11 Countervailing norms

Emanuela Carbonara, Francesco Parisi and Georg von Wangenheim

1. Introduction

There are some circumstances in which social norms counteract legal norms, reduce their effect and even produce results that are opposite to those intended by the lawmaker. In this sense, we say that social norms may be “countervailing”. Then, legal intervention that does not take into account existing social norms may not only be ineffective but may even have a negative impact. For instance, a law introducing a new sanction may either be less effective than expected or even cause an increase in the sanctioned behavior.

Such results contradict the predictions of traditional law and economics theories. According to such theories, the law is “an obligation backed by a sanction” (Raz, 1980). Economic analysts characterize legal sanctions as a price. Individuals consider sanctions as an external incentive and make decisions regarding legally relevant behavior by means of standard cost and benefits analysis. Alternatively, expressive-law theories argue that individuals tend to internalize the obligations expressed by the law, so that legal norms become social norms.

Both theories predict a unique outcome of lawmaking: stricter laws and larger sanctions reduce the prevalence of the regulated behavior. Conversely, more lenient laws and smaller sanctions increase its frequency.

Such a conclusion has recently been challenged by experimental evidence. Gneezy and Rustichini (2000) tested the effects of the introduction of a fine for latecomers on parents’ behavior in several day-care centers. In a field experiment, they found that the introduction of a fine produced an increase in the number of children picked up late by parents. According to the traditional economic model, since the fine increases the price of coming late, fewer parents should do so. Similarly, according to expressive-law theories, the fine should represent the obligation for parents to collect their children on time and should therefore induce parents to be timely. There could be several explanations for this phenomenon; the fine could have been interpreted as a price by parents who recognized it as a new service offered by the center. Alternatively, paying an additional fee may have assuaged the guilt associated with lateness. Whatever
the real motivation, the main point is that the new rule contradicted both traditional and expressive-law theories.

This result, albeit striking, is not the only one highlighting the unintended results of legal rules. Stuntz (2000) considers how criminal law can be self-defeating if it is not aligned with general values and norms. He illustrates how laws against vice crimes (gambling, drinking during the Prohibition) and highly divisive “moral” crimes (e.g. slavery) raised fierce opposition and generated disobedience rather than obedience. In these cases, opposition was made even stronger by the perception that these policies were biased against specific ethnic groups and social classes and this corroded the legitimacy of the law.

Empirical evidence showed that the enactment of restrictive copyright laws had unintended effects on the amount of peer-to-peer file sharing over the internet. After music companies started suing people for illegal downloading, in the summer 2003, the number of downloads increased significantly (Depoorter and Vanneste, 2005; Oksanen and Välimäki, 2007). Moreover, there is experimental evidence indicating that legal sanctioning of file sharing produces a backlash effect on behavior; exclusive reliance on punishment might be unfortunate if sanctions have a countervailing effect (Depoorter, Parisi and Vanneste, 2005).

Understanding both responses to legal rules is therefore crucial, in order to explain the huge variety of different and unexpected results of legal innovation. A regulator introducing new rules assessed on the basis of a wrong or incomplete model of human behavior may lead to outcomes far from the social optimum.

Human behavior is affected by a series of factors that go beyond the mere extrinsic incentives of sanctions and rewards. These factors include individual values and social norms and constitute the intrinsic motivation of behavior. Intrinsic and extrinsic motivations interact in a variety of ways. If the law reflects generally accepted values, then intrinsic motivation and extrinsic incentives push individual behavior in the same direction and the law produces the expected results. This occurs also when the law has a high expressive value, i.e. if the obligation contained in the law is internalized quickly by individuals. If, on the other hand, the law clashes with social norms, and intrinsic motivation exerts a stronger influence on individual behavior, then social norms may have a countervailing effect. Finally, the same type of extrinsic incentives introduced by the law may have a framing effect on behavior, determining the direction of the interaction with social norms. Particularly, designing incentives in the form of sanctions for violators rather than rewards for compliers may make a difference on how the law is perceived and may determine obedience or backlash.

An analysis of the effects of legal innovation must necessarily consider the interplay between social norms and the law. Particularly, we need to study how social norms are formed and enforced. Then we need to briefly characterize the expressive function of the law, to understand under which conditions the law
comports with social norms. This will help us understand when legal innovation may lead to a backlash due to countervailing social norms.

2. The Origins and Evolution of Social Norms

Social norms are defined as conventions accompanied by normative expectations (Sugden, 1989; Bicchieri, 2006). A convention is a pattern of behavior followed by the majority in a given social group (Young, 1993). Breaking such a pattern may result in substantial costs. Given that everybody else follows the convention, an individual following a different behavior most likely prevents coordination, thus imposing a cost on others. The classic example of this is choosing which side of the road to drive on, left or right. If one driver starts to drive on the left while everybody else drives on the right, an accident will almost surely happen. Hence, each person expects others to follow a convention and observing lack of compliance provokes anger and resentment and possibly leads to a social sanction. Such expectations, which might lead to anger and sanctions, are termed “normative expectations”, because they trigger the enforcement that transforms a behavior followed by a majority into a norm.

The enforcement of social sanctions follows two distinct channels. The first channel we described above and termed second-party and third-party enforcement. The direct victims of disobedience implement second-party enforcement, while other individuals who may be harmed indirectly apply third-party enforcement strategies, such as public disapproval, ostracism, or even physical punishment. The second channel is more subtle and can be termed “first-party enforcement” and, in this case, punishment is applied by violators themselves: since individuals dislike being disapproved of, they suffer a (psychological) cost (guilt).

How is a convention (and then, possibly, a social norm) established? The process closely resembles a bandwagon effect, that is, a situation in which a given behavior becomes more attractive the more people follow the same behavior (Leibenstein, 1950). A bandwagon effect presupposes the existence of a threshold number of followers, above which following the convention becomes optimal. In other words, a bandwagon effect implies that the number of followers needs to reach a critical mass.

Becoming the object of resentment and social sanctions reduces the individual’s utility from breaking with convention, while obedience to social customs yields extra utility, as people feel they are approved and are part of a community. This mechanism of expectations and social sanctions can pull individual behavior in socially undesirable directions. The literature establishes that people do not necessarily enact the most efficient social norm when more than one is available (Sugden, 2004). Once the number of people following a social norm reaches a given threshold, more and more people will be encouraged to follow it, regardless of whether such a norm is just and efficient. The possible
inefficiency and unjustice of social norms can be found also in Cesare Beccaria’s words “human conventions […] are expressly formulated or simply assumed in view of common necessity and utility” (Cesare Beccaria, 1764).

Conventions typically spread because of past experience (the last time I encountered another driver, he kept to his side of the road), common background and ideas, and by analogy. There is, therefore, a strong path dependency in the adoption of a convention (Sugden, 1989). Individuals choose behavior based on what they did or what they saw others doing in similar past circumstances. They take decisions following a backward-looking, myopic approach, the exact opposite of a forward-looking, rational expectations approach. Forward-looking individuals are able to spot the convention that maximizes total surplus among the many possible conventions upon which they may converge. On the contrary, backward-looking agents are not able to assess the effect of their choices on total surplus and can adopt and maintain the wrong convention. Similarly, conventions and social norms can be unjust. One should consider the many social conventions turned into very strict social norms that have limited the freedom, restricted the rights and diminished the dignity of women and ethnic minorities (Carbonara and Pasotti, 2010).

This is one of the reasons why a lawmaker may choose to enact laws that do not align with pre-existing social norms, in an attempt to correct those which are inefficient or unjust. We will state other possible explanations below, after describing how the law can bend moral and social values.

3. The Expressive Power of the Law
According to expressive law theories, the law can influence behavior independently of the sanction it imposes (McAdams, 2000a). Expressive-law scholars have provided several examples of how the law can change individual preferences and tastes and, in some cases, affect social norms and values. For instance, the ban on public smoking reduced the number of young black Americans who smoke by changing the social meaning of smoking from rebellion to dirtiness and foolishness (Sunstein, 1996). Then, the lawmaker can use the law to provoke changes in the social meaning of behavior (Lessig, 1995 and 1996).

The law can also create focal points that help people solve coordination problems in environments characterized by a multiplicity of equilibria (McAdams, 2000a). Driving conventions, once again, provide a good example. People can either drive on the right-hand or on the left-hand side of the road. Without a coordinating device, some individuals may choose to drive on the right, some on the left and several collisions may ensue. By stating that drivers should keep to the right, the law may solve the coordination problem, since now people receive an indication of the correct behavior and form expectations on the actions followed by others. This may have the effect of changing previous
equilibria, as people may change their behavior and become unwilling to follow tradition. Then, expressive law may contribute to changing existing social norms. To illustrate how that could happen, McAdams presents the example of a community in which smokers traditionally follow very aggressive strategies towards non-smokers, to affirm their perceived right to smoke in public. A law prohibiting smoking in public places may change not only the behavior of non-smokers, who now have a basis for questioning whether the strategies followed in the past will continue to apply in the future, but also the expectations of smokers about non-smokers’ reaction. Before the enactment of the law they expected passive behavior, but now they may believe that non-smokers feel entitled to react strongly to their smoking in prohibited public places. If non-smokers fight against smoking in public places in sufficiently high numbers, “they will establish a [...] convention against smoking” (McAdams, 2000a) and the law will be enforced by second-party sanctions. Moreover, laws enacted through a democratic process are generally positively correlated with “popular attitudes” (McAdams, 2000b) and thus provide a signal of those attitudes, changing one’s beliefs about what others will think. Given that individual preferences are affected by what others approve or disapprove of, the law can influence behavior even without a legal sanction.

Through expression, the law can also induce individuals to internalize the values it embodies (Cooter, 1998 and 2000). This distinguishes expressive-law theories from traditional theories focused on the role of law as an instrument for creating external incentives, such as taxes, sanctions and rewards. According to expressive-law theories, internalized rules may trigger private enforcement mechanisms and change observed patterns of behavior, even in the absence of other external incentives. As mentioned above, private enforcement mechanisms include first-party, second-party and third-party enforcement. First-party enforcement implies that legal rules are also able to affect behavior through “internalization.” To explain how the internalization process takes place, Cooter refers to social psychological research that suggests that the majority of citizens obey laws out of internal respect for the law in general (see Tyler, 1990). Citizens have a preference for obeying the law, so that the (psychological) cost of performing a given action increases after that particular action is prohibited by the law. Thus people are induced to refrain from that action.

Legal rules differ in their expressive power. Laws with a strong expressive power are laws that citizens are willing to pay a lot to obey. In general, even if literature on this specific issue is almost non-existent, it seems that the expressive power of a law is inversely related to the degree of forcefulness of the social values the law tries to bend. Therefore, the alignment of legal precepts and decisions of authorities with current social norms and values has a positive influence on whether people comply with law, even when it is not in their self-interest to do so (Tyler, 1990 addresses this issue with respect to law;
and Tyler and Huo, 2002 do so with respect to the decisions of authorities. Legitimacy is undermined when the content of the law departs from social norms, be they based on moral, ethical, or merely cultural values. Tyler (1990) and Sunshine and Tyler (2003) provide support for the argument that the public’s perceptions of legitimacy impact people’s compliance with the law.

4. Why Do Laws Clash With Social Norms?

We have just seen that the expressive power of the law is stronger the closer the content of the law to existing social norms. Whenever the law differs from current values, we might observe a backlash when society tries to correct the law and align it with social norms. The effects of the law therefore depend on the social “reaction” triggered by legal innovation, whereas social reaction is likely to be proportional to the degree of consistency of the new law with prior social values.

The question then becomes why lawmakers and norm entrepreneurs (i.e. individuals interested in changing the law; see Sunstein, 1996; Ellickson, 2001) enact laws that depart from the values and norms shared by the majority of the citizenry. As we argued in Section 2, social norms can be unjust and/or inefficient. A welfare-maximizing lawmaker may exert a paternalistic role, in the attempt to change “wrong” social norms. This happened when more severe rape laws were passed in an attempt to change the behavior of men (and women) who may have acted according to the belief that “no sometimes means yes”. Similarly, there are numerous attempts to reduce drunk driving by enacting more restrictive laws and penalties (Kahan, 2000).

Sometimes, however, new laws may reflect the values held by the majority at the moment of enactment, values that may then change (possibly as a result of legal innovation itself), leading to a misalignment between the content of the law and the new social norm. Stuntz (2000) uses Prohibition as an enlightening example of this principle.

The majority of the U.S. population approved of laws enacted in 1919 that banned the use of alcohol in the United States. Alcohol was a perceived cause of many social problems; such problems predominately affected the poor. Many factory workers struggled with alcoholism and spent all their wages on alcohol, left their families with hardly anything to eat and were often physically abusive. That is why the white, protestant majority saw Prohibition as a paternalistic way to tackle a problem affecting mainly poor neighborhoods, generally populated by immigrants (typically Italian and Irish Catholics). This explains why Prohibition passed notwithstanding strong opposition by rich lobbies. In 1933, Prohibition was repealed and an overwhelming majority approved repeal (Kyvig, 1979). What explains such a dramatic shift in social preferences? According to scholars, the way Prohibition was enforced had a negative impact on its legitimacy. Alcoholism permeated society, both poor
and rich neighborhoods alike. However, the police chose to target only poor neighborhoods. The reason is rather straightforward, even without resorting to arguments of social favoritism. Scarce police forces were allocated where the problems of alcohol abuse were more severe. However, ethnic minorities in poor neighborhoods felt persecuted; while alcohol consumption was criminal for them, it was not for rich Anglo-Saxon protestants. This eroded the power of the law in those communities, eliciting an enormous rise in the phenomenon (additionally, the price of alcohol had increased significantly, making alcohol manufacturing and distribution attractive to criminal organizations).

When the expressive power of a law vanishes completely, the lawmaker faces two possibilities: either increase sanctions considerably (which is costly) or lift the legal ban altogether. In the early 1930s, America was suffering through the Great Depression and it could not afford the costs of a war against the consumption of alcohol. Alternatively, it may have been that the majority of citizens could not bear to see a large number of white immigrants in jail for selling beer. Whatever the reason, an overwhelming majority supported the abandonment of Prohibition. This illustrates how a legislature might enact laws which depart from social values regardless of the fact that they are elected to represent their constituents. The same dynamics were observed when laws punishing crack consumption were passed in the United States in the 1980s and 1990s (Tonry, 1995). In that case, enforcement targeted poor, black neighborhoods, where crack use rose to incredible heights as enforcement grew. Unlike Prohibition, however, lawmakers responded to the backlash by dramatically raising penalties, because drug abuse presents even more serious social implications.

Finally, laws departing from existing social norms may be enacted in situations of serious danger for the population or in times of war. For example, following the terrorist attack on the World Trade Center in New York on September 11, 2001, the US Congress passed the Patriot Act, which limited personal freedom and constitutional rights. Such law raised considerable criticism as it was deemed to excessively weaken the protection of civil liberties (Carbonara et al., 2010).

5. Social Norms and Legal Backlash
We now explicitly consider the case of laws that do not align with pre-existing moral norms and social beliefs. By inducing civil disobedience, such laws might reinforce social values contrary to law, producing a countervailing effect.

Particularly, when the law is perceived as too strong, people tend to approve disobedience and try to stop enforcement of an “unjust law”. In addition to that, they might publicly protest and engage in civil disobedience. When the law is more lenient than socially desired, people tend to show reprobation when confronted with an infringement of the (weak) law. They might even try themselves to apply the larger penalty they believe appropriate. Hence social
reaction may reinforce or undermine the effect of legal intervention, depending on whether the law is too strict or too weak compared to the social norms.

Kahan (2000) considers the case in which a law is passed that tries to change an existing social norm. Modern sexual assault laws were passed that attempted to change the “no sometimes means yes” social norm. Those laws eliminated the physical resistance requirement traditionally needed to prove that intercourse was non-consensual. Such legal innovation has been highly ineffective, since prosecutors have been reluctant to charge men who disregarded verbal protestations and juries have rarely convicted them (Schulhofer, 1998). Kahan (2000) calls this the “sticky-norm problem”: social norms so rooted in common behavior that enforcers are reluctant to apply laws contrary to such norms. Enforcers face a tradeoff between their desire to do their civic duty by correctly applying the laws of the state and their aversion to enforcing rules that punish what is socially accepted behavior. He then argues that the sticky-norm problem affects laws that punish far more severely than enforcers believe is right. In such cases, in fact, their aversion to punish too severely will dominate their disposition to apply the law, and they will refuse to proceed.

The decision of some enforcers may affect other enforcers, who may follow suit. This “cascade effect” may also be driven by the legal system itself: if the enforcer is a judge, then her decision will constitute a precedent in common-law countries and this may bias future decisions toward the non-application of the controversial rule. Such an effect may occur in civil-law countries too, where judges base their interpretation of the law on precedent decisions (especially those by Supreme Courts).

As a result, legal innovation that departs too much from current social norms is incapable of reducing incidences of the prohibited actions. Kahan (2000) argues that it might even increase incidence. In fact, not applying the law (for instance, acquitting a defendant accused of raping a woman who protested but did not physically resist) reasserts the norm the law was trying to change, perpetuating and even encouraging the crime.

The solution to the sticky-norm problem then is to proceed with “gentle nudges” in lawmaking. If the increase in severity is limited in size, enforcers will apply the law, eroding the social norm, as individuals will refrain from the sanctioned behavior if they observe consistent punishment. “Hard shoves” (meaning sudden and conspicuous increases in severity) are therefore counterproductive.

Carbonara, Parisi and von Wangenheim (2008a) explicitly analyze the effect of the enactment of the law on social norms and their evolution. They assume that people behave according to prevailing social norms (Sugden, 1989; Sunstein, 1996; Cooter, 1998, 2000). There is an initial distribution of individual beliefs about the appropriateness of a given behavior, and social norms reflect such distribution (to continue with the previous example, if the majority believes
that “no sometimes means yes”, then rape when a woman protested but did not physically resist will not be the object of social sanctions and will generally be tolerated).

How does the initial distribution of beliefs originate a social norm? Individuals in society interact and exchange opinions, influencing each other’s beliefs (Bikhchandani, Hirschleifer and Welch, 1992 and 1998). Depending on the number of people holding a given belief and on the strength with which each belief is held and publicly defended, other individuals may be influenced. Then, some of the beliefs in society will function as “attractors”. If a sufficiently high number of individuals end up sharing the same belief, that belief will become a social norm. If the majority converges on a unique belief, then society will hold a shared social norm. If, however, individuals cluster around different beliefs, multiple social norms will coexist. It may even happen that a polarized society results from such a process, with people clustering around opposite social norms.

In this evolving scenario, in which people converge on social norms through interaction and mutual influence, the law exerts a twofold impact. On one hand, it is capable of attracting individual opinions towards the value embodied in the law (expressive function). On the other hand, it could repel those who hold strong contrary beliefs (countervailing effect).

If the second effect is strong enough, legal innovation can generate a social norm different from the norms embodied in the law. It may also generate social divide even in the face of a previously balanced society. Social divide is particularly problematic, since it provides grounds for discrimination, segregation or other forms of social conflict and violence (Carbonara and Pasotti, 2010). Social divide more likely occurs when individuals feel very strongly about discrepancies between their beliefs and the values embedded in the new law, which act as a powerful “repellent” for those whose beliefs are different enough, but as an effective attractor for those whose views are similar enough.

So far, we have considered how the law influences social norms. The analysis can be pushed forward to include the impact of the law on individual behavior. There are multiple determinants of behavior, each possibly exerting a different force. First of all, behavior is affected by external incentives, like the monetary gains from performing an action (or, more generally, the utility derived from it) and the sanction imposed by the law. Generally, in this respect, a higher sanction would increase the relative price of the sanctioned activity, leading to a substitution toward other non-sanctioned activities. The prevalence of a given action in the society also affects behavior: if many people perform it, no matter whether it is illegal or contrary to social norms, “wrongdoers” feel more comfortable. Finally, last but not least, behavior is influenced by social norms, which, as argued above, are not passive and react to legal innovation. When the expressive effect of the newly enacted law is strong enough, it might bend existing social norms. As mentioned above, a powerful social norm and one of
the main constituents of individual intrinsic motivation is the desire to obey the law (Cooter, 2000). Enacting strict laws that severely punish a given behavior clearly signals that the state considers of primary importance deterring citizens from performing the sanctioned action, and this tends to reduce prevalence. Conversely, a new law that is markedly contrary to current social values or more restrictive than what people would support triggers opposition, in the form of both open protest and civil disobedience and might bend social norms in the opposite direction.

Carbonara, Parisi and von Wangenheim (2010) study how these influences affect behavior. They find that negative reactions by a sufficiently high number of individuals may result in a reinforcement of other individuals’ dislike of the law, with countervailing effects. Such a result is even more likely if the expressive power of the law is not very strong and the behavior is relatively common. Then, the situation after legal change is characterized by high protest and an increased rate of violations compared to the initial situation.

From this, the following intuition results. When legal innovation is not far from pre-existing social values, sanctions and the expressive effect of the law move preferences in the same direction, forcing internalization. Conversely, when the law departs from current social values, the expressive and reaction effects pull behavior in opposite directions. The final outcome and the ability of the legislator to achieve her intended goal depends on the relative magnitude of external incentives, of expressive and reaction effects. In the following, we will provide an account of the results of our economic models. Thus, legal innovation may have the undesired consequence of increasing prevalence of the behavior it intends to ban. This result is similar to Kahan (2000), but is obtained through a completely new and more general mechanism. While in Kahan it was only a matter of enforcement, in Carbonara et al. (2010) the new law is enforced, notwithstanding possible disapproval. Hence, the countervailing effect is obtained in the face of opposing extrinsic incentives and not because such incentives are blocked by enforcers. It is therefore a much stronger result, obtained in a framework that seems more suitable to describe recent cases of countervailing effects, like the increase in music downloading presented above.

To see how the model works, consider a simple setting in which individuals decide whether to obey the law according to the magnitude of the sanction, the strength of the expressive power and expressed opinion (approval or disapproval) about a specific law. As it would be realistic to assume, individuals are more prone to obey the higher the sanction and the stronger the expressive power of the new law. The new element in the analysis is the effect of expressed opinion on compliance. The more individuals engage in social protest against the law, the less likely it is that people will obey the new law, since protest represents the public expression of dissent, which reveals that the law clashes with the values shared by those who protest. A very high level of protest shows that the
new law does not reflect social norms. Conversely, if many people express an opinion in favor of the new law, there would be no countervailing effect.

The more severe the law, the higher the likelihood that the law will be disapproved of and that protest will be high. In fact, consider an individual who believes that downloading music from the internet, infringing the copyright, is wrong (so that she would approve a law sanctioning such conduct) and that somebody downloading one song should be fined $10 (about ten times the price of a song purchased legally). Now assume that the law prescribes a fine equal to $2000. Our individual would disapprove of the excessive harshness of the law and empathize with the “wrongdoer”, notwithstanding her initial belief that illegal downloading should be punished.

To understand how the various forces illustrated above can impact behavior, consider the following simple model, introduced in Carbonara, Parisi and von Wangenheim (2008b). Define \( c \) as the percentage of people violating the law, and \( p \) as the level of protest. As explained above, \( c \) is an increasing function of protest, \( c'(p) > 0 \). Now assume that a new legal standard is introduced, which increases the sanction imposed on violators from \( S_A \) to \( S_B \). The function \( c(p) \) is drawn for a given sanction. Since a higher sanction reduces the incentive to violate the law, there will be fewer violators for all levels of protest. Hence, \( c(p) \) shifts downwards, as in Figure 11.1, where the increase from \( S_A \) to \( S_B \) moves the

![Figure 11.1](image)

**Figure 11.1** A more restrictive law with strong expressive power: the law has the desired effects
Production of legal rules

curve $c(p)$ from $c_A(p)$ to $c_B(p)$. The increase in the sanction also has the effect of raising protest, from $p_A$ to $p_B$. However, protest is directly influenced by the expressive power of the law, so that a law with a high expressive power will inhibit protest and generate a moderate increase in $p$. This is the case depicted in Figure 11.1, where the shift in protest is not high enough to compensate for the effect of extrinsic incentives and the new rate of violations, $c_B'$, is lower than the initial rate, $c_A$. Then, when the law has a strong expressive power, the combined impact of incentives and internalization will dominate and the increase in the severity of the law counterbalances the effect of protest and increases legal compliance. The enactment of a law backed by an enforceable sanction thus reduces the rate of violations, as normally expected.

Conversely, when the expressive power of the law is limited, the combined effect of incentives and internalization may not be sufficient to dominate the impact of protest. Figure 11.2 considers again the same increase in the sanction from $S_A$ to $S_B$, which causes $c(p)$ to shift from $c_A(p)$ to $c'_B(p)$. However, this time the expressive effect is low, so that social reaction predominates and protest increases from $p_A$ to $p'_B$, much more so than in Figure 11.1. Then, protest more than compensates for the incentives determined by the stronger sanction, and the new rate of violations, $c_B'$, is now larger than the initial rate,

![Graph showing the relationship between $c$, $p$, $c_A$, $c_B$, $p_A$, and $p_B$](image)

**Figure 11.2** A more restrictive law with weak expressive power: countervailing effects
Adding incentives by exacerbating legal sanctions may be counterproductive in this case.

Such an unintended effect of legal intervention would thus lead to a compliance paradox, where an increase in the strictness of the law leads to more violations, defeating the goals pursued by the lawmaker. That is why the law, in this case, is self-defeating.

Finally, one should also consider the case in which a very lenient law is passed. For instance, a behavior that was previously banned now becomes legal (the obvious example in this case is abortion, legalized in the US in 1973 by the Supreme Court’s decision *Roe v Wade*). This case too is likely to raise strong protest. In contrast to the above, however, protest in this case would signal disapproval of the now legal behavior and would inhibit it. The countervailing effect now would consist in a decrease in the prevalence of the behavior, notwithstanding liberalization. In the case of abortion, for instance, *Roe* and the recent battle over partial birth abortion played out in *Gonzales v Carhart* (2007) has sparked a political battle and still provokes much protest by pro-life groups. Protest ranges from open expression of dissent to violent assaults on women trying to access abortion services, doctors and abortion clinics. There is empirical evidence that such anti-abortion activities have had a significant impact on both the supply of and demand for abortion services in the US (Kahane, 2000).

### Faster Social Reaction and Social Outrage

We have seen above that any time the law has limited expressive power and social reaction is strong, the compliance paradox is more likely to occur. However, the model illustrated above implies that, in order to observe countervailing effects, the upsurge in protest has to be extreme and persistent in time. Generally, although legal innovation may provoke social reaction, protest tends to subside over time. Moreover, an extreme social reaction, of the kind needed for countervailing effects to materialize, is exceptional, and a theory capable of explaining only exceptional cases is of little use to the lawmaker. Therefore, we shall now present alternative explanations for the countervailing effects, explanations that do not require persistent increases in protest (this section follows Carbonara et al., 2010).

In general, protest is not triggered solely by the “absolute strictness” of the law but also by the relative change. Increasing a sanction from one year in jail to ten years in prison likely sparks much more protest than an increase from one to two years. The fact that protest is not influenced only by the absolute sanction $S$ but also by the relative change $\Delta S = S_B - S_A$, causes protest to rise more than it would if only the absolute level $S$ affected it. Protest may then subside over time, but if the initial rise is large enough, it might stabilize at a high enough rate to justify an increase in violations. Countervailing effects
are therefore explained also by “social outrage”, an initial, strong reaction that generally fades away over time. Due to such reaction, protest changes much faster than behavior, which facilitates the insurgence of countervailing effects (as Figure 11.2 shows).

A social reaction may occur much faster than a change in behavior. Also, when changes in the law are announced some time before actual enactment, it can hardly produce any deterrent effect on current behavior until it is in force. It can, however, trigger protest if it diverges from current social norms. When the new law eventually comes into force, the regulated behavior may be so prevalent that society converges toward an equilibrium with more violations than before the enactment of the law.

7. Choosing Optimal Instruments of Legal Innovation

The analysis in the previous sections shows that, when designing new laws, lawmakers should take into account several, often contrasting factors. Particularly, they should balance extrinsic incentives, internalization and reaction effects. First of all, the existence of a compliance paradox indicates that laws intending to induce substantial shifts from current norms may have to proceed in a gradual fashion. Moving the law in the desired direction in small, consecutive steps allows for the gradual adaptation of individual values to the values expressed in the law, and avoids countervailing effects, since the disobedience rate will be small at every step. The compliance paradox suggests another interesting solution. Legal deterrence could be increased by reducing the strictness of the law. This would induce approval of the law and reveal that society condemns the regulated behavior, thus discouraging it and magnifying the deterrence effect. When gradual adjustments are not possible (e.g. due to the discrete nature of legal change or where they are not viable on political grounds), legal instruments other than sanctions may be required to obtain the desired result. In some circumstances, taxes might be a better option, as they are perceived as a “price” that individuals are free to pay if they want to indulge in a given, socially costly, behavior (like polluting). However, taxes may also be unacceptable in some circumstances. It may be then be possible to induce the desired behavior by means of positive incentives (i.e. rewards and prizes).

Typically, the literature has shown the equivalence of positive and negative (i.e. sanctions) incentives (Fudenberg and Tirole, 1992). The reward of a non-violation may be seen as equivalent (in incentive terms) to the sanctioning of a violation. Experimental studies have, however, found that rewards and sanctions are perceived differently by people (Kahneman and Twersky, 1979). Reaction to incentives is sensitive to framing, and in some contexts individuals are much more motivated if they face the perspective of receiving a prize for good performance rather than a sanction for bad behavior. In a contracting environment, for instance, it has been shown that the introduction of a penalty
A clause in the contract significantly increases the incidence of contract breach and decreases the productivity of specific investments (Fehr and Rockenbach, 2003). Moreover, this equivalence may be lost when the expressive and countervailing effects are taken into account. While positive and negative incentives may have similar expressive effects, they may have opposite countervailing effects. The giving of a reward is in fact less likely to trigger reaction than the imposition of a punishment.

It would, however, be hard to maintain that criminal law should be enforced by rewarding those who do not commit crimes rather than punishing criminals. Perhaps criminal law is not the best environment in which to test positive incentives. However, many other legal fields, such as torts and contracts, may benefit from such a perspective. Previous literature has stressed that positive incentives are much more expensive to provide than negative ones (Dari-Mattiacci and De Geest, 2010). If a policymaker is willing to motivate individuals with prizes, they have to reward all those who behave correctly and do so repeatedly to avoid misbehavior once the prize has been received. With a sanction, the threat of applying it if deviations from correct behavior are detected is sufficient to guarantee compliance. A “dictator with one bullet” is enough to subdue an entire country. However, this argument should be amended to take into account the costs of the enforcement of sanctions. Even if, in equilibrium, no one will be sanctioned when negative incentives are designed properly, it is true that the policymaker will have to invest in “sanctioning capacity” (e.g. building prisons) and that the threat of the sanction will have to be credible (meaning that there will have to be a conspicuous investment in detection capacity, which is likely to be very costly).

Once these costs are taken into account, the difference in efficiency between positive and negative incentives may be considerably reduced. In circumstances in which countervailing effects are important, rewards may become socially preferable and more effective in taming certain types of socially undesirable behavior. It is interesting to notice that the balance between prizes and sanctions was explored back in 1767 by the Italian philosopher Giacinto Dragonetti, in his Essay On Virtues and Rewards, which was a response to the far more famous Essay on Crimes and Punishments by Cesare Beccaria (1764).

Finally, the analysis of countervailing norms sheds new light on the debate between the magnitude and the probability of a fine, a debate dating back to Becker (1968). As is well known in the literature, the same expected fine can be generated with infinite combinations of probability versus magnitude of a sanction. Given that increasing the probability of a sanction entails high costs and that increasing the magnitude of the fine is not as costly, the conclusion supported in the past is that the tradeoff between probability and magnitude should be resolved by setting the maximum fine as high as possible and reducing the probability of application. In the light of our results, the choice of
probability and magnitude combinations may have different effects with respect to countervailing norms. This might change quite substantially the policy recommendations made in the literature inspired by Becker’s model.

**Bibliography**


202  Production of legal rules


Countervailing norms


