COSTLIER AFTER OR BEFORE?
The alleged enforcement cost advantage of ex-post over ex-ante controls

GABRIEL DOMÉNECH
University of Valencia Law School

Abstract: Some authors have argued that liability for harm and, in general, other legal instruments of inducing behavior that are triggered by the occurrence of harm have an enforcement cost advantage over those carried out independently of such occurrence. This paper shows that harm-based (i.e. ex post) controls do not have such an intrinsic advantage over non-harm-based (i.e. ex ante) ones. The key factor with that regard is whether the control at issue is triggered or not on the relevant authority receiving a signal of noncompliance, and also whether potential injurers, by complying with the law, can lower the likelihood of such signal. It also shows that some non-triggered-by-harm regulatory interventions can have an enforcement cost advantage over modes of inducing behavior dependent on the occurrence of harm. Moreover, it argues that some triggered-by-a-signal-of-noncompliance-different-from-harm controls have several of the best features of ex ante and ex post worlds.

Keywords: enforcement costs, liability for harm, ex ante regulation

JEL Classification: K42, K13

I. INTRODUCTION

In order to induce people to abide by the law, legal systems establish control mechanisms through which compliance is examined and different measures (e.g. punishment for offenders or rewards for abiders) are to be taken depending on the result of the exam.

Those controls may be placed at different moments:

(i) Sometimes, they are carried out before the actual start of the activity subject to control. It is, for instance, the case of those authorization procedures (e.g. for marketing medicinal products) by which government agents verify whether the projects presented by applicants meet the requirements established by the law.

(ii) Other controls take place after the corresponding activity has been launched, but before it appears to have actually caused harm. An example would be the inspections carried out as a result of such activity having been reported as illegal by a third party.

(iii) And others are performed after a harmful accident has occurred as a consequence of that activity. The prime example would be liability for harm.

There is a considerable law and economics literature on the comparative advantages and disadvantages of those timing alternatives and, in particular, on the pros and the cons of enforcing the law either before or after the activities subject to control appear to have caused an accident.

SHAVELL (1984) compares regulation of safety (which –he says– is “public in character” and operates “through requirements that are imposed before, or at least
Gabriel DOMÉNECH, Costlier after or before?

independently of, the actual occurrence of harm”) and tort liability, which is “private in nature” and “works through the deterrent effect of damage actions that may be brought once harm occurs”. Firstly, liability could have an advantage over regulation, or vice versa, if there is a difference in knowledge about risky activities between private parties and regulatory authorities. Secondly, the greater the probability of the injurer being incapable of paying for the full magnitude of the harm caused is, the greater the relative desirability of regulation will be. Thirdly, the appeal of regulation is also positively correlated with the likelihood of injurers not being either sued or held liable for the harm done.

GAROUPA and OBIDZINSKI (2010) have noted that both act-based [ex ante] and harm-based [ex post] sanctions might eventually generate the same level of deterrence. However, as the latter sanctions should be higher than the former ones, they are riskier and imply a higher likelihood of judgment-proofness by offenders. Harm-based sanctions give individuals incentives to acquire information and to take safety measures in order to efficiently prevent harmful accidents. Furthermore, they provide a more stable legal framework, insofar as act-based sanctions need to be adjusted when the government misestimates the expected harm caused, whereas such adjustments are not necessary under the alternative regime.

FRIEHE and TABBACH (2013) have noted that preventive enforcement, by thwarting the successful commission of offences and precluding offenders from obtaining private gains from crime, avoids social harm and, in addition, becomes a relatively more effective instrument of deterrence than non-preventive enforcement. However, since preventive enforcement is applied indiscriminately, it may prevent not only inefficient but also efficient offenses, i.e. those in which the perpetrator’s gain is greater than the social harm. Non-preventive enforcement, by contrast, is a better instrument of screening potential offenders, as it allows efficient ones to breach the law.

GIVATI (2014 and 2016) has shown that licensing [a kind of regulatory intervention taking place before the activity at issue has been launched] can eliminate two problems associated with adjudication [which takes place after the start of such activity]. The first one is the hold-up problem. It arises because after individuals have acted they cannot easily undo their actions, thereby losing “bargaining power”. This makes it tempting for regulatory authorities to adopt a strict position when adjudicating. However, expecting to be held up in adjudication, interested parties will avoid those actions, choosing instead activities that may be less desirable from the social welfare point of view. The second problem is the leniency one. After interested parties have acted, they may find it impossible or too difficult to comply with a strict legal position in adjudication. This may force regulatory authorities to adopt a lenient legal position when adjudicating, in order to avoid the costs of ex post compliance. And since those parties can anticipate that by acting –i.e. infringing the law– they can force the State to adopt a lenient position in adjudication, they will do so. Licensing makes it mandatory for individuals to have the corresponding act preapproved, thereby allowing regulatory authorities to commit ex ante to a certain legal position and thus preventing both problems. Advance ruling also allows the State to commit ex ante to a certain policy, but given that this regulatory instrument may be triggered only on request of interested parties, it can eliminate the hold-up problem, but not the leniency problem, as these individuals benefit from authorities being lenient.

HOEPPNER and KIRCHNER (2015) have recently discussed the pros and the cons of monitoring –which they call ex post– and incentive contracting –ex ante– strategies in order to solve the agency problem, mostly in the context of corporate governance and
from a behavioral point of view. They point out that ex-post strategies are behaviorally dysfunctional. Self-serving, hindsight and confirmation biases can easily distort the ex-post assessment of compliance. We could add that both the so called status quo and omission biases might arguably have similar effects as well\(^1\).

This paper focuses on the administrative costs of both regimes. The model presented by GAROUPA and OBIDZINSKI (2010) suggests that, as the magnitude of harm-based actions should be greater than that of act-based sanctions, the former generate fewer prosecutions and convictions than the latter ones and, therefore, are cheaper.

SHAVELL (2013) has also tried to show that liability and, in general, other modes of inducing behavior which are triggered by the occurrence of harm (ex-post controls) have an enforcement cost advantage over means of control carried out independently of such occurrence (ex-ante). The explanation is as follows. Ex-ante regulation induces individuals to take due precaution, but it does not reduce the likelihood of regulatory authorities examining whether such precaution has been taken. That likelihood is fixed; it is independent from individuals abiding by the law. Compliance induced by ex-post controls, on the contrary, lowers the probability of accidents and, therefore, the number of instances where the relevant authorities examine whether the required safety measures have been adopted.

The present paper adds to the abovementioned literature by criticizing, extending and modifying the results of SHAVELL (2013) in three ways.

First, it shows that the crucial distinction with respect to the referred enforcement cost advantage is not that between triggered-by-harm and non-triggered-by-harm controls, but that between triggered-by-a-signal-of-noncompliance and non-triggered-by-a-signal-of-noncompliance controls. Put in other words, the crucial distinction approximately refers to what MOOKHERJEE and PNG (1992) call “monitoring” and “investigation”. These authors give the name of monitoring to “enforcement activity in which the regulator must commit resources before receiving information about the offence, if any”. “By contrast, in enforcement by what [they] call investigation, the regulator can condition the resources committed on information... about the severity of the offence”. Investigation would be intrinsically more cost-effective than monitoring.

In our opinion, ex post controls are often cheaper than ex ante ones insofar as the former are usually carried out by means of an investigation –i.e. accidents often send a signal of noncompliance to the relevant public authorities–, unlike the latter, which are mostly performed through monitoring. Nevertheless, it must be pointed out that there is no perfect correlation between both pairings. Noncompliance might well be investigated before the occurrence of harm –e.g. as a result of a whistleblower reporting an offence– or monitored ex post –e.g. since certain accidents and the agents responsible for them cannot be freely detected–\(^2\).

The distinction proposed by MOOKHERJEE and PNG (1992) might be slightly modified and extended. The key factor with respect to the abovementioned cost advantage is not exactly whether public intervention is or not conditioned on information about the severity of the infringement. The key is whether such intervention depends or not on the relevant authority receiving information indicating the probability

\(^1\) See SAMUELSON and ZECKHAUSER (1988); ANDERSON (2003).

\(^2\) As INNES (2004) notes, this is not uncommon in the food, drug, product safety, and environmental contexts, wherein the occurrence of contamination or adverse effects is often not readily observable.
of an infringement in a certain case is higher than normal, and also whether potential injurers, by complying with the law, can lower the likelihood of such a signal.

Second, the paper proves that, under certain circumstances, some non-triggered-by-harm regulatory interventions can have an enforcement cost advantage over modes of inducing behavior dependent on the occurrence of harm, and vice versa.

Third, the paper shows that the magnitude of the advantage mentioned in the precedent paragraph mostly depends on just two variables: (i) on the conditional probability of the relevant authority receiving a signal of noncompliance given that compliance has actually taken place (say likelihood of false positives); and (ii) on the conditional probability of that authority being conveyed a signal of non compliance given that the law has been breached (say likelihood of true positives).

The paper is organized as follows. Section I illustrates those three propositions with an example. In section III, I try to demonstrate them by presenting a formal model based on Shavell (2013). Section IV discusses the legal and policy implications of the three aforementioned propositions. Section V suggests some extensions. Section VI concludes.

II. AN ILLUSTRATIVE EXAMPLE

Let us illustrate the abovementioned propositions with an example based on Shavell (2013). Suppose that the owners of restaurants can install sprinklers there in order to prevent the risk of fire. Imagine that a sprinkler system would reduce such risk from 4% to 1%, that a fire would cause harm of $100,000, and that a sprinkler system would cost $2,000.

The goal of the legal system is to minimize social costs and, in particular, to induce a precaution whenever the benefits it would yield—equal to the reduction in expected harm—would outweigh its social costs—equal to the costs of the precaution itself and the costs of law enforcement. It is obvious, therefore, that the abovementioned sprinkler systems should be installed insofar as enforcement costs do not exceed $1,000 per restaurant.

Assume that the State can choose one of three following instruments in order to induce such precaution. In all three cases, public agents verify whether sprinklers systems have been installed, and impose a fine of $100,000 to the restaurant owners if they have not. The difference lies in the fact triggering verification and the stage at which this is carried out.

In the first case, verification takes places before a fire has occurred and independently from whether those agents have received or not information suggesting that the probability of the sprinklers not having been installed is especially high. Assume that every restaurant is checked.

In the second case, verification is carried out before an accident has occurred and as a consequence of law enforcers being conveyed the information (e.g. by means of a report filed by a third party) that the probability of the affected owners not having taken due care is higher than normal. Imagine that in 4% of the times those devices have not been installed at least one of the customers attending the restaurant reports such an omission. Imagine that in 0.5% of the times the required precaution has been taken one customer, intentionally or not, falsely reports that omission.

In the third case, verification is carried out after the occurrence of an accident. Let us suppose that victims always sue injurers before the relevant court.
It is assumed that all three types of verification are free of error, so that State agents always determine perfectly whether required precaution was taken or not, and the enforcement cost per instance of verification is exactly the same for all three types, $500, and borne by the State.

One can see that inducing the precaution at issue is worthwhile in all three scenarios, as its benefits ($3,000 = 4\% \cdot $100,000 – 1\% \cdot $100,000) always exceed its social costs ($2,000 + enforcement costs).

Nonetheless, it must be pointed out that the enforcement costs are lowest in the second scenario. Indeed, in the first one, where everyone is checked, such costs are $500 per restaurant. In the second case, as every owner installs the sprinklers and only one out of 200 restaurants are checked (as a result of a false report), those costs are just $2.5 per restaurant. In the last scenario, where one out of 100 restaurants are examined (as a result of an accident), enforcement costs are $5 per restaurant.

III. A FORMAL MODEL

1. Investigation versus monitoring

Let us assume that individuals, who are supposed to be identical and risk neutral, can take care in order to lower the risk of causing a harmful accident. Let $x$ be the amount and the cost of care, where $x \geq 0$; $p(x)$ be the probability of an accident, where $0 < p(x) < 1$, $p'(x) < 0$, and $p''(x) > 0$; $h$ be the magnitude of the harm eventually caused, where $h > 0$; and $a$ be the assets of each individual, where $a > h$.

The State may require a certain level of care and control it by using two different means. On the one hand, it can perform checks indiscriminately, without targeting certain instances where the likelihood of noncompliance appears to be particularly high. I will call this method of enforcement monitoring. Compliance may be checked either with certainty (i.e. it is verified whether each and every one has abide by the law) or randomly (i.e. not everybody is actually subject to scrutiny, but only some people). Let $x_a$ equal that required care, where $x_a \geq 0$; $r_a$ equal the probability of the State actually verifying whether the required standard of care has been met, where $0 < r_a \leq 1$; and $s$ equal the sanction if $x < x_a$, where $0 < s \leq a$.

On the other hand, the State can carry out controls targeting –either some or all of– those cases where the relevant authorities are conveyed information suggesting the likelihood of due care having been omitted is higher than normal, that is, the conditional probability of noncompliance in one case, given that information, is greater than the unconditional probability of such event. Let us imagine, for instance, that such an offence has been reported to a regulatory agency by a whistleblower, or that an accident has occurred and the victim has subsequently sued the injurer before a court. Both events constitute signals of noncompliance. It must be underlined that the likelihood of the relevant authority receiving such signal is lower when the potential injurer has complied with the law than when he or she has not. I will call this method of enforcement investigation. Let $x_p$ equal the required care, where $x_p \geq 0$. Such $x_p$ might be higher or lower than that established for indiscriminate controls, $x_a$. I denote by $q(x)$ the probability of the relevant authority receiving such information given the level of care taken by individuals, where $0 < q(x) < 1$, $q'(x) < 0$, and $q''(x) > 0$. One can thus see that if compliance is controlled by means of liability for harm –under negligence rule– and victims always sue injurers, then $p(x) = q(x)$. 
Here, too, compliance can be verified either with certainty –i.e. it is checked in every single case where the State has observed a signal of noncompliance– or probabilistically –i.e. it is randomly examined only in some of those cases–. Let $r_p$ equal the probability of the State actually verifying whether the required standard of care has been met, given that it has received such a signal, where $0 < r_p \leq 1$.

With the aim of comparing both instruments of enforcement on a level playing field, I assume that the sanction for noncompliance is the same for both of them. Therefore, the magnitude of the fine to be imposed if $x < x_p$ is $s$ as well, where $s < 0$, and $a \geq s$.

In order to verify whether the required level of care has been reached, the State has to incur some enforcement cost per instance, $k > 0$, which is also assumed to be the same for every type of control. I also suppose that the standards of care prescribed by the State, $x_a$ and $x_p$, are always met.

If controls are carried out indiscriminately –i.e. they are not triggered by signals of noncompliance–, social costs will therefore be:

$$C_a = x_a + p(x_a)h + r_a k$$

For the people to be induced to take due care, $x_a$, the expected sanction for noncompliance must be at least the cost of that care, $r_as \geq x_a$. In order to minimize social costs, the State will impose the maximum available sanction, which equals the assets of individuals, $a$, and choose the minimum probability of random enforcement that leads them to meet the prescribed standard of care, which is $r_a^* = \frac{x_a}{a}$. Let us also suppose that the State prescribes the level of care that minimizes such social costs, which it is denoted by $x_a^*$. Social costs are thus

$$C_a = x_a^* + p(x_a^*)h + \frac{x_a}{a} k$$

As [2] is convex, optimal care will equal zero, $x_a^* = 0$, and, therefore, the State will also choose a zero probability of enforcement, $r = 0$ if and only if the following condition is met:

$$1 + p'(0)h + \frac{k}{a} \geq 0$$

That is, the standard will not be enforced if enforcement costs, $k$, reach some critical point $k^*$ so that

$$k \geq -a[1 + p'(0)h] = k^* > 0$$

If this is not the case and, therefore, $k < k^*$, then the optimal level of care $x_a^*$ will be determined by the first-order condition.
If controls are triggered by a signal of noncompliance, social costs are:

\[ C_p = x_p + p(x_p)h + q(x_p)r_pk \]  \[ \text{[6]} \]

For individuals to be induced to take due care, \( x_p \), the expected sanction for noncompliance must be equal or greater than the cost of that care, \( q(0)r_pa \geq x_p \). As in the previous case, the State will impose the maximum available sanction, \( a \), and choose the minimum probability of enforcement that leads individuals to meet that standard, which is \( r_p^* = \frac{x_p}{q(0)a} \). Social costs are thus

\[ C_p = x_p + p(x_p)h + q(x_p)\frac{x_p}{q(0)a}k \]  \[ \text{[7]} \]

If we evaluate this expression at the level of care that minimizes social costs under a system of indiscriminate controls, that is if \( x_p = x_a^* \), then we will have:

\[ C_p = x_a^* + p(x_a^*)h + q(x_a^*)\frac{x_a^*}{q(0)a}k \]  \[ \text{[8]} \]

By comparing expressions [2] and [8], and given that \( x_a^* > 0, q'(x) < 0 \) and, therefore, \( q(x_a^*) < q(0) \), one can see that \( C_a(x_a^*) > C_p(x_a^*) \). In other words, the same level of care that would be optimal under a system of monitoring –i.e. non-triggered-by-a-signal-of-noncompliance controls– can be induced with less social costs through investigation –i.e. triggered-by-a-signal-of-noncompliance checks–. And it must be pointed out that social costs might still be reduced even more under the latter system if the standard of care is elevated up to the point \( x_p^* \), determined by the first-order condition

\[ 1 + p'(x_p^*)h + q'(x_p^*)\frac{x_p^*}{q(0)a}k + q(x_p^*)\frac{k}{q(0)a} = 0 \]  \[ \text{[9]} \]

2. **Ex-ante investigation versus ex-post investigation**

Let us suppose enforcement costs, which are still denoted by \( k \), are the same for both non-triggered-by-harm-but-by-another-signal-of-noncompliance controls –which I will call ex-ante investigation– versus triggered-by-harm controls –ex-post investigation–.

In the former case, compliance is verified in –either some or all of– those instances where the relevant authorities are conveyed some information suggesting the likelihood of due care having been omitted is higher than normal, although no harmful accident has occurred yet. Let us imagine, for instance, that such an infringement has
been reported to a regulatory agency by a third party. Let $x_r$ equal required care, where $x_r \geq 0$. Let $f(x)$ be the probability of the relevant authority receiving such information given the level of care taken by individuals, where $0 < f(x) < 1$, $f'(x) < 0$, and $f'''(x) > 0$; and $r_r$ be the probability of such authority actually verifying whether required care has been taken, where $0 < r_r \leq 1$.

In order to induce due care, the expected sanction for noncompliance must be equal or greater than the cost of abiding by the law, $f(0)r_r a \geq x_r$. Let us suppose the State imposes the maximum available sanction, $a$, and chooses the minimum probability of enforcement that leads individuals to take such care, which is $r_r^* = \frac{x_r}{f(0)a}$.

Social costs are thus

$$C_r = x_r + p(x_r) h + f(x_r) \frac{x_r}{f(0)a} k$$  \[10\]

Let us assume there is a level of required care $x_r^*$, where $x_r^* > 0$, that minimizes such costs.

When it comes to harm-based controls, compliance is examined, either with certainty or probabilistically, whenever the State is conveyed the information that an accident has happened. Let $x_h$ be required care, where $x_h \geq 0$; $g(x)$ be the probability of the State being informed on the occurrence of such an event, given the level of care taken by injurers. This probability might well equal that of an accident taking place, $p(x)$, but it is not necessary to do so. The likelihood of state agents verifying whether injurers took due care, once that signal has been received, is denoted by $r_h$, where $0 < r_h \leq 1$. In order to induce required care, the State imposes the maximum available sanction, $a$, and chooses the minimum probability of enforcement that leads individuals to comply with the law, which is $r_h^* = \frac{x_h}{g(0)a}$. Social costs are here

$$C_h = x_h + p(x_h) h + g(x_h) \frac{x_h}{g(0)a} k$$  \[11\]

Let $x_h^*$, where $x_h^* > 0$, be the level of required care that minimizes those costs.

By comparing equations [10] and [11] one can easily see that neither of both types of controls has an intrinsic enforcement cost advantage over the other. It all depends on the ratios between the conditional probabilities of the relevant authorities being conveyed the signal triggering verification of required care, given that such care has actually been taken –$g(x_h^*)$ and $f(x_r^*)$ respectively– and the conditional probabilities of those authorities receiving the referred signal given that due care has been omitted – i.e. $g(0)$ and $f(0)$ respectively–. Ex-ante investigation can induce the same level of required care, $x$, with less enforcement costs than ex-post investigation when:

$$\frac{f(x)}{f(0)} < \frac{g(x)}{g(0)}$$  \[12\]

On the contrary, ex-post investigation have ceteris paribus an enforcement cost advantage over ex-ante investigation with that regard when:
\[
\frac{f(x)}{f(0)} > \frac{g(x)}{g(0)}
\]  

[13]

It must also be noted that triggered-by-signals-of-noncompliance controls will be more costly, ceteris paribus: (i) the higher the likelihood of let us say “false positives” –i.e. \(g(x^*_n)\) or \(f(x^*_r)\)– is; and (ii) the lower the likelihood of “true positives” –i.e. \(g(0)\) or \(f(0)\)– is.

The results are quite intuitive. On the one hand, false positives force the State to “unnecessarily” incur the costs of investigation, \(k\), with some probability, \(x^*_h\frac{g(0)}{g(0)a}\) or \(x^*_r\frac{f(0)}{f(0)a}\), in cases where individuals have abided by the law. A reduction in the number of false positives, other things being equal, will therefore lower the social costs of inducing legal compliance. On the other hand, if the likelihood of the relevant authorities receiving a true signal of noncompliance, \(g(0)\) or \(f(0)\), increases, the State may reduce the probability of investigation and, therefore, law enforcement costs without affecting deterrence.

IV. DISCUSSION AND POLICY IMPLICATIONS

It must be pointed out that harm-based –i.e. ex post– controls do not have an intrinsic cost advantage over non-harm-based –i.e. ex ante– ones. The crucial factor with that regard is not the timing of the legal intervention, but whether this depends on a signal of noncompliance or not. Verification of compliance triggered by such signals is ceteris paribus less costly than verification carried out independently from them.

There is no perfect correlation between both dichotomies. Ex ante controls are typically carried out independently from any signal of noncompliance, but they could also be precipitated by such a signal. In fact, sometimes they are. An investigation conducted as a result of a whistleblower reporting some act-based crime might well illustrate this point. On the other hand, ex post controls usually take place in those cases where the relevant public agents are conveyed information suggesting the probability of an offence is higher than normal –e.g. when the victims of an accident sue the alleged injurers–, but they could well be performed indiscriminately, without specifically targeting such cases –let us think of routine medical check-ups performed with the aim of detecting certain occupational or environmental diseases–.

1. Advantages of ex-ante investigation

In this paper we want to draw attention to ex-ante investigation, thereby meaning controls triggered by a signal of noncompliance different from harm. First because they are relatively common. Second because they were missing in the previous analysis of Shavell (2013) as well as not explicitly considered by Mookherjee and PNG (1992). Third, and most importantly, because they combine some of the best features of two worlds.

Indeed, ex-ante investigation has the abovementioned fundamental cost enforcement advantage, which is typically –although not necessarily– associated to harm-triggered interventions.

As it can be applied before the causation of harm, it could also have the main pros of preventive enforcement. Ex-ante investigation might thwart the successful commission of harmful offences and preclude potential offenders from deriving private benefits from crime, thereby producing more intense deterrent effects than ex post
checks. Moreover, since investigation is not applied indiscriminately, but targeting those cases where the likelihood of socially undesirable behavior is relatively high instead, it is less costly than pure preventive enforcement in terms of precluding efficient offences.

Insofar as ex-ante investigation is usually carried out after the risky activity at issue has been launched, it is less prone to suffering a typical disorder of public interventions taking place at an earlier stage—e.g. administrative authorizations—: the excess volume.

Some empirical studies show that countries with heavier regulation of entry have higher corruption and larger unofficial economies, but not better quality of public or private goods, whereas countries with more democratic and limited governments have lighter regulation of entry. That evidence supports the view that politicians and bureaucrats benefit from regulation of entry. On the one hand, this regulation can keep out the competitors and raise incumbents’ profits at the expense of consumers. Industry incumbents are able to get those regulations as they usually face lower informational and organizational costs than consumers. On the other hand, public agents can extract some of those profits through contributions or bribes received in return for either providing access to the market or keeping potential entrants away from it. That might explain why the volume of entry controls tends to exceed the one that would be optimal for society as a whole.

Politicians and bureaucrats could also use other type of controls different from entry ones for the purpose of preventing potential competitors from entering the market, raising the profits of incumbents and obtaining some contributions from them in exchange. Public agents might indeed threaten to carry out those controls in order to deter entry of new players and extract rents from the incumbents.

It must be noted, however, that entry controls have an important advantage over other type of public interventions for the purpose of advancing the interests of bureaucrats. When it comes to preventing entry, they are typically less costly. Let us take as an example the entry controls par excellence, namely authorization schemes where newcomers cannot enter the market unless they apply for and are given a permit. Here, access is overall banned, although such ban can be lifted on a case-by-case basis, if the relevant administrative body verifies that the applicant meets the legal requirements to perform the corresponding activity and subsequently grants the authorization. In this scenario, bureaucrats do not need to make any effort in order to block applicants from entering the market. They can just do nothing. Inaction is enough. In particular, they do not need to verify whether applicants fulfill the legal requirements to be granted the permits. By contrast, to drive competitors out of the market by means of posterior controls, bureaucrats will probably need to make some effort and incur some costs, especially if those controls are unlawfully carried out with that purpose and the affected competitors have already entered the market. When performed lawfully and by benevolent public agents, entry controls are more resource-intensive than ulterior interventions, but the opposite is paradoxically true when carried out unlawfully and by self-interested bureaucrats who try to deter entry in order to create and extract rents.

Moreover, since entry controls, when properly carried out, tend to be more costly and require more resources than interventions taking place at later stages, bureaucrats

---

3 DJANKOV, LA PORTA, LÓPEZ-DE-SILANES and SHLEIFER (2002); DJANKOV (2009).
4 See, for instance, SCHAUMANS and VERBOVEN (2008); THORNTON and TIMMONS (2013).
might well promote the former ones in order to maximize their budget and influence\(^5\), regardless of whether they actually carry out those controls or end up slacking.

Finally, triggered-by-a-signal-of-noncompliance controls, performed after the start of the activity at issue, allow assessing both the risks posed by that activity and the corresponding optimal safety measures more accurately than entry controls. Indeed, the later the verification of compliance takes place, the more and better information can be collected in order to assess both those risks and measures properly. The information obtained at the earliest stage, before the risky activity is actually being performed, is particularly poor. On the one hand, because there could be no empirical evidence on the real consequences of that activity yet. On the other hand, because the relevant circumstances can change after that information has been collected. Let us consider, for example, the marketing of new medicines. Despite rigorous studies and trials carried out before the approval of a new medicinal product, relatively little is known about its benefits and risks until the medicine is marketed and widely used, because the circumstances of clinical practice differ notably from those of premarketing testing. Indeed, clinical trials often involve a limited number of patients, which usually precludes identifying rare—but sometimes serious—adverse effects. The duration of clinical trials is often much shorter than the expected length of treatment in clinical practice. Adverse drug reactions appearing after exposure for prolonged periods or only long after exposure are, consequently, improbable to be detected before being put on the market. Clinical trials usually involve relatively healthy patients, who may not be entirely representative of the future target population, i.e. the medicine’s consumers. The treatment of patients involved in clinical trials is usually limited to only one drug. It is therefore unlikely to identify adverse drug-drug interactions before marketing. In short, the full safety (and efficacy) profile of medicinal products can be known only after they have been placed on the market.

1. **Disadvantages of ex-ante investigation**

The main disadvantage of triggered-by-a-signal-of-noncompliance controls and, therefore, of ex-ante investigation lies in the fact that the ratio between false positives and false negatives, \( \frac{f(x^*)}{f(0)} \), can be too high. It could happen, firstly, if the likelihood of the relevant public agent receiving a true signal of noncompliance is relatively low. This might make necessary either to impose very serious sanctions in order to deter noncompliance, which in turn may be unfeasible or increase the probability of offenders being judgment-proof, or increase the probability of verification and, therefore, enforcement costs. Secondly, the likelihood of State agents being conveyed a true signal of noncompliance could be relatively high, which increases enforcement costs as well.

One might consider several policies aimed at facing this problem. Firstly, the government can increase the likelihood of law enforcers being conveyed true signals of noncompliance, for instance by establishing big data systems, imposing sanctions on the people who do not report certain offences, rewarding whistleblowers, protecting them against retaliation or facilitating –i.e. decreasing the private costs of—such reporting in some other way\(^6\), etc.

\(^5\) See NISKANEN (1971).

\(^6\) See, for instance, FELDMAN and LOBEL (2010).
Secondly, with the aim of lowering administrative costs, the State may also reduce the likelihood of those public agents receiving false signals of noncompliance, for instance by sanctioning individuals who deliberately file false reports.

Thirdly, the State can make verification dependent on a different signal of noncompliance with a better –i.e. lower– ratio between likelihood of false positives and likelihood of true positives. For instance, instead of fully investigating every report of noncompliance, State agents can conduct a thorough inquiry only on those reports supported by stronger evidence.

Fourth, the government can combine ex ante investigation with other modes of law enforcement, such as entry controls, liability for harm, etc.7

In fact, all of those four referred policies are often implemented at the same time, in one way or another.

V. POSSIBLE EXTENSIONS

The formal model presented above could be substantially modified and extended in order to discuss under which circumstances, how and to what extent the aforementioned strategies –increasing the probability of the relevant authorities receiving true signals-of-noncompliance, decreasing the likelihood of those public agents obtaining false signals-of-noncompliance, making investigation dependent on

7 There is already a considerable law and economics literature discussing under which circumstances ex ante safety regulation and ex post liability for harm should jointly employed. See SHAVER (1984a, 1984b, 1993 and 2013). KOLSTAD, ULEN and JOHNSON (1990) consider that regulatory standards could inform potential injurers on the minimum due care when there is uncertainty about the level of care required by courts. BURROWS (1999) shows that the effects of the joint use of regulation and liability under uncertainty are context specific. SCHMITZ (2000) shows that joint use of liability and regulation might be optimal if wealth varies among injurers. GAREAPA and GÓMEZ-POMAR (2004) show that agency costs when delegating law enforcement, legal errors and collusion between regulatory agents and offenders might, under certain conditions, justify the joint use of criminal sanctions and regulatory sanctions. HIRIART, MARTIMORT and POUYET (2004) assume that the regulator is ex ante uninformed about the harm level, which is verifiable ex post by the judge, and show that socially optimal behavior can be induced if the regulator may contract ex ante with potential injurers. Then, the regulator sets ex ante the optimal standard, while ex post liability provides incentives to potential injurers for revealing information. HIRIART, MARTIMORT and POUYET (2008) show that a certain mix of both types of controls could be optimal if it is difficult to verify ex post whether injurers have taken due care and also to determine ex ante the value of their assets for liability payments. ROUILLO (2008) shows that the joint use of liability and regulation is optimal when harm varies in the population and the probabilities of injurers escaping liability (namely not being sued) are also variable and not observable by courts. DE GEEST and DARI-MATTIACCI (2007) argue that, under certain conditions, regulatory standards can induce potential injurers to take higher, socially optimal levels of care, if the use of the liability system alone is undermined by judgment proof or disappearing injurers. BHOLE and WAGNER (2008) analyze the joint use of ex ante regulation and strict liability in cases where care is multidimensional, some of which is unobservable and therefore not being amenable to regulation, and there is some likelihood of accidents not being detected or injurers not being identified. LENNTORP (2009) analyzes the joint use of licensing and strict liability, taking into account the rules (either British or US ones) governing the allocation of litigation costs and paying special attention to the question of whether a compliance defense should be granted or not. HIRIART, MARTIMORT and POUYET (2010) show that combining ex ante and ex post intervention and splitting the power to intervene between different public bodies reduces the risk of capture. BENTATA (2013 and 2014) presents empirical evidence suggesting that civil liability complements regulation by both focusing on aspects of care that cannot be observed ex ante and reducing the risk of regulators being captured by injurers.
signals supported by stronger evidence, or combining different types of controls—should be implemented.

For instance, Mookherjee and PNG (1992) present a model where –ex post– monitoring and investigation are eventually combined, and the signal triggering investigation provides information about the severity of the offence. In this model: (i) fines enforced by investigation should be maximal; (ii) the optimal enforcement policy involves monitoring only if reporting is too low or the costs of investigation are sufficiently high; and (iii) the least severe offences are monitored whereas the most severe are investigated.

This model certainly provides very valuable insights. Nevertheless, it could be argued that monitoring and investigation are usually combined in a different way. Both methods of law enforcement are often employed not simultaneously, but successively, in a kind of multistage process, with respect to the same type of offences. In the first stage, authorities monitor the activities at issue. Given that such monitoring is not thorough, but rather relatively cursory, it will yield a noisy signal of noncompliance in each case. Then, if the evidence suggesting noncompliance exceeds some threshold, the case will be investigated and a more profound verification, requiring more resources per instance, will be carried out. If it does not, the case will be closed with no further action.

VI. CONCLUSION

This paper adds to the existing literature on the timing of law enforcement by making basically four points.

First, contrary to what Shavell (2013) has argued, it shows that ex post, triggered-by-harm instruments of inducing legal compliance do not have an intrinsic enforcement cost advantage over ex ante, non-triggered-by-harm instruments.

Second, it shows that the crucial distinction in that regard is whether verification of compliance depends or not on the relevant authority receiving information indicating the probability of an infringement in some case is higher than normal —i.e. the conditional probability of noncompliance, given that information, is greater than the unconditional probability of such event—, and also whether potential injurers, by complying with the law, can lower the likelihood of being conveyed such information. Instruments of law enforcement triggered by such a signal of noncompliance are cheaper than instruments not triggered by it.

Third, the paper shows that ex post, harm-based, interventions can be costlier than interventions dependent on noncompliance signals different from the occurrence of harm, and vice versa. Which kind of instrument of law enforcement has a cost advantage over the other depends on: (i) their respective conditional probabilities of the relevant authority receiving a signal of noncompliance given that compliance has actually taken place, i.e. likelihood of false positives; and (ii) their respective conditional probabilities of that authority being conveyed a signal of non compliance given that the law has been breached, i.e. likelihood of true positives. The higher the likelihood of false positives and the lower the likelihood of true positives, the costlier legal enforcement.

Fourth, the paper draws attention to interventions triggered by a signal of noncompliance different from the occurrence of harm, mostly because they combine some of the best features of other law enforcement instruments —e.g. the

---

8 See, mutatis mutandis, Heyes (2002); Kaplow (2013).
abovementioned cost advantage and their pure prevention effects—, while avoiding some of their most serious disadvantages.

REFERENCES


