simultaneously.

This paper mainly focuses on the interaction between ‘norms’ and ‘behavior’, thereby leaving ‘values’ aside. In general, values – the fundamental ethical principles of a society – are too abstract to prescribe a particular kind of behavior. Moreover, the reason for the recent public debate seems to be the ‘norm-less’ behavior of citizens rather than disagreement about fundamental values. The problems are not related to disagreement about values, but to the
socially undesirable behavior originating from non-compliance to the common norms – the prescribed guides for conduct or action (Ullmann-Margalit 1977).

The structure of this paper is as follows. Section 2 gives a short overview of the economic and legal literature on social norms. Section 3 analyses four types of incentives that operate in every society as coordination mechanisms of individual behavior. This is a static analysis. The dynamic analysis is subject of section 4, which describes the shifts in relative importance of different incentives that took place in the last decades: material incentives have gained ground compared to immaterial incentives. A simple formal model of norm dynamics is developed, which enables the analysis of the interrelationships between different coordination mechanisms. With this model we show in section 5 how a shift in the incentive structure can unintentionally lead to a downward spiral in which a rise in the incidence of norm violation results in a downfall of the support for norms. Section 6 investigates under what conditions the original situation can be restored, when such a process of norm deterioration has taken place. There appears to be no simple recipe to revive the old levels of norm internalization and norm abiding behavior. Section 7 concludes.

2 SOCIAL NORMS IN ECONOMIC THEORY AND LEGAL THEORY

In the last decade a growing number of economists have come to realize that the ‘classical’ conception of the economic agent as a rational, selfish homo economicus is far beside reality. People making economic decisions appear to be motivated not solely – and often not even primarily – by narrow self-interest, but also by other-regarding preferences. E.g., experiments in a laboratory setting show that in a classical prisoner’s dilemma game the majority of players display cooperative behavior instead of the self-centered opportunistic behavior that standard economic theory predicts. Even a simple market transaction is hard to understand if one assumes that people act only out of direct self-interest.

A similar development is taking place in the field of legal studies, where there is increasing awareness that most people do not obey the law primarily because they are deterred by the
risk of being caught and punished, but because they subscribe to the social norms and values underlying the law (Tyler 1990).

So, both economic and legal scholars are gradually widening their concept of the average human agent from a person basing his acts solely on a weighing of financial or other material incentives towards a person that is also motivated by social norms. Indeed, there is a growing literature that purports that social norms are crucial to understanding people’s behavior in practice, both in the economic and in the legal realm.

Two early economic advocates of a richer concept of economic agents than the traditional *homo economicus* are Nobel laureates Amartya Sen and George Akerlof. Sen has repeatedly emphasized that the self-centered *homo economicus* is in fact a modern invention, which, contrary to common thinking, is almost absent in the work of the classical economists (e.g. Sen 1980). Adam Smith, in particular, stressed the importance of social commitment and sympathy in human behavior.

Whereas Sen discusses commitment only in general terms, Akerlof, in his paper ‘A theory of social custom, of which unemployment may be one consequence’ (1980), has addressed the issue more formally by including social norms (or ‘social customs’ in his words) in the utility function of a utility maximizing agent. In his model, breaking a social norm causes the agent who has internalized the norm to feel guilt. But even if the agent does not endorse the norm, he may harm himself by non-compliance, due to a loss of reputation. In Akerlof’s model, this loss of reputation is proportional to the share of the population that subscribes to the norm. Akerlof shows that there are two stable equilibriums: one in which almost everyone obeys the norm and subscribes to the underlying values, and another in which hardly anyone obeys the norm nor believes in the values underlying it.
More recently, with the rise of experimental economics, numerous researchers have found that economic agents are moved at least as strongly by other-regarding or even altruistic motives as by selfish motives. Both laboratory experiments, e.g., of ultimatum, dictator and public good games, and field studies have shown that most people do not act purely selfish, but are also concerned with fairness or reciprocity (e.g. Fehr & Gächter 2000, Fehr & Fischbacher 2004). People are, for example, prepared to punish someone who has defected in a cooperative game, even if punishing is costly for the punisher and does not bring him any advantage (e.g. in one-shot games, where one does not meet the other player again). In a public good game most people put some of their money in a public fund in order to increase the common pool, instead of holding all their money for themselves, as the rational self-interested agent would.

Since the 1970s the economic analysis of legal problems has gradually gained ground. Starting with Gary Becker’s classic analysis of crime and punishment (1968), numerous authors have attempted to apply the neoclassical economic toolkit to individual behavior with respect to the law. The economic analysis of law starts from the rational person who weighs the payoffs of law abiding behavior and breaking the law. According to this approach, people will only be deterred from criminal behavior if the expected sanction, calculated as the product of the odds of being caught and the resulting penalty, is larger than the expected gain from violating the law.

During the 1990s the economic analysis of law has increasingly become subject to criticism. Cooter (1995) points to the weakness of the economic approach to law. The economic assumption of a rational, utility maximizing individual does not seem to fit reality: “Much of what people do that lands them in court, however, is uncalculated or even irrational” (Cooter 1995: 55). Besides, economists neglect the important interactions between
people: “the ‘meat’ of economics lacks some essential nutriments to nourish social science" (idem, 59). According to Cooter, the most important weakness of the economic approach is, however, that it cannot explain how preferences arise. He argues that people internalize values which change their preferences. Whereas standard economic theory is based on a concept of “thin self-interest”, which “looks only to objective payoffs in wealth or power”, he proposes the alternative concept of “thick self-interest”, which “modifies objective payoffs to encompass the subjective value of morality.” He purports that social norms which facilitate coordination can develop spontaneously as a result of repeated interactions in a community.

While Cooter presupposes that the motivating force of social norms springs from the fact that they are internalized by people, McAdams (1997) develops a theory of social norms which stem from the weight that people attach to esteem from others. He summarizes his theory as follows: “If many people agree that a behavior deserves disapproval, there is an inherent risk that the behavior will be detected, and if this agreement and risk are well-known, then the pattern of disapproval itself creates costs to the behavior. When sufficiently large, these costs produce a norm against the behavior.” An important characteristic of McAdams’ esteem theory is that it does not require people to exert a costly effort to sanction those who violate a norm. “The key feature of esteem is that individuals do not always bear a cost by granting different levels of esteem to others.” It suffices that norm violators are sensitive to the esteem or disapproval of other people. Nevertheless, the power of esteem may be strengthened if those who are willing to impose a material sanction on norm violators are bestowed even more esteem by others.

Although esteem forces in themselves may be weak, the resulting norms may be strong, because of a positive feedback effect. “Because esteem is relative, the intensity of disesteem directed at those who engage in a disapproved behavior is partly a function of the total
number of people who are thought to engage in that behavior. (...) Thus, because individuals value esteem relatively, the more a behavior negatively distinguishes them from others, the more costly it is.” There can be a “tipping point” (the term is Thomas Schelling’s), beyond which a small change in the cost of norm violation can cause a huge effect.

The idea of an esteem-based theory of social norms is elaborated by Geoffrey Brennan and Philip Pettit in their recent, path-breaking book The Economy of Esteem (Brennan and Pettit 2004). In this book they present the economy of esteem as a third coordination or control mechanism, besides the invisible hand of the market and the iron hand of the state. The forces of esteem create a mechanism which they call the intangible hand, which characterizes civil society. Like McAdams, Brennan and Pettit restrict themselves to the external reward of esteem and the external sanction of disesteem, which are conferred upon people who behave in accordance with or contrary to more or less generally accepted norms. Hence, they refrain from incorporating the possibility of internalization of norms into their theory, although they do not deny that norm internalization may also be important.

Brennan and Pettit also explain how esteem may give rise to the emergence of a social norm. If a particular behavior, which is only displayed by very few people, is regarded as highly laudable, then this will confer high esteem on these few performers, which may even give them the status of hero. This may allure others, who observe the praise they receive, to display the same behavior, in order to share in their status as a hero. Although the esteem bestowed on performers may fall as the number of performers increases, at the same time the particular behavior may become a regularity which is considered to be normal behavior. Hence, while the people who display this behavior receive less esteem, people who do not display this behavior may be increasingly disesteemed. This sanction may add to the incentives for people to behave in accordance with the social norm that thus emerges.
A limitation of the contributions to the literature on social norms, discussed in this section, is that most confine themselves to social norms and do not take other mechanisms that influence people’s behavior into account. Hence, they hardly pay attention to the interaction between social norms, on the one hand, and ‘ordinary’ self-centered behavior and legal sanctions, on the other hand. In this paper we present a framework to analyze behavior in a context in which these different motives, i.e. individual self-interest, formal sanctions, internalized social norms and social norms that are enforced by external sanctions, play their role simultaneously.

3 COORDINATION MECHANISMS FOR INDIVIDUAL BEHAVIOR

3.1 Norms, Coordination Problems and Incentives

A norm can be defined as “a prescribed guide for conduct or action which is generally complied with by the members of a society” (Ullmann-Margalit 1977: 13). Norms are the concrete elaborations of the group’s values, which are the abstract, ethical principles that lie at the root of a culture. Norms are often inherently ‘good’ in the sense that they exhibit positive externalities. They tell you to be honest, to be fair, to be trustworthy, to care for the weak, to use your voting right, to help people in need, to obey the law, to bring the wallet you found to its rightful owner, not to steal, not to harm other people, and so forth and so on. Since these norms have positive externalities, there cannot be too much of them and they cannot be obeyed too much.

However, not all social norms are by definition beneficial for society. Elster (1989) gives some examples of norms that are harmful for society at large, but are nonetheless observed by most people. E.g., norms of etiquette and of dress are often costly to comply with, without any apparent benefit for society. Nevertheless, in this paper we will restrict ourselves to norms
that are beneficial for society. The beneficial effects of norms can be analyzed in the well-known game theoretic framework of the prisoners’ dilemma, see Table 1. The societal optimum would be reached if both agents A and B would behave in line with the norms of civic cooperation, i.e. to produce instead of to steal. However, it would be rational for both individuals to discard the norms, i.e. to steal, with the result that this small society ends up in the Pareto sub optimal Nash equilibrium in which both agents steal.

[TABLE 1]

It is not necessary however that norms are ‘inherently good’ to be beneficial for society. Another type of norms helps to lower transaction costs in society. These are the ‘pure’ coordination norms. For example, there lies no ethical or economical principle behind the norm to drive on the right or left side of the road in a country, but the efficiency and safety of transportation are well served by obedience to a norm that settles this issue. The same applies for the specific language that is spoken in a country. From an economic point of view, we cannot say whether it is more efficient to speak Dutch, French, or English. However, it is evident that transaction costs (e.g. regarding translation and misinterpretation) can be reduced if people agree to speak just one language in a specific region, e.g. the European Union. So, the capacity of ‘coordination norms’ to create positive external effects is independent of their specific content, but hinges on the proportion of people that comply with them. In game theoretical terms, there is a social coordination problem with multiple optimal solutions. All optimal solutions have in common that all players choose the same strategy. Social norms act as guidelines that coordinate people’s behavior in reaching an optimal solution. In short, norms function as a coordination mechanism for people’s behavior that prevents them from getting stuck in the sub optimal Nash equilibrium of the prisoners’ dilemma (Den Butter and
Putting this in a broader perspective, we can see informal social norms as one of four types of incentives that affect people’s behavior, see Table 2.

We can distinguish between the kind of incentive and its source. Starting with the former, incentives can be material and immaterial. Material incentives are ‘objective’, formal rewards and punishments, e.g. in the form of a monetary reward, a business profit, a loss, a fine or imprisonment. Immaterial incentives have a ‘subjective’ and moral character. Examples are public praise, looks of disapproval from bystanders, a heartening welcome in a group, ostracism, a feeling of delight and pride, et cetera. The second distinction considers the source of the incentive, extrinsic or intrinsic. An extrinsic incentive originates from outside the individual. This can be a formal authority (in the case of material incentives: quadrant 4 in Table 2), but also members of the informal peer group of the individual (in the case of immaterial incentives: quadrant 3). The source of an intrinsic incentive is the individual himself. The *homo economicus* of the standard economic theory (quadrant 1) only weighs the material pros and cons of his behavior for his own well-being without taking the approval or disapproval of others into account. A purely moral person, on the contrary, acts solely on the basis of his own moral judgments about good and evil, and is, in that sense, also insensitive to the opinions of others (quadrant 2).

Of course, this distinction between four types of incentives is too rigid for any practical purposes, but it is useful for the clarity of analysis at this point. In reality, people are influenced by all types of incentives simultaneously. Moreover, the different incentive types overlap and influence each other mutually. E.g., social norms and properly designed and
upheld legal norms may reinforce each other, while badly designed formal norms may crowd-out informal ones. We will come back to this subject.

3.2 Safety and Enforcement: Material Incentives

The debate on unsafety on the streets and enforcement of legal norms focuses on the left hand column of Table 2. In this column people are seen as rational agents who are guided by material incentives. People behave as the *homo economicus* in his pure form (quadrant 1). Each individual calculates the costs and benefits in utility terms of different behaviors, and acts accordingly. Maximization of material self-interest is his only objective – he is not ‘distracted’ by emotions or social pressures. This extreme (and unrealistic) situation is comparable to the Hobbesian state of nature, which involves a war of all against all. This basic incentive structure can lead to the coordination problems sketched in Table 1 and Table 2.

To solve these coordination problems, the state can intervene by creating a legal system that constrains the behavior of the rational utility maximizing agent (quadrant 4). To accomplish this, the government (Hobbes’ Leviathan) issues *formal norms* (laws and regulations), which affect the *behavior* of the citizens by imposing sanctions (fine, imprisonment) on those who violate the norms. These norms can be seen as a contract between the citizens and the state which establishes the desired behavior. This kind of norms helps to solve both prisoners’ dilemma problems and pure coordination problems. Legal enforcement is, in general, only needed in the case of prisoners’ dilemmas, since rational individuals have no incentive to violate the norm in case of a pure coordination problem (there is no benefit in driving on the right side of the road or speaking Dutch in England).

This system of legally enforced norms changes the incentive structure for individuals (Table 3). Thieves that are caught are sent to jail or have to pay a fine. Their expected
punishment is equal to $S$ (legal action). If both the probability of detection and the sanction are high enough ($2 > 3 - S$ and $0 > 1 - S$, which implies $S > 1$), it is rational for every individual to choose the strategy ‘to produce’ instead of ‘to steal’. As a result, society reaches its high, optimal equilibrium.

### TABLE 3

The assumptions underlying this table are quite similar to those of the recent debate on unsafety on the streets and crime. The pleas for more police force and stronger sanctions are aimed at raising the probability of detection and strengthening the punishment, both of which raise the expected sanction $S$ in our model. The higher the expected sanction, the lower the number of people that will violate the norm.

#### 3.3 Values and Norms: Immaterial Incentives

The purported erosion of values and norms relates to the right column of Table 2. Immaterial incentives are at stake here. This corresponds to the usual approach to human behavior in sociology and psychology. Individuals are guided by incentives that originate from internalized norms (quadrant 2) or social pressure (quadrant 3).

Even a person who has internalized a norm might break it under certain conditions, e.g. when one cannot resist the temptation of a material benefit. In general, however, this will lead to feelings of discomfort due to regret, remorse or guilt ($D$), regardless of whether other people notice the norm violation or disapprove of it. Those who resist the temptation and behave in accordance with the norm, might have feelings of pride or contentment ($P$). Hence, even without social pressures a person might observe a norm although breaking it would lead to personal advantage.
The fact that a norm is internalized does not imply that it is a strictly individual norm. On the contrary, internalized norms are often deduced from or identical to the social norms that are held by one’s social group, e.g. the family, the neighborhood, friends, colleagues, or countrymen (quadrant 3). Individuals copy norms from (nearby) role-models, like parents, teachers, friends and idols. ‘Self-attribution’ plays an important role in this process. It refers to the psychological mechanism that people who behave according to the social norms in a community, tend to subscribe to these norms in the end. People’s need to have a coherent and clear self-interpretation is most easily satisfied by explaining their own behavior in terms of norms that they subscribe to (Schlicht 2002). Norm compliance thus leads to norm internalization. When people’s behavior is determined by the incentives of common social norms, these norms will be internalized after some time, and hence the difference between social and individual norms evaporates.

Informal norms do not only affect behavior because breaking or upholding them leads to regret or pride, but also because the social group can punish or reward people. These informal norms can be regarded as an implicit, social contract between citizens about how to behave in society (cf. the social contract of Locke). The incentive structure changes as follows (see Table 4).

| TABLE 4 |

In this case the sanctions and rewards related to norm breaking and norm abiding behavior have a social origin. When a norm is broken, the trespasser will be confronted with the sanction $R$, e.g. a loss of reputation or social exclusion. The coordination of behavior in the column on the right hand side of Table 2 thus runs parallel to the coordination in the left column, although the latter has a formal instead of an informal character. Note that both
formal and informal norms have a substantial as well as an enforcement aspect. The substantial aspect determines what is and what is not allowed or ‘normal’ in society, the enforcement aspect refers to the probability of detection of norm violation and the severity of the punishment. With respect to informal norms this last aspect relates to the ability and the willingness of group members or bystanders to intervene when they witness norm breaking behavior.

The systems of material and immaterial incentives differ in that in the latter case the victim of norm violation can be offered some kind of compensation \( C \) in close communities with strongly developed informal norms. For example, compassionate neighbors might help by replacing the stolen goods. It also happens that respected, norm abiding citizens receive a ‘reward’ \( L \) in the form of praise, a medal, a certificate of outstanding behavior, et cetera. When \( D, R, P \) and \( L \) are large enough, the sub optimal, low equilibrium of Table 1 (bottom right corner) tips to the high equilibrium (top left corner). The necessary condition is: \( 2 + P + L > 3 - D - R \), which implies \( P + L + D + R > 1 \). Moreover, a high \( C \) softens the discomfort of the victims.

4 SHIFTS IN THE INCENTIVE STRUCTURE

In the previous section we distinguished two types of behavior, guided by material and immaterial incentives, respectively. The corresponding coordination mechanisms – formal and informal norms, respectively, and the sanctions and rewards related with them – can affect both types of behavior in such a way that social dilemmas are dissolved. In this section we leave the sharp distinction in four quadrants behind. In practice, human behavior is usually influenced by both material and immaterial incentives that can have either an intrinsic or an extrinsic origin. What kind of incentive exerts the strongest influence on people’s behavior in
a particular society is not fixed, but may change over time. This section focuses on the character of these changes and the role that the government may play in this process.

Underlying most debates about unsafety and values and norms is the idea that there has been a shift from immaterial to material incentives (i.e. a move from the right to the left column of Table 2) and within the group of material incentives a move from extrinsic to intrinsic incentives (i.e. from quadrant 4 to quadrant 1). In other words, it is purported that people are influenced less by internalized norms and by social pressure of their peer group than they used to be in the past. Nowadays, their behavior is primarily affected by material incentives. Moreover, people seem to be influenced less by formal sanctions and rewards than they used to be. Their direct self-interest now seems to be the prime focus of their choices. Thus, the weight has moved from quadrant 2 and 3 via quadrant 4 to quadrant 1. This purported shift is generally regarded as undesirable, leading to pleas for a reverse shift towards quadrant 3 and 2, or at least to quadrant 4. Below, we investigate which changes in society might have caused these shifts.

4.1 The Move from Immaterial to Material Incentives

The coordination mechanisms based on immaterial incentives seem to have lost weight compared to coordination mechanisms based on material incentives, implying a move from the right column to the left column of Table 2. Which factors can explain this move? It is useful to distinguish between developments related to a lack of internalization of norms (quadrant 2) and those related to the decrease of social pressure for norm abiding behavior (quadrant 3).

The process of individualization in society is the first usual suspect for the purported decline in norm internalization (Beck & Beck-Gernsheim 2002, De Beer 2007). The values and norms one subscribes to tend to diffuse when they are less determined by traditional
social bonds, like the family, the church or the neighborhood. If every individual ‘chooses’ his own norms, these norms will lose their function of coordinating the behavior of individuals.

Immigration, too, might foster differentiation and fragmentation of norms. When newcomers adhere to other norms than the indigenous population, it is far from evident that they will adopt the incumbent norms and give up their own norms. This will only happen when immigrants assimilate fully in the dominant culture of the host country. However, assimilation tends to be a lengthy process, which may not be completed before very long. Immigrant groups sometimes stick to their ‘dissenting’ culture for several generations. Thus a large group of immigrants weakens the coordination function of ‘native’ norms.

The replacement of older by younger generations is a third mechanism through which the internationalization of norms may be weakened. Younger generations may not subscribe to norms that are self-evident for older generations if these norms are no longer transferred by parents, teachers or through the mass media. This might, however, go hand in hand with the rise of new norms, to which the younger generations more readily subscribe than older generations. Thus, in general, one should not speak of norm decline, but of norm replacement (cf. Inglehart 1990).

Until the 1960s those who broke norms often were corrected immediately by bystanders. Social control restricted individual freedom. However, the strength of social control deteriorated rapidly in the next decades. On the one hand, this fostered the emancipation of minority groups like feminists, non-religious people and homosexuals. On the other hand, it involved the less positive consequence that undesirable behavior was no longer automatically corrected.

Interfering with and trying to stop norm violation can be regarded as a prisoners’ dilemma itself. Interference is costly and dangerous for the one who interferes, while the benefits (the
upholding of the norm) accrue to society as a whole. There are two major, mutually related reasons why citizens tend to choose more and more often an opportunistic strategy of not getting involved when they notice that someone violates a norm: spatial expansion and rising mobility.

As the size of cities and other social structures grows, individuals more often meet people they do not know in person. In a small village it is noticeable when a stranger walks around, while in big cities it is more striking to meet someone familiar. Anonymity grows. This reduces the probability that people’s behavior is affected by social control, for two reasons.

First, bystanders are less eager to try and correct a stranger’s behavior in an anonymous surrounding, because they feel less responsible for what happens in public and because they feel uncertain about what reaction to expect. People do not feel the duty anymore to step in when something happens. Government officials are supposed to fulfill this task. One runs even a considerable risk that the norm violator will react aggressively when one intervenes. If that happens, intervention even has a perverse effect for the intervener. The enforcement of norms has thus become more risky. The people who have probably internalized the norms the strongest – viz. older people – are often physically the least able to correct the ones who are most likely to violate the norms – younger people, who also seem to have become more aggressive than they used to be. The demand for respect of gray hairs is less effective than a knife. Moreover, one cannot any longer be confident (ex ante) that bystanders will help when the situation gets out of hand. This is sometimes even stimulated by government policies that advise citizens to refrain from direct action, except for calling the police. People sometimes even risk legal punishment when they try to correct norm violators by using force themselves. Hence, government policies may actively discourage norm enforcement by the citizens.

Secondly, the probability declines that a person one talks to about his conduct will give in, and thus stop breaking the norm. He might simply not accept the norm and claim that other
norms apply to his situation. Moreover, a reprimand by a stranger, in general, causes less feelings of guilt or shame than a rebuke by a familiar person.

The effects of spatial expansion are enhanced by the rise in (physical) mobility. As a consequence, people get in contact with ever more strangers and spend less time among their acquaintances. This raises the probability of one-shot contacts, and hence lowers the risk of adverse future repercussions when one breaks a norm. In game theoretical terms: as the number of repeated games with the same player declines, the dominant strategy shifts from cooperative towards opportunistic behavior.

4.2 Declining Effectiveness of Material Incentives

The fact that the effectiveness of norm internalization and social control has declined during the last decades has increased pressure on extrinsic material incentives to guide people’s behavior. Unfortunately, the formal institutions entrusted with this task seem increasingly to have trouble in fulfilling this task. E.g. registered crime figures have risen strongly while the number of cases solved by the police has declined in many countries.

Research has shown that the deterring effect of formal control and sanctions is determined largely by the probability of detection and only for a small part by the severity of the sanction. Hence, it is not surprising that the number of people that is deterred from committing a crime by the threat of a formal sanction, has fallen. In terms of Table 3, the size of sanction $S$ is perceived to be so low that the coordination mechanism of formal (legal) norms and enforcement is no longer capable of ensuring the optimal societal equilibrium ($2 > 3 - S$ and $0 > 1 - S$). The material incentives of the legal apparatus are too weak to deter norm violation. It is only because a large part of the citizens is still affected by internalized norms and social pressures, that the proportion of the population that actually violates (legal) norms remains relatively small.
5 A NEGATIVE SPIRAL OF NORM VIOLATION AND NORM INTERNALIZATION

The previous section described the shift from immaterial to material incentives and the declining effectiveness of material incentives as if they are more or less autonomous developments. Of course, these developments influence and reinforce each other mutually. As a consequence, a negative spiral might emerge, in which fewer and fewer people subscribe to and comply with a particular norm, which finally disappears. This section analyses the way this process might unfold by means of a simple model. The annex provides the mathematical derivation of the model.

5.1 Formal Control and Norm Violation

The analysis starts with the usual (neoclassical) assumption of individual utility maximization. Utility captures everything that a person values positively or negatively. Well-being, satisfaction of needs and happiness can be considered as equivalent terms of utility. Utility partly depends on whether an individual commits a forbidden act. The proportion of the population that complies with the norm is denoted by $\gamma$. If a person violates the norm his utility rises with $G$ (e.g. the profit of a theft) on the one hand, but also raises the probability of being caught and punished, which decreases utility by $S$. Then the *homo economicus* will violate the law if $G > S_0$, or $G - S_0 > 0$. In this case, norm violation can only be fought by increasing the detection rate and/or the severity of the penalty so that the net benefit of norm violation becomes negative: $G - S_1 < 0$. This situation applies to quadrant 4 in Table 2.
5.2 Internalization of Norms

Next, we suppose that individuals internalize norms to some extent. Following Akerlof (1980) we suppose that part of the population has internalized the norm and subscribes to it, while another part does not. $\beta$ is the proportion of the population that subscribes to the norm (quadrant 2 in Table 2). Someone who subscribes to the norm does not automatically comply with it. However, violating the norm causes the person to feel regret, remorse or guilt. If breaking the norm reduces utility with $D_0$, an individual who subscribes to the norm will still break it when $G - S_0 > D_0$, i.e. if the extrinsic benefit of breaking the norm is larger than the intrinsic regret.

This is shown in Figure 1, where the regret $D_0$ one feels if one breaks the norm, varies between those who subscribe to the norm. The population has been sorted according to descending regret (the dotted line $D_0$): the individuals on the left side of the figure incur the largest regret when they break the norm (with a value of 7 ‘utility units’; note that these units are purely illustrative and have no practical meaning). Regret decreases when moving to the right. The last 30 percent of the population do not subscribe to the norm and, hence, do not feel any regret when they break the norm. A further assumption is that the extrinsic benefit of norm violating behavior, $G - S_0$, is positive and equal (a value of 2) for everyone.

[FIGURE 1]

In Figure 1 the extrinsic benefit of norm violating behavior is that large that all people who do not subscribe to the norm, break it ($G - S_0 > 0$). At the same time, the extrinsic benefit of norm breaking is too small to entice people who do subscribe to the norm into breaking it ($G - S_0 < D_0$). Hence, the proportion of the population that complies with the norm, $\gamma_0$, is equal to the proportion of the population that subscribes to it, $\beta_0$. This is a stable equilibrium ($E_0$):
those who abide by the norm are not tempted to break it in the future, while those who violate it are not tempted to abide by it in the future.

Next, assume that the extrinsic benefit of norm violation increases, for example because formal control is relaxed and, hence, the expected sanction declines. Thus, $S_1 < S_0$, or $G - S_1 > G - S_0$. In the short term, this does not affect the percentage $\gamma_0$ of the population that abides by the norm. However, if the expected sanction $S$ decreases even more, at some point some people who subscribe to the norm will be tempted to break it. As a result, the proportion of the population who comply with the norm drops from $\gamma_0$ to $\gamma_1$, in Figure 1 from 70 to 50 percent ($E_1$).

In point $E_1$ the number of people that comply with the norm is less than the number that subscribe to it. This could affect the belief in the norm. Akerlof (1980) makes the crucial but plausible assumption that the proportion of the population subscribing to the norm will ultimately adjust to the proportion complying with it. When the number of people acting in accordance with the norm is less than the number subscribing to it, the belief in the norm erodes (curve $D_1$ in Figure 2). We assume that those who feel the least regret are the first to lose their faith in the norm. After some time, a new equilibrium emerges in which the share of the population that subscribes to the norm equals the proportion that abides by the norm: $\beta_1 = \gamma_1$ (50% in Figure 2). (The model does not determine the time involved with the adjustment process. This could take one or two years, but also ten years or more.) When the net benefit of norm breaking behavior does not change any further, i.e. $G$ and $S_1$ stay the same, a new stable equilibrium arises ($E_1$).

[FIGURE 2]

Let us assume that the government wants to return to the old level of norm compliance: $\beta_0$ (70% in Figure 1). The detection rate and/or severity of punishment are therefore raised to the
original level. However, this has no effect on the proportion of the people complying with the norm, since the belief in the norm has eroded: those who lost faith in the norm are insufficiently deterred by the higher expected sanction to start behaving in accordance with the norm again. The expected sanction on norm violation has to be larger than the benefit, i.e. $G - S < 0$, to persuade the egocentric homo economicus who does not subscribe to the norm, to live up to it. In that case the total population will comply with the norm: $\gamma_2 = 1$. As a consequence, the belief in the norm will start to rise again until the whole population subscribes to it in the end: $\beta_2 = 1$.

An important conclusion follows from this simple analysis, viz. that the decline and the restoration of norm compliance are asymmetrical processes. If, as a consequence of the reduction of formal sanctions, the proportion of norm abiding people falls, it does not suffice to restore the former level of control and punishment in order to return to the original situation. When the belief in the norm has eroded, the expected sanction on norm violation has to be much higher than the original sanction to restore the original equilibrium.

5.3 Informal Social Control

In the previous section, individuals who did not have internalized the norm, only complied with the norm if rational calculations showed that the expected formal sanction outweighs the benefits of breaking the norm. However, people who do not believe in the norm, may also obey it because of informal, social control and sanctions (quadrant 3 in Table 2). Breaking the norm may cause feelings of shame or loss of reputation, which decrease the offender’s utility.

Following Akerlof (1980), we assume that the shame or the loss of reputation is related to the proportion $\beta$ of the population that subscribes to the norm. This is a plausible assumption, since it does not seem likely to be corrected by someone who does not subscribe to the norm himself. So the number of people who subscribe to the norm has a positive impact on the
probability of social control and on the size of the loss of reputation. A further assumption is that the degree of shame one feels, varies from person to person. For reasons of simplicity, we assume that those who have most strongly internalized the norm, also feel most shame when they break the norm and are detected, while those who do not believe in the norm feel the least shame. The loss of reputation a person incurs when he violates the norm is therefore equal to \( \beta R^i \), in which \( \beta \) is the proportion of the population that subscribes to the norm and \( R^i \) is the magnitude of personally felt shame if one is caught.

[FIGURE 3]

Figure 3 shows a possible scenario which is comparable to Figure 1. The line \( R_0 \) reflects the amount of shame or the loss of reputation caused by norm violation, and the line \( D_0 \) depicts the amount of regret or guilt. Total disutility of norm breaking behavior is then equal to \( D_0 + R_0 \) (discomfort + loss of reputation). In Figure 3 all people who subscribe to the norm also comply with the norm (\( \beta_0 = \gamma_0 = 70\% \) of the population), while all people who do not subscribe to it break the norm. Figure 3 differs from Figure 1 in that a part of the people who subscribe to the norm (the group between 58% and 70%), only obeys the norm because of the loss of reputation they would incur if they would break it. Without this loss of reputation, the regret they feel when breaking the norm would not be large enough to counterbalance the benefits of norm violation. Thus, for this group of people: \( G - S_0 > D^i \), but \( G - S_0 < D^i + \beta_0 R^i \). For those who do not subscribe to the norm, the loss of reputation is too small to prevent them from norm violation: \( G - S_0 > D^i + \beta_0 R^i \).

As in the previous section, we assume that the formal sanction on norm violation declines over time. After a while, norm violation becomes so attractive that some people who subscribe to the norm are tempted to break it. However, contrary to the previous section, the new equilibrium that emerges at \( \gamma_1 \) (50%) is not stable (\( E_1 \)), since the loss of reputation is
related to the percentage of people who subscribe to the norm. When $\beta$ adjusts to the proportion of people who violate the norm and, hence, declines ($\beta_1 = \gamma_1 < \beta_0$), the loss of reputation diminishes too. Therefore, curve $R_1$ in Figure 4 lies below curve $R_0$ in Figure 3.

[FIGURE 4]

As a consequence, some of those who subscribe to the norm stop living up to it, thereby moving the equilibrium even more to the left ($E_2$ in Figure 4). In the next period, the belief in the norm will decline further, inducing another reduction of the loss of reputation, et cetera. In general, it is not certain where this process of norm deterioration ends. If the loss of reputation falls faster (as a result of the erosion of the belief in it) than the belief of the “marginal subscriber” increases, then, ultimately, the norm vanishes completely (see Annex). In that case even the strongest believers only abstain from breaking the norm out of fear for loss of reputation. Hence, when the loss of reputation becomes smaller and smaller, in the end no-one subscribes to the norm anymore. However, a new stable internal equilibrium arises if for some people the norm internalization is strong enough, so that they comply with the norm even if there is no external control or sanction at all. In this equilibrium, part of the population subscribes to and complies with the norm, while the rest of the population does not subscribe to and behaves in accordance with the norm. This situation is displayed in Figure 5, in which in the stable equilibrium ($E_3$) 36.5 percent of the population subscribes to ($\beta_2$) and complies with the norm ($\gamma_2$).

[FIGURE 5]

Once again, it does not suffice to restore the original level of formal control and sanctions to return to the original situation. The sanctions have to be raised to a level that is high enough
to deter also those who do not believe in the norm from breaking it \((G - S_2\) in Figure 6). If this is successful, i.e. if the proportion of the population abiding by the norm returns to the original size \((\gamma_3 = 70\%\) in Figure 6), then the belief in the norm will also recover. Consequently, informal control regains ground. Ultimately, it will be possible to relax the formal control somewhat, because informal control can take over part of the job.

This second model shows the asymmetry between norm decline and norm restoration even more clearly than the first model. If formal control of norm violation is relaxed too much, a negative spiraling effect might be set in motion, which will only come to an end when the norm has disappeared completely or when only a small group of strong believers remains. Then, a huge effort is required to return to the original situation.

The model also shows the complex relation between formal and informal control. Although the two types of control can be substitutes to some extent, they can also reinforce each other. The consequences of a relaxation of formal control can be temporarily mitigated by informal control. However, in the long run the informal control will weaken too, and will no longer be strong enough to prevent people from breaking the norm. Then, a strong reinforcement of formal control is necessary to restore the norm. If one succeeds in doing this, then informal control has the opportunity to recuperate, which opens up the possibility to relax formal control to some extent once more. However, one has to be careful not to relax formal control too much, since otherwise the equilibrium may be disturbed once again.

A corollary of this analysis is that one has to be careful in judging the effectiveness and the efficiency of the police force and the judiciary. The effectiveness of deterrence by legal forces may vary over time independently of the intensity of enforcement. A relatively minor effort of
the police force and the legal department may suffice to keep crime rates low in a society in which the legal norms are generally internalized and violation results in a severe loss of reputation. A much larger effort may be less effective in a society in which the belief in the norms has eroded and, as a result, social control has deteriorated. Although it may seem that the effectiveness of the police in the second case is much smaller, which could easily lead to complaints about sloppy police activities or overdone bureaucracy, the real cause are the different societal circumstances.

A second corollary is that it is not necessarily alarming if part of the population does not subscribe to and comply with the dominant norms. There may be a stable equilibrium in which a fixed proportion of the population breaks the rules without having a subversive influence on the norm internalization and norm abiding behavior of the majority. There might even be a stable minority that upholds its own specific norms (cf. vegetarians or orthodox religious people) in a generally “norm-less” society. Note, however, that these conclusions hinge on the assumption that the belief in a norm adjusts to the proportion of the population that acts in accordance with the norm.

5.4 A Preliminary Empirical Test

For a number of reasons it is not simple to test this model of the waning and restoration of social norms empirically. One would need data on the adherence to and the compliance with social norms, and the strength of both formal and informal control and sanctions, for quite a long period. Unfortunately, we only have data on the compliance with particular norms and on formal sanctions at our disposal. These enable us to perform a very partial test of the model only. Nevertheless, it is worthwhile to examine whether these data support our model. We will use data on violent crimes and on the number of prisoners in the United States.
Figure 7 shows the annual number of reported violent crimes per 100,000 population and the number of sentenced prisoners per 1,000 violent crimes, since 1960. Of course, not all prisoners are convicted of violent crimes, but figures of convicted prisoners by type of offence are only available since 1980. These data are also shown in figure 7.

[FIGURE 7]

The number of prisoners per 1,000 crimes can be considered to be an indicator for the formal sanctions a person can expect to experience, on average, when conducting a crime ($S$ in the model). If only these formal sanctions would matter, we would expect a simple inverse relationship between the level of these sanctions and the rate of violent crime. Figure 7 does indeed show that, from 1961 to 1975, the expected sanction for a violent criminal dropped considerably and the violent crime rate more than tripled. However, although the expected sanction stabilised during the second half of the 1970s and started to rise in the 1980s, the crime rate continued to rise. In 1990 the expected sanction had returned to the level of 1967, but the crime rate was still almost three times as high. It thus seemed that the strengthening of the sanctions during the 1980s had not been effective. In our model, this is explained by the fact that the rise of the violent crime rate during the 1970s resulted in a deterioration of the adherence to the norms against violence. So, as in figure 6, in the 1980s much more control and sanctions were needed to reduce crime to the same rate as in the early 1960s. Finally, in the 1990s the crime rate started to fall, after all, probably because by that time the expected sanctions had become so large, that they also deterred many potential criminals who did not adhere to the norm. If our model is correct, this would mean that norm adherence is improving again since the 1990s. Unfortunately, we lack the data to test this conjecture. If it is true, however, this might show itself in the future if the sanctions would be relaxed again somewhat and the violent crime rate would not start to rise again shortly after.
Although the evolution of the violent crime rate and the number of prisoners seems to be in accordance with our model, this does, of course, not prove that our model is correct. The same evolution of crime rate and imprisonment could be expected if there was a time lag between changes in the expected sanction and norm compliance. However, in that case one would have to assume that it took ten years before potential criminals realised that sanctions were strengthened in the 1980s and consequently changed their behavior. In our model we do not have to rely on such a slow reaction to explain why the crime rate did not react sooner to the strengthening of sanctions.

6 RESTORATION OF NORMS AND NORM COMPLIANCE

The analysis in the previous section shows that downward and upward shifts in norm belief and norm violation are asymmetrical processes: if a norm has deteriorated, it costs much more effort to restore it than it originally did to maintain it. In this analysis we assumed that one can only move from quadrant 1 of Table 2 (absence of norms) to quadrant 4 (guidance through material, extrinsic incentives) by increasing formal control and toughening sanctions, as was demonstrated by figures on violent crime in the USA. If one succeeds in doing this, then the immaterial incentives (both from informal social control and from norm internalization) will catch up again.

An important question is whether it is also possible to combat norm violation directly through immaterial incentives. In other words, is it possible to foster norm compliance by enhancing social control and stimulating the transfer of values and norms, resulting in norm internalization? The problem here is that direct manipulation of immaterial incentives by the government is impossible. Government action, by definition, has a formal, material character. Nevertheless, the government can try to enforce the role of social institutions in transferring values and norms.
6.1 Strengthening Social Control

Is it possible to foster social control? It seems to be illusionary to try to develop policies that restore its past role in preventing norm violation in the public space. It is difficult to change the forces that have led to the erosion of social control – individualization, urbanization and probably immigration without integration. The literature on social capital has shown that it is very hard to undo a change from spontaneous, informal cooperation (social capital) to formally enforced coordination (cf. Ostrom 2000). These processes are asymmetric: it is easier to disrupt old social ties and customs than to restore them. It is naïve to assume that informal social control automatically returns when formal coordination mechanisms are relaxed. Most likely the result will be a complete lack of coordination. Informal social control requires a certain level of mutual trust among citizens. For example, whether people will try to stop norm breaking behavior when they encounter it, depends on whether they have confidence that bystanders will help them. If this confidence has deteriorated, as a consequence of lack of formal control, most citizens will only interfere in order to prevent norm violation when they are victim themselves.

Nevertheless, it is conceivable that social control would regain some ground in modern society, if it would get a more organized and formal character. One way to accomplish this is by organizing control tasks of citizens in cooperation with formal organizations. An example are neighborhood prevention projects, in which citizens perform surveillance tasks in their neighborhood in cooperation with the local police. Eventually, such types of organized social control might evolve into spontaneous social control, when the mutual trust between the citizens in the neighborhood has recovered.
6.2 Strengthening the Transfer of Norms

Another way to try to foster norm abiding behavior directs attention at norm internalization (quadrant 2 in Table 2). Still, many people act more or less in accordance with social and legal norms without formal or informal coordination, because they have internalized these norms and would feel shame or regret if they break them. Internalization of norms is an important – arguably the most important – part of the socialization process of every child and adolescent, in his family, at school, among friends, in associations, and at work. In these communities one learns how to behave in different circumstances. Consequently, most people act according to these rules and customs, without constantly being monitored and corrected.

One of the causes of rising norm violation might be that the transfer of norms in people’s socialization process fails. Thus, improving this transfer could help to push back norm violation. However, it is far from easy to point out how the government could improve this transfer. Direct government intervention in the raising of children is generally regarded as unacceptable. Most citizens probably prefer a somewhat flawed raising by parents to a state upbringing by professionals. Nevertheless, this principle of non-interference in the raising of children, withholds the society the opportunity of exerting substantial influence on a crucial phase in life with regard to norm internalization. In our opinion, the principle of non-interference should be waived in cases in which the lack of appropriate raising is sufficiently clear. It thus deserves serious consideration to legally coerce parents of children that regularly commit crimes, to accept professional help in the raising of their children or even to deprive them of parental custody. This serious infringement on parental rights should be weighed against the losses society would incur otherwise, and in particular those citizens who would be victims. Experiments with early help in raising children for citizens of low social classes in the United States show that these children attain more self discipline and perseverance, tend to
do better at school and at the labor market and have a lower criminal record than the children in the control group whose parents did not receive support (Heckman 2003).

The government can more easily influence the transfer of norms at schools. However, here, too, the general opinion seems to be that direct government interference with respect to the transfer of norms should be restricted. Again, there are sound arguments to plead against an overly strict principle of non-interference, considering that the lack of norm transfer at school is probably one of the causes of norm violation among youngsters.

A third domain in which government could try to foster norm transfer are the mass media. Despite the differing opinions about the impact of the mass media on the behavior of the public, it is very likely that there is a substantial influence. However, in general, broadcasting companies do not seem to focus on the transfer of broadly shared norms, considering the large share of programs containing violence, abusive language and sex. But again, an appeal for more government influence on the content of programs stumbles across a fundamental right, viz. freedom of speech. So, in this domain too, the infringement of civil rights has to be weighted against the loss that society suffers due to the impact of norm breaking behavior broadcasted by the media.

7 CONCLUSION

The present debates on crime and unsafety on the streets on the one hand and the purported waning of norms and values on the other hand touch upon two types of coordination mechanisms for human behavior. Unsafety and crime are related to a lack of formal control and sanctions that can guide the behavior of calculating, self-interested citizens. Erosion of values and norms is related to a lack of norm internalization and of informal social control. This paper has tried to connect these two types of coordination mechanisms – based on
material and immaterial incentives, respectively. Most people’s behavior is guided not only by a rational calculation of material costs and benefits (cf. the *homo economicus*), but neither are most people purely “moral” beings who base their decisions exclusively on ethical considerations. In practice, most people are influenced by both material and immaterial incentives. Increasing norm violation, which fuels the debate on insecurity and values and norms, is related to both the shortcomings of formal control mechanisms and sanctions, and a lack of social control and norm internalization. In this paper, we developed a theoretical model describing the mutual relationship between these coordination mechanisms. This dynamic model shows that a process of increasing norm violation can invigorate itself, by eroding social control and norm internalization, and diminishing the effectiveness of formal control and sanctions. The model also shows that a return to the original equilibrium, in which the majority of the population complies with the norm, cannot be accomplished by simply restoring the original level of the formal control mechanisms. Erosion and restoration are asymmetrical processes, in which the restoration of eroded norms involves a much larger effort than was needed to maintain the norm in the original situation. Hence, strengthening of formal control and sanctions is not an easy way to reduce norm violation. A better way is to try to reinforce formal control, promote social control and enhance norm internalization simultaneously. It is, however, not easy for the government to boost these alternatives, since it either lacks the means to do this or is “hindered” by fundamental rights. One can argue that this last point is raised to soon in situations in which these fundamental rights (like freedom of speech and freedom of education) are in conflict with other vital social values.

Finally, we want to stress that this paper simply presumes the desirability of certain, broadly shared norms. This presumption implies that a rise in norm breaking behavior and a decline of the number of persons that subscribe to the norm, are regarded as undesirable. However, we do not want to suggest that every form of norm violation or every change of
norms is, by definition, undesirable. In a dynamic, open society norms change all the time. Old-fashioned norms are replaced by new ones. There is no reason to judge this, a priori, either positively or negatively. Usually, the replacement of a “traditional” norm by a “modern” norm is interpreted as progress. E.g. the changes of the norms with respect to the role of women in society in the past decades are generally regarded as an advancement. In this case, the downward spiral of the old norm, as described in this paper, should be considered rather as a positive development. Restoration of the old norm will then not be an option favored by the large majority of the population, although, of course, a small orthodox minority may still adhere to the old norm and regret its demise.
MATHEMATICAL APPENDIX: FORMAL DERIVATION OF THE MODEL OF NORM
COMPLIANCE AND NORM VIOLATION

An individual’s utility can be written as:

\( U_i = \bar{U}_i + (1-\gamma_i) G_i - (1-\gamma_i) S_i - \beta_i (1-\gamma_i) D_i - \bar{\beta} (1-\gamma_i) R_i \)

in which \( \bar{U}_i \) is a constant and:

\( \gamma_i = 1 \) if the individual acts in accordance with the norm and \( \gamma_i = 0 \) if the individual breaks the norm;

\( \beta_i = 1 \) if the individual subscribes to the norm and \( \beta_i = 0 \) if the individual does not subscribe to the norm;

\( G_i \) is the subjective valuation (utility) of the benefit of breaking the norm;

\( S_i \) is the subjective valuation (disutility) of the formal sanction on breaking the norm;

\( D_i \) is the feeling of discomfort as a result of regret or guilt if one breaks a norm that one subscribes to;

\( R_i \) is the subjective valuation of the loss of reputation or shame one incurs because of social control if one breaks the norm and the entire population subscribes to the norm;

\( \bar{\beta} \) is the proportion of the population that subscribes to the norm (the average value of \( \beta_i \)).

Individual \( i \) chooses \( \gamma_i \) so, that his utility \( U_i \) is maximized. Hence, an individual acts in accordance with the norm \( (\gamma_i = 1) \) if the (subjective) benefit of norm violation is smaller than the (subjective) costs:

\( G_i < S_i + \beta_i D_i + \bar{\beta} R_i \)
The individual $i$ breaks the norm if the right hand side is smaller than the left hand side of the equation.

Assume that everybody attaches the same values to the benefit and the sanction of norm violation, i.e. $G_i = G$ and $S_i = S$ for all $i$. Moreover, assume that the disutility of norm breaking in terms of discomfort $B_i$ and loss of reputation $R_i$ varies across the population in the following way:

$$D_i = D - d \frac{i}{N} \quad \text{for} \quad i = 1, 2, 3, \ldots, \bar{\beta} \; N$$
$$= 0 \quad \text{for} \quad i = \bar{\beta} \; N + 1, \bar{\beta} \; N + 2, \ldots, N$$

$$R_i = R - r \frac{i}{N} \quad \text{for all} \; i \; \text{if} \; r \leq R \; \text{and for} \; i < \frac{RN}{r} \; \text{if} \; r < R$$
$$= 0 \quad \text{for} \; i \geq \frac{RN}{r} \; \text{if} \; r < R.$$ 

$N$ is the size of the population. The population is sorted according to descending discomfort and descending loss of reputation when breaking the norm.

Condition (2) for individual utility maximization with norm compliance can now be written as:

$$G - S < \beta_i (D - d \frac{i}{N}) + \bar{\beta} (R - r \frac{i}{N}) = \beta_i D + \bar{\beta} R - (\beta_i d + \bar{\beta} r) i/N$$

This leads to three types of (short term) equilibrium.

The first equilibrium concerns the situation in which all people who subscribe to the norm also comply with it, i.e. $\gamma = \bar{\beta}$ (in which $\gamma$ is the average of $\gamma_i$). Thus, in this equilibrium, in which $i = \bar{\beta} \; N$:

$$G - S - D < (R - d - \bar{\beta} r) \bar{\beta}$$
$$G - S > (R - \bar{\beta} r) \bar{\beta}$$
Condition (4a) refers to the people at the margin who still subscribe to and comply with the norm and (4b) to the people at the margin who do not subscribe to the norm but nevertheless comply with it.

In the second equilibrium, part of the people who subscribe to the norm do not comply with it (i.e. \( \gamma < \beta \)), because the risk of a formal sanction, the discomfort felt and the loss of reputation do not weigh up against the benefit of norm violation. For the marginal person who complies with the norm \((i = \gamma N \text{ en } \beta_i = 1)\), this implies:

\[
G - S - D - \beta R = -(d + \beta r) \gamma
\]  

Or:

\[
\gamma = \frac{D + \beta R - G + S}{d + \beta r}
\]

In the third short-term equilibrium, part of the population complies with the norm although they do not subscribe to it (i.e. \( \gamma > \beta \)), because the disutility of the formal sanction, regret and loss of reputation with regard to norm violation is larger than the utility of the related benefit. For the marginal person who complies with the norm \((i = \gamma N \text{ en } \beta_i = 0)\), this means:

\[
G - S - \beta R = -\beta r \gamma
\]

Or:

\[
\gamma = \frac{\beta R - G + S}{\beta r}
\]
The second and third type of equilibrium, in which $\gamma < \beta$ and $\gamma > \beta$, respectively, are not stable in case the proportion of the population subscribing to the norm, $\beta$, adjusts to the proportion complying with it, $\gamma$. In the long term, then $\gamma = \beta$. If originally $\gamma < \beta$, the long term equilibrium can be derived by substituting $\beta$ for $\gamma$ in equation (6):

$$\beta^* = \frac{-(d - R) + \sqrt{(d - R)^2 - 4r(G - S - D)}}{2r}$$

This equilibrium $\beta^*$ is a positive number if $G - S < D$ and $d < R$. The first condition means that those who subscribe to the norm most strongly, will never break it, even if no loss of reputation is involved. The second condition means that, when the share of “believers” in the population declines, the loss of reputation related to norm violation decreases more rapidly ($R$) than the intensity of belief among the remaining believers increases ($d$).

If originally $\gamma > \beta$, the belief in the norm and norm compliance rise until the group that subscribes to the norm coincides with the group that complies with the norm, i.e. $\gamma = \beta$. In this long term equilibrium, the loss of reputation for the marginal person who complies with the norm but does not subscribe to it, is just big enough to prevent him from norm violation. This long term equilibrium can be determined by substituting $\beta$ for $\gamma$ in equation (7):

$$\beta^* = \frac{R + \sqrt{R^2 - 4r(G - S)}}{2r}$$

In general, it cannot be ascertained whether in this equilibrium the entire population will subscribe to and comply with the norm or just a large part of the population. If $R \leq r$, then not everyone will comply with the norm in the equilibrium, since some people are not sensitive to a loss of reputation.
Finally, we analyze a situation in which a group of immigrants enter the society, who do not subscribe to the prevailing norm and neither comply with the norm, since they are not very sensitive to a loss of reputation. Assume that immigration increases the population with a factor $\alpha > 1$. Assume further that the society is in the equilibrium point of equation (9), in which everybody who subscribes to the norm also complies with it. Hence, the proportion of the population subscribing to and complying with the norm declines by a factor $\alpha$: $\beta_1 = \beta^* / \alpha$. This reduces the potential loss of reputation for those who do comply with the norm. As a result, some indigenous people start breaking the norm. The new long term equilibrium $\beta_1^*$ is therefore smaller than $\beta^*/\alpha$; immigration of “non-believers” also disrupts the compliance of the indigenous citizens.

This can be shown as follows. The slope of the discomfort curve, $b$, and the reputation curve, $r$, increase with a factor $\alpha$. Hence, in the long term equilibrium:

\[
(11) \quad (b - R + \beta^* r) \beta^* = (\alpha d - R + \beta_1^* \alpha r) \beta_1^*
\]

Assume that $\beta_1^* \geq \beta^*/\alpha$. This implies:

\[
(\alpha d - R + \beta_1^* \alpha r) \beta_1^* \geq (d - R/\alpha + \beta^* r/\alpha) \beta^* = (d - (R - \beta^* r)/\alpha) \beta^* = (d - R + \beta^* r) \beta^* + (1 - 1/\alpha)(R - \beta^* r) \beta^* > (d - R + \beta^* r) \beta^*
\]

since $\alpha > 1$ and $R - \beta^* r > 0$. This contradicts equation (11), hence $\beta_1^* < \beta^*/\alpha$. 

37
REFERENCES


Figures

Figure 1. Costs and benefits of norm violating behavior, internalization and formal control (t=1)

Figure 2. Costs and benefits of norm breaking behavior, internalization and formal control (t=1)
Figure 3. Costs and benefits of norm breaking behavior, norm internalization, formal and informal control (t=0)

Figure 4. Costs and benefits of norm violating behavior, norm internalization, formal and informal control (t=1)
Figure 5. Costs and benefits of norm violating behavior, norm internalization, formal and informal control (t=2)

Figure 6. Costs and benefits of norm violating behavior, norm internalization, formal and informal control (t=3)
Figure 7. Violent crime rate and number of sentenced prisoners in the United States (1950-2005)

Source: U.S. Department of Justice, Bureau of Justice Statistics (http://www.ojp.usdoj.gov/bjs), Sourcebook of Criminal Justice Statistics (http://www.albany.edu/sourcebook/wk1/t628004.wk1)
## Tables

### Table 1. Prisoners’ dilemma

<table>
<thead>
<tr>
<th>Behavior A</th>
<th>Behavior B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Produce</strong></td>
</tr>
<tr>
<td><strong>Produce</strong></td>
<td>2, 2</td>
</tr>
<tr>
<td><strong>Steal</strong></td>
<td>3, 0</td>
</tr>
</tbody>
</table>

### Table 2. Four types of incentives that affect people’s behavior

<table>
<thead>
<tr>
<th>Source of incentives</th>
<th>Character of incentives</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Material</strong></td>
</tr>
<tr>
<td><strong>Extrinsic</strong></td>
<td>(4) Formal norms: legal sanctions</td>
</tr>
<tr>
<td><strong>Intrinsic</strong></td>
<td>(1) Absence of norms: material self-interest</td>
</tr>
</tbody>
</table>

### Table 3. Incentive structure with legally enforced extrinsic norms

<table>
<thead>
<tr>
<th>Behavior A</th>
<th>Behavior B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Produce</strong></td>
</tr>
<tr>
<td><strong>Produce</strong></td>
<td>2, 2</td>
</tr>
<tr>
<td><strong>Steal</strong></td>
<td>3–S, 0</td>
</tr>
</tbody>
</table>
Table 4. Incentive structure with informal norms

<table>
<thead>
<tr>
<th>Behavior A</th>
<th>Behavior B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Produce</td>
</tr>
<tr>
<td>Produce</td>
<td>2+P+L, 2+P+L</td>
</tr>
</tbody>
</table>