Retroactivity in Criminal Law*

Gabriel Doménech-Pascual[1] and Fernando Peinado[2]

September 15th, 2025

[1]Professor of Law. University of Valencia, [2]Professor of Economics. CUNEF Universidad

Abstract

This paper examines the economic rationale of the two key rules governing retroactivity in criminal law: the prohibition of *ex post facto* laws, and the *lex mitior* principle. The former precludes the retroactive imposition or aggravation of criminal liability, while the latter mandates the retroactive application of laws decriminalizing conduct or reducing the punishment for it. The study presents a formal model analyzing how these principles impact social welfare under conditions of risk neutrality and risk aversion, as well as when the implementation of sanctions and the adjudication of cases entail significant social costs, and when the legislature's utility is not aligned with social welfare. Our model shows that, if individuals are risk averse, retroactivity is not optimal. If implementing penalties entails costs, the results depend on the function of these costs. If it is strictly convex (e.g., sanctions show diseconomies of scale, which is arguably the case with the most severe penalties) or strictly concave, non-retroactivity is socially optimal. If it is linear (as in the standard Law and Economics model of law enforcement), both retroactivity and non-retroactivity can maximize social welfare. If adjudicating cases entails social costs, the retroactivity (non-retroactivity) of either *ex post facto* laws or *lex mitior* might increase, lower or have no effect on such costs, while preserving the same level of deterrence. In all but one scenario, retroactivity is suboptimal or non-necessary to maximize social welfare. If some legislatures overestimate the harm caused by potentially harmful conduct, then the constitution should forbid retroactive increases in sanctions, and either allow or mandate the retroactivity of milder laws. If establishing a milder law with retroactive effect entails a political cost for the legislature, a constitutional provision mandating the retroactivity of *lex mitior* would induce the legislature to impose suboptimal sanctions. Overall, the findings of this paper provide strong support for the prohibition of *ex post facto* laws, and weak or no support for the application of *lex mitior*.

Key words: Law and economics; Criminal law; Retroactivity; *Lex mitior*; *Ex post facto* laws; Risk aversion.
JEL Codes: K14, K42.

*Corresponding author: . E-mail: . Authors wish to thank financial support from the Spanish Ministry of Science and Innovation under research project PID2021-125155NB-100.

1 Introduction

According to the Oxford Dictionary of Law, retroactive (or retrospective) legislation means "legislation that operates on matters taking place before its enactment, e.g. by penalizing conduct that was lawful when it occurred". From an economic standpoint, we can consider that a law is retroactive when it changes the costs or benefits of a decision made before the law was enacted.

Retroactive legislation has long been a focal point of legal scholarship, generating extensive analysis across diverse fields, namely private law, criminal law, constitutional law, procedural law, administrative law, tax law, environmental law, etc. This body of research largely centers on three core questions: (1) under what circumstances a law should be construed as having retroactive effect; (2) when retroactivity renders a law invalid due to its conflict with overriding legal or constitutional principles; and (3) whether retroactive application is justifiable from a policy perspective.

The Law and Economics literature on retroactive legislation remains relatively limited[1]. A handful of studies examine the attenuation of legal change, that is, whether the costs imposed by retroactive legal reforms should be mitigated through measures such as: (1) adopting less stringent rules than would otherwise be optimal[2], (2) grandfathering existing situations[3], or (3) compensating individuals adversely affected by the new rules[4]. Non-retroactivity itself serves as one potential mechanism (among others) for attenuating such legal transitions.

In this body of literature, however, there are no specific studies on retroactive legislation in the realm of criminal law. The present paper tries to fill this gap. It analyzes the rationale of the most relevant rules governing the retroactivity of criminal legislation: (1) the prohibition of the so called *ex post facto* laws, and (2) the *lex mitior* principle. According to the former rule, no one may be punished for any offense on account of any act or omission which did not constitute an offense under the law at the time when it was committed, nor may a heavier penalty be imposed than the one that was provided by the law at the time the offense was committed. According to the latter one, if after the commission of an offense a new law provides for a lighter penalty or no penalty at all, this law is to be applied.

This paper explores in particular the role that risk aversion and the social costs of punishment and adjudication play in order to justify both rules. The paper is organized as follows. Section 2 describes the content, the scope, and the rationale of both rules in different legal systems. Section 3 presents the baseline model, showing how these rules impact social welfare given risk neutrality. Section 4 extends the previous model to the case where potential offenders are risk-averse. Section 5 shows how both rules affect social welfare when sanctions are socially costly. Section 6 extends the basic model to take into account that adjudication is costly. In Section 7 we present different sequential games where the legislature's utility is not aligned with social welfare. Section 8 discusses the results and their implications.

2 The law

The prohibition of legislation that retroactively imposes or aggravates criminal liability

A cornerstone principle of modern criminal law is the prohibition of legal rules that either retroactively criminalize conduct that was lawful when committed or impose harsher penalties than those applicable at the time of the offense.

[1] See Kaplow (1986; 2003); Levmore (1999); Shaviro (2000); Adler (2003); Epstein (2003); Nash and Revesz (2007); Shavell (2008; 2014); Masur and Nash (2010); Revesz and Westfahl Kong (2011); Trebilcock (2014); Franzoni (2019); Damon et al. (2019).

[2] See Shavell (2014) and Franzoni (2019).

[3] See Shavell (2008).

[4] See, for instance, Miceli and Segerson (1994).

This prohibition, often encapsulated in the Latin maxim *nulla poena sine lege* (no punishment without law) and, in some countries, known as the prohibition of *ex post facto* laws, has long been a widely established principle across the world. It was already enshrined under Article 8 of the 1789 Declaration of the Rights of Man and of the Citizen, which states that "no one may be punished except by virtue of a Law drawn up and promulgated before the offense is committed, and legally applied".

In 2007, 162 out of 192 United Nations member states recognized this principle in their constitutions. Beyond that, virtually every state has accepted it through a combination of constitutional provisions, statutory law, or treaty obligations (Gallant, 2009).

This principle is indeed enshrined in the most relevant international treaties on human rights, such as the Universal Declaration of Human Rights (Article 11), the International Covenant on Civil and Political Rights (Article 15), the American Convention on Human Rights (art. 9), the African Charter on Human and Peoples' Rights (art. 7.2), the Arab Charter on Human Rights (art. 15), and the Charter of Fundamental Rights of the European Union (Article 49.1). For instance, Article 7.1 of the European Convention on Human Rights establishes that "no one shall be held guilty of any criminal offense on account of any act or omission which did not constitute a criminal offense under national or international law at the time when it was committed. Nor shall a heavier penalty be imposed than the one that was applicable at the time the criminal offence was committed". This prohibition is considered both a general principle of law accepted by the community of nations and a rule of customary international law (Gallant, 2009).

It is not entirely clear whether this principle admits any "exceptions". Note that the European Court of Human Rights has underlined the fact that no derogation from said prohibition is permissible in time of war or other public emergency (under Article 15 of the Convention)[5].

However, in the realm of transitional justice, courts have occasionally imposed severe penalties for actions that did not constitute offenses under domestic law when committed. In such cases, these courts have argued that said acts were to be considered crimes under the principles of international law. Domestic statutory provisions that would have otherwise exempted the perpetrators from criminal liability were deemed null and void[6]. Similarly, Article 7.2 of the European Convention on Human Rights establishes that the prohibition of *ex post facto* criminal laws "shall not prejudice the trial and punishment of any person for any act or omission which, at the time when it was committed, was criminal according to the general principles recognised by the civilised nations"[7].

It is also unclear whether the prohibition of *ex post facto* laws extends to procedural legal rules, such as those modifying evidentiary rules, reducing time-limits for appeals or extending statutes of limitations, etc.[8].

Another controversial issue is whether the prohibition of retroactivity of criminal laws also extends to judicial interpretations. Courts sometimes change their interpretation of legal provisions, expanding their scope over time. While not technically *ex post facto* (since the law existed at the time of the conduct under consideration), broad judicial interpretations can have a similar retroactive effect[9].

Various arguments have been advanced in support of said prohibition.

---

[5] Judgment of the European Court of Human Rights of 12 February 2008 (*Kafkaris v. Cyprus*, application no. 21906/04, para. 137).

[6] See, for instance, the Judgments of the European Court of Human Rights of 22 March 2001 (*Streletz, Kessler and Krenz v. Germany*, applications nos. 34044/96, 35532/97 and 44801/98) and 22 March 2001 (*K.-H. V. v. Germany*, application no. 37201/97), regarding the convictions of former East German officials for their involvement in the deaths of individuals attempting to flee across the Berlin Wall.

[7] Article 49.2 of the European Union Charter of Fundamental Rights.

[8] See the Judgment of the European Court of Human Rights of 19 September 2009 (*Scoppola v. Italy*, application no. 10249/03), where the Court "reiterates that the rules on retrospectiveness set out in Article 7 of the Convention apply only to provisions defining offences and the penalties for them".

[9] In favor of judicial retroactive lawmaking in criminal law, see, for example, Krent (1997) and Kahan (1997). Against it, see, for instance, the Judgment of the European Court of Human Rights of 21 October 2013 (*Del Río Prada v. Spain*, application no. 42750/09). However, the Judgment of the European Court of Human Rights of 22 November 1995 (*C.R. v. United Kingdom*, application no. 20190/92), seems to accept the retroactivity of case law criminalizing marital rape. For a critical discussion of this judgment, see Lawrence (2012)

(1) *Protection against arbitrary use of public power.* As the European Court of Human Rights has declared, "the object and purpose of [that prohibition] is to ensure that no one is subjected to arbitrary prosecution, conviction or punishment."[10] The referred prohibition indeed protects disfavored individuals or groups against being targeted by public agents for disadvantageous treatment. It prevents people who have little political influence from being singled out based on prior conduct and opportunistically punished. As Krent (1996) has suggested, this would arguably explain why the prohibition of retroactivity is much stricter in criminal law than in civil law, where retrospective lawmaking is usually accepted. Those who are imposed criminal sanctions usually lack the necessary influence or money to defend themselves against eventual punitive abuses from the legislature or other public agents. By contrast, potential victims of abusive civil legislation usually have more political influence and resources to prevent these abuses.
(2) *Principle of culpability.* According to this principle, "no person ought to be punished in the absence of a guilty mind". Therefore, no one ought to be punished for violating a legal rule that was non-existent at the time he or she acted. In such cases, the actor lacks a guilty mind, although she or he was aware of the factual nature of her or his act and although it was actually bad, because she or he "neither knew nor reasonably should have known that the relevant political community would regard the act as warranting the punishment at issue" (Westen, 2007, p. 285).
This argument is tautological. One can reasonably argue that if it is sufficiently evident that an act is harmful enough to society, the actor should refrain from carrying it out even if the political community has not yet established it as a crime, since the actor should anticipate that the community will consider it a punishable offense.
(3) *Principle of legal certainty.* The prohibition of *ex post facto* laws promotes predictability in judging the legal consequences of one's conduct. It ensures that individuals can foresee what acts are criminal and what penalties apply before they act, allowing them to avoid criminal liability when planning their future activities (Gallant, 2009, p. 20).

The *lex mitior* principle: the obligation to retroactively apply legal rules that decriminalize conduct or impose milder penalties

The so called *lex mitior* (Latin for "the milder law") principle requires the retroactivity of legal rules that either decriminalize conduct altogether or reduce the punishment for it.

There is an intriguing asymmetry between this mandate and the prohibition of *ex post facto* criminal laws. The *lex mitior* principle has been recognized later, with less constitutional support, more exceptions, and less global consensus than said prohibition.

We find divergent approaches in different legal systems regarding such principle. In 2007, only 48 UN member states had mandated the application of the milder criminal law in their constitutions, and at least 21 additional countries have established this principle in their statutory law (Gallant, 2009).

In the United States, twenty two jurisdictions wholly embrace it, ten jurisdictions partly embrace it, and eighteen jurisdictions entirely reject it (Westen, 2015). In any case, legislatures have discretion to give or not give retroactive effect to more lenient criminal laws. The question is how to interpret in each case what the legislature's intent was in this regard. On this point, U.S. jurisdictions have adopted the following positions: "A) a rebuttable presumption against *lex mitior* in all cases; (B) a rebuttable presumption against *lex mitior* regarding lessened penalties, but acceptance of it regarding decriminalizations; (C) a rebuttable presumption against *lex mitior* regarding decriminalizations, but acceptance of it regarding lessened penalties; (D) no presumption against *lex mitior*, but an unbiased inquiry into legislative intent in every case; (E) a rebuttable presumption in favor of *lex mitior* in all cases" (Westen, 2015).

[10] Judgment of the European Court of Human Rights of 22 March 2001 (*Streletz, Kessler and Krenz v. Germany*, applications nos. 34044/96, 35532/97 and 44801/98, paragraph 87).

Some international conventions on human rights have also established the *lex mitior* mandate. This is the case, for instance, of the International Covenant on Civil and Political Rights (art. 15.1, although the United States, Italy, Trinidad and Tobago, and Germany have made reservations to this provision); the American Convention of Human Rights (art. 9); the Arab Charter on Human Rights (art. 15); and the Charter of Fundamental Rights of the European Union (Article 49.1). For instance, the latter provides that "if, subsequent to the commission of a criminal offense, the law provides for a lighter penalty, that penalty shall be applicable".

Other international conventions on human rights do not explicitly mention said principle. This is the case of the Universal Declaration of Human Rights, the African Charter on Human and Peoples' Rights, and the European Convention on Human Rights. However, the European Court of Human Rights have declared that "a consensus has gradually emerged in Europe and internationally around the view that application of a criminal law providing for a more lenient penalty, even one enacted after the commission of the offence, has become a fundamental principle of criminal law"[11], and affirmed that "Article 7 § 1 of the Convention guarantees not only the principle of non-retrospectiveness of more stringent criminal laws but also, and implicitly, the principle of retrospectiveness of the more lenient criminal law. That principle is embodied in the rule that where there are differences between the criminal law in force at the time of the commission of the offence and subsequent criminal laws enacted before a final judgment is rendered, the courts must apply the law whose provisions are most favorable to the defendant"[12].

In those jurisdictions where it is recognized, the *lex mitior* principle applies subject to many qualifications, restrictions, and exceptions.

First, in many jurisdictions, the principle does not apply when the case has already been adjudicated by means of a final decision made under a previous, harsher statute. In other words, such cases may not be reopened to apply the new, more lenient law, even if the convicted person has not yet completed serving their sentence. For instance, Section 2(3) of the German Criminal Code provides that "if the law in force at the time of the completion of the act is amended before judgment, the most lenient law is to be applied". A *contrario sensu*, if the law in force at the time of the completion of the criminal act is amended after the judgment, the new, more lenient law is not retroactively applicable.

Second, the principle does not apply either when the convicted person has already served his or her sentence under the more severe law. Such cases may not be reviewed to apply the more lenient law retroactively. For instance, individuals who paid a fine for an offence that was later decriminalized are not entitled to a refund, and those who served a prison sentence under a law that was subsequently repealed have no right to compensation for the losses incurred during their imprisonment.

Third, the principle does not apply if a criminal provision is repealed not due to a legislative reassessment of the act's wrongfulness, but because punishment is no longer deemed necessary, such as when social circumstances have changed. The legislature may still consider the act wrongful, yet consider that there is no further need for a punishment. This explains Section 2(3) of the German Criminal Code, which provides that "a law which was intended to be in force only for a determinate time is, as a rule, still to be applied to acts committed whilst it was in force even after it ceases to be in force. This does not apply to the extent that a law provides otherwise".

Fourth, in some countries the *lex mitior* principle does not apply during the *vacatio legis*, i.e. during the period between the law's publication and its entry into force[13]. Consequently, legislatures may strategically extend this interim period to avoid the principle's application.

Fifth, legislatures may also circumvent this principle by deliberately delaying the enactment of

[11] Judgment of 17 September 2009 (*Scoppola v. Italy*, application no. 10249/03, para. 106).

[12] Ibídem, para. 109. See also the Judgments of the ECHR of 5 February 2013 (*Martirosyan v. Armenia*, application no. 23341/06); 18 July 2013 (*Maktouf and Damjanovic v. Bosnia*, applications no. 2312/08 and 34179/08); 12 July 2016 (*Ruban v. Ukraine*, application no. 8927/11); 24 January 2017 (*Koprivnikar v. Slovenia*, application no. 67503/13); 18 February 2020 (*Jidic v. Romania*, application no. 45776/16); 24 May 2022 *(Sinan Çetinkaya and Ağyar Çetinkaya v. Turkey*, application n. 74536/10); and 18 October 2022 (*Mørck Jensen v. Denmark*, application no. 60785/19).

[13] See, for instance, the Judgment of the Spanish Constitutional Court 6/2010, of 7 May 2010.

more lenient criminal legislation.

It has been argued that the principle of proportionality supports the *lex mitior* rule. If the legislature reduces the sanction for an act, it implies that the former, harsher penalty was excessive. Continuing to impose such a penalty would be disproportionate once its social costs outweigh its benefits. Likewise, if punishing conduct occurring after a law's repeal is no longer justified, then punishing acts committed before the repeal –but adjudicated afterwards– is equally untenable.

In a similar vein, Westen (2015) points out that a legislative decision to mitigate a penalty or decriminalize an act reflects the judgment that the new, milder approach is sufficient to adequately serve the legitimate goals of criminal law. If a legislature decides that a lesser penalty, or no penalty at all, is enough to deter, incapacitate, and rehabilitate for conduct committed after the new law enters into force, it logically follows that the harsher, older penalty is no longer necessary or justified for conduct committed before the new law. If the future benefits of the older penalty (deterrence, incapacitation, and rehabilitation) do not justify punishing a person for conduct committed after an ameliorative repeal, they are unlikely to justify still punishing a person who acted before the repeal.

However, it could be argued that the *lex mitior* rule may undermine deterrence. If potential offenders perceive –for instance, during legislative debates– a substantial likelihood that conduct will be decriminalized or penalties reduced, the *lex mitior* principle lowers the expected sanction for violating current laws, which might remain unchanged.

3 Baseline model

Let G be a set representing the group of all individuals in a society. The net gain each individual obtains from committing some potentially harmful act is denoted $g \geq 0$. We assume that all costs individuals may incur to commit the harmful acts before and during the actions are exogenous to the model. In this paper, we deal with the consequences after the criminal action took place. The probability density of gains is $z(g)$. Let $h \geq 0$ represent the harm inflicted on G due to the harmful act. Assume that harms are equally distributed among G. Also, assume that G has a sufficiently large number of elements so that the fraction of h suffered by an individual is negligible. The probability of the individual who commits h being sanctioned for it is $p \in [0,1]$.[14] We denote s as the sanction for committing h. In the baseline model, s is a socially costless fine collected and distributed among G in a way analogous to h. We aim to maximize social welfare, W, choosing the most appropriate sanction.

At the time of deciding whether or not to commit the harmful act, individuals do not know with certainty the magnitude of the harm it will cause. They can only estimate that such act will cause harm h_1 with probability $(1-q)$ or harm h_2 with probability q. They can also estimate that if the act causes h_1 there will be no legal change and those who commit such act will receive the sanction s_1, while if the act causes h_2 the legislature will establish the sanction s_2 for them.

After the individuals have made their decision, the uncertainty is resolved, i.e. everybody is able to know whether the harmful act causes either h_1 or h_2. If it causes h_2 the legislature provides for a sanction s_2 for those who committed the act in the past. Note that, if $s_1 = s_2$, then the new law is not retroactive. If, on the contrary, $s_1 \neq s_2$, then the new law is retroactive.

Recall that our framework is completely different from the standard law and economics model of law enforcement, where there is no uncertainty about the actual harm inflicted. The s_2 found through our model only applies for offenders who experienced *ex ante* some uncertainty about the harm. A different optimal sanction s_3 that only accounts for h_2 will be imposed to individuals that commit the act after such uncertainty has been resolved.

Assume individuals are risk-neutral; then, they commit the harmful act if

$$g \geq p \cdot [s_1 \cdot (1-q) + s_2 \cdot q], \tag{1}$$

[14] Results are the same if we believe individuals are not sure about the actual p, i.e. they know its probability density $r(p)$.

that is, if the gain for committing the harmful act exceeds some critical threshold

$$\tilde{g} = p \cdot [s_1 \cdot (1-q) + s_2 \cdot q]. \tag{2}$$

Social welfare derived from a g that safisfies (1) is[15]

$$W = \int_{\tilde{g}}^{\infty} [g - (1-q) \cdot h_1 - q \cdot h_2] \cdot z(g)\, dg. \tag{3}$$

The social welfare problem is to choose the s_1 and s_2 to maximize (3).

First-order conditions are

$$\frac{\delta W}{\delta s_1} = 0 \Leftrightarrow -\{p \cdot [s_1 \cdot (1-q) + s_2 \cdot q] - (1-q) \cdot h_1 - q \cdot h_2\} \cdot z(\tilde{g}) \cdot p \cdot (1-q) = 0 \tag{4}$$

$$\frac{\delta W}{\delta s_2} = 0 \Leftrightarrow -\{p \cdot [s_1 \cdot (1-q) + s_2 \cdot q] - (1-q) \cdot h_1 - q \cdot h_2\} \cdot z(\tilde{g}) \cdot p \cdot q = 0. \tag{5}$$

We know $z(\tilde{g}) > 0$ and, if $p, q > 0$, then $p \cdot (1-q) > 0$ and $p \cdot q > 0$, hence, it should be the case that

$$\begin{aligned} & p \cdot [s_1 \cdot (1-q) + s_2 \cdot q] - (1-q) \cdot h_1 - q \cdot h_2 = 0 \\ & \Leftrightarrow s_1 \cdot (1-q) + s_2 \cdot q = \frac{(1-q) \cdot h_1 + q \cdot h_2}{p} \end{aligned} \tag{6}$$

for $\frac{\delta W}{\delta s_1} = 0$ and $\frac{\delta W}{\delta s_2} = 0$.

Retroactivity is unnecessary if $s_1 = s_2 = s$ exists for a single $s = s^*$. If such s exists, social welfare can be maximized by imposing s_2 with retroactive effect or without it. Assume $s_1 = s_2 = s$, then (6) is

$$s = \frac{(1-q) \cdot h_1 + q \cdot h_2}{p} \tag{7}$$

that provides a single optimal value $s = s^*$ for any given value of $p \in (0, 1]$. Subsequently, we can conclude that retroactivity can be avoided under risk neutrality.

> ***Proposition 1*** *If offenders are risk-neutral, social welfare could be maximized either by giving retroactive effect the new laws imposing a new sanction or giving them no retroactive effect.*

Moreover, one can see that sanctions $s_1 = \frac{h_1}{p}$ and $s_2 = \frac{h_2}{p}$ satisfy condition (6). In other words, when offenders are risk neutral and there is uncertainty on whether the act causes harm h_1 or harm h_2, social welfare is maximized if the law ignores this uncertainty. That is, the law initially sets the sanction that would be socially optimal under certainty of harm h_1 and, if the harm later materializes as h_2, retroactively imposes the same sanction that would have been optimal had the harm been h_2 with certainty.

Note that according to the standard law and economics model of law enforcement, the optimal fine for committing an act that causes harm h with certainty is

$$s = \frac{h}{p}. \tag{8}$$

[15] Harmful acts that provide a $g < p \cdot [s_1 \cdot (1-q) + s_2 \cdot q]$ do not change social welfare because they do not take place. Formally, social welfare would be $\int_{\tilde{g}}^{\infty} [g - (1-q) \cdot h_1 - q \cdot h_2] \cdot z(g)\, dg + \int_0^{\tilde{g}} 0 \cdot z(g)\, dg$ where $\int_0^{\tilde{g}} 0 \cdot z(g)\, dg = 0$, hence, (3).

Proposition 2 *If individuals are risk neutral and there is uncertainty on whether an act causes harm h_1 or harm h_2, social welfare is maximized if the law initially sets the sanction that would be socially optimal under certainty of harm h_1 and, if the harm later happens to be h_2, retroactively imposes the same sanction that would have been optimal had the harm been h_2.*

4 Model with risk aversion

Assume that the preferences of individuals committing some potentially harmful acts are well described by a utility function U that is increasing and concave. However, society as a whole is risk-neutral. Additionally, assume individuals have some initial wealth, w. Hence, $U(w)$ is the reserve utility.

The harmful act is committed if

$$U(w+g)\cdot(1-p)+U(w+g-s_1)\cdot p\cdot(1-q)+U(w+g-s_2)\cdot p\cdot q\geq U(w), \tag{9}$$

or equivalently if

$$U(w+g)\geq\frac{U(w)-U(w+g-s_1)\cdot p\cdot(1-q)-U(w+g-s_2)\cdot p\cdot q}{1-p}. \tag{10}$$

Thus, a harmful act is committed if $U(w+g)$ is above some critical threshold

$$\widetilde{U(w+g)}=\frac{U(w)-U(w+g-s_1)\cdot p\cdot(1-q)-U(w+g-s_2)\cdot p\cdot q}{1-p}. \tag{11}$$

Denote $k(U(w+g))$ the probability density of $U(w+g)$. The social welfare derived from a $U(w+g)$ that satisfies (10) is

$$W = \int_{\widetilde{U(w+g)}}^{\infty}\left[U(w+g)-(1-q)\cdot h_1-q\cdot h_2\right]\cdot k\left(U(w+g)\right)dU(w+g). \tag{12}$$

The social welfare problem is to choose the s_1 and s_2 to maximize (12).

First-order conditions are

$$\begin{aligned}\tfrac{\delta W}{\delta s_1}=0\Leftrightarrow\ & -\Bigg[\frac{U(w)-U(w+g-s_1)\cdot p\cdot(1-q)-U(w+g-s_2)\cdot p\cdot q}{1-p}\\ & -(1-q)\cdot h_1-q\cdot h_2\Bigg]\cdot k\left(\widetilde{U(w+g)}\right)\cdot\left(\frac{-p\cdot(1-q)}{1-p}\cdot\frac{\delta U(w+g-s_1)}{\delta s_1}\right)=0\end{aligned} \tag{13}$$

and

$$\begin{aligned}\tfrac{\delta W}{\delta s_2}=0\Leftrightarrow\ & -\Bigg[\frac{U(w)-U(w+g-s_1)\cdot p\cdot(1-q)-U(w+g-s_2)\cdot p\cdot q}{1-p}\\ & -(1-q)\cdot h_1-q\cdot h_2\Bigg]\cdot k\left(\widetilde{U(w+g)}\right)\cdot\left(\frac{-p\cdot q}{1-p}\cdot\frac{\delta U(w+g-s_2)}{\delta s_2}\right)=0.\end{aligned} \tag{14}$$

We know $k\left(\widetilde{U(w+g)}\right)>0$ and, if $p,q>0$, then $\frac{-p\cdot(1-q)}{1-p}<0$, $\frac{-p\cdot q}{1-p}<0$ and, by definition $U'(\cdot)>0$, which implies that $\frac{\delta U(w+g-s_1)}{\delta s_1}<0$ and $\frac{\delta U(w+g-s_2)}{\delta s_2}<0$, hence, $\left(\frac{-p\cdot(1-q)}{1-p}.\frac{\delta U(w+g-s_1)}{\delta s_1}\right)>0$ and $\left(\frac{-p\cdot q}{1-p}.\frac{\delta U(w+g-s_2)}{\delta s_2}\right)>0$, dealing

$$\frac{\delta W}{\delta s_1} = \frac{\delta W}{\delta s_2} = \frac{U(w) - U(w+g-s_1)\cdot p\cdot(1-q) - U(w+g-s_2)\cdot p\cdot q}{1-p} - (1-q)\cdot h_1 - q\cdot h_2 = 0$$

$$\Leftrightarrow U(w+g-s_1)\cdot(1-q) + U(w+g-s_2)\cdot q = \frac{U(w) - (1-p)\cdot[(1-q)\cdot h_1 + q\cdot h_2]}{p}. \quad (15)$$

Assuming $s_1 = s_2 = s$, then (15) is

$$U(w+g-s) = \frac{U(w) - (1-p)\cdot[(1-q)\cdot h_1 + q\cdot h_2]}{p}, \quad (16)$$

or equivalently

$$U(w+g-s) - \frac{U(w) - (1-p)\cdot[(1-q)\cdot h_1 + q\cdot h_2]}{p} = 0. \quad (17)$$

For specific values of w, p, q, h_1 and h_2, we can conclude that $\left[-\frac{U(w)-(1-p)\cdot[(1-q)\cdot h_1+q\cdot h_2]}{p}\right]$ is a constant, that we can denote n, and $U(w+g-s)$ depends on s, specifically $\frac{\delta U(w+g-s)}{\delta s} < 0$. Therefore, we can call

$$F(s) = U(w+g-s) + n = 0. \quad (18)$$

Notice that

$$F'(s) = \frac{\delta U(w+g-s)}{\delta s} + 0 = \frac{\delta U(w+g-s)}{\delta s} < 0 \quad (19)$$

so we can conclude that $F(s)$ is strictly decreasing in s. Knowing that $F(s)$ is strictly decreasing, $F(s) = 0$ is only true for one value of $s = s^*$, which implies that retroactivity can be avoided under risk aversion. Notice that this result is also true if the victims suffered harms following $U(h_1)$ and $U(h_2)$.

> ***Proposition 3*** *If offenders are risk-averse, and there is uncertainty about the actual harm, there exists a single optimal sanction that accounts for this uncertainty; therefore, social welfare cannot be maximized if the new sanction is applied with retroactive effect.*

Looking at (16), the value s^* will be a positive real number if $n < 0$ and $(w+g) > U^{-1}(-n)$, assuming $U(\cdot)$ takes the form of a "power", "logarithmic (Bernoulli)", "negative exponential" or "quadratic" utility function. Looking at (17), we can observe: a negative correlation between p and s, a positive correlation between h_1, h_2 and s, a positive correlation between g and s, and a positive correlation between w and s until some threshold $\bar{w}$ above which they are negatively correlated.

5 Incorporating implementation costs of sanctions

Assume implementing the applicable penalty entails costs for society (Kaplow, 1990). This is definitely the case, for instance, with imprisonment (Becker, 1968; Polinsky and Shavell, 1984), as well as relatively large fines (Polinsky and Shavell, 1984). Assume these costs are well described by a function $c(s) > 0$ that is strictly increasing and strictly convex –doubling the penalty requires more than double the implementation costs, i.e. diseconomies of scale.

A fine will show diseconomies of scale because a risk-averse individual suffers more disutility from a potential large fine (one that could, for instance, bankrupt them) than a simple monetary valuation would suggest. This adds a social cost beyond the mere transfer. But even assuming risk-neutral

individuals, when a fine exceeds an offender's ability to pay, it imposes significant social costs. These include the costs of bankruptcy proceedings, the distortion of economic activity as individuals take actions to shield assets, and the deadweight loss of forgoing the fine entirely, which may necessitate the use of more socially costly sanctions like imprisonment.

Imprisonment will presumably imply diseconomies of scale once a prison system reaches capacity because the marginal cost of adding an inmate skyrockets. It is no longer a marginal operating cost but the massive, lumpy cost of building a new prison. In the short run, overcrowding within existing facilities can lead to exponentially rising costs associated with violence, health crises, and litigation. Moreover, as a society expands its use of incarceration, it moves from imprisoning the most dangerous and hardened criminals to those who have committed less severe offenses. The social cost of imprisoning a high-frequency, violent offender is largely offset by the social benefit of incapacitation. The social cost of imprisoning a low-level drug offender, who might otherwise be a productive worker and family provider, is far higher. Levitt (2004) has shown that the marginal benefit of incapacitation declines as the prison population grows. This implies an increasing marginal social cost for each additional prisoner, a form of convexity.

Without risk aversion, individuals still commit the harmful act if (1) is satisfied. We can rewrite social welfare derived from a g that safisfies (1) incorporating implementation costs of penalty like

$$W = \int_{\tilde{g}}^{\infty} \left[g - (1-q) \cdot h_1 - q \cdot h_2 - p \cdot \left((1-q) \cdot c\left(s_1\right) + q \cdot c\left(s_2\right) \right) \right] \cdot z\left(g\right) dg. \tag{20}$$

The social welfare problem is to choose the s_1 and s_2 to maximize (20).

First-order conditions are

$$\begin{aligned} \tfrac{\delta W}{\delta s_1} = 0 \Leftrightarrow \quad & -\{p \cdot [s_1 \cdot (1-q) + s_2 \cdot q] - (1-q) \cdot h_1 - q \cdot h_2 \\ & -p \cdot \left((1-q) \cdot c\left(s_1\right) + q \cdot c\left(s_2\right) \right) \} \cdot z\left(\tilde{g}\right) \cdot p \cdot (1-q) \\ & -p \cdot (1-q) \cdot c^{'}\left(s_1\right) \cdot \int_{\tilde{g}}^{\infty} z\left(g\right) dg = 0; \end{aligned} \tag{21}$$

$$\begin{aligned} \tfrac{\delta W}{\delta s_2} = 0 \Leftrightarrow \quad & -\{p \cdot [s_1 \cdot (1-q) + s_2 \cdot q] - (1-q) \cdot h_1 - q \cdot h_2 \\ & -p \cdot \left((1-q) \cdot c\left(s_1\right) + q \cdot c\left(s_2\right) \right) \} \cdot z\left(\tilde{g}\right) \cdot p \cdot q \\ & -p \cdot q \cdot c^{'}\left(s_2\right) \cdot \int_{\tilde{g}}^{\infty} z\left(g\right) dg = 0; \end{aligned} \tag{22}$$

or equivalently

$$\begin{aligned} \tfrac{\delta W}{\delta s_1} = 0 \Leftrightarrow \quad & -\{p \cdot [s_1 \cdot (1-q) + s_2 \cdot q] - (1-q) \cdot h_1 - q \cdot h_2 \\ & -p \cdot \left((1-q) \cdot c\left(s_1\right) + q \cdot c\left(s_2\right) \right) \} \cdot z\left(\tilde{g}\right) = c^{'}\left(s_1\right) \cdot \left(1 - Z\left(\tilde{g}\right)\right); \end{aligned} \tag{23}$$

$$\begin{aligned} \tfrac{\delta W}{\delta s_2} = 0 \Leftrightarrow \quad & -\{p \cdot [s_1 \cdot (1-q) + s_2 \cdot q] - (1-q) \cdot h_1 - q \cdot h_2 \\ & -p \cdot \left((1-q) \cdot c\left(s_1\right) + q \cdot c\left(s_2\right) \right) \} \cdot z\left(\tilde{g}\right) = c^{'}\left(s_2\right) \cdot \left(1 - Z\left(\tilde{g}\right)\right); \end{aligned} \tag{24}$$

where $Z\left(\tilde{g}\right)$ is the cumulative distribution function of $z\left(g\right)$ evaluated at $\tilde{g}$.

First-order conditions imply that $c^{'}\left(s_1\right) = c^{'}\left(s_2\right)$. Implementation costs are strictly increasing and strictly convex, hence, $c^{'}\left(s\right)$ is strictly increasing. Strictly increasing functions are injective, thus, a single solution $s_1 = s_2 = s^*$ exists. This result is also true if implementation costs of penalty are strictly concave –i.e. economies of scale.

Proposition 4 *Incorporating implementation costs of penalty that show diseconomies of scale or economies of scale, retroactivity will not be socially optimal.*

Results are different if implementation costs of sanctions are strictly increasing and linear (as in the standard model of law enforcement[16]) since $c^{'}(s) = k$ where $k > 0$ is a constant. Increasing the sanctions does not change the marginal cost of enforcement. First-order conditions become identical. Therefore, we have a single equation with two variables, which provides infinitely many combinations of s_1 and s_2 that yield the same social welfare. The only restriction is that the combination of s_1 and s_2 satisfies $\tilde{g} = p \cdot [s_1 \cdot (1-q) + s_2 \cdot q]$.

Proposition 5 *If the costs of implementing penalties are linear, a non-retroactive sanction is as good as a retroactive one.*

6 Model with adjudication costs

Agents commit the harmful act if (1). Adjudication costs are denoted a. Social welfare depends on the scenario that takes place.

Scenario 1: criminalization of conduct

In this scenario, $h_1, s_1 = 0$, $h_2 > 0$. If $s_2 = 0$, i.e., no retroactivity of *ex post facto* law, the critical threshold of the gain for committing the harmful act is $\tilde{g} = 0$. Then, social welfare is

$$W_{1,n} = \int_0^{\infty} [g - q \cdot h_2] \cdot z(g)\, dg. \tag{25}$$

If $s_2 > 0$, i.e., retroactivity of *ex post facto* law, then $\tilde{g} = p \cdot s_2 \cdot q$. Hence, social welfare is

$$W_{1,r} = \int_{\tilde{g}}^{\infty} [g - q \cdot (h_2 + p \cdot a)] \cdot z(g)\, dg. \tag{26}$$

We can find s_2^* using F.O.C.

$$\frac{\delta W_{1,r}}{\delta f_2} = 0 \Leftrightarrow \quad -\{[p \cdot s_2 \cdot q - q \cdot (h_2 + p \cdot a)] \cdot z(p \cdot s_2 \cdot q)\} \cdot p \cdot q = 0 \tag{27}$$

where $p \cdot q$, $z(p \cdot s_2 \cdot q) \neq 0$, therefore,

$$s_2^* = \frac{q \cdot (h_2 + p \cdot a)}{p \cdot q}, \tag{28}$$

into (26):

$$W_{1,r}(s_2^*) = \int_{q \cdot (h_2 + p \cdot a)}^{\infty} [g - q \cdot (h_2 + p \cdot a)] \cdot z(g)\, dg. \tag{29}$$

We cannot simplify (29) anymore without knowing the exact $z(g)$. At least, we can compare the masses of $W_{1,n}$ and $W_{1,r}(s_2^*)$. On the one hand, $W_{1,n}$ has a purely negative mass between $g = 0$ and $g = q \cdot h_2$, and it has a purely positive mass between $g = q \cdot h_2$ and $g = \infty$. On the other hand, $W_{1,r}(s_2^*)$ only has positive mass. We can conclude that, if $z(g)$ is concentrated below $g = q \cdot h_2$ –i.e., due to strong bias towards gains for the offender that are smaller than the new weighted harm h_2–, and/or a is small or p is small, then $W_{1,r}(s_2^*) > W_{1,n}$, which means that retroactivity of *ex post facto* law is optimal. Otherwise, $W_{1,n} > W_{1,r}(s_2^*)$, this is, no retroactivity of *ex post facto* law is optimal.

Proposition 6 *A trade-off exists between deterrence and minimization of adjudication costs when conduct is criminalized and the cases have not yet been adjudicated.*

[16] See, for instance, Garoupa (1997) and Polinsky and Shavell (2000).

Scenario 2: decriminalization of conduct, and the case has not been adjudicated yet

In this scenario, $h_1, s_1 > 0$, $h_2 = 0$. If $s_2 = s_1$, i.e., no retroactivity of *lex mitior*, the critical threshold of the gain for committing the harmful act is $\tilde{g} = p \cdot s_1$. Then, social welfare is

$$W_{2,n} = \int_{p \cdot s_1}^{\infty} [g - (1-q) \cdot (h_1 + p \cdot a) - q \cdot p \cdot a] \cdot z(g)\, dg. \tag{30}$$

We can find s_1^* using F.O.C.,

$$\begin{aligned} \tfrac{\delta W_{2,n}}{\delta f_1} = 0 \Leftrightarrow \quad & -\{[p \cdot s_1 - (1-q) \cdot (h_1 + p \cdot a) - \\ & -q \cdot p \cdot a] \cdot z(p \cdot s_1)\} \cdot p = 0 \end{aligned} \tag{31}$$

where p, $z(p \cdot s_1) \neq 0$, therefore,

$$s_1^* = \frac{(1-q) \cdot (h_1 + p \cdot a) + q \cdot p \cdot a}{p}, \tag{32}$$

into (30):

$$W_{2,n}(s_1^*) = \int_{(1-q) \cdot (h_1 + p \cdot a) + q \cdot p \cdot a}^{\infty} [g - (1-q) \cdot (h_1 + p \cdot a) - q \cdot p \cdot a] \cdot z(g)\, dg. \tag{33}$$

If $s_2 = 0$, i.e., retroactivity of *lex mitior*, then $\tilde{g} = p \cdot s_1 \cdot (1-q)$. Hence, social welfare is

$$W_{2,r} = \int_{p \cdot s_1 \cdot (1-q)}^{\infty} [g - (1-q) \cdot (h_1 + p \cdot a)] \cdot z(g)\, dg. \tag{34}$$

We can find s_1^* using F.O.C.,

$$\begin{aligned} \tfrac{\delta W_{2,r}}{\delta f_1} = 0 \Leftrightarrow \quad & -\{[p \cdot s_1 \cdot (1-q) - (1-q) \cdot (h_1 \\ & +p \cdot a)] \cdot z(p \cdot s_1 \cdot (1-q))\} \cdot p \cdot (1-q) = 0 \end{aligned} \tag{35}$$

where $p \cdot (1-q)$, $z(p \cdot s_1 \cdot (1-q)) \neq 0$, therefore,

$$s_1^* = \frac{(1-q) \cdot (h_1 + p \cdot a)}{p \cdot (1-q)}, \tag{36}$$

into (34):

$$W_{2,r}(s_1^*) = \int_{(1-q) \cdot (h_1 + p \cdot a)}^{\infty} [g - (1-q) \cdot (h_1 + p \cdot a)] \cdot z(g)\, dg. \tag{37}$$

Both $W_{2,n}(s_1^*)$ and $W_{2,r}(s_1^*)$ have purely positive mass, but the integrand of $W_{2,n}(s_1^*)$ is always below the integrand of $W_{2,r}(s_1^*)$. Additionally, $W_{2,r}(s_1^*)$ always includes the summation from $(1-q) \cdot (h_1 + p \cdot a)$, while $W_{2,n}(s_1^*)$ only from $(1-q) \cdot (h_1 + p \cdot a) + q \cdot p \cdot a$. Hence, it is always true that $W_{2,r}(s_1^*) > W_{2,n}(s_1^*)$, i.e., retroactivity of *lex mitior* is optimal regarding decriminalization of conduct when the case has not yet been adjudicated.

> ***Proposition 7*** *Retroactivity of* lex mitior *is optimal when conduct is decriminalized and the cases have not yet been adjudicated.*

Scenario 3: decriminalization of conduct, and the case has already been adjudicated

In this scenario, $h_1, s_1 > 0$, $h_2 = 0$. If $s_2 = s_1$, i.e., no retroactivity of *lex mitior*, the critical threshold of the gain for committing the harmful act is $\tilde{g} = p \cdot s_1$. Then, social welfare is like (33), so

$$W_{3,n}\left(s_1^*\right) = \int_{(1-q)\cdot(h_1+p\cdot a)+q\cdot p\cdot a}^{\infty} \left[g - (1-q)\cdot(h_1 + p\cdot a) - q\cdot p\cdot a\right]\cdot z\left(g\right) dg. \tag{38}$$

If $s_2 = 0$, i.e., retroactivity of *lex mitior*, then $\tilde{g} = p \cdot s_1 \cdot (1-q)$. Hence, social welfare is

$$W_{3,r} = \int_{p\cdot s_1\cdot(1-q)}^{\infty} \left[g - (1-q)\cdot(h_1 + p\cdot a) - q\cdot p\cdot 2\cdot a\right]\cdot z\left(g\right) dg. \tag{39}$$

We can find s_1^* using F.O.C.,

$$\begin{aligned} \tfrac{\delta W_{3,r}}{\delta f_1} = 0 \Leftrightarrow \quad & -\{[p\cdot s_1\cdot(1-q) - (1-q)\cdot(h_1 + p\cdot a) \\ & -q\cdot p\cdot 2\cdot a]\cdot z\left(p\cdot s_1\cdot(1-q)\right)\}\cdot p\cdot(1-q) = 0 \end{aligned} \tag{40}$$

where $p\cdot(1-q)$, $z\left(p\cdot s_1\cdot(1-q)\right) \neq 0$, thus,

$$s_1^* = \frac{(1-q)\cdot(h_1 + p\cdot a) + q\cdot p\cdot 2\cdot a}{p\cdot(1-q)}, \tag{41}$$

into (39):

$$W_{3,r}\left(s_1^*\right) = \int_{(1-q)\cdot(h_1+p\cdot a)+q\cdot p\cdot 2\cdot a}^{\infty} \left[g - (1-q)\cdot(h_1 + p\cdot a) - q\cdot p\cdot 2\cdot a\right]\cdot z\left(g\right) dg. \tag{42}$$

Both $W_{3,n}\left(s_1^*\right)$ and $W_{3,r}\left(s_1^*\right)$ have purely positive mass, but $q\cdot p\cdot 2\cdot a > q\cdot p\cdot a$, hence, the integrand of $W_{3,r}\left(s_1^*\right)$ is always below the integrand of $W_{3,n}\left(s_1^*\right)$, and $W_{3,n}\left(s_1^*\right)$ includes the summation from $(1-q)\cdot(h_1 + p\cdot a) + q\cdot p\cdot a$, while $W_{3,r}\left(s_1^*\right)$ only from $(1-q)\cdot(h_1 + p\cdot a) + q\cdot p\cdot 2\cdot a$. Therefore, $W_{3,n}\left(s_1^*\right) > W_{3,r}\left(s_1^*\right)$, i.e., no retroactivity of *lex mitior* is optimal regarding decriminalization of conduct when the case has already been adjudicated.

> ***Proposition 8*** *Non-retroactivity of* lex mitior *is optimal when conduct is decriminalized and the cases have already been adjudicated.*

Scenario 4: the new law either aggravates or attenuates criminal liability, and the case has not been adjudicated yet

In this scenario, $h_1, h_2, s_1, s_2 > 0$. If $s_2 = s_1$, i.e., no retroactivity, the critical threshold of the gain for committing the harmful act is $\tilde{g} = p \cdot s_1$. Then, social welfare is

$$W_{4,n} = \int_{p\cdot s_1}^{\infty} \left[g - (1-q)\cdot(h_1 + p\cdot a) - q\cdot(h_2 + p\cdot a)\right]\cdot z\left(g\right) dg. \tag{43}$$

We can find s_1^* using F.O.C.,

$$\begin{aligned} \tfrac{\delta W_{4,n}}{\delta f_1} = 0 \Leftrightarrow \quad & -\{[p\cdot s_1 - (1-q)\cdot(h_1 + p\cdot a) \\ & -q\cdot(h_2 + p\cdot a)]\cdot z\left(p\cdot s_1\right)\}\cdot p = 0 \end{aligned} \tag{44}$$

where p, $z\left(p\cdot s_1\right) \neq 0$, thus,

$$s_1^* = \frac{(1-q)\cdot(h_1+p\cdot a)+q\cdot(h_2+p\cdot a)}{p}, \tag{45}$$

into (43):

$$W_{4,n}(s_1^*) = \int_{(1-q)\cdot(h_1+p\cdot a)+q\cdot(h_2+p\cdot a)}^{\infty} [g-(1-q)\cdot(h_1+p\cdot a)-q\cdot(h_2+p\cdot a)]\cdot z(g)\,dg. \tag{46}$$

If $s_2 \neq s_1$, i.e., retroactivity, then $\tilde{g} = p\cdot[s_1\cdot(1-q)+s_2\cdot q]$. Then, social welfare is

$$W_{4,r} = \int_{p\cdot[s_1\cdot(1-q)+s_2\cdot q]}^{\infty} [g-(1-q)\cdot(h_1+p\cdot a)-q\cdot(h_2+p\cdot a)]\cdot z(g)\,dg. \tag{47}$$

Using F.O.C.,

$$\begin{aligned} \tfrac{\delta W_{4,r}}{\delta f_1} = 0 \Leftrightarrow \quad & -\{[p\cdot[s_1\cdot(1-q)+s_2\cdot q]-(1-q)\cdot(h_1+p\cdot a) \\ & -q\cdot(h_2+p\cdot a)]\cdot z(p\cdot[s_1\cdot(1-q)+s_2\cdot q])\}\cdot p\cdot(1-q)=0 \end{aligned} \tag{48}$$

$$\begin{aligned} \tfrac{\delta W_{4,r}}{\delta f_2} = 0 \Leftrightarrow \quad & -\{[p\cdot[s_1\cdot(1-q)+s_2\cdot q]-(1-q)\cdot(h_1+p\cdot a) \\ & -q\cdot(h_2+p\cdot a)]\cdot z(p\cdot[s_1\cdot(1-q)+s_2\cdot q])\}\cdot p\cdot q=0 \end{aligned} \tag{49}$$

where p,$(1-q)$, q, $z(p\cdot[s_1\cdot(1-q)+s_2\cdot q]) \neq 0$, thus,

$$\frac{\delta W_{4,r}}{\delta f_1} = \frac{\delta W_{4,r}}{\delta f_2} \Leftrightarrow s_1^*\cdot(1-q)+s_2^**q = \frac{(1-q)\cdot(h_1+p\cdot a)+q\cdot(h_2+p\cdot a)}{p}, \tag{50}$$

into (47):

$$W_{4,r}(s_1^*,s_2^*) = \int_{(1-q)\cdot(h_1+p\cdot a)+q\cdot(h_2+p\cdot a)}^{\infty} [g-(1-q)\cdot(h_1+p\cdot a)-q\cdot(h_2+p\cdot a)]\cdot z(g)\,dg. \tag{51}$$

Under this scenario, $W_{4,n}(s_1^*) = W_{4,r}(s_1^*,s_2^*)$, therefore, society is indifferent between retroactivity and no retroactivity.

Proposition 9 *Retroactive application of a law which aggravates or attenuates criminal liability has no effect on social welfare, provided that the case has not yet been adjudicated.*

Scenario 5: the new law either aggravates or attenuates criminal liability, and the case has already been adjudicated

In this scenario, $h_1, h_2, s_1, s_2 > 0$. If $s_2 = s_1$, i.e., no retroactivity, the critical threshold of the gain for committing the harmful act is $\tilde{g} = p\cdot s_1$. Then, social welfare is like (43), so

$$W_{5,n}(s_1^*) = \int_{(1-q)\cdot(h_1+p\cdot a)+q\cdot(h_2+p\cdot a)}^{\infty} [g-(1-q)\cdot(h_1+p\cdot a)-q\cdot(h_2+p\cdot a)]\cdot z(g)\,dg. \tag{52}$$

If $s_2 \neq s_1$, i.e., retroactivity, then $\tilde{g} = p\cdot[s_1\cdot(1-q)+s_2\cdot q]$. Then, social welfare is

$$W_{5,r} = \int_{p\cdot[s_1\cdot(1-q)+s_2\cdot q]}^{\infty} [g-(1-q)\cdot(h_1+p\cdot a)-q\cdot(h_2+p\cdot 2\cdot a)]\cdot z(g)\,dg. \tag{53}$$

Using F.O.C.,

$$\begin{aligned}
\tfrac{\delta W_{5,r}}{\delta f_1} = 0 \Leftrightarrow \quad & -\{[p\cdot[s_1\cdot(1-q)+s_2\cdot q]-(1-q)\cdot(h_1+p\cdot a) \\
& -q\cdot(h_2+p\cdot 2\cdot a)]\cdot z\left(p\cdot[s_1\cdot(1-q)+s_2\cdot q]\right)\}\cdot p\cdot(1-q)=0 \qquad (54)\\
\tfrac{\delta W_{5,r}}{\delta f_2} = 0 \Leftrightarrow \quad & -\{[p\cdot[s_1\cdot(1-q)+s_2\cdot q]-(1-q)\cdot(h_1+p\cdot a) \\
& -q\cdot(h_2+p\cdot 2\cdot a)]\cdot z\left(p\cdot[s_1\cdot(1-q)+s_2\cdot q]\right)\}\cdot p\cdot q=0 \qquad (55)
\end{aligned}$$

where p,$(1-q)$, q, $z\left(p\cdot[s_1\cdot(1-q)+s_2\cdot q]\right)\neq 0$, thus,

$$\frac{\delta W_{5,r}}{\delta f_1}=\frac{\delta W_{5,r}}{\delta f_2}\Leftrightarrow s_1^*\cdot(1-q)+s_2^**q=\frac{(1-q)\cdot(h_1+p\cdot a)+q\cdot(h_2+p\cdot 2\cdot a)}{p}, \qquad (56)$$

into (53):

$$W_{5,r}\left(s_1^*,s_2^*\right)=\int_{(1-q)\cdot(h_1+p\cdot a)+q\cdot(h_2+p\cdot 2\cdot a)}^{\infty}[g-(1-q)\cdot(h_1+p\cdot a)-q\cdot(h_2+p\cdot 2\cdot a)]\cdot z(g)\,dg. \qquad (57)$$

Both $W_{5,n}\left(s_1^*\right)$ and $W_{5,r}\left(s_1^*,s_2^*\right)$ have purely positive mass, but $q\cdot p\cdot 2\cdot a > q\cdot p\cdot a$, hence, the integrand of $W_{5,r}\left(s_1^*,s_2^*\right)$ is always below the integrand of $W_{5,n}\left(s_1^*\right)$, and $W_{5,n}\left(s_1^*\right)$ includes the summation from $(1-q)\cdot(h_1+p\cdot a)+q\cdot(h_2+p\cdot a)$, while $W_{5,r}\left(s_1^*,s_2^*\right)$ only from $(1-q)\cdot(h_1+p\cdot a)+q\cdot(h_2+p\cdot 2\cdot a)$. Therefore, $W_{5,n}\left(s_1^*\right) > W_{5,r}\left(s_1^*,s_2^*\right)$, i.e., no retroactivity is optimal if the case has already been adjudicated.

> ***Proposition 10*** *Retroactive application of a law which aggravates or attenuates criminal liability has a negative effect on social welfare, provided that the case has already been adjudicated.*

Summing up, the retroactivity of criminal laws sometimes entails additional adjudication costs, since it implies that a case that would not have been tried, or that has already been tried under the previous law, will be tried for the first time or retried under the new law. This happens when a conduct is criminalized *ex novo* and the corresponding cases have not yet been adjudicated (scenario 1), and when the new law aggravates or mitigates criminal liability in cases that have already been adjudicated (scenario 5). Under these circumstances, the non-retroactivity of the corresponding law makes it possible to avoid such additional adjudication costs while maintaining the optimal level of deterrence. As we have seen in section 3, this optimal level can be achieved under a rule of non-retroactivity by adjusting the initial sanction s_1. Therefore, in these cases, the socially optimal solution is non-retroactivity.

Scenario 1 (criminalization of conduct) deserves special consideration. Here, the new law imposes a positive sanction on actions or omissions for which the previous law provided a sanction $s_1 = 0$. Here, it would be possible to avoid the adjudication costs that retroactivity of the new law entails while maintaining the same (optimal) level of deterrence. To achieve this, the initial sanction s_1 would have to be set above zero. But, obviously, in that case we would no longer be in scenario 1, where $s_1 = 0$.

According to our model, scenario 1 constitutes a suboptimal solution. If adjudication is costly, the optimal solution is to establish an initial sanction greater than zero and to preclude the retroactive application of the new law. Nevertheless, in real life there are surely circumstances (not captured by our model) that justify the legislator setting a zero sanction for many types of conduct, despite the probability that such conduct may later prove harmful and become criminalized —for example, the costs of enacting and enforcing the corresponding laws may be excessive. Indeed, in practice scenario 1 is frequently observed.

Note that in scenario 1 there is a trade-off. Non-retroactivity of the *ex post facto* law avoids adjudication costs but weakens deterrence. If the resulting loss of deterrence outweighs the adjudication savings —that is, when h_2 and q are relatively high, and g, a, and p are relatively small— retroactivity becomes optimal.

In other circumstances, the retroactivity of criminal laws does not entail additional adjudication costs. This occurs when the new law either increases or decreases the sanction prescribed for certain conduct and the corresponding cases have not yet been adjudicated (scenario 4). In such situation, retroactivity does not generate additional adjudication costs, since the cases in question would have to be adjudicated in any event, even if the new law were not applied retroactively. Under these circumstances, both retroactivity and non-retroactivity can be socially optimal.

There is only one situation in which the retroactivity of a criminal law reduces adjudication costs: when conduct is decriminalized and the corresponding cases have not yet been adjudicated (scenario 2). In this situation, the retroactive application of the decriminalizing law implies that cases which would otherwise have to be adjudicated under the prior law no longer require adjudication. In such circumstances, the optimal solution —which avoids those adjudication costs while preserving the optimal level of deterrence— is to apply the new law retroactively.

7 Constitutional protection against excessive punishment

In this section, we assess the need to establish—through constitutional provisions—the prohibition of *ex post facto* laws and the *lex mitior* principle in order to prevent public agents (and, in particular, legislatures) from abusing their sanctioning power. We examine specifically whether these rules are necessary when the legislature has punitive preferences because the private costs and benefits of its decisions are misaligned with social welfare.

Let h^{soc} be the social harm generated by the harmful act. We propose different sequential games with complete information for *ex post facto* laws and *lex mitior*.

Ex post facto laws

The sequence is the following:

- **Time 1.** The constitution-maker chooses between three rules: (i) retroactivity is allowed in any case; (ii) retroactivity is always prohibited; or (iii) retroactivity of lex mitior is allowed, but *ex post facto* laws are prohibited.

- **Time 2.** A benevolent legislature, that tries to maximize social welfare, establishes the socially costless sanction s_1 for the individuals who commit the harmful act under consideration.

- **Time 3.** Individuals decide whether to commit the act. An individual commits the act if their private gain, g, exceeds the expected sanction $\tilde{g}$.

- **Time 4.** Elections are held. With probability q, a new political party wins. If the incumbent party wins, the law remains s_1. If the new party wins, it sets a socially costless sanction s_2 for acts committed before. When determining the magnitude of this sanction, the new party maximizes its own private utility, which does not coincide with social welfare. The difference lies in the fact that, in its private utility function, the harm caused by the act $\left(h^{priv}\right)$ is overestimated, i.e., $h^{priv} > h^{soc}$. The model would generate analogous results if the difference between the social welfare function and that private utility function lied in the fact that the new political party does not fully internalize the gains individuals derive from committing the harmful act. However, we assume for the sake of simplicity that the new political party fully internalizes g.

The game is solved by backwards induction:

Time 4. If the incumbent party wins the election, s_1 remains. If retroactivity is fully allowed and the new political party wins, it imposes retroactively the socially costless sanction $s_{2,r}$ that maximizes its own utility U_r, where

$$U_r = \int_{\tilde{g}}^{\infty} \left[g - h^{priv}\right] \cdot z\left(g\right) dg, \tag{58}$$

that is,

$$s_{2,r}^{*} = \frac{\frac{h^{priv}}{p} - (1-q) \cdot s_1}{q}, \tag{59}$$

which leads to $\tilde{g} = h^{priv}$. In other words, the new party sets the sanction $s_{2,r}^{*}$ so that individuals commit the act only when the benefits derived from it exceed $\tilde{g} = h^{priv}$.

Note that the new party aggravates criminal liability retroactively. As we will see, the socially optimal sanction is $s_2 = s_1 = h^{soc}/p$. If we assume that the benevolent legislature establishes this optimal sanction $s_{1,r} = h^{soc}/p$, then, $s_{2,r}^{*} > s_{1,r}$. Indeed,

$$s_{2,r}^{*} - s_{1,r} = \frac{h^{priv} - h^{soc}}{p \cdot q} > 0$$

since $h^{priv} > h^{soc}$.

If retroactivity is fully prohibited and the new party wins, it cannot change the sanction for past acts, so $s_{2,n}^{*} = s_{1,n}$.

If retroactivity is prohibited only in cases of *ex post facto* laws and the new party wins, it is constrained to $s_{2,l} \leq s_{1,l}$. If the initial sanction was socially optimal, $s_{1,l} = h^{soc}/p$, then the new party will set $s_{2,l}^{*} = s_{1,l}$ because, in this case, its private utility would be

$$U_r = \int_{h^{soc}}^{\infty} \left[g - h^{priv}\right] \cdot z\left(g\right) dg, \tag{60}$$

while, if it establishes a lower penalty, $s_{2,l} < s_{1,l}$, its private utility would be

$$U_r = \int_{\tilde{g}}^{h^{soc}} \left[g - h^{priv}\right] \cdot z\left(g\right) dg + \int_{h^{soc}}^{\infty} \left[g - h^{priv}\right] \cdot z\left(g\right) dg. \tag{61}$$

One can see that, since $g \leq h^{soc} < h^{priv}$, the first integral is negative. Therefore, the private utility of the new party is lower than when $s_{2,l}^{*} = s_{1,l}$. The new party would prefer to impose a higher sanction, but the prohibition of *ex post facto* laws prevents it from doing so. Setting a lower sanction ($s_{2,l} < s_{1,l}$) would induce more harmful acts (where $g < h^{priv}$), reducing its utility. Conclusion: if retroactivity is prohibited only in cases of *ex post facto* laws and the new party wins, it will impose $s_{2,l}^{*} = s_{1,l}$.

Time 3. Individuals commit harmful acts if $g > \tilde{g} = p \cdot [s_1 \cdot (1-q) + s_2 \cdot q]$. If $s_2 = s_1$ (the legal change does not have retroactive effect), the expected sanction is $p \cdot s_1$, that is, $\tilde{g} = p \cdot s_1$.

Time 2. The benevolent legislature chooses s_1 to maximize social welfare W, where

$$W = \int_{\tilde{g}}^{\infty} \left[g - h^{soc}\right] \cdot z\left(g\right) dg. \tag{62}$$

If retroactivity is fully prohibited, then $s_{2,n}^{*} = s_{1,n}$. In this scenario, the legislature chooses the socially optimal sanction

$$s^*_{1,n} = \frac{h^{soc}}{p}. \tag{63}$$

If retroactivity is allowed only in cases of *lex mitior*, the outcome is the same,

$$s^*_{1,l} = \frac{h^{soc}}{p}. \tag{64}$$

As we have seen, in this scenario the new party will establish the same penalty provided for by the incumbent party: $s^*_{2,l} = s_{1,l}$. Given that, the incumbent choses the $s_{1,l}$ that maximizes social welfare, this is, $s^*_{1,l}$.

If retroactivity is always allowed, every $s_{1,r}$ produces the same level of social welfare. Whatever $s_{1,r}$ is chosen by the incumbent party, the risk of the new party winning the elections, observing $s_{1.r}$ and setting $s^*_{2,r}$ allows this party to make the expected sanction equal to h^{priv}.

Time 1. The constitution-maker decides after comparing the resulting social welfare under each of the three rules. If retroactivity is fully forbidden or prohibited only in the case of *ex post facto* laws, social welfare will be

$$W_n = \int_{h^{soc}}^{\infty} [g - h^{soc}] \cdot z\,(g)\, dg. \tag{65}$$

If retroactivity is always allowed, social welfare will be

$$W_r = \int_{h^{priv}}^{\infty} [g - h^{soc}] \cdot z\,(g)\, dg. \tag{66}$$

Since $h^{priv} > h^{soc}$, it follows that $W_n > W_r$.

It must be underlined that, to maximize social welfare, the constitution-maker does not need to prohibit the retroactivity of *lex mitior* (i.e., laws that attenuate criminal liability), but only *ex post facto* laws (i.e., laws that retroactively increase sanctions). In this scenario, the *lex mitior* principle is irrelevant, as the new party has no incentive to attenuate criminal liability. The key is prohibiting retroactive increases in sanctions.

> ***Proposition 11*** *If the legislature does not fully internalize the social costs of sanctions, then in order to maximize social welfare, the constitution-maker must forbid retroactive increases in sanctions.*

Lex mitior (I)

This is a sequential game with complete information:

- **Time 1.** The constitution-maker chooses between the following rules: (i) retroactivity is allowed, but *ex post facto* laws are prohibited; (ii) retroactivity is mandatory for *lex mitior*, but *ex post facto* laws are prohibited; (iii) retroactivity is prohibited in all cases.
- **Time 2.** The legislature sets the initial sanction s_1 for any individual who commits some harmful act. When determining s_1, the legislature tries to maximize its own private utility, thereby overestimating the harm caused by the harmful act $\left(h^{priv}\right)$, so that $h^{priv} > h^{soc}$.
- **Time 3.** Individuals decide whether to commit the act. An individual commits the act if their private gain, g, exceeds the expected sanction $\tilde{g}$.
- **Time 4.** Elections are held. With probability q, a new political party wins. If the incumbent party wins, the law remains s_1. If the new party wins, it establishes a sanction s_2 for the people who committed the harmful act. When determining this sanction, the new party tries to maximize social welfare.

The game is solved by backwards induction:

Time 4. If the incumbent party wins the election, s_1 remains. If retroactivity is either mandatory or just allowed (save in cases of *ex post facto* laws, that are prohibited) and the new party wins, it imposes the sanction $s_{2,r}$ hat maximizes social welfare W_r where

$$W_r = \int_{\tilde{g}}^{\infty} [g - h^{soc}] \cdot z(g)\, dg. \tag{67}$$

Thus, new party's best response function is

$$s^*_{2,r} = \frac{\frac{h^{soc}}{p} - (1-q) \cdot s_1}{q}. \tag{68}$$

Note that $s^*_{2,r}$ is so that $\tilde{g} = h^{soc}$. In other words, the new party sets the sanction $s^*_{2,r}$ so that individuals commit the harmful act only when the gains derived from it exceed the social harm caused by this act, $g > h^{soc}$.

If retroactivity is always forbidden, $s^*_{2,n} = s_{1,n}$.

Time 3. Individuals commit the harmful act if the gain from it exceeds the expected sanction $\tilde{g} = p \cdot [s_1 \cdot (1-q) + s_2 \cdot q]$.

Time 2. The incumbent party tries to impose the sanction s_1 that maximizes its own utility U, where

$$U = \int_{\tilde{g}}^{\infty} \left[g - h^{priv}\right] \cdot z(g)\, dg. \tag{69}$$

If retroactivity is fully prohibited, then $s^*_{2,n} = s_{1,n}$, and the incumbent party chooses

$$s^*_{1,n} = \frac{h^{priv}}{p}. \tag{70}$$

Note that this $s^*_{1,n}$ is higher than the sanction that would maximize social welfare, which is $s_1 = h^{soc}/p$.

If retroactivity is either mandatory or just allowed (save in cases of *ex post facto* laws, that are prohibited), the new party is constrained to $s_{2,r} \leq s_{1,r}$. Given that, the incumbent party will never set $s_{1,r} < h^{soc}/p$ because if the incumbent party sets $s_{1,r} = h^{soc}/p$, the new party will not change it $\left(s^*_{2,r} = s_{1,r}\right)$, as this is the sanction that maximizes social welfare. However, if the incumbent party sets a lower sanction $s_{1,r} < h^{soc}/p$, the new party will not change it $\left(s^*_{2,r} = s_{1,r}\right)$. The new party would prefer to establish a higher sanction, but the prohibition of *ex post facto* laws prevents it from doing so. Given that, the incumbent party will not set a sanction lower than the socially optimal one. If it chooses $s_{1,r} = h^{soc}/p$, its private utility will be

$$U_r = \int_{h^{soc}}^{\infty} \left[g - h^{priv}\right] \cdot z(g)\, dg. \tag{71}$$

If, on the contrary, it establishes a lower sanction $(s_{1,r} < h^{soc}/p,)$, its private utility would be:

$$U_r = \int_{p \cdot s_{1,r}}^{h^{soc}} \left[g - h^{priv}\right] \cdot z(g)\, dg + \int_{h^{soc}}^{\infty} \left[g - h^{priv}\right] \cdot z(g)\, dg. \tag{72}$$

Since $g \leq h^{soc} < h^{priv}$, the first integral is negative. Therefore, the private utility of the incumbent party is lower than when $s_{1,r} = h^{soc}/p$. Setting a lower sanction would induce more harmful acts (where $g < h^{priv}$), thereby reducing its utility.

In this scenario where retroactivity is either mandatory or just allowed (save in cases of ex *post facto* laws, that are prohibited), every $s_{1,r} \geq h^{soc}/p$ produces the same level of private utility for the incumbent party. Whatever $s_{1,r}$ is chosen by it, the possibility of the new party winning the elections, observing $s_{1,r}$ and setting $s_{2,r}$ following its best response function $\left(s^*_{2,r}\right)$ with retroactive effect allows this party to make the expected sanction equal to h^{soc}. It must be noted, nevertheless, that if the incumbent party chooses a sanction higher than the socially optimal one, $s_{1,r} > h^{soc}/p$, then the new party passes a *lex mitior* with retroactive effect, as the new sanction is milder than the incumbent one, $s^*_{2,r} < s_{1,r}$. Indeed,

$$s_{1,r} - s^*_{2,r} = \frac{s_{1,r} - \frac{h^{soc}}{p}}{q} > 0.$$

Time 1. The constitution-maker compares the resulting social welfare to choose a rule. If retroactivity is fully forbidden, social welfare will be

$$W_n = \int_{h^{priv}}^{\infty} [g - h^{soc}] \cdot z\,(g)\, dg. \tag{73}$$

If retroactivity of *lex mitior* is either allowed or mandatory (but ex *post facto* laws are forbidden), social welfare will be

$$W_r = \int_{h^{soc}}^{\infty} [g - h^{soc}] \cdot z\,(g)\, dg. \tag{74}$$

Since $h^{priv} > h^{soc}$, social welfare is higher in the latter scenario: $W_r > W_n$. That is, social welfare is higher when retroactivity of *lex mitior* is allowed or mandatory than when it is prohibited, because the lower cutoff h^{soc} leads to more socially efficient acts (whose benefits exceed their costs, $g - h^{soc} > 0$).

Two results should be underlined. First, to maximize social welfare, the constitution-maker does not need to impose on the legislature the obligation to give retroactive effect to the laws that attenuate criminal liability; to achieve this goal, it just needs to allow the legislature to give retroactive effect to such laws. Second, this rule is fully compatible with the prohibition of ex *post facto* laws.

> ***Proposition 12*** *If the legislature does not fully internalize the social costs of sanctions and* ex post facto *laws are prohibited, then to maximize social welfare it is enough for the constitution-maker to allow retroactive reductions in sanctions.*

Lex mitior (II)

This is a sequential game with complete information:

- **Time 1.** The constitution-maker chooses between the following rules: (i) retroactivity is allowed, but *ex post facto* laws are prohibited; (ii) retroactivity is mandatory for *lex mitior*, but *ex post facto* laws are prohibited; (iii) retroactivity is prohibited in all cases.

- **Time 2.** The legislature sets a socially costly sanction s_1 for committing some potentially harmful act. When setting s_1, the legislature tries to maximize its own utility. At this time, nobody knows with certainty the magnitude of the harm; individuals can only estimate that such act will cause harm h_1 with probability $(1 - q)$ or harm h_2 with probability q, where $h_1 > h_2$. They can also anticipate that if the act causes h_1 there will be no legal change and those who commit such act will receive the sanction there will be no legal change and those who commit such act will receive the sanction s_1, while if the act causes h_2 the legislature will set the sanction s_2 for them.

- **Time 3.** Individuals decide whether to commit the act. An individual commits the act if their private gain, g, exceeds the expected sanction $\tilde{g} = p \cdot [s_1 \cdot (1-q) + s_2 \cdot q]$.
- **Time 4.** Uncertainty is resolved and everybody is able to know whether the act causes h_1 or h_2. If it causes h_2 the legislature sets the sanction s_2 for those who committed the act in the past and sanction s_3 for those who will commit it in the future. When setting both sanctions, the legislature tries to maximize its own utility. Establishing a milder law with retroactive effect ($s_1 > s_2 = s_3$) has a political cost c for the legislature. Let us assume that such cost derives from the fact that, unlike the authors and the victims of future crimes, both the authors and the victims of crimes already committed are "identifiable", and, *ceteris paribus*, people tend to be more punitive (sympathetic) towards identifiable wrongdoers (victims) than toward unidentified wrongdoers (victims) —see Jenni and Loewenstein (1997) and Small and Loewenstein (2005).
- **Time 5.** Individuals decide whether to commit the act. An individual commits the act if their private gain g from it exceeds the expected sanction after uncertainty is resolved, $\tilde{\tilde{g}}$. Let us suppose, for simplicity, that the population is the same before and after the uncertainty is resolved (i.e. in both the first and the second periods).

The game is solved by backwards induction:

Time 5. Individuals decide whether to commit the act. An individual commits the act if their private gain g from it exceeds the expected sanction for the second period, $\tilde{\tilde{g}} = p \cdot s_3$.

Time 4. If the act causes h_1, then the law does not change and s_1 remains, which is equivalent to $s_3 = s_2 = s_1$.

If the act causes h_2, the legislature sets s_2 and s_3, trying to maximize its own utility. If the retroactivity of *lex mitior* is mandatory and $s_3 < s_1$, then $s_2 = s_3$. The legislature has then two alternatives. First, it can set a *milder sanction with retroactive effect* ($s_{1,r} > s_{2,r} = s_{3,r}$), in which case its private utility will be:

$$U_r = \int_{p\cdot[(1-q)\cdot s_{1,r}+q\cdot s_{2,r}]}^{\infty} [g - (1-q)\cdot h_1 - q\cdot(h_2+c)]\cdot z(g)\,dg + \int_{p\cdot s_{2,r}}^{\infty} [g-h_2]\cdot z(g)\,dg. \tag{75}$$

To maximize U_r, the legislature would set

$$s^*_{1,r} = \frac{h_1}{p} + \frac{q\cdot c}{p\cdot(1-q)} \tag{76}$$

and

$$s^*_{2,r} = \frac{h_2}{p}. \tag{77}$$

Second, it may choose *not to change the law*, i.e. to keep the *status quo* ($s_{3,sq} = s_{2,sq} = s_{1,sq}$), in which case its private utility will be:

$$U_{sq} = \int_{p\cdot s_{1,sq}}^{\infty} [g - (1-q)\cdot h_1 - q\cdot h_2]\cdot z(g)\,dg + \int_{p\cdot s_{1,sq}}^{\infty} [g-h_2]\cdot z(g)\,dg. \tag{78}$$

To maximize U_{sq}, the legislature would set

$$s^*_{1,sq} = \frac{(1-q)\cdot h_1 + (1-q)\cdot h_2}{2\cdot p}. \tag{79}$$

Notice that $U_r\left(s_{1,r}^*, s_{2,r}^*\right) \gtrless U_{sq}\left(s_{1,sq}^*\right)$, but the larger (smaller) c and q, the more (less) probable that $U_{sq}\left(s_{1,sq}^*\right) > U_r\left(s_{1,r}^*, s_{2,r}^*\right)$.

If the act causes h_2 and the retroactivity of *lex mitior* is just allowed, then legislature has one additional alternative, besides the two abovementioned ones. It may also provide for a *milder sanction with no retroactive effect*, that is, $s_{1,n} = s_{2,n} < s_{3,n}$, in which case its private utility will be:

$$U_n = \int_{p \cdot s_{1,n}}^{\infty} [g - (1-q) \cdot h_1 - q \cdot h_2] \cdot z(g)\, dg + \int_{p \cdot s_{3,n}}^{\infty} [g - h_2] \cdot z(g)\, dg. \tag{80}$$

To maximize U_n, the legislature would set

$$s_{1,n}^* = \frac{(1-q) \cdot h_1 + q \cdot h_2}{p} \tag{81}$$

and

$$s_{3,n}^* = \frac{h_2}{p}. \tag{82}$$

Note that, if retroactivity is allowed, the legislature never chooses not to change the law, as $U_n\left(s_{1,n}^*, s_{2,n}^*\right) > U_{sq}\left(s_{1,sq}^*\right)$.

It must also be noted that, if retroactivity of *lex mitior* is just allowed, the legislature never chooses to pass a milder law with retroactive effect, as $U_n\left(s_{1,n}^*, s_{2,n}^*\right) > U_r\left(s_{1,r}^*, s_{2,r}^*\right)$. Indeed, in this scenario, the sanction $s_{3,n}^*$ that maximizes its own private utility is so that the expected sanction for the second period is $\tilde{\tilde{g}} = p \cdot s_{3,n}^* = h_2$. For the first period, as we have already seen (baseline model) any level of deterrence that can be achieved by means of $s_1 \neq s_2$ (i.e., retroactivity) can also be achieved by means of $s_1 = s_2$ (i.e. non-retroactivity). Therefore, the level of deterrence that maximizes the private utility of the legislature during the first period if $s_1 \neq s_2$ (which is so that $\tilde{g} = p \cdot \left[s_{1,r}^* \cdot (1-q) + s_{2,r}^* \cdot q\right] = (1-q) \cdot h_1 + q\,(h_2 + c)$) can also be achieved by means of $s_1 = s_2$ (this is if $\tilde{g} = p \cdot s_{1,n}^* = (1-q) \cdot h_1 + q \cdot h_2$). And the latter solution has the advantage over the former one that it avoids the political cost of retroactivity c. Indeed,

$$\int_{p \cdot s_{1,n}^*}^{\infty} [g - (1-q) \cdot h_1 - q \cdot h_2] \cdot z(g)\, dg > \int_{p \cdot \left[(1-q) \cdot s_{1,r}^* + q \cdot s_{2,r}^*\right]}^{\infty} [g - (1-q) \cdot h_1 - q \cdot (h_2 + c)] \cdot z(g)\, dg. \tag{83}$$

Therefore, if the retroactivity of *lex mitior* is just allowed, the legislature will always pass a milder law with no retroactive effect.

If the act causes h_2 and *retroactivity is fully prohibited*, the legislature also provides for a *milder sanction with no retroactive effect*, this is $s_{1,n}^*$ and $s_{3,n}^*$.

If the act causes h_2 and the *retroactivity of lex mitior is mandatory*, the decision of the legislature will depend on the political cost of retroactivity c, and the probability of a legal change being needed q. If c and q are high enough, the utility of not changing the law will be greater than that of passing a retroactive milder law, $U_{sq}\left(s_{1,sq}^*\right) > U_r\left(s_{1,r}^*, s_{2,r}^*\right)$.

Time 3. Individuals decide whether to commit the act. An individual commits the act if their private gain g from it exceeds the expected sanction before uncertainty is resolved, $\tilde{g}$.

Time 2. The legislature sets the initial sanction s_1.

Time 1. The constitution-maker compares the resulting social welfare to choose a rule.

If the social welfare function coincides with the legislature's private utility function, social welfare will obviously be maximized if the retroactivity of *lex mitior* is just allowed. Under this rule, the legislature will pass a milder law without retroactive effect, which is the decision that maximizes both the legislature's utility and social welfare.

Interestingly, the result is the same if there is a difference between the legislature's private utility function and the social welfare function, being that the political cost of retroactivity is not included in the latter one. In this scenario, if retroactivity of *lex mitior* is mandatory and its political cost exceeds $\bar{c}$, the legislature will choose not to change the law $\left(s^*_{3,sq} = s^*_{2,sq} = s^*_{1,sq}\right)$, in which case social welfare will be

$$W_{sq} = \int_{p \cdot s^*_{1,sq}}^{\infty} \left[g - (1-q) \cdot h_1 - q \cdot h_2\right] \cdot z\left(g\right) dg + \int_{p \cdot s^*_{1,sq}}^{\infty} \left[g - h_2\right] \cdot z\left(g\right) dg. \tag{84}$$

If retroactivity of lex mitior is mandatory and its political cost does not reach $\bar{c}$, the legislature will pass a milder law with retroactive effect $\left(s^*_{1,r} > s^*_{2,r} = s^*_{3,r}\right)$, and social welfare will be

$$W_r = \int_{p \cdot \left[(1-q) \cdot s^*_{1,r} + q \cdot s^*_{2,r}\right]}^{\infty} \left[g - (1-q) \cdot h_1 - q \cdot h_2\right] \cdot z\left(g\right) dg + \int_{p \cdot s^*_{2,r}}^{\infty} \left[g - h_2\right] \cdot z\left(g\right) dg, \tag{85}$$

where $s^*_{1,r} > h_1/p$, thus, $s_{1,r}$ will be larger than it should be to maximize social welfare.

If retroactivity of *lex mitior* is just allowed or prohibited, the legislature will pass a milder law with no retroactive effect $\left(s^*_{1,n} = s^*_{2,n} < s^*_{3,n}\right)$, and social welfare will be

$$W_n = \int_{p \cdot s^*_{1,n}}^{\infty} \left[g - (1-q) \cdot h_1 - q \cdot h_2\right] \cdot z\left(g\right) dg + \int_{p \cdot s^*_{3,n}}^{\infty} \left[g - h_2\right] \cdot z\left(g\right) dg. \tag{86}$$

One can see that $W_n\left(s^*_{1,n}, s^*_{3,n}\right) > W_{sq}\left(s^*_{1,sq}\right)$, and $W_n\left(s^*_{1,n}, s^*_{3,n}\right) > W_r\left(s^*_{1,r}, s^*_{2,r}\right)$, hence, allowing retroactivity of *lex mitior* or *lex mitior* prohibited maximizes social welfare. Also, $W_r\left(s^*_{1,r}, s^*_{2,r}\right) \gtrless W_{sq}\left(s^*_{1,sq}\right)$, but if q was sufficiently large to make $U_{sq}\left(s^*_{1,sq}\right) > U_r\left(s^*_{1,r}, s^*_{2,r}\right)$, this difference is even larger for $W_r\left(s^*_{1,r}, s^*_{2,r}\right)$ and $W_{sq}\left(s^*_{1,sq}\right)$, i.e. $W_{sq}\left(s^*_{1,sq}\right) = U_{sq}\left(s^*_{1,sq}\right) > U_r\left(s^*_{1,r}, s^*_{2,r}\right) > W_r\left(s^*_{1,r}, s^*_{2,r}\right)$.

Indeed, when retroactivity of *lex mitior* is allowed or prohibited, the legislature will pass a milder law with no retroactive effect, setting $s^*_{1,n}$ and $s^*_{2,n}$. In this way, individuals commit the harmful during the first (second) period if and only if the gains from it exceed the expected harm. Note that these sanctions maximize not only the legislature's private utility function, but also the social welfare function, as both functions coincide with each other when the new law is not retroactive.

By contrast, when the retroactivity of *lex mitior* is mandatory, the sanctions set by the legislature do not maximize social welfare. If the legislature chooses not to change the law, the expected sanction for the first period, $p \cdot s^*_{1,sq}$, will be lower than the optimal one for the same period, $p \cdot s_1 = (1-q) \cdot h_1 + q \cdot h_2$. And the expected sanction for the second period, $p \cdot s^*_{2,sq} = p \cdot s^*_{1,sq}$, will be higher than the optimal one for this period, $p \cdot s_2 = h_2$.

If the legislature passes a milder law with retroactive effect, the sanction it sets for the second period is the socially optimal one: $s^*_{2,r} = s^*_{3,n} = h_2/p$. However, the expected sanction for the first period is harsher than the socially optimal expected sanction: $p \cdot \left[(1-q) \cdot s^*_{1,r} + q \cdot s^*_{2,r}\right] = (1-q) \cdot h_1 + q \cdot h_2 + q \cdot c > (1-q) \cdot h_1 + q \cdot h_2$.

Summing up, to maximize social welfare, the constitution-maker should provide that the retroactivity of laws reducing sanctions is either allowed or prohibited. Both rules will induce the legislature to pass such laws with no retroactive effects. The constitutional rule that laws reducing sanctions must have retroactive effect would induce the legislature to impose socially suboptimal sanctions.

> ***Proposition 13*** *If establishing a milder law with retroactive effect entails a political cost for the legislature, then the constitutional-maker, in order to maximize social welfare, should not mandate the retroactivity of laws reducing sanctions, but should merely allow it.*

8 Discussion

This paper analyzes how the social costs and benefits of the retroactivity of criminal laws vary depending on individuals' risk aversion, the costs of implementing sanctions, and the costs of adjudicating cases.

Our model shows that if individuals are risk-averse, the non-retroactivity of criminal laws is preferable to their retroactive application. The welfare-maximizing solution is that when criminal law changes as a result of a better assessment of the harm caused by certain conduct, the new law should not have retroactive effects. Since most individuals are plausibly risk-averse in most situations, this finding supports the prohibition of *ex post facto* criminal laws but not the *lex mitior* principle.

The model further demonstrates that when the implementation of sanctions entails significant social costs, the results depend on the shape of the cost function.

If the function is convex (meaning that imposing penalties exhibits diseconomies of scale –a plausible assumption for very severe sanctions–) or concave, non-retroactivity maximizes social welfare. In such these cases, criminal laws –regardless of whether they criminalize or decriminalize conduct, or whether they increase or reduce penalties– should not apply retroactively.

If the function of social costs is linear (as assumed in the standard Law and Economics model of law enforcement), both retroactivity and non-retroactivity can be socially optimal. That is, social welfare could be maximized either by giving retroactive effect to new criminal laws (regardless of whether they criminalize or decriminalize conduct or whether they raise or lower sanctions) or giving them no retroactive effect.

Our analysis also highlights the importance of adjudication costs. Non-retroactivity is optimal whenever retroactivity would generate additional adjudication costs—specifically, when new laws (i) criminalize previously lawful conduct or (ii) modify liability in cases that have already been adjudicated. In these situations, the non-retroactivity of the new law allows avoiding unnecessary adjudication costs while preserving optimal deterrence.

There is only one case where retroactivity reduces adjudication costs without undermining deterrence: when conduct is decriminalized and the relevant cases have not yet been adjudicated. In such circumstances, retroactively applying the decriminalizing law eliminates the need to adjudicate cases that would otherwise proceed under the old law. Here, the optimal solution is to apply the *lex mitior* principle.

By contrast, when a new law alters the level of sanctions (either increasing or decreasing them) and the relevant cases remain pending, its retroactivity neither raises nor reduces adjudication costs. In this situation, both retroactivity and non-retroactivity can be socially optimal.

Finally, we evaluate the necessity of establishing—through constitutional provisions—the prohibition of *ex post facto* laws and the *lex mitior* principle when the legislature shows punitive preferences because its private costs and benefits are not aligened with social welfare. If some legislatures overestimate the harm caused by potentially harmful acts, then the constitution-maker, to maximize social welfare, should forbid *ex post facto* laws and allow the retroactivity of milder laws. That is, in this scenario, to maximize social welfare, it is sufficient for the constitution to forbid the retroactivity of laws that increase sanctions and to allow (without making it constitutionally mandatory) the retroactivity of laws that reduce sanctions.

If establishing a milder law with retroactive effect entails a political cost for the legislature, then, in order to maximize social welfare, the constitution should not mandate the retroactivity of laws reducing sanctions, but should merely allow it. In this scenario, a constitutional rule requiring that milder laws be retroactive would induce the legislature to impose socially suboptimal sanctions. Paradoxically, such a constitutional provision could even discourage the legislature from passing new milder laws, even though this legal change would be socially desirable.

In sum, our findings strongly support the prohibition of *ex post facto* laws, insofar as: (i) retroactivity is suboptimal when individuals are risk-averse; (ii) if sanctioning is socially costly and exhibits economies or diseconomies of scale, non-retroactivity is required to maximize welfare; (iii) even with

linear sanction costs, non-retroactivity remains a viable (though not necessary) welfare-maximizing rule; (iv) when adjudication is costly, retroactivity is suboptimal or not necessary to maximize social welfare in all but one scenario; (v) if some legislatures overestimate the harm caused by harmful acts, the constitutional prohibition of *ex post facto* laws is necessary to maximize social welfare.

By contrast, our model offers only weak or no support for the *lex mitior* principle, as: (i) *lex mitior* creates a costly risk for individuals who anticipate favorable legal changes that may never materialize; (ii) if the costs of implementing sanctions exhibit economies or diseconomies of scale, retroactivity is suboptimal; (iii) if the cost function of implementing sanctions is linear, *lex mitior* is not necessary to maximize social welfare; (iv) if adjudication entails costs, retroactivity of *lex mitior* is welfare-maximizing only when conduct is decriminalized and pending cases have not yet been adjudicated; in all other scenarios, it is either suboptimal or unnecessary; (v) if some legislatures overestimate the harm caused by harmful acts or establishing a milder law with retroactive effect entails a political cost for the legislature, then, a constitutional provision mandating the retroactivity of *lex mitior* could be either suboptimal or not necessary to maximize social welfare; social welfare can always be maximized if the retroactivity of *lex mitior* is merely allowed and not mandatory for the legislature.

The results of our paper account thus for the observed asymmetry between both rules. Our findings somehow explain why the prohibition of *ex post facto* law was established earlier, more strictly and categorically, and with more constitutional support and greater global consensus than the *lex mitior* principle.

Table 1: Summary of results

Benevolent legislature				
Scenario	Rule	Risk aversion	Implementation costs linear (non)	Adjudication costs
1. Criminalization and not adjudicated yet	R (ex post facto)	Opposes	Indifferent (Opposes)	Trade-off*
	N (ex post facto)	Supports	Indifferent (Supports)	
2. Decriminalization and not adjudicated yet	R (lex mitior)	Opposes	Indifferent (Opposes)	Supports
	N (lex mitior)	Supports	Indifferent (Supports)	Opposes
3. Decriminalization and already adjudicated	R (lex mitior)	Opposes	Indifferent (Opposes)	Opposes
	N (lex mitior)	Supports	Indifferent (Supports)	Supports
4a. Increases sanction and not adjudicated yet	R (ex post facto)	Opposes	Indifferent (Opposes)	Indifferent
	N (ex post facto)	Supports	Indifferent (Supports)	Indifferent
4b. Reduces sanction and not adjudicated yet	R (lex mitior)	Opposes	Indifferent (Opposes)	Indifferent
	N (lex mitior)	Supports	Indifferent (Supports)	Indifferent
5a. Increases sanction and already adjudicated	R (ex post facto)	Opposes	Indifferent (Opposes)	Opposes
	N (ex post facto)	Supports	Indifferent (Supports)	Supports
5b. Reduces sanction and already adjudicated	R (lex mitior)	Opposes	Indifferent (Opposes)	Opposes
	N (lex mitior)	Supports	Indifferent (Supports)	Supports

Preferences of some legislatures are misaligned with social welfare		
Scenario	Rule	Constitution-maker choice
Potential new legislature overestimates harm	Retroactivity is always allowed	Opposes
	Retroactivity is always prohibited	Supports
	Only retroactivity of EPFL is prohibited	Supports
Incumbent legislature overestimates harm	Only retroactivity of LM is allowed	Supports
	R is mandatory (prohibited) for LM (EPFL)	Supports
	Retroactivity is always prohibited	Opposes
Retroactivity of LM entails a political cost	Retroactivity of LM is allowed	Supports
	Retroactivity of LM is mandatory	Opposes
	Retroactivity is always prohibited	Opposes

*No retroactivity minimizes adjudication costs, but there is a trade-off between deterrence and minimizing adjudication costs. If deterrence costs are higher than the reduction in adjudication costs, retroactivity becomes optimal.

References

Adler, M. D. (2003). Legal transitions: Some welfarist remarks. *Journal of Contemporary Legal Issues*, 13(1):5–28.

Becker, G. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):169–217.

Damon, M., Cole, D. H., Ostrom, E., and Sterner, T. (2019). Grandfathering: Environmental uses and impacts. *Review of Environmental Economics and Policy*, 13(1):23–42.

Epstein, R. A. (2003). Beware of legal transitions: Presumptive vote for the reliance interest. *Journal of Contemporary Legal Issues*, 13(1):69–92.

Franzoni, L. A. (2019). Legal change in the face of risk-averse subjects: A generalization of the theory. *American Law and Economics Review*, 21(2):394–430.

Gallant, K. S. (2009). *The Principle of Legality in International and Comparative Criminal Law*. Cambridge University Press, Cambridge.

Garoupa, N. (1997). The theory of optimal law enforcement. *Journal of economic surveys*, 11(3):267–295.

Jenni, K. and Loewenstein, G. (1997). Explaining the identifiable victim effect. *Journal of Risk and Uncertainty*, 14(3):235–257.

Kahan, D. M. (1997). Some realism about retroactive criminal lawmaking. *Roger Williams University Law Review*, 3(1):95–118.

Kaplow, L. (1986). An economic analysis of legal transitions. *Harvard Law Review*, 99:509–617.

Kaplow, L. (1990). A note on the optimal use of nonmonetary sanctions. *Journal of Public Economics*, 42(2):245–247.

Kaplow, L. (2003). Transition policy: A conceptual framework. *Journal of Contemporary Legal Issues*, 13:161–209.

Krent, H. J. (1996). The puzzling boundary between criminal and civil retroactive lawmaking. *Georgetown Law Journal*, 84:2143–2184.

Krent, H. J. (1997). Should bouie be buoyed?: Judicial retroactive lawmaking and the ex post facto clause. *Roger Williams University Law Review*, 3(1):35–94.

Lawrence, I. (2012). Punishment without law: How ends justify the means in marital rape. *The Denning Law Journal*, 18(1):37.

Levitt, S. D. (2004). Understanding why crime fell in the 1990s: Four factors that explain the decline and six that do not. *Journal of Economic Perspectives*, 18(1):163–190.

Levmore, S. (1999). Changes, anticipations, and reparations. *Columbia Law Review*, 99:1657–1700.

Masur, J. S. and Nash, J. R. (2010). The institutional dynamics of transition relief. *New York University Law Review*, 85:101–162.

Miceli, T. J. and Segerson, K. (1994). Regulatory takings: When should compensation be paid? *Journal of Legal Studies*, 23:749–776.

Nash, J. R. and Revesz, R. L. (2007). Grandfathering and environmental regulation: The law and economics of new source review. *Northwestern University Law Review*, 101:1677–1732.

Polinsky, A. and Shavell, S. (1984). The optimal use of fines and imprisonment. *Journal of Public Economics*, 24(1):89–99.

Polinsky, A. M. and Shavell, S. (2000). The economic theory of public enforcement of law. *Journal of economic literature*, 38(1):45–76.

Revesz, R. L. and Westfahl Kong, A. L. (2011). Regulatory change and optimal transition relief. *Northwestern University Law Review*, 105:1581–1634.

Shavell, S. (2008). On optimal legal change, past behavior, and grandfathering. *The Journal of Legal Studies*, 37(1):37–85.

Shavell, S. (2014). Risk aversion and the desirability of attenuated legal change. *American Law and Economics Review*, 16:366–402.

Shaviro, D. (2000). *When Rules Change: The Economics of Retroactivity*. University of Chicago Press, Chicago.

Small, D. A. and Loewenstein, G. (2005). The devil you know: the effects of identifiability on punishment. *Journal of Behavioral Decision Making*, 18(5):311–318.

Trebilcock, M. J. (2014). *Dealing with Losers: The Political Economy of Policy Transitions*. Oxford University Press, Oxford.

Westen, P. (2007). Two rules of legality in criminal law. *Law and Philosophy*, 26(3):229–305.

Westen, P. (2015). Lex mitior: Converse of ex post facto and window into criminal desert. *New Criminal Law Review*, 18(2):167–213.