LAW ENFORCEMENT, INFRINGERS HETEROGENEITY AND OPTIMAL LEGAL STANDARDS.

Juan José Ganuza*and Fernando Gomez†
Universitat Pompeu Fabra, Barcelona
Ramon Trias Fargas 25-27
08005 Barcelona, Spain

APRIL 2016.‡

ABSTRACT

Optimal law enforcement policy should try to maximize compliance levels with legal requirements taking as given the costs of such compliance for the target population. Heterogeneity in the abilities to comply with the law, and the resulting costs, should be given proper weight, as people and firms differ widely on such variables. We derive results for the optimal setting of standards of compliance and enforcement effort at the individual level and at the level of a population of potential infringers. We also explore, in settings in which several regulators/enforcers are in place, how to match different groups of potential infringers to regulators. These findings show some additional inadequacies of private enforcement of law that had not been previously identified in the literature dealing with the debate of public vs. private enforcement: the close link between setting standards of compliance and enforcement, and increased misalignment of the choices of private enforcers with socially optimal policies.

KEYWORDS: Law Enforcement, Optimal Legal Standards and Incentives.

JEL classification numbers: C72, D82.
In the functioning of societies and legal systems, law enforcement covers a wide range of institutions, agents, decisions and actions. Courts, judges and juries, prosecutors, police officers, regulatory agencies and their employees, and also, in certain areas, private agents, deal with various dimensions (sometimes distinct, sometimes overlapping) of enforcing legal rules in nearly all areas of activity. The actions of all those institutions and agents, the decisions they adopt and the policies they design and/or pursue compose the bulk of law enforcement in a given legal system.

The firms, organizations and individuals who are subject to this complex law enforcement apparatus come in many forms and kinds. Some potential infringers are individuals, others are organizations. Within each category, their respective ability and cost to comply with the requirements of the law who may be enforced upon them varies greatly. The same is true for their wealth or ability to face monetary penalties for law infringement, and the benefits they may derive from breaking the law or departing from the legally mandated behavior.

In sum, the agents whose behavior the law enforcement apparatus intends to influence are heterogeneous. Heterogeneity of potential infringers has been analyzed in the law enforcement literature only in certain contexts. First, the standard model of law enforcement (Polinsky and Shavell 2007) considers a probability distribution of gains of the population of potential law-breakers. Second, heterogeneous wealth levels and their impact on optimal fines has also been analyzed (Polinsky and Shavell 1991; Garoupa 2000).

The issue of agents’ differential costs of complying with the requirements posed by the law has not been consistently analyzed in the law enforcement setting. Perhaps this is due to the natural tendency to consider, at least implicitly, that law enforcement focuses on a set of actions for which the direct cost to the agent of complying with the legally required behavior is zero, often even negative. Arguably, for many of the everyday crimes, the direct cost –not the foregone gain, naturally- of observing the legally required behavior, were it to be deemed positive, would
be, at most, trivial, since the direct costs would simply consist of refraining to carry out the action that constitutes the criminal offence. It is true though, that the standard economic model of law enforcement explicitly contemplates that gains from breaking the law may come as utility increases or as savings in the costs of law abiding behavior.¹

The former assumption (no direct costs of abiding with legal mandates in the law enforcement settings), is not realistic for most conduct subject to law enforcement concerns. Activities required to comply with environmental laws, industrial and product safety rules, market disclosure regulations, antitrust regulation, consumer protection legislation, but also many activities in which individuals are frequently involved, such as construction, and even everyday activities such as vehicle driving or house repairing, do offer settings in which satisfying the legal mandates as to behavior or outcomes entails substantial direct costs for agents. Moreover, in these settings law enforcement typically deters behavior that falls short of certain standards of conduct legally defined, and does not operate as a quasi-automatic response to a certain observed action or outcome. In other words, the law enforcement apparatus very often enforces a fault-based liability regime and not a strict liability one.²

This is the setting in which we explore the relevance of the heterogeneity of potential infringers for the definition of standards of behavior and for designing law enforcement policy. This set of issues has been to a significant extent overlooked in the law enforcement context, although it has been deemed relevant, and has led to a significant amount of economically oriented literature in other areas, especially in that of Tort. Heterogeneity of potential victims has been explored in a series of papers: Ganuza and Gomez, 2004; Friehe, 2007, 2008, 2009. Heterogeneity of potential injurers has given rise to even a more extensive literature, considering both the differential levels of assets and the differential costs of care among injurers: Schwartz, 1989; Arlen, 1992; Ganuza and Gomez, 2008, 2011; Ben-Shahar and Porat, 2016 (forthcoming).

When heterogeneity among potential infringers is taken seriously, the separation in the Law

¹See Polinsky and Shavell 2007.
²This difference is acknowledged in the standard economic models of law enforcement: Polinsky and Shavell 2007.
between goal-setting – defining the required behavior – and enforcement – detecting and sanctioning those who depart from the required behavior becomes blurred and intermingled if one is interested in inducing socially desirable behavior from the target population. Severing standard setting from enforcement allows only partial understanding of relevant law enforcement policy questions.

In many areas of law enforcement, both sides to the process are explicitly present: some type of legal standard or goal governing an underlying behavior, and some kind of negative (sanction) incentive tool linked to the outcome of the process of actually failing to reach the relevant standard. Sanctions vary widely, and go from the explicit to the implicit, from the monetary to the non-monetary. Obviously, purely intrinsic (guilt) motivators may also be present.

In Law and Economics, in addition to some of the ones cited above in the Tort context, relevant contributions surely exist dealing with setting certain behavioral standards (the literature on the Hand rule and other formulas for setting due care is vast\(^3\)) and on the general structure of the incentive system embedded in legal rules (Shavell, 2004). There is also literature on the relative advantages of having positive or negative legal incentives (rewards or sanctions) to enforce legal standards (Dari Mattiacci and De Geest, 2010; De Geest and Dari-Mattiacci, 2013).

In this paper we intend to present a basic economic model of law enforcement with potential infringers who vary in their ability and cost of compliance with legally required behavior. Using a simple principal-agent framework, we are able to characterize the level of standards and law enforcement effort, given the costs of complying with the law, and the size of the explicit negative incentive (the sanction). We show that, in this setting, it may happen that by lowering the standard of behavior from a level that was too ambitious for the ability to comply of the infringer, one may actually observe an increase in the level of legal compliance with law chosen by the potential infringers. Counterintuitively, being softer on those who have higher costs of abiding with the law pays in terms of incentives.

We then extend this basic framework to settings of enforcing legal standards in groups and populations that are not homogeneous in ability and costs of compliance with the law, or equiv-

\(^3\)We have contributed to that literature in a paper that anticipates in a more limited scenario some of the findings of this paper: Ganuza and Gomez (2008).
ently, in settings of asymmetric information concerning such variables. In our basic setting we robustly find that legal standards should optimally be decreasing in the individual costs of compliance and increasing in the size of the sanction and the enforcement expenditure. For large populations subject to uniform legal standards (as is commonly the case in law enforcement, since informational and equality reasons very often prevent individually tailored requirements and/or sanctions) we characterize the conditions for the optimal policies in terms of the distributions of compliance costs in the relevant population.

The above results have some bearing on a classical topic in the law enforcement literature, namely that of public vs. private enforcement of the Law. Since the pioneering piece by Becker and Stigler (Becker and Stigler, 1974), the relative efficiency advantages of the allocation of law enforcement powers to public and private institutions has provided a fruitful caput controversum for providing insights into the theory of law enforcement. Many dimensions of the problem have been brought to the debate: the level of enforcement effort when private enforcers are to gain what sanctioned infringers are forced to pay\(^4\); economies of scale in law enforcement\(^5\); collusion between infringers and enforcers\(^6\); setting compensation for enforcers\(^7\); monopolistic vs. competitive enforcement\(^8\); exercise of discretion in the enforcement effort\(^9\); the goals of public enforcers, whether aligned or misaligned with the maximization of social welfare\(^10\).

When heterogeneity of offenders is added to the picture, and leaving aside the ability to pay the sanction (assuming it is monetary and the wealth constraint is not binding), our results suggest some additional disadvantages of law enforcement entrusted to private organizations.

First, the close link between the level of the standard of legal compliance and the law enforcement policy makes it generally less desirable to allocate both tasks to separate institutions. This provides a theoretical justification\(^11\) for the concentration of regulatory powers—in our setting

\(^5\)See, Polinsky 1980.
\(^7\)See, Landes and Posner 1975; Polinsky 1980; Friedman, 1999.
\(^8\)See, Landes and Posner 1975; Friedman 1984; Garoupa and Klerman 2002.
\(^11\)Obviously this justification is not of universal application, or unqualified. We believe, however, that it is fairly general for settings of fault-based liability, where the costs of compliance with legal requirements is a relevant factor.
this means standard-setting- and enforcement powers –detection and sanctioning of infringers- to
the same agency or institution. This makes delegating the task of identifying and punishing to
a private enforcer separate from the lawmaker or policymaker –the institution setting the goals
to be satisfied by the potential infringers- less desirable than otherwise. Second, if goal-setting is
also entrusted to a private institution, standards and enforcement will be tougher on the poten-
tial infringers who face relatively higher costs of legal compliance, just the reverse of the optimal
policy. When the standard is to be set for a population of potential infringers, a profit-motivated
enforcer will set population standards that are not based on the distribution of compliance costs,
but on the distribution of income, in order to maximize revenue.

Our model also allows us to derive implications concerning the allocation of populations of
potential infringers to different law enforcers (if for some reason, such as diseconomies of scale, or
returns to specialization, it is desirable to have more than one). In some settings, an enforcer may
face competing enforcers, each imposing its own standard of behavior and perhaps its own level
of sanction. It may make sense in such a scenario to allocate each enforcer to a given sub-group of
the population of infringers, and thus, to potentially divergent standards of compliance with the
law. As to grouping of agents, we show that perfect matching of infringers, who vary in their costs
of satisfying legal requirements, to regulator/enforcer is not the optimal policy from the point of
view of maximizing compliance with the law.

Our focus in the paper is the provision of incentives, so we disregard other relevant dimensions
of setting both standards of legal compliance and enforcement effort, such as, inter alia, their
informational and expressive properties (as explored by McAdams, 2015). Similarly, we do not
consider the level of information that individuals may have of factors relevant to social welfare
compared to that of the legal policymaker, nor the existence of positive or negative compliance
externalities, dimensions explored in Shavell, 2012. It is also important to bear in mind that we
are not addressing the specifics of any given area of the law (crime, environmental law, antitrust,
etc.), and thus, many of the sector-specific factors and circumstances are naturally ignored in an abstract approach such as the one we attempt here.

The paper is organized as follows: Section 2 presents the basic model for individual infringers. Section 3 extends the model to heterogeneous populations. Section 4 deals with optimal grouping of agents for inducing effort in satisfying legal requirements. Section 5 presents some qualifiers, and additional implications and potential extensions.

2 Setting Policies at the Individual Level

We study a standard principal-agent environment where sanctions for not reaching certain compliance levels over a certain relevant variable are set and enforced. We start by assuming that the principal (the policy maker or law maker) wants to maximize the agent’s compliance with the law, $x$. In section 5 we will consider enforcement costs and we will define a total welfare function, and revise the analysis accordingly. Let $C(x, \theta)$ be the agent’s cost of legal compliance effort, where $C(x, \theta)$ is increasing and convex in $x$. The parameter $\theta$ is the ability of the agents to comply with the law.\footnote{This ability to comply with the law captures, in the case of firms, their technology, organization and the like, factors that clearly affect the cost of legal compliance (think of enviromental or safety regulation and even market disclosure). In the case of individuals, it captures the features or qualities that influence the effort or expenses involved in adhering to a given level of legally required behavior.} The larger the ability, the lower the cost and the marginal cost of legal compliance effort, i.e $C_2(x, \theta) < 0$ and $C_{12}(x, \theta) < 0$. The agent’s incentives to try to comply with the law depend on a standard set by the principal, $\pi$. If the level of compliance achieved by the agent is equal or larger than the standard, the agent is not sanctioned. If the agent fails to meet the required standard of compliance, she receives the sanction $D$. The agent has also an intrinsic motivation to comply with the law, $g(x) \geq 0$ with $g(x)$ being weakly increasing and concave (we do not exclude that $g(x) = 0$). The agent’s preferences are:

$$
\pi(x, \pi, \theta) = \begin{cases} 
  g(\pi) - C(x, \theta) & \text{if } x \geq \pi \\
  g(x) - p(e)D - C(x, \theta) & \text{otherwise.}
\end{cases}
$$

Where $p(e)$ is the probability of detecting the lack of compliance and $e$ is the enforcement policy. Later we will consider enforcement costs, but we start the analysis taking $e$ as exogenous.
Given this incentive scene,\textsuperscript{13} the equilibrium effort exerted by the agent is characterized by the following expression:

$$x^*(\theta, \bar{x}) \in \arg\max\{\pi(x, \bar{x}, \theta)\}$$

**Lemma 1** *The equilibrium level of compliance is increasing in $\theta$ and non monotonic in $\bar{x}$.*

The intuition of the first part is immediate: given the same incentives, agents who face lower cost of complying with the law exert more effort and reach higher levels of compliance. The second part of the Lemma is due to the fact that if the standard of compliance is too high, the agent will give up and abandon the provision of all effort in complying. Figure 1 below characterizes equilibrium compliance as a function of the standard.

Let $x_I^*$ be the equilibrium effort in compliance of the agent when she decides not to attain the legal standard, and thus her effort is only driven by intrinsic motivation (or by extrinsic motivations other than the law, if they are in place), $x_I^*(\theta) \in \arg\max\{g(x) - p(e)D - C(x, \theta)\}$. The equilibrium effort will be equal to the standard until the net utility of satisfying the standard, $g(\bar{x}) - C(\bar{x}, \theta)$, is equal to the net utility of exerting $x_I^*(\theta)$.

From the above characterization, it is immediate that the optimal standard which maximizes the exerted legal compliance effort is interior, and in particular is implicitly defined by the following expression.

$$g(\bar{x}^*) - C(\bar{x}^*, \theta) = g(x_I^*) - p(e)D - C(x_I^*, \theta)$$

Notice that $\bar{x}^* > x_I^*$. The next figure characterizes the optimal standard.

\textbf{Lemma 2} *The optimal standard $\bar{x}^*(\theta)$ is increasing in $\theta$ and in $D$ and $e$.*

The optimal individual standard of legal compliance depends on the cost of compliance level of the agent, and on the size of the sanction for failing to reach the standard. Notice that if we set

\textsuperscript{13}This objective function does not provide incentives to reach levels of legal compliance above the standard. We think that this is a realistic feature for most of the settings in which law enforcement operates.
a higher standard than $\overline{x}^*(\theta)$, the agent will reduce her compliance effort to $x_I^*(\theta)$. In other words, it is possible that by reducing the standard and adapting it to the agent’s abilities and costs, we can increase her effort. Counterintuitively, by requiring a lower level of legal compliance, the principal actually gets more compliance.

Also, the required level of compliance effort is increasing in the size of the sanction and the investment in enforcement.

3 Setting Policies for Populations of Infringers

In this section, we proceed to characterize the optimal policies for an entire population of agents or, equivalently, for a situation in which there is asymmetric information regarding the abilities/cost of compliance of the potential infringer.

Assume that abilities in a population of agents are distributed according to $F(\theta, s)$, where $s$ is a parameter indexing populations by different measures. Following the findings of the previous section, and with a slight abuse of notation, we denote the level of compliance effort exerted by an agent of ability $\theta$ facing a standard $\overline{x}$ as

$$x(\overline{x}, \theta) = \begin{cases} \overline{x} & \text{if } \overline{x} \leq \overline{x}^*(\theta) \\ x_I^*(\theta) & \text{otherwise.} \end{cases}$$

Let $X(\overline{x}, s)$ be the expected compliance effort of a population characterized by an ability distribution $F(\theta, s)$ facing a standard $\overline{x}$

$$X(\overline{x}, s) = \int x(\overline{x}, \theta) f(\theta, s) d\theta.$$

We denote as $\overline{x}(s)^*$ the optimal standard for compliance effort, the one that maximizes expected effort in compliance of the target population,

$$\overline{x}(s)^* \in \arg \max X(\overline{x}, s)$$

The next proposition shows the monotone relationship between total expected effort of a population and its abilities/cost.
Proposition 1 If the parameter $s$ indexes the ability distributions according to First Order Stochastic Dominance (FOSD), such that $s > s'$ then $F(\theta, s) \leq F(\theta, s')$ for all $\theta$, then the expected effort $X(\pi(s)^*, s)$ is weakly increasing in $s$.

Following the comparative statics methods developed by Milgrom and Shannon (1994), we can address an important policy question: What is the relationship between the level of ability (for instance, the technological or organizational sophistication of firms who may be subject to the legal requirements) of a population and the optimal standard for it? Building upon the results of the previous section, one could conjecture that since the optimally required level of care is non decreasing in the level of ability, better endowed (and thus, with lower compliance costs) populations of infringers should face higher legal standards. This is not generally true, however.

We can show this counterintuitive finding using a simple counter-example in which only two types with positive mass in the population exist: $\theta_L < \theta_H$. Consider that all agents except a negligible portion are low ability types, and that the low types generate almost no compliance effort. In this case, although the mass of high types is very small, the optimal standard is the one maximizing the effort exerted by these high types, since it is impossible to induce effort from the low types. Compare it now with a population having the same proportion of low and high types, but in which the low types have larger ability. The optimal standard for this second population, which is better endowed than the first according to FOSD, must be lower, since the standard should be set in order to maximize effort by the now overwhelming majority, the low types.

To guarantee robust results on the relationship between the optimal standard and distribution of ability, we need a stronger notion of what better endowed means.

Definition 1 A population $s$ is better endowed than population $s'$ in terms of the Monotone Likelihood Ratio (MLR), if $\frac{f(\theta, s)}{f(\theta, s')}$ $\geq$ $\frac{f(\theta', s)}{f(\theta', s')}$ for all $l > l'$.

MLR order is a refinement of FOSD, and in fact implies that the distributions, conditional on any interval, are ordered by FOSD. For instance, in our setting, this ordering asserts that when comparing two distributions, the expected ability of agents with abilities larger than a given
threshold, is larger in the wealthier distribution among the pair.

**Proposition 2** If the parameter $s$ indexes the ability distributions according to MLR, then the optimal standard $\pi(s)^* \text{ is weakly increasing in } s$.

We have thus shown that, under some technical conditions, a better endowed (with lower costs of compliance) population of agents should be subject to higher legal standards than a more poorly endowed population. Note, however, that such technical conditions are linked not to the average endowment level of the population, but to the distribution of abilities is the group.

As far as we know, there are no empirical studies examining how standards of legal compliance in certain areas of legal and regulatory enforcement depend on relevant features of the population of potential infringers. We believe this is a relevant question, both for intellectual and policy reasons. As an exception to the lack of empirical illustrations of the variation of legal and regulatory standards across populations and legal systems we should cite Immordino and Pagano (2010), who show that a result that, although not identical with Proposition 2, points in the same direction: when comparing similar countries (in terms to the distribution of income) there is a positive relationship between the strictness of regulatory environmental standards and wealth.

![Figure 1 around here](image)

It is clear that wealth is not a proxy for ability/cost of compliance (though arguably it could be in certain settings), it shows that heterogeneity of populations seems to matter for the level of legal standards.

4 Matching Infringers and Regulators/Enforcers

In this section we address a research question possibly with important practical implications in some contexts. What is the optimal way to organize the members of a given population into different groups (for regulatory/enforcement purposes) in order to maximize the total level of legal compliance. Take a population of agents with known abilities described by the distribution
Consider partitions generated by dividing the interval \((\theta, \varpi]\) into subintervals: i.e. \(A := \{A_1, \ldots, A_k\}\) is a partition of \([\theta, \varpi]\) and for all \(A \in A\), \(\exists a, b \in (\theta, \varpi]\) \(A = (a, b]\) and \(i > j\) implies for all \(\theta \in A_i, \theta' \in A_j\), \(x > x'\). Given two partitions \(\mathcal{A}\) and \(\mathcal{B}\), \(\mathcal{B}\) is finer than \(\mathcal{A}\) if for all \(B \in \mathcal{B}\), there is \(A \in \mathcal{A}\) such that \(B \subseteq A\).

**Proposition 3** If we set an optimal legal standard for each subinterval of every partition, the finer the partition, the larger the total exerted compliance effort.

This Proposition is very intuitive (and simple): Finer partitions allow us to set standards that fit better with the endowments of the individuals within a partition of the population. In the limit, the optimal solution will be to set optimal individual standards, one for each member of the group.

It is clear that finer partitions involve larger costs in design and enforcement of the law. They may, perhaps, be unfeasible on other grounds (they cannot be based on prohibited criteria to segregate, for instance). Then, in some situations we have to take the number of subgroups as given: For instance, think of budget constraints or legal limits that set a cap on the number of different regulators operating in a given area (environmental agencies, safety authorities, districts and so on) Consider, for example, that we have to divide the population in two groups of a given size (half and half, for example), is it optimal, as in the previous proposition, to have two complete ordered subintervals?

**Proposition 4** In general it is not optimal to have complete ordered subintervals.

The intuition and the direct proof of this proposition is that if we take an ordered population, the optimal legal standard may involve that the low types of that subgroup do not attain the standard given their costs of satisfying it. If we have an ordered population divided in two groups in such situation, by switching the set of agents that do not reach the standard in their initial group, from their initial group to the other group, we increase total compliance effort. This is because the agents that do not reach the standard in the better endowed group will perhaps do it in the other group.
This result leads us to a characterization of the optimal solution for the situation in which we have a given the number of groups, but we can adapt the size of each group.

**Proposition 5** For a given number of subintervals, the solution that maximizes the total compliance effort is one in which all the agents are ordered and the subgroups are such that for all but the less endowed subgroup (with the lower types) the optimal standard is satisfied by all the agents in the subgroup.

The intuition of this proposition requires the following two arguments. In the optimal solution, agents have to be ordered. Suppose that it is not the case, then this means that an agent with lower ability than another is facing a higher compliance standard than the latter. By switching these two agents, the total compliance effort must weakly increase. The second argument is due to Proposition 4: All agents but the ones in the first interval, have to satisfy the standard. If we have a solution such that in a subinterval the low types do not reach the standard, by allocating these agents to a subinterval with a lower standard, we weakly increase total compliance effort.

5 Optimal Enforcement Policies

In this section, we consider the cost of enforcement policies, and more importantly, we characterize the optimal enforcement policies. The objective of the principal (the policy maker or law maker) is to maximize a social or total welfare function using both instruments: a standard \( \overline{x} \), and the investment in enforcement, \( e \). In particular, the principal wishes to characterize the optimal enforcement policy that requires to set an optimal standard and an optimal enforcement level.

\[
(x^*, e^*) \in \operatorname*{arg \, max} W(\overline{x}, e, \theta) = H(x^*(\overline{x}, e^*, \theta)) + g(x) - C(x^*(\overline{x}, e^*, \theta), \theta) - \Delta(e)
\]

Where \( H(x) \) is increasing in \( x \) and relates to the principal’s benefits (or social benefits, if one prefers) of increasing the agent’s compliance with the law; \( g(x) \) is agent’s intrinsic motivation to comply with the law; \( C(x, \theta) \) is the agent’s private cost of this compliance effort; and \( \Delta(e) \) is the enforcement cost. Finally, the optimal compliance effort exerted by agent \( x^*(\overline{x}, e^*, \theta) \) depends on...
the enforcement policy and the agent’s type. As the principal wishes to maximize total welfare, it is important to define the efficient level of compliance, \( x^*_E(\theta) \in \arg \max H(x) + g(x) - C(x, \theta) \), where \( x^*_E(\theta) \) is an increasing function of \( \theta \).

The timing is as follows. First, the principal sets the enforcement policy \((\pi, e)\) and second, the agent chooses compliance effort. For a given enforcement policy chosen by the principal, \((\pi, e)\), the agent’s problem coincides with the one analyzed above in Section 2 and Lemma 1 holds. Then, the equilibrium level of compliance \( x^*(\pi, e, \theta) \) increases in \( \theta \) and \( e \), and is non monotonic in \( \pi \).

Most importantly, the optimal legal standard also coincides with the one analyzed in Section 2, and then is characterized by equalizing, for a given enforcement level \( e \), the agent’s utility when not satisfying the standard and the cost of reaching the standard,

\[
g(\pi^*) - C(\pi^*, \theta) = g(x^*_I) - p(e)D - C(x^*_I, \theta)
\]

where (as above) \( x^*_I \) is the equilibrium effort in compliance of the agent when she decides not to attain the standard, \( x^*_I(\theta) \in \arg \max \{g(x) - p(e)D - C(x, \theta)\} \).

The intuition of this result is as follows. The equilibrium level of compliance \( x^*(\pi, e, \theta) \) increases in \( e \). Then, we can increase total welfare with an arbitrary standard \( \pi^* \), by replacing this standard with the optimal one, and decreasing the enforcement level \( e \) until the same compliance effort is exerted by the agent, and by doing that, save enforcement cost, and increase total welfare.\(^{14}\) Lemma 2 also holds implying that the optimal standard is increasing in \( e \) (since by the envelope theorem the right hand side of the above equality decreases in \( e \), and the left hand side decreases also in \( \pi^* \)).

Then, for each level of enforcement investment, \( e \), an optimal standard \( \pi^* \) exists. It is characterized by making the agent indifferent between complying with the law and disobeying it. The optimal standard is designed in such a way that, in equilibrium, determines the equilibrium compliance effort, \( x^* = \pi^* \). This optimal standard \( \pi^*(e, \theta) \) is an increasing function of \( e \). Then, the principal’s problem becomes

\[
e^* \in \arg \max W(\pi, e, \theta) = H(\pi^*(e, \theta)) + g(\pi^*(e, \theta)) - C(\pi^*(e, \theta), \theta) - \Delta(e)
\]

\(^{14}\)Noticeably this result has the flavor of the notorious beckerian result, Becker (1968).
It is not clear if increasing $\theta$ translates into lower or higher enforcement investment, since higher $\theta$ makes it more desirable to increase the level of compliance, but it also makes easier to do it. For the same argument higher $\theta$ leads to higher legal standards and compliance effort. Next Proposition states that the optimal standard, and consequently, the equilibrium level of legal compliance is always below the efficient level.

**Proposition 6** The optimal standard $\bar{x}^*(e^*, \theta)$ (and the level of compliance $\bar{x}(\bar{x}^*, e^*, \theta)$) is never higher than the efficient compliance effort $x^*_E(\theta)$.

The intuition of this result is that if $\bar{x}^*(e^*, \theta)$ is higher than the efficient level $x^*_E(\theta)$, total welfare is decreasing in the compliance effort, and we can reduce compliance effort by decreasing the enforcement level. This, on the other hand, would decrease costs and then, unambiguously, total welfare increases.

Therefore, when we set standards at the individual level there is never overcompliance. We expect, however, that overcompliance may be present when setting standards for an heterogenous population. In such a case, although the issue depends on the distribution of types, it is very likely that in order to increase the compliance effort of the non marginal types (those with very high $\theta$), we have to increase the standard in such a way that marginal types exert higher levels of compliance than the efficient ones for them.

Thus, the results of Section 3 hold under the qualification that, given the social welfare function and the population of potential infringers, overcompliance with the Law will not emerge, or will not have any negative consequences in terms of welfare (for instance, we may want to disregard certain opportunity costs, e.g. sex or hate crimes). Then, setting optimal policy implies maximizing incentives, and the analysis in Section 3 follows.

6 Qualifications Implications and Conclusions

Our model of law enforcement and incentives for agents with diverse abilities and costs linked to compliance with the law tries to capture the fundamental features of many –but not all-scenarios
in which setting legal requirements and enforcing them is at stake.

We consider that the emphasis on the costs of compliance with legal requirements, and how these costs are dependent on certain traits or factors that vary across the different agents subject to the law’s prescriptions is justified in many settings. In our model, such traits are combined in a single variable which we label as ability to comply with the law, and that affects the costs of reaching a given level of compliance with the law, in its various forms: due care, maximum emission levels, minimum precaution to be adopted, mandated level of disclosure, minimum quality level, and so forth.

In the world of the model, ability is exogenous, or innate, if one prefers. Although the exogenous nature of such a variable seems to a large extent natural in several contexts (for instance, strength, sharpness, or physical ability for individuals) it is clear that in many others the agents could affect (enhance) their ability by taking certain measures, typically at a cost to themselves. Think of technological and organizational features of firms that influence the costs of compliance with a legal standard (emission levels, quality of disclosure). They are not fixed or permanent, but may be improved by capital investments or adequate improvements that are typically costly. The additional complexity that endogeneity would entail is not currently addressed in our model, although our setting could be altered to accommodate it. With asset levels of potential injurers, we showed that endogenizing levels of wealth did not alter our results as to optimal care levels (Ganuza and Gomez, 2011): Using standards that are tailored to the characteristics of each potential injurer, the legal system, counter to intuition, induces firms to choose higher levels of assets than the ones that they would be inclined to adopt under more exacting rules. Relaxing liability requirements for those firms that may decide to be smaller or to hold less assets, instead of creating a rush toward undercapitalization, actually improves the choice of size and assets, other things being equal, compared to alternative liability rules. In the present setting the moral hazard problem may be more severe (decreasing abilities not only reduces the legal standard, it also implies a reduction in direct costs), and thus in the endogenous case matters are likely to be more complex in the current setup compared to one in which heterogeneity only concerns asset
levels. This suggest that addressing heterogeneity is costs through legal standards is a tougher job than handling heterogeneity of assets levels.

In section 2, when dealing with the individual level, we assume that the ability of the potential infringer or offender can be taken into account by the lawmaker when setting the policy tools at its disposal. Naturally, this assumption will not hold in reality in many situations, perhaps in most of them, since the amount of information needed for this perfect tailoring or subjectifying of the legal requirements to each individual infringer almost defies the wildest imagination about the capabilities of a legal system. Even imperfect observation may be unfeasible in many situations. Imagine, for instance, that the source of costs of compliance with the legal requirements are of moral or psychological character, as is the case in the area of conscientious objection to legal duties or obligations, to some of them at least.\textsuperscript{15} Making the level of legal standards depend on ability and cost of compliance may be easier in some situations, most notably when they are linked to the external and readily observable properties of the infringer, such as size, age, economic sector, and so on. In any case, we do not take an issue with the plausibility of our assumption, but merely inquire into how to best provide incentives for legal compliance if the assumption holds. In fact, section 3 is devoted to optimal policy when individual tailoring is not possible, either because requirements have to be set at a given (high) level of generality, or due to unobservability of the relevant variable on an individual basis. One may conjecture, however, that vague standards (a pervasive phenomenon in the Law) may be an avenue for allowing ex-post tailoring of legally required behavior to the individual features of potential infringers.

We do not consider heterogeneity of wealth levels of potential infringers. The reasons are threefold. First, other papers have already considered the law enforcement model when wealth varies (Polinsky and Shavell, 1991; Garoupa, 2000). Second, we have extensively explored the issue ourselves in the tort context, both with exogenous and endogenous asset levels (Ganuza and Gomez, 2008, 2011). Third, in some areas ability may be correlated with proxies for wealth (size, technological quality, education), and thus the separate consideration of wealth levels would add

\textsuperscript{15}See on the efficiency analysis of conscientious objection, Domenech and Puchades, 2014.
relatively little to our current analysis. Similarly, we do not consider variations in the amount of the sanction. Such a limitation in the complexity of the model does not seem to us to be crucial. First, because a maximal fine of a fixed amount would be optimal if potential offenders do not vary in their asset level — and we have just stated our reasons for leaving aside variation in the level of assets as a helpful addition to the insightfulness of our approach. And when wealth varies and thus the optimal fine may also be variable, the tailoring of the fine to the asset level presents problems of its own that seem to be alien to our main concerns. Finally, in section 2 we do derive results (applicable also in the more general setting of section 5) on how other tools would optimally depend on the size of the sanction.

With all the above qualifications, we believe our results are helpful to illuminate certain dimensions of law enforcement that had not been fully recognized previously. The ability (and the ensuing cost) of potential offenders in adjusting behavior to the requirements set out in the law is a factor that should be given proper weight in determining the substantive standards of behavior mandated by the law under the penalty of a given sanction. Compliance with the law is not for free, even leaving aside the investment and expense of detecting and punishing offences. When what one may call “legal effort” differs greatly across potential infringers given the endowments of each of them, optimal legal policy should be more responsive to the various endowment levels. At the extreme, if feasible (and we remain agnostic about the feasibility range in the real world), this would call for “personalizing” legal requirements, borrowing the expression from Ben-Shahar and Porat (2016, forthcoming). Even when perfect individualization is beyond the sensible reach of a legal system, the endowment level, and how it is distributed among the relevant population of potential infringers is crucial for the design of optimal legal policy.

Turning now to the private enforcement saga, the implications of our results point to a closer connection between setting the required legal behavior and enforcing the legal mandates than is commonly conceived. This is most obvious in the extreme case of perfectly personalized legal standards. Providing the best incentives requires fixing standards that adapt to the endowments
and costs of the potential offenders, and to the expected sanction, and this can hardly be achieved at a level, and by an institution, that is not also involved in the enforcement activities directed against those who break their individualized standards. It does not seem possible to entirely segregate the determination of the standard and its enforcement, by leaving the first entirely in the hands of a “public” lawmaker, and the expected sanction, with its ensuing costs –of detecting and/or sanctioning– entirely in the hands of a “private” enforcer.16

Thus, at least for a number of areas of possible legal offenses, concentrating upon one single institution the powers to set the legal standards, and the powers to set –even indirectly through expenditures in detection and sanctioning– the expected sanctions seems a desirable strategy. This leaves little room for separate private enforcement while leaving the “lawmaking” dimensions in the hands of public institutions. And if the alternative of choosing a private agency as the one holding both sets of powers comes to mind, the increased divergence of their expected outcomes with respect to the socially desirable ones becomes apparent. Standards and enforcement will not be softer on the less endowed offenders, but would actually be the reverse of the optimal policy: the tougher the private enforcer would be, the lower the ability –the higher the cost– to comply with the legal requirements of a given offender.

For the same token, when a private enforcer is entrusted with setting standards and enforcement effort for a population of potential offenders, the level and distribution of compliance abilities in the relevant population, contrary to what social welfare maximization would dictate, will not constitute the guiding factors.

As to the existence of various agencies (with their own standards of compliance and expected sanctions) potentially dealing with a population of potential offenders, we have obtained the result that dividing the whole population among them based on a perfect ordering (the tougher agency for the better endowed sub-group and so on) is not the optimal policy. One can improve the provision of incentives through somewhat more “mixed” partitions.

16The Tort context is not really an example of such segregated system, since the filing of suits cannot properly be considered law enforcement in its fullest sense as the plaintiff does not determine the expected sanction, although it can make it zero by not filing suit. It could eventually be observed in areas of secondary or “add-on” private litigation following public sanctioning, as happens often in the antitrust area.
We believe that thinking about the costs of legal compliance and the factors influencing it is relevant for legal policy in the law enforcement area. Also, taking into account heterogeneity in such dimensions is almost inescapable if trying to do the job seriously and properly. This paper is just a first step in this direction. Richer models, and more information and detail concerning sector-specific (characteristics of the population of potential offenders in a given setting, of the enforcement technology, of the type of sanction, etc.), and institution-specific (how the enforcement agency behaves) should bring additional insights into these issues.

A Appendix

Proof of Lemma 1: (i) The utility function of the agent, \( \pi(x, \bar{x}, \theta) \), is supermodular in \((x, \theta)\), because \( C_{12}(x, \theta) \geq 0 \). Then the monotonicity of the equilibrium effort in \( \theta \) follows from the Theorem 4 of Milgrom and Shannon (1994). (ii) follows from the characterization of the equilibrium effort in the discussion of the Lemma.

\[
x(x, \bar{x}, \theta) = \begin{cases} 
\bar{x} & \text{if } \bar{x} \leq \bar{x}^*(\theta) \\
x_1^*(\theta) & \text{otherwise.}
\end{cases}
\]

Notice that effort is monotonic in the standard for \( \bar{x} \leq \bar{x}^*(\theta) \), and by construction \( \bar{x}^*(\theta) \geq x_1^*(\theta) \).

Proof of Lemma 2:

i) Consider \( \theta' > \theta \). Then by definition, \( x_1^*(\theta) \in \arg \max \{g(x) - p(e)D - C(x, \theta)\} \), \( x_1^*(\theta') > x_1^*(\theta) \). If \( \bar{x}^*(\theta) < x_1^*(\theta') \), then \( \bar{x}^*(\theta') \geq \bar{x}^*(\theta) \) since \( \bar{x}^*(\theta') > x_1^*(\theta') \). Assume now that \( \bar{x}^*(\theta) > x_1^*(\theta') \). From the definition of optimal standards we have:

\[
U(\bar{x}^*(\theta), \theta) = U(x_1^*(\theta), \theta) - p(e)D
\]

\[
U(\bar{x}^*(\theta'), \theta') = U(x_1^*(\theta'), \theta') - p(e)D
\]

Where, \( U(x, \theta) = g(x) - C(x, \theta) \). Subtracting the first equation from the second, we obtain:

\[
U(\bar{x}^*(\theta'), \theta') - U(\bar{x}^*(\theta), \theta) = U(x_1^*(\theta'), \theta') - U(x_1^*(\theta), \theta)
\]
Therefore

\[ U(x^*_l(\theta'), \theta') - U(x^*_l(\theta), \theta) < U(x^*_l(\theta'), \theta') - U(x^*_l(\theta), \theta) < U(\pi^+(\theta), \theta') - U(\pi^*(\theta), \theta) \]

The first inequality comes from \( x^*_l(\theta') > x^*_l(\theta) \) and \( U(x, \theta) \) being concave. The second is implied by \( U_{12}(x, \theta) > 0 \). Then

\[ U(\pi^*(\theta'), \theta') - U(\pi^+(\theta), \theta) = U(x^*_l(\theta'), \theta') - U(x^*_l(\theta), \theta) < U(\pi^+(\theta), \theta') - U(\pi^*(\theta), \theta) \]

This implies that \( U(\pi^*(\theta'), \theta') < U(\pi^+(\theta), \theta') \Rightarrow \pi^*(\theta') > \pi^*(\theta) \). This concludes the proof of the first part. ii) The second part of the Lemma follows from \( U(x^*_l(\theta), \theta) - U(x, \theta) \) being increasing in \( x \) if \( x^*_l(\theta) < x \) and the definition of \( \pi^*(\theta) \), \( p(e)D = U(x^*_l(\theta), \theta) - U(\pi^*(\theta), \theta) \). Then, the higher \( p(e)D \), the higher \( \pi^*(\theta) \). ■

**Proof of Proposition 1:** Notice that \( x(\pi, \theta) \leq \pi \) is weakly increasing on \( \theta \). Then using the following property of the FOSD order:

**Lemma 1** if \( X \geq_{FOSD} Y \) then for all increasing functions \( \psi \), \( E[\psi(X)] \geq E[\psi(Y)] \).

We obtain that \( X(\pi, s) = \int x(\pi, \theta)f(\theta, s)dl \geq \int x(\pi, \theta)f(l, s')dl = X(\pi, s') \). We finish the proof with the inequalities

\[ X(\pi(s)^*, s) \geq X(\pi(s')^*, s) \geq X(\pi(s')^*, s') \]

where the first inequality is due to the fact that \( \pi(s)^* \) is the optimal standard for a population characterized by the ability distribution \( F(l, s) \). ■

**Proof of Proposition 2:** We start with a minimal and intuitive characterization of optimal standards. Let \( \theta_{\text{max}}(s) \) be the ability of the most talented agent in a population \( s \). A necessary condition for a standard to be optimal is that it must be achievable at least by this agent. That is: \( \pi^*(s) \in [0, \pi(\theta_{\text{max}}(s))^*] \). Otherwise, by lowering the standard to \( \pi(\theta_{\text{max}}(s))^* \) we would increase the effort of this agent and we would not negatively affect the incentives of the rest.

Following the approach of Milgrom and Shannon (1994), in order to demonstrate that \( \pi^*(s) \) is (weakly) increasing on \( s \), we simply need to show that \( X(\pi, s) \) satisfies the single crossing property
(SCP) for the relevant standards, as have been just characterized in the previous paragraph: For all $\overline{x} > \overline{x}'$ and $s > s'$

$$X(\overline{x}, s') - X(\overline{x}', s') \geq (>)0 \implies X(\overline{x}, s) - X(\overline{x}', s) \geq (>)0$$

In words, if a higher standard is superior in the less endowed population, it must also be superior in the more endowed one. Now we focus on the function

$$X(\overline{x}, s) - X(\overline{x}', s) = \int (x(\overline{x}, \theta) - x(\overline{x}', \theta)) f(\theta, s) d\theta$$

where

$$x(\overline{x}, \theta) - x(\overline{x}', \theta) = \begin{cases} 0 & \text{if } \theta : \overline{x}' > \overline{x}(\theta)^* \\ x_\theta^*(\theta) - \overline{x}' & \text{if } \theta : \overline{x} \geq \overline{x}(\theta)^* \geq \overline{x}' \\ \overline{x} - \overline{x}' & \text{if } \theta : \overline{x} < \overline{x}(\theta)^*. \end{cases}$$

Let $\theta^+$ be the level of ability in which $\overline{x}' = \overline{x}(\theta^*)$. $x(\overline{x}, \theta) - x(\overline{x}', \theta)$ is negative at $\theta^+$, but it is weakly increasing for all $\theta \geq \theta^+$. We can rewrite $X(\overline{x}, s) - X(\overline{x}', s)$ as follows:

$$X(\overline{x}, s) - X(\overline{x}', s) = (1 - F(\theta^+, s))E_s\{x(\overline{x}, \theta) - x(\overline{x}', \theta)|\theta \geq \theta^+ \}$$

The parameter $s$ indexes the distributions $F(s, \theta)$ according to MLR, and this implies that the distributions, conditional on any interval, are ordered by FOSD. This implies, on the one hand,

$$(1 - F(\theta^+, s)) \geq (1 - F(\theta^+, s')) > 0,$$

where $(1 - F(\theta^+, s')) > 0$ derives from $\overline{x}' < \overline{x} \leq \overline{x}(\theta_{\max} (s))^*$.

On the other hand, as $x(\overline{x}, l) - x(\overline{x}', l)$ is increasing for all $\theta \geq \theta^+$, by Lemma 1

$$E_s\{x(\overline{x}, \theta) - x(\overline{x}', \theta)|\theta \geq \theta^+ \} \geq E_s\{x(\overline{x}, \theta) - x(\overline{x}', \theta)|\theta \geq \theta^+ \}$$

Thus $X(\overline{x}, s') - X(\overline{x}', s') \geq (>)0$ implies that $E_s\{x(\overline{x}, \theta) - x(\overline{x}', \theta)|\theta \geq \theta^+ \} \geq (>)0$ which also implies $E_s\{x(\overline{x}, \theta) - x(\overline{x}', \theta)|\theta \geq \theta^+ \} \geq (>)0$ and $X(\overline{x}, s') - X(\overline{x}', s') \geq (>)0$. Therefore $X(\overline{x}, s)$ satisfies SCP. Hence, applying Theorem 4 of Milgrom and Shannon (1994), $\overline{x}^*(s)$ is weakly increasing on $s$. 

$\blacksquare$
Proof of Proposition 3: Consider two partitions $A$ and $B$, $B$ is finer than $A$. Use the optimal standards for the subgroups of $A$ in the subgroups of the partition $B$. Take a subgroup $B$ of $B$ which is fully content in a subgroup $A$ of $A$, $B \subset A$. By setting the optimal standard for this subgroup, we make larger the exerted effort in the partition $B$. ■

References


Figure 8: Environmental Legislation (UNCED data) and log per capita GDP

FIGURE 1.