Criminal Records*

Giuseppe Dari-Mattiacci† and Piero Gottardi‡

June 6, 2025

Abstract

We present a dynamic model of the labor market, where workers may commit crimes and employers can gather information about workers' criminal history from a publicly available record and set wages accordingly. We characterize the socially optimal duration of the record, which balances two conflicting objectives: deter inefficient crimes for workers without a record and keep the share of the population with a record low to reduce recidivism. We also show that, when the social harm from crime is neither too high nor too low, it is optimal to impose finite nonmonetary sanctions followed by a finite criminal-record period.

JEL codes: D83, J70, K14.
Keywords: criminal record, optimal expungement, recidivism, employment of ex convicts, incapacitation.

*The authors would like to thank Ronel Elul, Eric Langlais and the participants in the 2024 International Meeting in Law & Economics at the College of Europe in Bruges and the Barcelona School of Economics 2024 Summer Forum for valuable comments, and Sophie Austin, Antonia Grundl, and Melek Redzheb for excellent research and editorial assistance. Piero Gottardi acknowledges financial support from MUR, PRIN 2022.

†University of Amsterdam.
‡University of Essex.

1 Introduction

Since Becker (1968), the law-enforcement literature has focused on the type, magnitude, and probability of criminal sanctions (Polinsky and Shavell, 2000). Yet, the most pervasive and long-lasting consequences of criminal violations do not stem from the imposition of a sanction, which is often relatively short-lived. Rather, they derive from the fact that an official record of the individual's criminal history is kept for an often indefinite period of time. In the United States, there are 80 million individuals with a criminal record (Goggins and DeBacco, 2022; National Employment Law Project, 2024), which amounts to 24% of the population. By comparison, 5.5 million (1.7%) individuals are under the supervision of adult correctional systems, including 1.7 million (0.5%) incarcerated (Kluckow and Zeng, 2022).[1] Criminal records provide information on past felony convictions only in minor part. Most entries are about nonviolent misdemeanor violations and often include charges and arrests that did not lead to a conviction (Stevenson and Mayson, 2018).

Next to an array of official legal consequences — including licensing restrictions for a variety of occupations and bans from welfare and food stamps (Yang, 2017) — and discrimination in housing and education (Clean Slate Initiative, 2023), individuals with a criminal record face stigma in the labor market (Kuehn and Vosgerau, 2024). Employers make widespread use of background checks to screen applicants (Agan et al., 2024) and experiments on job applications in the U.S. found that applicants with a record received significantly fewer callbacks than those without a record (Pager, 2003; Uggen et al., 2014; Agan and Starr, 2017; Leasure, 2018; Cerda-Jara, Harding, and Cohort, 2024). Other studies using both experimental and observational data document the long-term consequences of criminal records on employment and wages (Nagin and Waldfogel, 1998; Leasure and Andersen, 2016; Natapoff, 2018). Finlay, Mueller-Smith, and Street (2023: 2206) point to the "scarring effects of criminal records" as the most likely channel for the negative economic effects of criminal convictions on families.

To guarantee individuals with a record a second chance, the Fair Credit Reporting Act[2] and numerous "ban-the-box" laws[3] have limited the use of criminal history information by employers. Similarly, an increasing number of "clean-slate" laws[4] has

[1] The report by Goggins and DeBacco (2022), prepared for the U.S. Department of Justice, counts 114 million individual criminal history files in 2020 (the latest available estimate), noting that the number of individuals with a record is likely less than the number of files because some individuals have a record in multiple states and some records may concern deceased individuals. To account for these issues, the Fact Sheet by the National Employment Law Project (2024) reduces the estimate by 30% to roughly 80 million. The data on correctional systems is taken for the same year. Percentages on the population are calculated using as a basis 331.5 million, the total number of US residents in 2020 according to the 2020 Census data, available at https://www.census.gov/programs-surveys/decennial-census/decade/2020/2020-census-results.html, last accessed January 31, 2025.

[2] The Fair Credit Reporting Act (15 U.S.C § 1681) limits reporting of information on arrests, indictments, and other records older than 7 years. Some states have also enacted more restrictive provisions (Agan et al., 2024).

[3] A majority of U.S. states and numerous cities and counties have enacted some form of ban-the-box laws, prohibiting employers from asking about a job applicant criminal history in job applications (Agan and Starr, 2018; Avery and Lu, 2021).

[4] According to a recent study by Prescott and Starr (2020: 2488-2510) conducted in Michigan, 5 years after becoming eligible, only 10% of individuals with a criminal record had ap-

aimed at facilitating the expungement[5] of criminal records. Evidence suggests that such policies increase the probability of employment for applicants with a criminal record (Selbin, McCrary, and Epstein, 2018; Craigie, 2020). Removing information, however, is not easy. Even if the official record is erased or sealed, it may remain easily accessible in media and web outlets (Lageson, 2020). In Europe, the European Convention on Human Rights (ECHR) and the General Data Protection Regulation (GDPR)[6] have enhanced privacy rights but do not have a global reach. In a 2019 landmark case, the European Court of Justice compelled Google to remove references to three news articles about a convicted murderer but noted that the "right to be forgotten" did not apply outside EU borders and hence links remained accessible to American IP addresses.[7] Recent research also points out that the removal of information may not be enough when employers can infer criminal history from gaps in an applicant's employment history, and hence a criminal record may keep haunting individuals for years after expungement (Agan et al., 2024).

When effective, the relief provided by expungement needs to be weighed against the positive value of criminal-history information for assessing an individual's propensity to re-offend and, more generally, its effects on the crime rate. On the upside, Selbin, McCrary, and Epstein (2018) and Prescott and Starr (2020) found that expungement of criminal records contributed to improved job opportunities and lower recidivism rates. Even individuals who were wrongfully convicted—and therefore never actually committed a crime—face a higher risk of offending after exoneration if their records are not expunged (Shlosberg et al., 2014). Agan, Doleac, and Harvey (2023) observe that individuals charged with a nonviolent misdemeanor, even if never convicted, still have a criminal record, have "less to lose from engaging in criminal activity", and are twice as likely to (re)offend compared to individuals who were never charged.[8] On the downside, critics of expungement remark that it suppresses valuable information, with negative repercussions on honest individuals, including statistical discrimination (Kogon and

plied for expungement. Consequently, over the last 5 years, 12 states have enacted clean-slate laws, which automatize, facilitate and expand the expungement of criminal records. See https://www.cleanslateinitiative.org/states#states, last accessed January 31, 2025. Expungement is automatic in most European countries, with some exceptions. In Italy, for instance, an application is needed.

[5]"Expungement" proper is the legal elimination of the criminal record, but with the same term we also refer to lesser forms of intervention like "sealing", which makes the record available only for selected law-enforcement purposes of future crimes. Such terms, and the term "setting aside" are often used interchangeably, although each jurisdiction has its own specific rules. A convenient comparison of the various expungement policies in force in the United States can be found on the website of the Collateral Consequences Resource Center, which manages the Restoration of Rights Project. See https://ccresourcecenter.org/state-restoration-profiles/50-state-comparisonjudicial-expungement-sealing-and-set-aside-2-2/, last accessed May 21, 2025.

[6]Regulation (EU) 2016/679.

[7]CJEU, C-507/17 Google LLC vs. CNIL, 2019.

[8]Advocates of more effective expungement policies stress that criminal records unduly restrict an individual's job opportunities and encourage recidivism, a strong correlate with unemployment (Pettler and Hilmen, 1967; Roberts, 2015; Mungan, 2017a; Murray, 2021). Other reasons given in favor of expungement include the fact that individual attitudes may change over time (Galle and Mungan, 2020), the possible decriminalization of the recorded offense (Rosen, 2019), the fact that expungement allows to enhance workers' productivity (Wurie, 2012), privacy (McIntyre and O'Donnell, 2017), and the proportionality of punishment (Corda, 2016).

Loughery Jr., 1970; Franklin and Johnsen, 1981; Snow, 1992; Funk and Polsby, 1997; Tobin and Walz, 2015). In a recent study, Agan and Starr (2018) found that, prior to the introduction of "Ban the Box" policies in New York City and New Jersey, white job applicants received 7% more callbacks than similar black applicants. The introduction of the policies made this gap grow 6-fold to 43%, suggesting that the suppression of information at the early stages of the hiring process made employers discriminate against the group they perceived to be more likely to have a criminal record, with severe consequences for black applicants without a criminal record.[9] Further evidence is provided by Honigsberg and Jacob (2021), who show that expungement of the Broker-Check record resulted in a higher probability of misconduct by expunged brokers, relative to those who did not have their record expunged.

Extant theoretical work provides little guidance as to how policies should balance these opposite effects. While some of the trade-offs associated with expungement have been highlighted in the literature, it is important to analyze them within the context of a model where the effects of alternative policies on labor market outcomes, crime, and overall welfare can be examined and proper counterfactual analysis can be conducted. To date, little to no work has been done in this dimension.

In this paper, we consider a dynamic model of the labor market that allows us to study the consequences of the presence of criminal records for the lifetime profile of wages of heterogeneous individuals, and for how many and which types of crimes are committed. We then characterize the socially optimal expungement rate, that is, the length of time a crime should remain in the record before (if ever) it is removed. Keeping a record of an individual's criminal history is a punishment from the individual's perspective, which adds to more often studied non-monetary sanctions, such as, but not exclusively, imprisonment. We compare these two forms of punishment and investigate how they can be optimally combined. Our notion of expungement captures more generally the effects of any policy aimed at suppressing information about an individual's criminal history. Thus, our analysis has implications not only for criminal law and expungement policies strictly defined, but also more broadly for privacy regulations, data protection laws, and labor laws.

In the model, at any point in time workers are offered a wage by firms and a fraction of them may receive a crime opportunity. We allow for a rich variety of crime opportunities, featuring different private benefits to the worker. Crimes are costly for society, and in particular for the firm employing the worker, due to bad publicity, litigation, or actual losses in revenue. The key choice in the model is the workers' decision of which crimes to commit, when the opportunity arises. This decision will be taken weighing the benefits of a specific crime against the punishment faced. We consider first the case where the only available (and informal) punishment for committing a crime is the *shaming*[10] resulting from the fact that the crime is recorded in a worker's criminal

[9] Bushway (2004) also argues that the availability of criminal records increases wages for non-convicted individuals and prevents statistical discrimination.

[10] Shaming is no different from stigma in our model, but the two are considered differently in the literature. While the term "stigma" is usually used for the negative consequences of a criminal record

record. The cost of shaming is endogenously determined, and is given by the difference between the wage firms choose to offer to workers without a criminal record and that offered to those with a record.[11] This difference is positive and reflects the higher propensity to commit crimes of workers who have a criminal record and the associated cost. Since they face a lower additional punishment[12] for any further crime they may commit, as their wages are already discounted, they have less to lose and so they will offend more often.

Hence, the model endogenously generates higher incentives to commit crimes for individuals with a criminal record as compared to those without; that is, the rate of recidivism is higher than the crime rate in the overall population. This effect is reinforced by the fact that a criminal record also conveys information to employers on individual traits of workers, which determine the likelihood they will receive crime opportunities in the future. The fact that crimes are recorded in a public registry thus generates deterrence of first-time offenders but this comes at the cost of facilitating recidivism among convicted individuals. The length of time until a crime is expunged matters. As we show, the shorter this time, the lower the *general* deterrence of workers without a criminal record, because the cost of committing a crime decreases when a crime is kept in the record for a shorter period, during which the worker incurs a loss in wage earnings. This worsening of incentives drives down the wage of workers without a criminal record, as detractors of expungement policies lament. At the same time, since workers with a criminal record commit more crimes, a speedier expungement increases the *specific* deterrence of these individuals, as it is costlier to commit a crime for workers after regaining a clean record; such workers are then less likely to re-offend after expungement than they were before expungement.

Thus, the duration of a criminal record before expungement has a variety of contrasting effects on the overall crime rate and also affects the types of crimes committed. Moreover, the length of the record and the crime rate of first-time offenders determine, given the dynamic nature of our model, the share as well as the composition of the population that at any point in time has no criminal record. This in turn influences the wages offered by firms, further affecting the number and types of crimes committed. These properties are summarized in Proposition 1 where we show how, in a labor market equilibrium, the overall crime rate and the types of crimes committed depend on the expungement policy — that is, the duration of a criminal record — as well as on the magnitude of the social cost of crime and on the share of the population who may receive crime opportunities.

The welfare maximizing expungement policy is then characterized in Proposition 2. This policy optimally addresses a trade-off, which we identify in this paper, between

accessory to another sanction, "shaming" is seen as a punishment in itself, independently of other sanctions, as in Kahan and Posner (1999).

[11]Although outright discrimination may be illegal, it is nevertheless an empirical reality that workers with a criminal record face disproportionate hurdles in the labor market, ranging from longer job searches to under-skilled employment.

[12]We restrict attention to binary records, featuring the presence or absence of any crime committed in the past.

deterrence and recidivism. That is, it has to balance two, typically conflicting objectives: deter inefficient crimes — those for which the employer's cost exceeds the private benefit — for workers without a criminal record, and keep the share of the population with a criminal record as low as possible to reduce recidivism. Expunging the record reduces recidivism, by bringing individuals back to a state where deterrence is higher so that they are less prone to offend (as found by Selbin, McCrary, and Epstein, 2018; Prescott and Starr, 2020; Agan, Doleac, and Harvey, 2023). However, expungement also lowers the deterrence of workers without a record because, by shortening the duration of the stigma, it softens the negative economic consequences of committing the first offense. Such decrease in deterrence narrows the wedge between the wage of those with a record and those without, and hence further weakens incentives not to commit crimes. Indeed, Prescott and Rockoff (2011) found that sex-offender registration laws improved deterrence while also increasing recidivism.

We find that when the harm for employers from crimes committed by their workers is not too high, it is socially optimal to expunge criminal records after a finite period of time. This results in a relatively low level of deterrence for individuals without a record, in the sense that not all inefficient crimes are deterred, that is, there is under-deterrence. The benefit of doing so is that allowing a relatively fast *exit* from the criminal record contains the share of the population with a criminal record and hence reduces the cost of recidivism. When instead the harm from crime is large, it is optimal to never expunge a criminal record. In this case, full deterrence is induced when that is feasible. Hence, the level of deterrence of workers without a record is high and even some efficient crimes — with large private gains, higher than the harm to the employer — are deterred, that is, there is over-deterrence. This constitutes an alternative way to curb the costs of recidivism, by ensuring that as few individuals as possible commit a crime and hence *enter* the criminal record.

We then extend the analysis to expand the set of available punishments for crime to include nonmonetary sanctions. Nonmonetary sanctions are widely used both as a direct consequence of a criminal conviction and indirectly when individuals are unable to pay fines. They are, hence, the most relevant form of punishment for wealth-constrained individuals. Nonmonetary sanctions can take a variety of forms, from imprisonment, probation, electronic monitoring, and community service to measures targeted to specific crimes, such as the suspension of a driving license, a ban from public offices, a ban on living within a certain distance from a school or a park, the revocation of a passport, or the cancellation of a visa. Different from the merely shaming consequence of a criminal record, nonmonetary sanctions have both a deterrence and an incapacitation effect.

There are two kinds of implications of incapacitation. First, the ability to work is constrained, so there is a loss of productivity. Second, crime opportunities are also reduced. The prospect of reduced employment and crime opportunities after committing a crime also works as a deterrent to crime ex-ante.[13] Incapacitation is maximal in

[13] Note that some degree of incapacitation may also arise with a criminal record, when employers choose to restrict the tasks and responsibilities attributed to a worker, on the basis of the worker's

the case of imprisonment. With nonmonetary sanctions, it may then happen that the propensity of an individual to commit a crime rises when the punishment ends, the opposite as with a criminal record (in line with Honigsberg and Jacob, 2021). In the design of the optimal duration of sanctions, a different trade-off is faced, this time between deterrence and incapacitation. This trade-off is analogous to the one identified in the literature on credit ratings (Musto, 2004; Elul and Gottardi, 2015). There, a borrower with a low credit score is unable to obtain credit and hence cannot produce, but the prospect of ending up with a low score provides ex-ante incentives to exert effort, with an associated productivity loss.

In Propositions 3 and 4, we compare the level of deterrence (ex ante, by the threat of a sanction) that is attained with nonmonetary sanctions to the one with a criminal record. We then compare welfare under the two policies. This depends not only on the level of deterrence in the two cases, but also on what happens when the punishment is in place. The incapacitation induced by sanctions implies that now some crimes are avoided and there is also a productivity loss. The effect on crime of incapacitation, however, differs from that of deterrence in an important aspect: incapacitated individuals are prevented from committing all crimes, while individuals who are deterred choose to commit a crime if their personal gain is higher than the expected sanction, and refrain from it otherwise. Deterrence, therefore, filters the crimes that occur, while incapacitation does not.[14]

We find that when employers' harm from crime is sufficiently low, criminal records allow to attain a higher welfare level than nonmonetary sanctions. The reason is that a primary consideration in this case is filtering crimes, and deterrence is more effective than incapacitation at achieving that. When instead the costs of crimes for society are large, sanctions are preferable when they allow to achieve a higher level of deterrence (for instance in the case of maximal incapacitation, as with imprisonment). In that case, the main concern driving the socially optimal policy is to prevent the maximal number of crimes.

Finally, we examine the benefits of suitably combining both forms of punishments, so that both trade-offs are present at the same time. We show it is always beneficial to combine the two forms of punishment, by having a period of time, after the nonmonetary sanction, when only the criminal record is in place. The benefits of combining the two are particularly significant when employers' harm from crimes is sufficiently large and come from the possibility of inducing some deterrence also during the punishment phase, thus making the punishment more cost effective.

---

criminal history information. In that case both trade-offs are present with shaming sanctions alone.

[14] In our baseline model, the presence of a criminal record induces some deterrence of first-time offenders, thus filtering crimes ex-ante, but no deterrence of second-time offenders, and hence there is no filtering of crimes ex-post. We analyze this as a problem of recidivism.

Related Literature

The theoretical literature on reputational sanctions and expungement has two main limitations. First, extant models are static and hence (as noted by Harel and Klement, 2007: 370, fn. 26) cannot address key questions concerning the effects on the composition of job market applicants as well as the duration of the criminal record and of the nonmonetary sanctions that predate it, which are considered in this paper. Second, as we explain in detail below, they focus on a subset of the effects considered here or employ exogenously fixed parameters for variables that are, in fact, endogenous to the labor market.

This literature originates from the seminal contribution by Rasmusen (1996) on the stigmatization effect of criminal sanctions. In his model, stigma results in employers paying lower wages to workers with a criminal record because crime hurts the employers' net payoff. The same is true in our model. However, while we consider the welfare costs of stigmatizing individuals — specifically, recidivism — Rasmusen (1996: 536) explicitly assumes that stigma is socially costless and hence does not address the trade-offs we consider in this paper. In addition, his paper and the subsequent literature (Furuya, 2002; Iacobucci, 2014; Harel and Klement, 2007; Mungan, 2015, 2016)[15] do not consider the duration of stigma as a policy variable, which is central to our analysis.

The contributions that explicitly address expungement and (some of) the trade-offs we consider here focus only on some aspects of the problem only and hence can be seen as special cases of our more general model. Funk (2004) deploys a single-agent model, where wages are exogenously given and hence do not reflect the information produced by the criminal record and the induced beliefs of employers. Mungan (2019) does not consider recidivism, as he assumes that the crime rate of convicted and non-convicted individuals is the same. In contrast, we determine both wages and crime rates endogenously as part of a labor-market equilibrium.[16]

In our model, there is no age discrimination so that no information can be derived from the time a worker has spent free of a criminal record. This aspect is instead central to the literature on reputation (starting with Diamond, 1989). Compared to this literature, our model is less rich on the states in which a worker can be — they are only two, with or without a record — but is richer on the actions that individuals can take — the choice regarding crime is not a binary one, we have a continuum of crime opportunities which may arise, and for each of them a choice must be made entailing different social costs and benefits. Our approach reflects plausible assumptions on the effect of anti-age-discrimination laws and the coarse way in which employers process information. Relatedly, Ganuza, Gomez, and Robles (2016) and Baker and Choi (2018) study the interaction of reputational sanctions and tort or contract liability but do so in a model based on a repeated game with imperfect punishment, which is quite different

[15]See also, Mungan (2017b), arguing that making expungement costly may reduce crime because it may reduce recidivism of individuals with a criminal record without diluting deterrence of those without.

[16]There is of course a large informal literature discussing the issue of expungement. See, for instance, Blanchette and Johnson (2002); Blumstein and Nakamura (2009).

from ours.

This paper is organized as follows. In Section 2, we introduce our baseline model of the labor market with criminal records. In this model, the only sanction is the stigma resulting from the criminal record. We then turn, in Section 3, to the characterization of the socially optimal expungement policy. In Section 4, we expand the model to consider the imposition of a nonmonetary sanction as another form of punishment for crime and compare its effectiveness and welfare properties to criminal records. We then examine the benefits of combining the two forms of punishment. As a special case, we study imprisonment. In Section 5, we conclude. The Appendix contains an outline of the proofs. Technical details of the proofs and of our numerical specifications are in the Online Appendix.

2 A model of employment with criminal records

2.1 Setup

We consider a dynamic model of the labor market. At each point in time, there is a population of mass 1 of workers. Each worker, when hired by a firm, generates a flow of output $\pi \geq 1$ and faces random crime opportunities. For each worker, the end of life arrives according to a Poisson process with rate $\tau > 0$; hence workers live on average for $\frac{1}{\tau}$ units of time. (We will assume throughout that τ is finite).[17] There are two types of workers: a fraction $r \in (0, 1)$ of the workers are "dishonest" and randomly receive 1 crime opportunity per unit of time, where 1 is the arrival rate in a Poisson process; the remaining fraction $1 - r$ of the workers are "honest" and never commit a crime. Honesty can be interpreted as higher moral standards or as the inability to capture the gains from crime. When a worker dies, the individual is replaced by a newborn worker of the same type. Thus, population size and composition are constant over time.

We consider a rich set of possible crimes. Different crimes — say, theft versus embezzlement — are characterized by different values of the harm $h \in (0, 1)$ they impose on society (in our environment, on the employer as there is no other loss);[18] h is publicly observable. Each crime yields a private benefit $g \in [0, 1]$ to the worker who commits it. We assume g is randomly drawn from the uniform distribution on the unit interval and is privately observed by the employer. The arrival of crime opportunities, as well as the value of the private benefit of committing a crime, are assumed to be independently and identically distributed across workers.[19] When an opportunity to commit crime h

[17]There is no discounting in the model. However, the workers' utility would be the same if we interpreted τ as a discount rate rather than a probability of dying. Under this interpretation, other parts of the analysis, such as the dynamic process of the population, would have to be slightly amended but the qualitative properties of our results would remain valid.

[18]The model could be easily extended to consider harm to third parties in addition to the loss to the employer. As we explain below, the fact that employers directly suffer a loss from crime generates stigma in the labor market. The smaller the fraction of the social harm borne by the employer rather than third parties, the weaker is the deterrence effect of stigma. We discuss this issue in the Conclusions.

[19]In some situations it is natural to argue that the private benefit of crime g is independent of its social harm h: think of somebody stealing a bike for a joy ride (low g) versus stealing the same bike

materializes, a dishonest worker decides whether or not to commit the crime depending on the value g of the worker's private benefit. The net social loss from a crime is so $h-g$. Hence committing a crime is socially "desirable" when the private benefit g exceeds the social harm h, while crimes with $g < h$ should be deterred. The first-best threshold, $g^{FB} = h$, identifies then the crime opportunities for which total benefits equal total costs. The notion of "desirable" violations of the law is in line with previous literature (Becker, 1968; Polinsky and Shavell, 2000).

Criminal law allows prosecutors and judges to consider a wide array of aggravating or extenuating circumstances that may partially capture the private benefits g, but can do so only imperfectly. In our model, the feature that h is observable but no information is publicly available about the value of g implies that the sanctions and the duration of the record considered in our analysis will only be contingent on h.

We assume, for simplicity, that if a crime is committed, is detected by the law enforcement agency with probability 1.[20] In the basic setting of our model, the only punishment for crime consists in the fact that the law enforcement agency enters the name of who committed it into a criminal record, which is publicly available. The main focus is then on the deterrence effect of the stigma that endogenously arises in equilibrium from the presence of a criminal record. (In Section 4, we extend the analysis to the case where workers may also be subjected to a nonmonetary sanction with both deterrence and incapacitation effects.)

For simplicity, we also assume that the criminal record is binary and for every worker it only reports whether or not the worker has committed at least 1 crime in the past. No distinction is thus made with regard to the number of crimes committed. At any point in time, a worker can then be in either one of the following two states: with a criminal record (state C) or with no criminal record (state N). The record is kept for a random time, also modeled as a Poisson process with an exit rate from the record equal to $\sigma > 0$ (the expungement rate), so that after committing a crime, a worker expects to have a criminal record for $\frac{1}{\sigma}$ units of time, after which the record is expunged and the worker returns to the original situation with no criminal record.[21] The expungement rate σ is the core policy variable in our model: we will investigate how the rate and

to escape from kidnappers (high g). In other situations it seems more likely the values of g and h are correlated: for example, stealing a more expensive painting results in higher gains for the thief. We abstract from this possibility in our model, to preserve the simplicity of our analysis, though allowing for it would not alter our qualitative findings.

[20] We discuss the role of this assumption in the Conclusions.

[21] An important feature of our specification is that the commission of an additional crime by employees resets but does not increase the expected time to expungement, that is, σ is not a function of the number of crimes committed. This feature is in line with many actual expungement policies. For instance, the recent New York State's Clean Slate Act (S7551A/A1029C) entered into effect in 2024 provides that eligible convictions are automatically sealed after 3 years for misdemeanors and 8 years for felonies, and explicitly states that the waiting period starts over with the same length, if the perpetrator commits another crime. See https://www.nycourts.gov/FORMS/criminal-record-sealing.shtml, last accessed June 2, 2025. The randomness surrounding the actual expungement date captures factors such as the discretion used by the public officials involved, intervening events such as press coverage, the availability of privately-collected data on workers' criminal history even after expungement, or the time it takes for people to forget. See footnote 5 for a legal definition of expungement and a reference to expungement policies in the United States.

characteristics of crimes committed, and hence welfare, vary with σ, and what is then the optimal value of σ as a function of the harm from crime h and other parameters of the model.

We assume that employers are unable to write optimal long-term contracts with their workers, which defer pay and make it conditional on the number and, possibly, the severity of the crimes committed. This is because workers may be liquidity constrained, the employer may be short-lived or unable to commit, firms may go bankrupt, such arrangements may be unenforceable in court, among possibly many other reasons. Therefore, workers are employed "at will", so their wages are immediately modified to incorporate any change in available information, used to assess the expected gains from hiring a worker, net of the expected costs of crime.

We assume that the honesty-type of a worker is not observable by the employer, while a worker's criminal record is. Hence, the wage of a worker can only depend on the worker's criminal record. We show in the next section that workers with a criminal record are penalized in the labor market as employers are only willing to offer them a lower wage, hence the resulting stigma. This difference in wages reflects both the different likelihood that a worker in N and in C is honest and the different crime rate in the two states. Because of this wage difference, workers in state N face a cost if they choose to commit a crime, since when they do so they transition to state C and earn a lower wage. In contrast, workers in state C face no punishment. By virtue of the memoryless property of the Poisson process governing expungement, even though the commission of a new crime "resets the clock," the expected duration of the criminal record is unchanged. Consequently, when in C, a dishonest worker will commit all crimes, irrespective of the size g of their benefit. While there is some deterrence in state N, there is no deterrence in state C. This setup captures in stark terms the idea that individuals with a criminal record have "less to lose" if they commit a crime and will then be more inclined to do so as compared to the case where they have no criminal record.

Given the stationarity of the environment and the presence of a large population, in a steady state there is a constant distribution of workers between the two states at any given point in time and only two wages, constant over time, are offered: one to workers without criminal record, w_N, and one to workers with a criminal record, w_C. The analysis that follows will focus on steady-state equilibria. At such equilibria, while there is a dynamic process at the individual level, governing the transition of workers during their lifetime between the two states C and N, there is no dynamic change in the aggregate for the whole population. Also, deterrence in state N is described by a constant threshold g_N, endogenously determined at an equilibrium of the model, such that a dishonest worker commits all crimes that result in a private gain $g \geq g_N$.

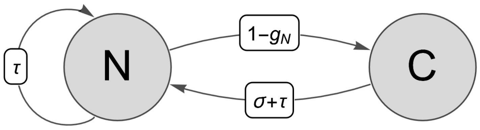


Figure 1: Population dynamics for dishonest workers

The dynamics of the flow between the two states for a dishonest worker is described by a simple Markov chain in continuous time (Figure 1). The transition rate from N to C is given by $1 - g_N$ since the worker receives 1 crime opportunity per unit of time and commits the crime only if g is above the threshold g_N, that is, with probability $1 - g_N$. Instead, the transition rate from C to N is independent of the choices of the worker regarding which crimes to commit and is simply equal to the expungement rate σ plus the death rate τ — since each death is followed by a new birth and newborns clearly start in state N. Hence, at each point in time the share of dishonest workers who are in C is equal to the ratio between the transition rate from N to C over the sum of the transition rates: $\frac{1-g_N}{1-g_N+\sigma+\tau}$. Similarly, the share of dishonest workers in N is $\frac{\sigma+\tau}{1-g_N+\sigma+\tau}$. Honest workers, in contrast, never commit a crime and always remain in N. In expectation, the fraction of the whole population that at any point in time is in state N is thus equal to $(1-r) + r\frac{\sigma+\tau}{1-g_N+\sigma+\tau}$.

2.2 Employment and wages

To determine the equilibrium wage level, we assume that workers have all the bargaining power, so firms make zero profits and workers are paid an amount equal to their productivity π — the same for all of them — minus the expected cost of crimes for the employer. This cost in turn depends on whether the worker has a criminal record.[22] Since $\pi \geq 1$ and the employer's expected cost of crimes does not exceed $h < 1$, equilibrium wages are positive and there is always full participation in the labor market. To find the equilibrium wages, we need to determine the crimes that workers choose to commit in state N and in state C.

We already noted that, in state C, dishonest workers will commit a crime every time an opportunity arises, no matter how small the gain g may be. Also, only dishonest workers are present in C. Hence the employers' cost of crimes committed equals h. As a consequence, workers in C are paid a wage equal to:

$$w_C = \pi - h \tag{1}$$

In contrast, in state N honest workers, who never commit crimes, are also present. In addition, in this state, dishonest workers commit a crime only if its gain g is larger than

[22]Both wages and productivity should be understood as per unit of time. Moreover, in the basic model there is no loss of productivity due to stigma; we will consider the possibility of productivity losses when nonmonetary sanctions are introduced in Section 4.

some threshold $g_N \geq 0$. It thus follows that the expected cost of crime for employers is smaller in N than it is in C and the wage of workers without a criminal record is:

$$w_N = \pi - \frac{r(\sigma+\tau)}{(1-r)(1-g_N)+\sigma+\tau}(1-g_N)h \tag{2}$$

where $\frac{r\frac{\sigma+\tau}{1-g_N+\sigma+\tau}}{(1-r)+r\frac{\sigma+\tau}{1-g_N+\sigma+\tau}} = \frac{r(\sigma+\tau)}{(1-r)(1-g_N)+\sigma+\tau}$ is the fraction of dishonest workers among all workers in state N and $1-g_N$ is the probability that a dishonest worker faces a crime opportunity $g > g_N$ and hence commits a crime.

It is then immediate to see that we always have $w_N \geq w_C$ and the inequality is strict as long as $r < 1$ (that is, if there are some honest workers) or if the value of g_N, endogenously determined in equilibrium, is strictly greater than 0. When $w_N > w_C$ workers in N face an informal sanction for committing a crime due to the fact that this induces a transition to state C where they will earn a lower wage.

2.3 Deterrence

As explained above, honest workers in N never commit a crime and dishonest workers in C are not deterred from committing any crime. Thus the only agents facing a real decision are the dishonest workers in N. Any such worker will choose to commit a crime if and only if the instantaneous payoff from committing it exceeds the resulting loss in continuation utility. Formally, if and only if $g > V_N - V_C$, where V_N is the expected value of the stream of wages and gains from crime that a dishonest worker anticipates to receive in the future starting in state N (in short, the value of a dishonest worker in state N), while V_C is the corresponding value starting in state C. Hence, $V_N - V_C$ represents the cost of transitioning, even if only temporarily, to state C after committing the crime. When the crime has value

$$g_N = V_N - V_C \tag{3}$$

a dishonest worker is indifferent between committing the crime, transitioning so to C, and remaining in N. Hence g_N defines the level of deterrence in state N: the worker commits all crimes such that $g > g_N$ and refrains from the ones with value $g \leq g_N$. The larger is g_N, the more crimes are deterred. In particular, we have partial deterrence when $g_N \in [0, 1)$ and full deterrence (no crime is committed) when $g_N = 1$, in which case the worker's optimality condition is given by:

$$1 \leq V_N - V_C \tag{4}$$

To complete the analysis of the model, we consider next the Bellman equation determining the value in state C:

$$\begin{aligned} V_C &= \mathrm{E}_{\nu\sim\exp[\sigma+\tau]}\left[\int_0^{\nu}\left(w_C+\tfrac{1}{2}\right)dt\right] + \tfrac{\sigma}{\sigma+\tau}V_N \\ &= \tfrac{1}{\sigma+\tau}\left(w_C+\tfrac{1}{2}\right) + \tfrac{\sigma}{\sigma+\tau}V_N \end{aligned} \tag{5}$$

which reflects the property that workers remain in C for a random time ν where, per unit of time, they earn a wage w_C plus the benefits from all crime opportunities, whose expected value is $\int_0^1 g dg = \frac{1}{2}$. Workers exit from C either because they die (at rate τ) or because their criminal record is expunged (at rate σ), whatever comes first, so that the time ν of exit from state C follows a Poisson process with rate $\sigma + \tau$. Hence the expected length of time a worker remains in C is equal to $\frac{1}{\sigma+\tau}$. If expungement occurs, the worker transitions to state N and switches to earning the value V_N. Exit is due to expungement with probability $\frac{\sigma}{\sigma+\tau}$. With the complementary probability, it is due to death and no additional value is earned.

The corresponding Bellman equation for dishonest workers in state N is then:

$$\begin{aligned} V_N &= \mathrm{E}_{\nu\sim\exp[1-g_N+\tau]}\left[\int_0^\nu \left(w_N + \frac{1-g_N^2}{2}\right) dt\right] + \frac{1-g_N}{1-g_N+\tau} V_C \\ &= \frac{1}{1-g_N+\tau}\left(w_N + \frac{1-g_N^2}{2}\right) + \frac{1-g_N}{1-g_N+\tau} V_C \end{aligned} \tag{6}$$

The expression captures the fact that the worker remains in N until either of two events occurs: the opportunity to commit a crime with value $g > g_N$ arrives (at rate $1 - g_N$) or the worker dies (at rate τ). Hence the worker remains in N for $\frac{1}{1-g_N+\tau}$ units of time. Only if exit is due to a crime opportunity, thus with probability $\frac{1-g_N}{1-g_N+\tau}$, will the worker transition to state C and earn V_C. With the complementary probability the workers dies.

Solving then (5) and (6) for (V_C, V_N) we obtain:

$$\begin{aligned} V_C &= \frac{1}{\tau}\frac{\sigma\left(w_N+\frac{1-g_N^2}{2}\right)+(1-g_N+\tau)\left(w_C+\frac{1}{2}\right)}{1-g_N+\sigma+\tau} \\ V_N &= \frac{1}{\tau}\frac{(\sigma+\tau)\left(w_N+\frac{1-g_N^2}{2}\right)+(1-g_N)\left(w_C+\frac{1}{2}\right)}{1-g_N+\sigma+\tau} \end{aligned} \tag{7}$$

These expressions are easy to interpret. Both values are scaled by $\frac{1}{\tau}$, the expected length of life of the worker. When in N, the worker earns the flow of wages w_N plus the expected payoff from a crime of high value (above g_N) that is committed when the opportunity arises and determines the transition to C. In turn, when in C, the worker earns the flow of wages w_C plus the expected benefit of all the crime opportunities that arise. The weights given to these payoffs in the above expressions are different depending on whether we consider the worker's value starting from N or C because death may occur (at rate τ) before the transition to the other state and this possibility weighs in favor of the starting state. Simple manipulations then yield:

$$g_N = V_N - V_C = \frac{w_N - w_C - \frac{g_N^2}{2}}{1 - g_N + \sigma + \tau} \tag{8}$$

which characterizes the optimal crime choice of dishonest workers, given the wages they face in the N and C states. Summing up: honest workers commit no crimes, remain in N, and earn a wage w_N for their entire lives. In contrast, dishonest workers commit some crimes in N, namely those crimes that earn them a benefit higher than g_N. If they do so, they transition to state C where they earn a lower wage, $w_C < w_N$, commit all crimes and remain there for an expected time equal to $\frac{1}{\sigma}$.

A (steady-state) *equilibrium* is thus given by a pair of wages (w_N, w_C) and workers' values in the two states (V_N, V_C), as well as a deterrence level g_N in N such that: (i) at those wages employers are willing to hire workers with records N and C, given their crime decisions (1 and 2); (ii) workers choose optimally which crimes to commit (3 or 4) and their values satisfy the Bellman equations (7). In equilibrium, employers make zero profits and their beliefs about the crime choices of workers are consistent. In the next section, we characterize the properties of equilibria for any given level of the expungement rate σ, our key policy parameter and derive the welfare maximizing level of σ.

3 Expungement of criminal records

3.1 Effects of expungement on deterrence

Note first that no deterrence in N — that is, $g_N = 0$ — can only occur in equilibrium when $\sigma = \infty$, that is, when crimes are immediately expunged. In that case, even though we still have $w_N > w_C$, workers face no punishment for the crimes committed. The wage difference is positive in spite of the fact that the crimes committed by dishonest workers are the same in N and in C, because there are also honest workers in N while there are none in C.

In contrast, when σ is finite the equilibria feature partial deterrence, characterized by the threshold $\underline{g_N} \in (0, 1)$, solving (1), (2), and (8), or full deterrence, denoted $\overline{g_N} = 1$, satisfying (1), (2), (4), and (7). Full deterrence obtains in equilibrium whenever the difference in wages between the two states is so large that the expected cost of transitioning from N to C exceeds the highest possible gain from crime ($g = 1$), as stated in (4). When this happens, dishonest workers are fully deterred and commit no crimes in N, thus $w_N = \pi$. It is still true that all crimes are committed by workers in C, and w_C reflects this property, but with no crime in N, no worker will ever transition to C. If the harm for employers from these crimes is large enough, the difference in wages generates enough deterrence that indeed no crime is committed in N. With partial deterrence, instead, some crimes — though not all possible crimes — are committed in N, and $\underline{g_N}$ is the value of a crime opportunity at which a worker in N is indifferent between committing and not committing the crime.

For some parameter values, multiple equilibria exist (as in Rasmusen, 1996). In that case, we focus on equilibria that are stable according to the following out-of-equilibrium adjustment process, in the spirit of the cobweb model. At any stage t of this process, wages are set consistently with the crime choices made by workers at $t - 1$, workers then revise their crime decisions given the new level of wages (assuming that wages will remain constant at the current level at all future dates), and so on. The following proposition characterizes the equilibria of the model described for all parameter values.

Proposition 1. *For any fraction $r \in (0,1)$ of dishonest workers and finite σ, we have:*

1. *If $h(1-r) < \frac{1}{2}$, there exists a threshold $\xi < h - \frac{1}{2}$, such that:*

   (a) *If $\sigma + \tau < \xi$, there is a unique equilibrium with full deterrence, $\overline{g_N} = 1$;*

   (b) *If $\xi \leq \sigma + \tau \leq h - \frac{1}{2}$, there are three equilibria: two of them with partial deterrence, $\underline{g_N}, \hat{g_N} \in (0,1)$, solving (1), (2), and (8), the third one with full deterrence, $\overline{g_N} = 1$; the equilibria $\underline{g_N}$ and $\overline{g_N}$ are locally stable, while $\hat{g_N} > \underline{g_N}$ is locally unstable;*[23]

   (c) *If $\sigma + \tau > h - \frac{1}{2}$, there is a unique equilibrium with partial deterrence, $\underline{g_N} \in (0,1)$.*

2. *If, instead, $h(1-r) \geq \frac{1}{2}$, we have:*

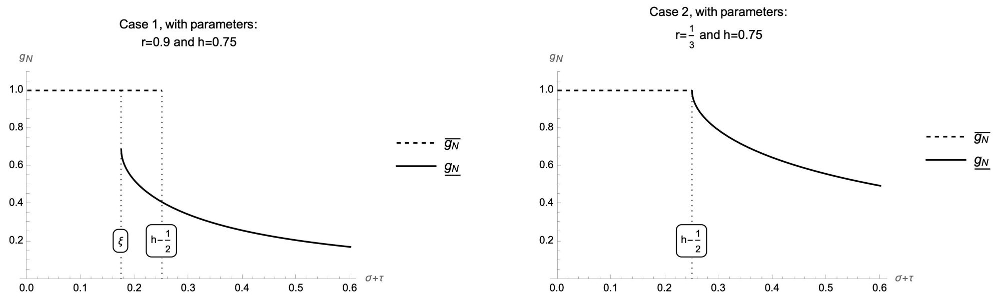
   (a) *If $\sigma + \tau \leq h - \frac{1}{2}$, there is a unique equilibrium with full deterrence, $\overline{g_N} = 1$;*

   (b) *If $\sigma + \tau > h - \frac{1}{2}$, there is a unique equilibrium with partial deterrence, $\underline{g_N} \in (0,1)$.*

The level of deterrence at a stable equilibrium with partial deterrence decreases in $\sigma + \tau$, $\frac{\partial \underline{g_N}}{\partial(\sigma+\tau)} < 0$ and converges to no deterrence in the limit: $\lim_{\sigma+\tau\to\infty} \underline{g_N} = 0$. Deterrence decreases also when the share of dishonest workers increases, $\frac{\partial \underline{g_N}}{\partial r} < 0$.

Hence, full deterrence obtains in equilibrium when $\sigma + \tau$ is sufficiently low, that is, workers are sufficiently long-lived and the criminal record lasts for a sufficiently long time. In particular, $\sigma + \tau$ should be lower than $h - \frac{1}{2}$, where h is the value of the wage difference $w_N - w_C$ and $-\frac{1}{2}$ the loss of workers due to crimes not committed in N when there is full deterrence. The numerator on the right hand side of (8) is, in fact, given by the sum of these two terms. In that case, the punishment induced by shaming is sufficiently long-lasting, and workers care sufficiently for the future that they prefer not to commit any crime. For higher values of $\sigma + \tau$, the punishment is instead insufficient to deter the high-benefit crimes, and so the equilibrium features partial deterrence. In addition, for intermediate values of $\sigma+\tau$, when the employers' cost of crime is low and/or most workers are dishonest, the complementarity between the wage level in state N and the crime choices of workers induces multiple equilibria to exist, both with full and with partial deterrence.

[23] When $\xi = \sigma + \tau$ we have $\underline{g_N} = \hat{g_N}$, while when $\sigma + \tau = h - \frac{1}{2}$ $\overline{g_N} = \hat{g_N}$.



(a) $h\,(1-r) < \frac{1}{2}$ (b) $h\,(1-r) \geq \frac{1}{2}$

Figure 2: Deterrence in equilibrium

Figures 2 illustrates the equilibrium level of deterrence as characterized in Proposition 1. The solid line depicts the level of deterrence at the (stable) partial-deterrence equilibrium as a function of $\sigma+\tau$, when that equilibrium exists. The dashed line depicts the full deterrence equilibrium, when that exists. In line with the result established in the proposition, we see that a marginal increase of the expungement rate always leads to a lower level of deterrence, since the punishment for crimes diminishes. The proposition shows that a similar effect also obtains when r increases: in populations where a larger share of workers receive crime opportunities, the crime rate in N is higher, for the same value of g_N, hence w_N is lower, and so is the deterrence effect.

3.2 Optimal expungement rate

In this section, we analyze the properties of the optimal expungement rate, that is, the value of σ that maximizes social welfare in equilibrium. Since employers and workers are assumed to be risk neutral, it is natural to take total surplus as the measure of social welfare. Furthermore, employers always break even in equilibrium, hence they are unaffected by changes in σ. Social surplus is thus equal to the sum of the lifetime expected utilities of all workers:

$$W \equiv rV_N + (1-r)\,V$$

where V_N and $V = \frac{w_N}{\tau}$ are the expected values of the payoffs, over their lifetime, of a dishonest worker and of a honest worker, respectively. After substituting the expression we found for V_N and few simple manipulations, social welfare can be written as $W = \frac{1}{\tau}\,(\pi - rL)$ with

$$L \equiv \frac{\sigma+\tau}{1-\underline{g_N}+\sigma+\tau}\left(1-\underline{g_N}\right)\left(h-\frac{1+\underline{g_N}}{2}\right)+\frac{1-\underline{g_N}}{1-\underline{g_N}+\sigma+\tau}\left(h-\frac{1}{2}\right) \tag{9}$$

denoting the net social loss from the crimes committed — per unit of time — by a dishonest worker in N (first term) and in C (second term). Hence, maximizing W is equivalent to minimizing L. The expression of L highlights the fundamental trade-off between recidivism and deterrence. We see that the direct effect of increasing the expungement rate σ (that is, keeping g_N fixed) is to raise the share of dishonest workers who are in N, because they transition at a faster rate out of state C to return to N, which in turn reduces recidivism because deterrence in N is higher than in C. The second, indirect effect of an increase in the expungement rate is that it increases workers' incentives to commit a crime in N: as shown in Proposition 1, faster expungement reduces deterrence in N, that is, g_N. This means not only that workers exit faster from N too but also that the net social loss from crimes committed in N changes. Since only some crimes are committed in N, both the total amount and the average type of crimes committed are affected by an increase in σ.

The balance between deterrence and recidivism is further complicated by the fact that the crime rate is not all that matters for welfare. Deterrence is, in fact, not always beneficial: some crimes, namely those with $g > h$, should not be deterred. Therefore, which crimes are deterred in equilibrium also matters. This can be seen from the expression of the net social loss from crimes in (9): $(1 - g_N)\left(h - \frac{1+g_N}{2}\right)$ is the value of the net loss in state N and $h - \frac{1}{2}$ is the net loss in state C. The latter is positive if and only if $h > \frac{1}{2}$, that is, as long as the social cost of crimes is greater than the average private gain from crime. In contrast, the social cost of crimes committed by dishonest workers in state N depends on the level of deterrence g_N induced by the expungement policy σ. It is immediate to verify that the value of g_N that minimizes this loss is $g_N = h$, thus the same as the first-best level of deterrence g^{FB}. We will then say that we have *over-deterrence* if $g_N > h$ and *under-deterrence* when the opposite inequality holds. This is with a slight abuse of terminology, since deterrence in N also affects the time workers spend in C, where there is no deterrence. The total loss L, in fact, also includes the loss due to crimes committed in C, which we can see is greater than the one in N if and only if $h - \frac{g_N}{2} > 0$.[24]

The socially optimal expungement policy is the value of σ that minimizes the expected net social loss from crimes, L, which obtains in equilibrium.[25] An important first step in finding the optimal policy is to determine whether or not full deterrence ($g_N = 1$) can be attained in equilibrium. We have shown in Proposition 1 that full deterrence arises in equilibrium only if $\sigma + \tau \leq h - \frac{1}{2}$. It can then be implemented by an expungement rate $\sigma = 0$ (that is, by never expunging crimes from a record) only if workers are sufficiently long-lived, that is, if $\tau \leq h - \frac{1}{2}$.[26]

[24]This inequality is always verified when the cost of crimes to employers is not too low, that is, when $h > \frac{1}{2}$. It is instead violated when we have a sufficiently high level of over-deterrence in state N, that is, when $g_N > 2h$.

[25]Note that in our model the optimal expungement policy, σ, is a function of the harm from crime, h. We implicitly assume that h is set for the whole population and constant over time, that is, successive crime opportunities have the same h. Considering the possibility that the next crime opportunity has a different h (that is, considering a sequence of potentially different crimes) would make the analysis more involved without changing the main driver of our results: that deterrence in C is less than deterrence in N. In line with our results, the New York State Clean Slate Act (see footnote 21 above), sets a longer waiting period before a record is automatically sealed for more serious convictions: felonies (8 years) versus misdemeanors (3 years).

[26]Note that when this inequality is strict, full deterrence can also be implemented with $\sigma > 0$ but sufficiently close to 0. All these values are equivalent to $\sigma = 0$ in terms of social welfare. This is because,

Partial deterrence ($g_N < 1$) can instead always be implemented in equilibrium by setting σ sufficiently high. Partial deterrence may be optimal even when full deterrence is implementable ($\tau \leq h - \frac{1}{2}$), if the social loss it induces is less than the social loss with full deterrence. The latter is always equal to $L = 0$ since, when all workers are deterred in N, no crime is committed in equilibrium in that state, hence no worker transitions to C. Therefore, when full deterrence is implementable, partial deterrence is optimal only if it yields a negative social loss from crimes, $L < 0$. A negative social loss arises because some crimes are "efficiently undeterred" as they generate private gains that are larger than the harm. It is immediate to verify that a necessary (but not sufficient) condition for $L < 0$ is $\sigma + \tau > h - \frac{1}{2}$,[27] that is, the expungement rate, σ, must be large enough compared to the employer's harm from crime h.

More precisely, we show in the next proposition that there exists a threshold level $\hat{h} > \frac{1}{2}$ such that, if $h \geq \hat{h}$, it is optimal never to expunge the record, $\sigma = 0$, when this policy induces full deterrence. A weakly positive expungement rate $\sigma \geq 0$, implementing partial deterrence, is instead optimal if $h < \hat{h}$. The intuition for this result is that, whenever $h < 1$, full deterrence induces over-deterrence in N and, at the same time, no worker ever transitions to C since no crime is committed. Partial deterrence, in turn, can be modulated to be closer to the first best level $g_N = h$, in state N, but may also result in some workers transitioning to C where they are undeterred. We thus see that a trade-off emerges in the choice between partial and full deterrence when the latter is implementable.

When the employers' cost of crimes h is sufficiently large, the benefits of implementing a level of deterrence in N that is close to the first best — that is, of allowing crimes whose private benefits are greater than the social costs — are small and outweighed by the costs of inefficiently too many crimes committed in C. This favors full deterrence, if achievable. Vice versa, when h is not too large, over-deterrence in N is more problematic than under-deterrence in C, favoring partial deterrence. In particular, we show that when the cost of crimes h is less than the average benefit of crime, $h < \frac{1}{2}$, at the optimum workers are always under-deterred, $\underline{g_N} < h$. When h is low, lowering deterrence below h via a faster rate of expungement, as we show in the proof[28] always reduces the time workers spend in C, which is beneficial. The situation is thus analogous to the one when h is large (greater than $\hat{h}$) where raising deterrence to full deterrence allows to reduce (to 0 in that case) the time spent in C, even though the effects on the crime rate in N work in opposite directions. Also, when τ is sufficiently low, so that incentives are strong, and the optimum features under-deterrence ($h < 1/2$), the optimal policy exhibits forgetting after some finite time ($\sigma > 0$).

---

even though with $\sigma > 0$ a worker can exit from C, this is irrelevant when there is full deterrence in N as nobody commits a crime and transitions to C. We can hence focus without loss of generality on $\sigma = 0$ as the value of the policy associated to full deterrence.

[27] Under this condition, as shown in Proposition 1, there is a unique equilibrium with partial deterrence. Hence, whenever at the optimal expungement policy we have a multiplicity of equilibria (and hence full deterrence is implementable), the optimum features full deterrence.

[28] See Fact OA.6 in the Online Appendix. Note also that first-best deterrence, $g_N = h$, is attainable as long as τ is small enough, which is easy to show.

The effects on the optimal policy of raising τ, so that workers become shorter-lived, are also of interest. We see the optimal level of deterrence remains constant as long as τ is small enough. When full deterrence is optimal, small changes in τ are irrelevant; when partial deterrence is optimal, a small increase in τ can be compensated by a reduction in σ, so that it is optimal to keep records for a longer period of time when individuals are short-lived.[29]

Proposition 2. *For any fraction of dishonest workers, r, there exists a cutoff level of the harm, $\hat{h} \in \left(\frac{1}{2}, 1\right)$, decreasing in r and such that:*

1. *If $h < \frac{1}{2}$, the optimal policy is a weakly-positive expungement rate, $\sigma^* \geq 0$, such that workers are under-deterred, $\underline{g_N}^* < h$. When τ is small enough, the optimal expungement rate is strictly positive, $\sigma^* > 0$.*
2. *If $h \geq \max\left\{\hat{h}, \frac{1}{2} + \tau\right\}$, the optimal policy is a null expungement rate, $\sigma^* = 0$, such that workers are fully deterred, and hence over-deterred, $\overline{g_N}^* = 1 > h$.*
3. *In the intermediate case $\frac{1}{2} \leq h < \max\left\{\hat{h}, \frac{1}{2} + \tau\right\}$, the optimal expungement rate, $\sigma^* \geq 0$, features partial deterrence (and we may have either over- or under-deterrence). When, in particular, $\frac{1}{2} + \tau < h < \hat{h}$, the optimal expungement rate is strictly positive, $\sigma^* > 0$.*

In all cases, the optimal level of deterrence is positive and weakly increasing in the longevity of workers, $\frac{1}{\tau}$, and the optimal expungement rate is finite and weakly decreasing in $\frac{1}{\tau}$.

It is easy to see that we always have some deterrence at the optimum, $\underline{g_N} > 0$ (and hence a finite expungement rate, σ), even when h is very close to 0. No deterrence, $\underline{g_N} = 0$, can only be attained by setting $\sigma = \infty$, so that dishonest workers commit all crimes in N, move then to C, and transition immediately back to N. But that would mean having under-deterrence both in C and in N and, in such a situation, a marginal reduction in σ is beneficial as it improves deterrence $\underline{g_N}$ in N.[30]

It is useful to briefly discuss also the distributional effects of expungement. Honest workers are always made worse-off by expungement, since w_N unambiguously goes down, in line with the remarks made by Funk (1995) and Funk and Polsby (1997). The effect on dishonest workers, on the other hand, is more complex as they are not only hurt by the decline of the wage in N but also benefit from the increase in crimes committed in N; the variation in the time spent in N relative to C then further affects their utility. Hence

[29] More precisely, when full deterrence is optimal small changes in τ do not change the optimal policy as long as $h \geq \max\left\{\hat{h}, \frac{1}{2} + \tau\right\}$. When instead partial deterrence is optimal, a small increase in τ can be compensated by a reduction in σ so that $\sigma + \tau$ remains constant. Keeping $\sigma + \tau$ unchanged is optimal since the equilibrium level of deterrence, $\underline{g_N}$, only depends on $\sigma + \tau$, and the social welfare loss L only depends on $\sigma + \tau$ and $\underline{g_N}$. This follows from the fact that in the equation determining the equilibrium value of g_N (see A.1 in the Online Appendix) only the sum $\sigma + \tau$ appears. This is however feasible only as long as the reduced value of σ is nonnegative. When τ becomes large enough so that the non-negativity constraint on σ binds, $\sigma + \tau$ must increase as τ increases and the level of deterrence $\underline{g_N}$ must decrease.

[30] Formally, we have $\lim_{\sigma \to \infty} \frac{dL}{d\sigma} = -h \frac{\partial \underline{g_N}}{\partial \sigma} > 0$.

whenever expungement is optimal, this is always driven by the benefits for dishonest workers.

Figures 3 and 4 illustrate our findings regarding the optimal policy for the case where $r = \frac{1}{4}$. We see that, in this case, under-deterrence is optimal not only for all $h < \frac{1}{2}$ but also for most values $h \geq \frac{1}{2}$ when full deterrence is not attainable. The extent to which under-deterrence occurs at the optimum can be clearly seen from Figure 4a, where the level of deterrence $\underline{g_N}^*$ at the optimal policy is reported as a function of τ and h. For instance, when $h = \frac{1}{2}$, the value of $\underline{g_N}^*$ is about half the value of h. We should stress that this is not due to the constraints faced in implementing deterrence: we see, in fact, in Figure 4b that as long as workers are sufficiently long-lived — $\tau \leq 0.7$ — the optimal level of σ is strictly positive (so it could be lowered) and implies that the length of time before a crime is expunged from the criminal record ranges from 1.5 to 10 periods.

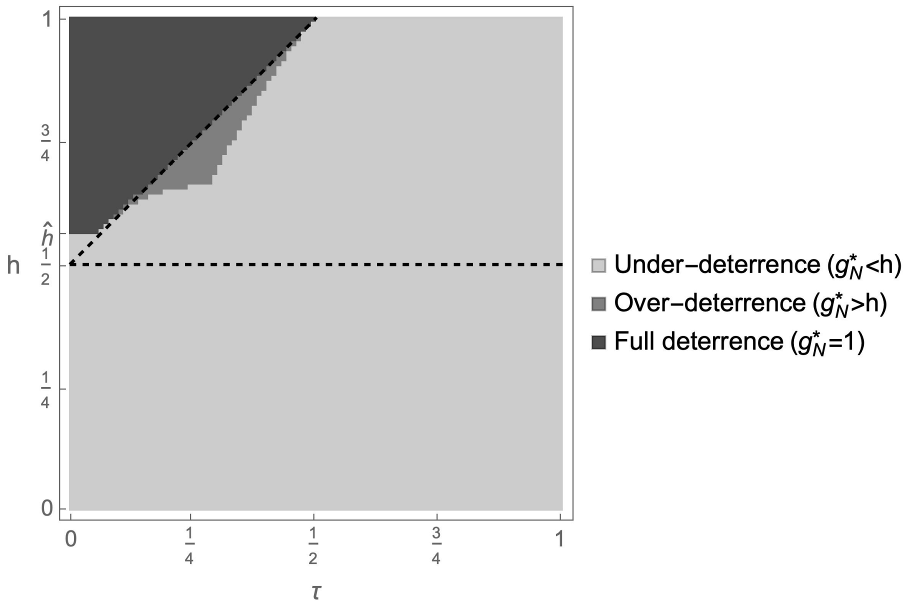
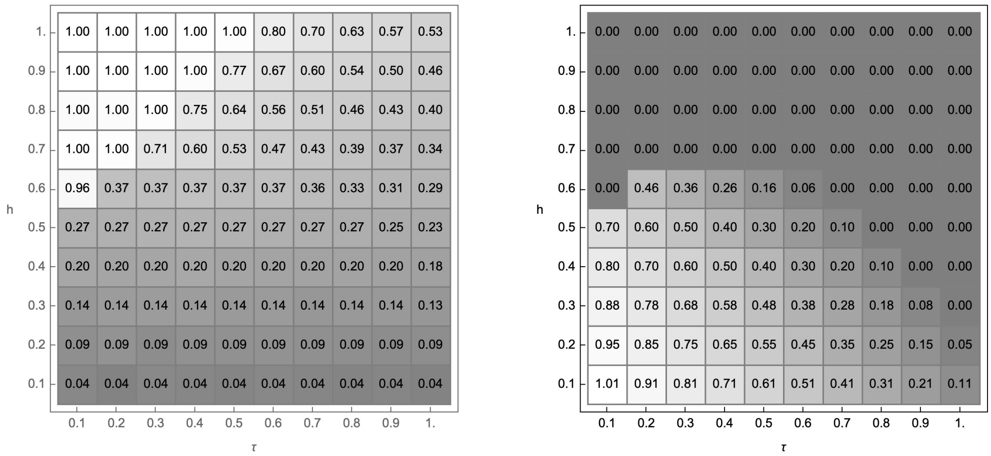


Figure 3: Over- and under-deterrence at the optimal policy (with $r = 0.25$)

It is also of interest to examine how the presence of a larger share of dishonest workers in the population affects the optimal policy. As shown in Proposition 1, an increase in r lowers deterrence. We find[31] that the policy response is to slow down expungement, though not too much, so that the level of deterrence at the optimal policy is still lower when r is higher.

[31] Figures OA.8 and OA.9 in the Online Appendix report the optimal policy and associated level of deterrence for various values of r higher than 0.25.



(a) Optimal deterrence, g_N^*

(b) Optimal expungement rate, σ^*

Figure 4: The optimal policy (with $r = 0.25$)

4 Criminal records and nonmonetary sanctions

Nonmonetary sanctions, given by a vast array of measures with varying degrees of coercion, are widely used as a direct consequence of a criminal conviction. They have an incapacitation effect, which reduces both the workers' ability to work and their crime opportunities. For instance, if an individual's driving license is suspended for a period of time, the individual cannot commit traffic violations and is unable to drive to or for work. Nonmonetary sanctions have a selective effect on crime ex ante (before a crime is committed), like the criminal record, as they deter crimes with low private benefits g. But they also have a uniform ex-post effect on crime (after a crime is detected and a sanction imposed), as incapacitation reduces crime opportunities irrespective of the benefit g.

In the next section, we examine the consequences of having nonmonetary sanctions as the only punishment for crime, thus in the absence of a criminal record. This allows us to relate to the large literature on law enforcement (Becker, 1968; Polinsky and Shavell, 2000). We can then compare the effects of nonmonetary sanctions to those of the criminal record derived in the previous section. As we will see, an additional trade-off arises in this environment, between the effects of the sanctions (when in place) on productivity and crime and the deterrence they induce prior to committing a crime. In the subsequent Section 4.3, we extend the analysis by allowing for the simultaneous presence of nonmonetary sanctions and a criminal record.

4.1 Effects of nonmonetary sanctions

The working of the labor market with nonmonetary sanctions is similar in several respects to the one we saw with criminal records. Workers are born in state N as before.

If they commit a crime, they are now punished with the imposition of a nonmonetary sanction and hence transition to the punishment state, denoted I (for incapacitation). In I, the worker is partially incapacitated so that crime opportunities arrive at a lower rate $1-\lambda<1$, and the worker's productivity is also reduced to $(1-\theta)\,\pi<\pi$.[32] When the term of the sanction ends — which happens at rate ζ, the release rate — the worker returns to N without any record of the crimes committed. Thus, there is no state C. Death still occurs at a rate τ. We assume that any new crime committed by a worker when in I resets the clock of the sanction. Then a new sanction period starts but, as for expungement in the previous section, due to the memoryless property of the Poisson process, the expected time to release is unchanged.[33] As a consequence, in I the worker commits all crimes whenever an opportunity arises (but, as we just said, there are fewer crime opportunities). In N, on the other hand, only the crimes whose benefits exceed the costs due to the imposition of a sanction are committed. With a slight abuse of notation, we still denote by $\underline{g_N}$ and $\overline{g_N}$ the level of deterrence at an equilibrium respectively with partial and full deterrence, by w_N and V_N the wage and the lifetime utility of a (dishonest) worker starting in N, and by L the social loss; differences will be clear from the text and the formulas.

The flow of motion for the population of dishonest workers in this scenario is described in Figure 5.

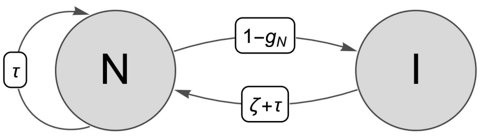


Figure 5: Population dynamics for dishonest workers with nonmonetary sanctions

The expression for the equilibrium wage in N as a function of g_N is then analogous to the one obtained in the previous section:

$$w_N = \pi - \frac{r\,(\zeta+\tau)}{(1-r)\,(1-g_N)+\zeta+\tau}\,(1-g_N)\,h \tag{10}$$

In contrast, the expression of the wage w_I in state I, after a crime is committed, is different from the one we derived for the wage w_C with a criminal record. Even though it is still true that all possible crimes are committed in I, the wage reflects both the lower productivity at work and the lower frequency of crime opportunities:

$$w_I = (1-\theta)\,\pi - (1-\lambda)\,h$$

[32]The parameter θ is also a proxy for any other cost of applying a nonmonetary sanction. These costs may include, for instance, the costs of electronic monitoring devices, probation officers, and prison security.

[33]This assumption helps to make the analysis more tractable and facilitates the comparison with the effects of the criminal record. Allowing for tougher sanctions for the crimes committed when individuals are already subject to a sanction would further reduce the crimes committed in I, but otherwise not change the main qualitative features of equilibria.

The expression of w_I presupposes that employers are aware of the fact that the worker is being subjected to a sanction. The value of the lifetime expected utility starting, respectively, in I and N can then be derived following a similar procedure as in the previous section:

$$\begin{aligned} V_I &= \frac{1}{\tau}\frac{\zeta\left(w_N+\frac{1-g_N^2}{2}\right)+(1-g_N+\tau)\left(w_I+\frac{1-\lambda}{2}\right)}{1-g_N+\zeta+\tau} \\ V_N &= \frac{1}{\tau}\frac{(\zeta+\tau)\left(w_N+\frac{1-g_N^2}{2}\right)+(1-g_N)\left(w_I+\frac{1-\lambda}{2}\right)}{1-g_N+\zeta+\tau} \end{aligned}$$

Similarly, the equilibrium level of deterrence in N equals the difference between the expected utilities in these two states, as in (8):

$$g_N = V_N - V_I = \frac{w_N - w_I + \frac{\lambda-g_N^2}{2}}{1-g_N+\zeta+\tau} \tag{11}$$

Note that now the equilibrium level of deterrence g_N depends on π, as so does the wage difference $w_N - w_I$, reflecting the fact that under a nonmonetary sanction the worker's productivity is proportionally reduced. A nonmonetary sanction not only conveys information to employers — as long as it is in place — about the commission of a crime, just like the criminal record does, but also reduces crime opportunities ($\lambda > 0$) and productivity ($\theta > 0$). Thus, we will typically have $w_I \neq w_C$ and the consequences of nonmonetary sanctions and criminal records will be different. In particular, when comparing the levels of deterrence obtained with nonmonetary sanctions and with criminal records, we have:

Proposition 3. *When nonmonetary sanctions are the only punishment for crimes:*

1. *Full-deterrence can be implemented in equilibrium for a broader range of values of τ than with a criminal record iff $\frac{\theta}{\lambda} > \frac{h-\frac{1}{2}}{\pi}$;*
2. *If the expected length of punishment is the same as with a criminal record, $\sigma = \zeta$, the equilibrium level of deterrence is higher iff $\frac{\theta}{\lambda} > \frac{h-\frac{1}{2}}{\pi}$;*
3. *If the equilibrium features partial-deterrence, the severity of incapacitation:*
   (a) *regarding crime opportunities, λ, increases deterrence, $\frac{\partial g_N}{\partial \lambda} > 0$, if $h < \frac{1}{2}$, and reduces it, $\frac{\partial g_N}{\partial \lambda} < 0$, if $h > \frac{1}{2}$;*
   (b) *regarding productivity, θ, always increases deterrence, $\frac{\partial g_N}{\partial \theta} > 0$.*

To gain some intuition for this result, it is useful to start from claim 3. The reduction in crime opportunities ($\lambda > 0$) induced by incapacitation has two opposite effects on deterrence. On the one hand, it increases the wage employers are willing to pay in state I proportionally to h, the harm from crime suffered by employers. On the other hand, it reduces the utility of a worker from the crimes committed in I proportionally to $\frac{1}{2}$, the expected gain from a crime. Thus, depending on which effect prevails, an increase in λ reduces or increases deterrence. In contrast, the loss of productivity ($\theta > 0$), which is also induced by incapacitation, has an unambiguously positive effect on deterrence because it always reduces the wage and hence the worker's payoff in I.

Next, compare equations (8) and (11), which determine the level of deterrence respectively with criminal records and with nonmonetary sanctions. After substituting the expression of the equilibrium wages we see that, for the same length of the punishment, $\frac{1}{\sigma} = \frac{1}{\zeta}$, the two equations only differ for the presence in the latter of the additional term $\pi\theta - \lambda(h - \frac{1}{2})$. This captures precisely the sum of the two effects of incapacitation on deterrence described above. We then show in the proof of claim 2 that deterrence is higher under sanctions when this term is positive, and is lower otherwise.

The same term also determines the strength of the incentives that allow to sustain full deterrence in equilibrium. We show that full deterrence obtains with nonmonetary sanctions when $\theta\pi + (1-\lambda)\, h + \frac{\lambda - 1}{2} \geq \zeta + \tau$. The corresponding condition with a criminal record (Proposition 1) is $h - \frac{1}{2} \geq \tau + \sigma$. Hence, the productivity loss $\pi\theta$ due to the imposition of a nonmonetary sanction positively contributes to achieving full deterrence, while the decrease in crime opportunities contributes by an amount $-\lambda\left(h - \frac{1}{2}\right)$, positive when $h < \frac{1}{2}$ and negative otherwise (see claim 1).

These two features of incapacitation also affect overall welfare. With nonmonetary sanctions, the net social loss is:

$$L = \frac{\zeta + \tau}{1 - g_N + \zeta + \tau}\,(1 - g_N)\left(h - \frac{1 + g_N}{2}\right) + \frac{1 - g_N}{1 - g_N + \zeta + \tau}\left((1-\lambda)\left(h - \frac{1}{2}\right) + \theta\pi\right) \tag{12}$$

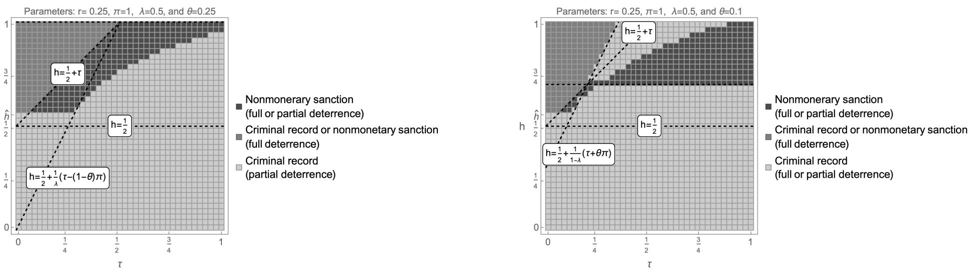
The first addendum is the loss due to crimes committed in N and takes the same form as with a criminal record. The second one is the loss in I, which, as already noted, is different from the expression of the loss in state C obtained in the previous section, as it includes the effects of incapacitation on crime opportunities and productivity in I.

The expression of the social loss allows us to see more clearly the different trade-offs arising when the duration of the sanction, as determined by ζ, is varied. In contrast to the case with criminal records, the direct effect on the overall crime rate of a faster release of offenders is now ambiguous since fewer crime opportunities arrive when a sanction is in place, while more arrive after release though only a fraction $(1 - g_N)$ of them are exploited. In addition to this direct effect, there is again an indirect effect, due to the fact that faster release lowers deterrence in N, which increases the crime rate in that state as well as the transition back to I. Release has also a novel — unambiguously positive — effect on productivity and hence on output. There is so a trade-off between incapacitation and deterrence when the duration of sanctions is varied.

Our main focus here is not on characterizing the optimal policy regarding nonmonetary sanctions, that is, the value of ζ that maximizes social welfare.[34] Rather, we aim to compare the relative welfare benefits of using nonmonetary sanctions versus criminal records. As we will see below, the minimum value of the loss under nonmonetary sanctions can be larger or smaller than the minimal loss attainable under criminal

[34]The optimal ς can be derived proceeding along similar lines as in Section 3.2.

records, depending on the magnitude of the employers' harm from crimes, h, and on the value of other parameters.



(a) Large productivity loss

(b) Small productivity loss

Figure 6: The optimal choice between criminal record and nonmonetary sanction

This is shown in Figure 6, where we report for each value of the social cost of crime h and of τ (the inverse of the workers' expected lifetime), whether welfare is higher with a nonmonetary sanction (for the optimally chosen release rate ζ), with a criminal record (again for the optimal expungement rate σ), or is the same with both for given values of the other parameters. We see from the figure that, when the harm from crime is low, $h < \frac{1}{2}$, a criminal record yields a higher level of welfare than a nonmonetary sanction. In this case, the reduction in crime opportunities caused by incapacitation entails a net cost because the average private gain from crime is equal to $\frac{1}{2}$ and hence is greater than the social harm of crime. The productivity loss associated with the nonmonetary sanction constitutes an additional welfare cost. The benefits consist in the deterrence level g_N that is induced. In contrast, with a criminal record deterrence is achieved without incurring any such costs. From claim 2 of Proposition 3 it follows that deterrence is higher with sanctions;[35] however, deterring crime is not a primary concern for welfare when h is low.

When instead $h > \frac{1}{2}$, reducing crime opportunities generates a net social gain. Also, the socially desirable level of deterrence could be higher than what can be achieved with a criminal record, and nonmonetary sanctions could prove more effective to achieve it. More specifically, Figures 6a and 6b consider two situations where the productivity loss θ with nonmonetary sanctions is, respectively, relatively large (25%) and small (10%). In both cases, we see that the level of social welfare attainable with a criminal record is strictly higher than the one achievable with nonmonetary sanctions not only for $h < \frac{1}{2}$ but, in fact, for all $h < \hat{h}$. As shown in Proposition 2, when $h < \hat{h}$ in the presence of a criminal record welfare is higher with partial than with full deterrence, so the crimes of higher value for the worker are committed. The facts that, for the parameter values of Figure 6, the welfare loss in the punishment state with sanctions (in I) is higher than with criminal records (in C) and that deterrence in the latter case is not constrained (expungement occurs in finite time) then imply that welfare is higher

[35]It is immediate to verify that the condition stated in claim 2 is always satisfied when $h < \frac{1}{2}$, for any θ, π, and λ.

with criminal records.

The situation is instead different for higher values of $h > \hat{h}$, for which full deterrence is optimal, if implementable, with a criminal record. In this case achieving a high level of deterrence is key, and if nonmonetary sanctions are more effective in this respect, they will be preferable. That is indeed the case in the situation considered in Figure 6a.[36] Even when neither measure can implement full deterrence, partial deterrence is still higher with nonmonetary sanctions.[37]

Instead, in the case of Figure 6b, given the lower value of θ, the inequality in claim 1 is not satisfied for h close to 1. We see from the figure that for those values, social welfare is higher with a criminal record because, when τ is sufficiently low, it allows to attain full deterrence while a nonmonetary sanction does not.[38] For the lower value of the productivity loss θ the situation is then the opposite, when h is high, to the one we found in Figure 6a.

The incapacitation effect considered in this section needs not be the result of the imposition of a formal sanction. Also a criminal record has an incapacitation effect when employers adjust the tasks assigned to workers with a criminal record, so as to curb crime opportunities at the cost of reduced productivity. For instance, a dishonest accountant could be given manual tasks instead of bookkeeping, or a teacher with a record for sex offenses could be cleaning the building after school hours instead of teaching. Therefore, the model of this section could also be interpreted as a model of criminal records where the incapacitation-deterrence trade-off is also operative.

4.2 Imprisonment

An interesting special case of nonmonetary sanctions is when their incapacitation effects are maximal: $\theta = \lambda = 1$. In this extreme case, a worker subjected to the sanction can neither work nor commit crimes. We can view this as a somewhat simplified description of imprisonment. Workers receive then a payoff of 0 when in prison since they receive no wage payment, $w_I = 0$, nor have any benefit from crime.[39] In this case, the properties of the welfare maximizing duration ζ of the sanction are quite stark:

[36] For those values, the inequality in claim 1 of Proposition 3 is satisfied for all h and so the set of values of τ — and h — for which full deterrence is implementable with a nonmonetary sanction is strictly larger and includes the set for which this is possible with a criminal record. In particular, in the triangle above the line $h = \frac{1}{2} + \tau$, both instruments allow to achieve full deterrence and hence the highest level of social welfare that is attainable is the same in the two cases. In the area between the lines $h = \frac{1}{2} + \tau$ and $h = \frac{1}{2} + \frac{\tau-(1-\theta)\pi}{\lambda}$, only a nonmonetary sanction achieves full deterrence and hence is preferable.

[37] At the same time, as we can see from the expression of the net social loss we derived in the two cases, τ has a direct effect on the level of the social loss. We then see from the figure that, when τ becomes sufficiently large — and so the workers' horizon shortens — criminal records become again preferable to nonmonetary sanctions.

[38] The same is true for higher values of τ (as long as τ is not too high), since a higher level of deterrence is attained with a criminal record, even though full deterrence is not implementable.

[39] There is of course a heavy disutility from imprisonment in reality. We disregard it in the model as it does not affect our qualitative results.

Proposition 4. *When $\theta = \lambda = 1$, for any fraction $r \in (0, 1)$ of dishonest workers:*

1. *If $h < \frac{1}{2}$ the optimal policy is immediate release, $\zeta \to \infty$, that is, no imprisonment, inducing no deterrence, $\underline{g_N} = 0$;*

2. *If $h \geq \frac{1}{2}$, when workers are sufficiently long-lived, $\tau \leq \pi$,*[40] *the optimal policy is never release, $\zeta = 0$, that is life imprisonment, inducing full deterrence, $\overline{g_N} = 1$.*

Hence, the optimal policy with imprisonment is always extreme, either sufficiently harsh to induce full deterrence or extremely lenient so that there is no deterrence. The choice between the two depends on whether the employer's cost of crime, h, is larger or smaller than the worker's benefit of committing any crime whenever the opportunity arises. The way to understand this finding is that the punishment given by imprisonment is quite effective in deterring crime (as long as τ is not too high ($\tau \leq \pi$), full deterrence is implementable for any level of the employers' harm from crime h), but is also quite costly since it entails a complete loss of productivity.[41] Thus, at the optimum, no worker ever ends up in prison either because the threat of imprisonment discourages all crimes, or because crimes are never punished. This all-or-nothing solution is in line with the findings obtained in static models, thus abstracting from deterrence, by Shavell (1987) and Kaplow (1990). When the punishment is imprisonment, the optimal policy results in under-deterrence when the cost of crime is low and in over-deterrence when such a cost is high (in accordance with Polinsky and Shavell 1984; Miceli 2010, 2012).[42]

On the basis of the above results, we can again examine whether a higher level of social welfare is attained using criminal records or imprisonment. The following Figure 7 reports, for each value of h and τ, whether welfare is higher under criminal records or imprisonment (for optimal σ and ζ).

[40] When instead $\tau > \pi$, full deterrence is not implementable even with the maximal punishment $\zeta = 0$. The optimum in that case could obtain at the maximal implementable level of (partial) deterrence, thus still with $\zeta = 0$, but might also be at no deterrence, $\zeta = \infty$.

[41] We see from 12 that the net social loss in I is equal to π and so is always greater than the one in C with criminal records, given by $h - \frac{1}{2}$.

[42] Unlike most of the literature, Miceli (2012) presents a dynamic model of the deterrence and incapacitation effect of imprisonment, as we do. However, in his model the private cost of imprisonment is an exogenous disutility from incarceration. In contrast, our focus is on the labor market and the private cost of imprisonment is an endogenous wage loss.

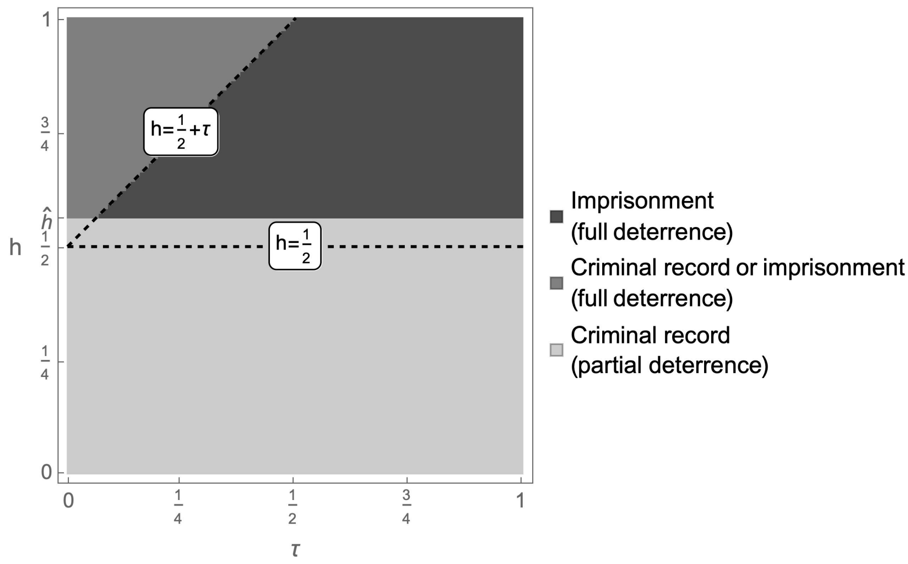


Figure 7: Optimal choice between criminal records and imprisonment (with $r = \frac{1}{4}$)

The figure shows that when the employers' loss from crimes is sufficiently low ($h < \hat{h}$), a criminal record is preferable, as we saw also happens in the case of partial incapacitation (see Figure 6). The reason is now more transparent. When $h < \frac{1}{2}$, at the optimal policy with imprisonment, nobody is ever sent to prison ($\zeta = \infty$), and all crimes, even those of very low value, are committed, so that $L = h - \frac{1}{2} < 0$. With a criminal record, no deterrence is always attainable, as we argued, but is always dominated by some positive level of deterrence.

When instead $h \in \left(\frac{1}{2}, \hat{h}\right)$, the relatively high cost of crime implies that the optimal imprisonment policy is the harshest possible ($\zeta = 0$), inducing full deterrence so that we have $L = 0$. For those values of h, the optimum under a criminal record features partial deterrence, so the benefits of some crimes for which the private gains exceed the social costs can be captured, yielding a negative level of the net social loss L. Conversely, when $h > \hat{h}$ crimes are more costly and full deterrence is always preferable to any level of partial deterrence with criminal records, but it is not always implementable. In contrast, full deterrence is always implementable with imprisonment for all $\tau \leq 1$. Hence, imprisonment is clearly superior (only weakly for τ sufficiently low),[43] due to its greater effectiveness in deterring crime. [44]

To sum up, the advantage of imprisonment relative to criminal records is that it is more effective in inducing full deterrence. It is then preferable in situations where the social cost of crime is sufficiently high and the discount rate is not too low, $\tau > h - \frac{1}{2} > 0$, which are cases in which the punishment is restricted under criminal records. The advantage of criminal records lies instead in the fact that they allow for better modulated

[43]When $\tau \leq h - \frac{1}{2}$, a (never expunged) criminal record also allows to deter all crimes, hence criminal records and imprisonment are equivalent, since no crime is committed in either case.

[44]This is not always the case, as we saw in Figure 6, when the incapacitation induced by sanctions is only partial.

levels of partial deterrence, and this is socially optimal when the cost of crime is not too high.

4.3 Optimal mix of expungement and nonmonetary sanctions

In the previous section we studied nonmonetary sanctions as an alternative to the shaming sanction induced by a criminal record. We investigate here whether and when it is beneficial to use a combination of both instruments. Here, individuals who commit a crime are first subjected to a non-monetary sanction (for an expected time $\frac{1}{\zeta}$) and then, after the sanction is lifted, their names are kept in a criminal record for some additional time (equal, in expectation, to $\frac{1}{\sigma}$).

Therefore, there are now 3 states. Workers are born in state N, where they work for a wage w_N and, if dishonest, face crime opportunities at rate 1. If they commit a crime — which happens if the gain from the crime, g, is larger than some threshold level g_N — they transition to state I where the application of a nonmonetary sanction reduces the rate of arrival of crime opportunities to $1-\lambda$ and the workers' productivity to $(1-\theta)\,\pi$. Note that in I workers are undeterred, hence they commit all crimes whenever an opportunity arises, and receive a wage w_I which reflects both their productivity and their propensity to commit crimes in that state, as well as the fact that a worker who is sanctioned is identified as being dishonest.

Then, at rate ζ, the sanction is lifted and workers transition from I to C, the criminal-record state. Here, there is no incapacitation but, differently from the basic model with only criminal records of Section 3, there is now some deterrence also in C because, if workers in that state commit a crime, they transition back to I. Therefore, they commit a crime only if $g > g_C$. If they do not commit crimes, they transition to N at rate σ. Employers thus adjust the wage paid in C, w_C, to take into account the fact that not all crimes are committed in this state.

It is easy to see that in equilibrium $g_C < g_N$, that is, the level of deterrence in C is less than in N, even though the consequences of a crime are the same in the two states. The reason is that the wage in C is lower than the wage in N due to the presence of honest workers in N. As a result, workers in N have more to lose from committing a crime and hence will be deterred more than when they are in C. In each state, if a worker dies (at rate τ), a new worker is born in N. Figure (8) depicts the flow of motion for the population of dishonest workers in the present environment.

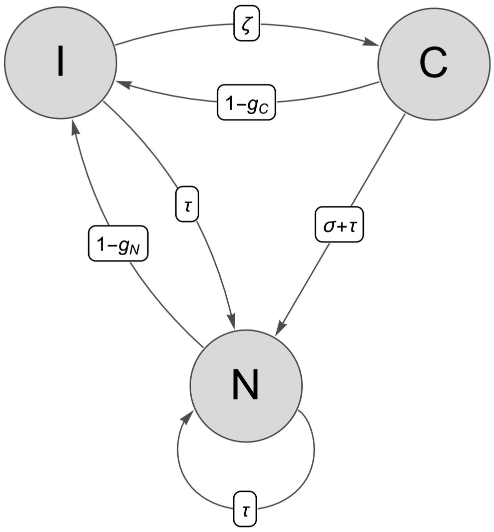


Figure 8: Population dynamics for dishonest workers with nonmonetary sanctions and criminal records

It is easy to verify that the net social loss from crimes, reflecting the loss in each state — given the level of deterrence, the rate of arrival of crime opportunities and, possibly, the productivity loss in that state — weighted by the distribution (p_N, p_I, p_C) of dishonest workers across the 3 states, is now:

$$\begin{aligned} L \quad = \quad & p_N \left(1 - g_N\right) \left(h - \tfrac{1+g_N}{2}\right) + p_I \left(\left(1 - \lambda\right) \left(h - \tfrac{1}{2}\right) + \theta\pi\right) \\ & + p_C \left(1 - g_C\right) \left(h - \tfrac{1+g_C}{2}\right) \end{aligned}$$

We will illustrate the possible benefits of suitably combining nonmonetary sanctions and criminal records in an environment where workers live on average for 5 periods, 60% of them are dishonest, labor productivity is equal to 1, and nonmonetary sanctions prevent 30% of the crimes and result in a 1% productivity loss; that is, $\tau = 0.2$, $r = 0.6$, $\pi = 1$, $\lambda = 0.3$, and $\theta = 0.01$. We allow the employers' harm from crime to take three possible values: (1) $h = 0.45 < \frac{1}{2}$, (2) $h = 0.52$ which is greater than 0.5 but smaller than $\hat{h}$ (equal to 0.55 in this setting) and (3) $h = 0.6 > \hat{h}$. For each of these values we find the optimal policy in the three scenarios we considered: the two cases with only criminal records or nonmonetary sanctions and the mixed case with both sanctions. Table 1 below reports the key findings [45].

When the harm is low, $h = 0.45$, we see the optimal policy in the mixed case yields a strictly lower net social loss than at the optimum with nonmonetary sanctions or criminal records. An analogous result is obtained when the harm is intermediate, $h = 0.52$, or high, $h = 0.6$. Thus, in all three cases the possibility of resorting to a suitably modulated combination of nonmonetary sanctions and criminal records proves

[45] Table 1 reports the first three decimal digits of the numerical specification outcomes without rounding. Full details are provided in Appendix Sections A.6 and A.7. Section OA.6 of the Online Appendix contains additional plots.

strictly superior, in terms of welfare, to the use of only one of these instruments. To gain some understanding of the determinants of this finding, it is useful to examine the equilibrium values of deterrence, wages, and workers' distribution across the states N, I, and C.

	Net social loss, L			Deterrence, g_N			Time spent in N, p_N		
	$h=.45$	$h=.52$	$h=.6$	$h=.45$	$h=.52$	$h=.6$	$h=.45$	$h=.52$	$h=.6$
Mixed case	$-.081$	$-.029$	.023	.200	.258	.348	.439	.413	.378
Only nonmonetary sanctions	$-.070$	$-.021$	.029	.168	.252	.358	.526	.441	.385
Only criminal record	$-.080$	$-.023$	.038	.202	.256	.344	.437	.431	.419

Table 1: Simulation results

Starting from the case where harm is intermediate ($h = 0.52$), we see the level of deterrence g_N in state N is *higher* in the mixed case — where 25.8% of crimes are deterred — than with nonmonetary sanctions or with criminal records. This happens in spite of the fact that the total expected length of the combined punishment in the mixed case — given by the expected release time from the nonmonetary sanction ($\frac{1}{\zeta} = 1$) plus the expected expungement time ($\frac{1}{\sigma} = 1$) — is equal to 2, and hence, is *less* than the expected duration of the punishment when there is only the nonmonetary sanction (2.56) or only the criminal record (2.75).

To reconcile these features, we should point out that in the mixed case, workers who commit a crime in C transition back to I, where they are subjected to a nonmonetary sanction, and then move back to C before — eventually — transitioning to N. The possible loop between states C and I reduces the time spent in N more than the release and expungement rates would suggest. We see in fact that, in the situation considered, dishonest workers spend *more* time out of N in the mixed case as compared to the other two cases and this contributes to explaining why deterrence in N is highest in the mixed case. To understand then the reason why welfare is higher in the mixed case, notice that dishonest workers also spend 19% of their time in state C where 10% of the crimes are deterred. Hence, we can say that altogether a larger fraction of the inefficient crimes (those of value less than $h = 0.52$) are deterred in the mixed case compared to the other two scenarios.

When crimes become more costly, $h = 0.6$, the welfare benefits of using sanctions and criminal records together increase further. Deterrence in state N is now higher with sanctions (and so is the time spent in N), but the level of deterrence in C in the mixed case is also considerably higher (more than 16% of the crimes are deterred). The determinants of the benefits of combining a nonmonetary sanction with a criminal record are then rather similar to the previous case. They reside in the fact that using both instruments allows the punishment to be more cost-effective. Under the parameter values used in this simulation, having workers in C, the criminal-record state, is not as expensive as having them in I, where incapacitation entails bigger welfare losses. At the same time, the losses in C are smaller when the criminal record is complemented by the use of nonmonetary sanctions because the threat of receiving them — that is, of transitioning back from C to I — generates some deterrence also in state C and this makes up for a possible decrease in deterrence in N.

When instead h is lower (0.45) and hence more crimes are socially efficient, the level of deterrence occurring in state C in the mixed case is negligible (1%). Thus, welfare is primarily driven by the time spent in N, which is now the highest in the mixed case, and by the level of deterrence in that state, which is slightly higher when only criminal records can be used. In this situation, the level of welfare is quite close to the optimum in those two scenarios and clearly higher than with only nonmonetary sanctions. This reflects the fact that the resort to sanctions to deter crimes is not so desirable when crimes are not too costly, in line with what we have seen already in Section 4.1. Thus, the marginal benefit of using sanctions on top of criminal records is also very limited in this case.

Lastly, note that as h increases across the three values considered, in the mixed case nonmonetary sanctions become more severe (higher expected length $\frac{1}{\zeta}$), while the duration of the criminal record is — weakly — reduced (lower $\frac{1}{\sigma}$), as indicated in Table 2.

	$h = 0.45$	$h = 0.52$	$h = 0.6$
Duration of the sanction, $\frac{1}{\zeta}$	0.1	1	2
Duration of the record, $\frac{1}{\sigma}$	2	1	1

Table 2: Optimal policies with a mix of nonmonetary sanctions and criminal records

Consider also that, when h is close to 0, the criminal record alone is much more efficient than the nonmonetary sanction. This is because incapacitation is very costly for society as it prevents crimes indiscriminately, thereby increasing rather than reducing the net social loss when the employer's cost of crime, h, is less than $\frac{1}{2}$ while the average private gain from crime is equal to $\frac{1}{2}$. The opposite obtains at the other end, where h is close to 1: sanctions are more effective at deterring crimes and are more often used to do so. This suggests that the main benefit of combining sanctions with a criminal record may be attained at intermediate levels of the social harm from crimes, h.

5 Conclusions

In this paper, we study the optimal duration of criminal records, shaped by the trade-off between deterrence and recidivism. This policy is then contrasted with nonmonetary sanctions, which rely on a different trade-off between deterrence and incapacitation. Our results show that the optimal expungement policy depends on the severity of the crime, as captured by its social harm, h. These stylized predictions match general trends in criminal law. Lastly, it is often remarked that individuals change over time. We show that our results depend on the individual's life expectancy, so that individuals with different ages could be optimally subjected to different policies.

We conclude by discussing three possible possible extensions of our model, which we leave to future research. In our analysis all workers are always employed. In reality, individuals with a criminal record may have a hard time finding a job. The effects of unemployment are somewhat similar to those of incapacitation: the unemployed are (at least partially) incapacitated from committing employment-related crimes and suffer a loss of productivity. Yet, these intuitions do not replace a full formal treatment.

We assumed that the probability of apprehension after a crime is 1. A lower probability would only rescale the model and hence would not change our results. What is crucial is that the probability of apprehension is exogenously given and the same for all h. This is realistic in a general enforcement framework (Shavell, 1991) where the probability of apprehension is the same for a broad class of crimes and hence cannot be set optimally for each type of crime, h. This is typically the case when enforcement authorities invest in methods and technologies to collect and process information about crimes in general, just like a police patrol detects both high speed and driving through a red light at the same time. In some, but probably a minority of cases, enforcement authorities are able to tailor the probability of apprehension to the specific crime. In that case, the probability of apprehension becomes a relevant policy variable and there is a trade-off between magnitude and probability of sanctions.

Another important feature of our model is that the harm a crime committed by a worker inflicts on the employer constitutes all the social harm from the crime. Is it easy, however, to conceive of situations in which crime harms third parties in addition to the employer, and to extend our model accordingly. Consideration of an additional cost of crime for third parties would have no effect on the equilibrium level of deterrence, because deterrence is induced by wages and these only reflect the harm to the employer. Yet, adding an additional cost clearly implies that a higher level of deterrence would be socially optimal. The optimal policy would then be different and feature longer periods before expungement and greater reliance on nonmonetary sanctions to induce higher deterrence.[46]

Finally, our analysis is also related to the literature on repeat offenders (starting with Polinsky and Rubinfeld, 1991). In these models, individuals with a criminal record may face a greater sanction than first-time offenders for the same crime. In our model, the state-imposed sanction is constant (though the stigma varies). Introducing greater sanctions for repeat offenders would reduce recidivism and hence lower the cost of stigmatization, but would also reduce the wedge between the wages of workers with and without a record — because those with a record commit fewer crimes now — which has a negative feedback effect on deterrence.

[46]More specifically, adding harm to third parties other than the employer in our framework would not affect Proposition 1 but would increase the optimal expungement rate derived in Proposition 2 because a higher deterrence is achieved by increasing σ (as shown in 1). Similarly, in the specification with nonmonetary sanctions, Proposition 3 would be unchanged, while full deterrence would become optimal for a larger set of parameters as in Proposition 4.

References

Agan, Amanda, Jennifer L Doleac, and Anna Harvey. 2023. "Misdemeanor Prosecution." *Quarterly Journal of Economics* 138:1453–1505.

Agan, Amanda, Andrew Garin, Dmitri K. Koustas, Alexandre Mas, and Crystal Yang. 2024. "Can you Erase the Mark of a Criminal Record? Labor Market Impacts of Criminal Record Remediation." *NBER Working Paper No. w32394*.

Agan, Amanda and Sonja Starr. 2017. "The Effect of Criminal Records on Access to Employment." *American Economic Review, Papers and Proceedings* 107:560–564.

———. 2018. "Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment." *Quarterly Journal of Economics* 133:191–235.

Avery, Beth and Han Lu. 2021. "Ban the Box: U.S. Cities, Counties, and States Adopt Fair Hiring Policies." *National Employment Law Project*.

Baker, Scott and Albert H. Choi. 2018. "Reputation and Litigation: Why Costly Legal Sanctions Can Work Better than Reputational Sanctions." *Journal of Legal Studies* 47 (1):45–82.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (2):169–217.

Blanchette, Jean-François and Deborah G Johnson. 2002. "Data Retention and the Panoptic Society: The Social Benefits of Forgetfulness." *The Information Society* 18 (1):33–45.

Blumstein, Alfred and Kiminori Nakamura. 2009. "Redemption in the Presence of Widespread Criminal Background Checks." *Criminology* 47 (2):327–359.

Bushway, Shawn D. 2004. "Labor Market Effects of Permitting Employer Access to Criminal History Records." *Journal of Contemporary Criminal Justice* 20 (3):276–291.

Cerda-Jara, Michael, David J. Harding, and The Underground Scholars Research Cohort. 2024. "Criminal Record Stigma in the Labor Market for College Graduates: A Mixed Methods Study." *Sociological Science* 11 (2):42–66.

Clean Slate Initiative, Annual Report. 2023. *Clean Slate Initiative*.

Corda, Alessandro. 2016. "More Justice and Less Harm: Reinventing Access to Criminal History Records." *Howard Law Journal* 60:1–60.

Craigie, Terry-Ann. 2020. "Ban the Box, Convictions, and Public Employment." *Economic Inquiry* 58 (1):425–445.

Diamond, Douglas W. 1989. "Reputation Acquisition in Debt Markets." *Journal of Political Economy* 97 (4):828–862.

Elul, Ronel and Piero Gottardi. 2015. "Bankruptcy: Is It Enough to Forgive or Must We Also Forget?" *American Economic Journal: Microeconomics* 7 (4):294–338.

Finlay, Keith, Michael Mueller-Smith, and Brittany Street. 2023. "Children's Indirect Exposure to the U.S. Justice System: Evidence From Longitudinal Links between Survey and Administrative Data." *Quarterly Journal of Economics* 138 (4):2181–2224.

Franklin, Marc A. and Diane Johnsen. 1981. "Expunging Criminal Records: Concealment and Dishonesty in an Open Society." *Hofstra Law Review* 9:733–774.

Funk, Patricia. 2004. "On the Effective Use of Stigma as a Crime-Deterrent." *European Economic Review* 48 (4):715–728.

Funk, T. Markus. 1995. "A Mere Youthful Indiscretion-Reexamining the Policy of Expunging Juvenile Delinquency Records." *University of Michigan Journal of Law Reform* 29:885–938.

Funk, T. Markus and Daniel D. Polsby. 1997. "Distributional Consequences of Expunging Juvenile Delinquency Records: The Problem of Lemons." *Washington University Journal of Urban and Contemporary Law* 52:161–186.

Furuya, Kaku. 2002. "A Socio-Economic Model of Stigma and Related Social Problems." *Journal of Economic Behavior and Organization* 48 (3):281–290.

Galle, Brian and Murat Mungan. 2020. "Predictable Punishments." *UC Irvine Law Review* 11:337–382.

Ganuza, Juan José, Fernando Gomez, and Marta Robles. 2016. "Product Liability versus Reputation." *Journal of Law, Economics and Organization* 32 (2):213–241.

Goggins, Becki R. and Dennis A. DeBacco. 2022. "Survey of State Criminal History Information Systems, 2020." *United States Bureau of Justice Statistics.*

Harel, Alon and Alon Klement. 2007. "The Economics of Stigma: Why More Detection of Crime May Result in Less Stigmatization." *Journal of Legal Studies* 36 (2):355–377.

Honigsberg, Colleen and Matthew Jacob. 2021. "Deleting Misconduct: The Expungement of BrokerCheck Records." *Journal of Financial Economics* 139 (3):800–831.

Iacobucci, Edward M. 2014. "On the Interaction Between Legal and Reputational Sanctions." *Journal of Legal Studies* 43 (1):189–207.

Kahan, Dan M. and Eric A. Posner. 1999. "Shaming White-Collar Criminals: A Proposal for Reform of the Federal Sentencing Guidelines." *Journal of Law and Economics* 42 (S1):365–392.

Kaplow, Louis. 1990. "A Note on the Optimal Use of Nonmonetary Sanctions." *Journal of Public Economics* 42:245–247.

Kluckow, Rick and Zhen Zeng. 2022. "Correctional Populations in the United States, 2020 - Statistical Tables." *United States Bureau of Justice Statistics*.

Kogon, Bernard and Donald L. Loughery Jr. 1970. "Sealing and Expungement of Criminal Records? The Big Lie." *Journal of Criminal Law, Criminology and Police Science* 61:378–392.

Kuehn, Sarah and Joachim Vosgerau. 2024. "The Public's Overestimation of Immorality of Formerly Incarcerated People." *Journal of Experimental Criminology* 20:269–295.

Lageson, Sarah Esther. 2020. *Digital Punishment: Privacy, Stigma, and the Harms of Data-Driven Criminal Justice*. Oxford University Press.

Leasure, Peter. 2018. "Misdemeanor Records and Employment Outcomes: An Experimental Study." *Crime and Delinquency* 65:1850–1872.

Leasure, Peter and Tia Stevens Andersen. 2016. "Recognizing Redemption: Old Criminal Records and Employment Outcomes." *NYU Review of Law and Social Change* 41:271.

McIntyre, T. J. and Ian O'Donnell. 2017. "Criminals, Data Protection, and the Right to a Second Chance." *The Irish Jurist* 58:27–55.

Miceli, Thomas J. 2010. "A Model of Criminal Sanctions That Incorporate Both Deterrence and Incapacitation." *Economics Letters* 107 (2):205–207.

Miceli, Thomas J. 2012. "Deterred or Detained? A Unified Model of Criminal Punishment." *Review of Law and Economics* 8 (1):1–20.

Mungan, Murat C. 2015. "Stigma Dilution and Over-Criminalization." *American Law and Economics Review* 18 (1):88–121.

———. 2016. "A Generalized Model for Reputational Sanctions and the (Ir)relevance of the Interactions Between Legal and Reputational Sanctions." *International Review of Law and Economics* 46:86–92.

———. 2017a. "Gateway Crimes." *Alabama Law Review* 68:671–706.

———. 2017b. "Reducing Crime Through Expungements." *Journal of Economic Behavior and Organization* 137:398–409.

———. 2019. "On the Optimality of Sealing Criminal Records and How It Relates to Adverse Selection, Productivity Reduction, and Stigma." *Supreme Court Economic Review* 26:135–149.

Murray, Brian M. 2021. "Retributive Expungement." *University of Pennsylvania Law Review* 169:665–716.

Musto, David K. 2004. "What Happens When Information Leaves a Market? Evidence from Postbankruptcy Consumers." *Journal of Business* 77 (4):725–748.

Nagin, Daniel and Joel Waldfogel. 1998. "The Effect of Conviction on Income Through the Life Cycle." *International Review of Law and Economics* 18 (1):25–40.

Natapoff, Alexandra. 2018. *Punishment Without Crime: How Our Massive Misdemeanor System Traps the Innocent and Makes America More Unequal.* Basic Books.

National Employment Law Project, Fact Sheet. 2024. *National Employment Law Project.*

Pager, Devah. 2003. "The Mark of a Criminal Record." *American Journal of Sociology* 108:937–975.

Pettler, Peter D. and Dale Hilmen. 1967. "Criminal Records of Arrest and Conviction: Expungement from the General Public Access." *California Western Law Review* 3:121–133.

Polinsky, A. Mitchell and Daniel L. Rubinfeld. 1991. "A Model of Optimal Fines for Repeat Offenders." *Journal of Public Economics* 46 (3):291–306.

Polinsky, A. Mitchell and Steven Shavell. 1984. "The Optimal Use of Fines and Imprisonment." *Journal of Public Economics* 24 (1):89–99.

———. 2000. "The Economic Theory of Public Enforcement of Law." *Journal of Economic Literature* 38 (1):45–76.

Prescott, James J. and Jonah E. Rockoff. 2011. "Do Sex Offender Registration and Notification Laws Affect Criminal Behavior?" *Journal of Law and Economics* 54:161–206.

Prescott, James J. and Sonja B. Starr. 2020. "Expungement of Criminal Convictions: An Empirical Study." *Harvard Law Review* 133:2460–2555.

Rasmusen, Eric. 1996. "Stigma and Self-Fulfilling Expectations of Criminality." *Journal of Law and Economics* 39:519–543.

Roberts, Jenny. 2015. "Expunging America's Rap Sheet in the Information Age." *Wisconsin Law Review* 2015:321–348.

Rosen, Alana E. 2019. "High Time for Criminal Justice Reform: Marijuana Expungement Statutes in States with Legalized or Decriminalized Marijuana Laws." *Working paper.*

Selbin, Jeffrey, Justin McCrary, and Joshua Epstein. 2018. "Unmarked? Criminal Record Clearing and Employment Outcomes." *Journal of Criminal Law and Criminology* 108:1–72.

Shavell, Steven. 1987. “The Optimal Use of Nonmonetary Sanctions as a Deterrent.” *American Economic Review* 77:584–592.

———. 1991. “Specific versus General Enforcement of Law.” *Journal of Political Economy* 99 (5):1088–1108.

Shlosberg, Amy, Evan J. Mandery, Valerie West, and Bennett Callaghan. 2014. “Expungement and Post-exoneration Offending.” *Journal of Criminal Law and Criminology* 104:353–388.

Snow, Carlton J. 1992. “Expungement and Employment Law: The Conflict Between an Employer’s Need to Know About Juvenile Misdeeds and an Employee’s Need to Keep Them Secret.” *Washington University Journal of Urban and Contemporary Law* 41:3–73.

Stevenson, Megan and Sandra Mayson. 2018. “The Scale of Misdemeanor Justice.” *Boston University Law Review* 98:731–777.

Tobin, D. Charles and Christine N. Walz. 2015. “Right to Be Forgotten, Expungement Law Raise New Challenges on the 40th Anniversary of Cox Broadcasting v. Cohn.” *Communications Lawyer* 31:4–10.

Uggen, Christopher, Mike Vuolo, Sarah Lageson, Ebony Ruhland, and Hilary K. Whitham. 2014. “The Edge of Stigma: An Experimental Audit of the Effects of Low-Level Criminal Records on Employment.” *Criminology* 52:627–654.

Wurie, Zainab. 2012. “Tainted: The Need for Equity Based Federal Expungement.” *Southern Region Black Law Students Association Law Journal* 6:31–57.

Yang, Crystal S. 2017. “Does Public Assistance Reduce Recidivism?” *American Economic Review* 107 (5):551–555.

Appendix

A.1 Proof of Proposition 1

It is convenient to rewrite equation (8) as

$$\gamma(g_N) = 0, \tag{A.1}$$

where:

$$\gamma(g_N) \equiv V_N - V_C - g_N \tag{A.2}$$

Using (7) and substituting the expressions for the wages from (1) and (2), we obtain:

$$\begin{aligned} \gamma(g_N) \quad = \quad & \frac{((1-r)(1-g_N)+\sigma+\tau)\left(h-\frac{g_N^2}{2}\right)}{(1-g_N+\sigma+\tau)((1-r)(1-g_N)+\sigma+\tau)} \\ & - \frac{r(\sigma+\tau)(1-g_N)h}{(1-g_N+\sigma+\tau)((1-r)(1-g_N)+\sigma+\tau)} \\ & - \frac{(1-g_N+\sigma+\tau)((1-r)(1-g_N)+\sigma+\tau)g_N}{(1-g_N+\sigma+\tau)((1-r)(1-g_N)+\sigma+\tau)} \end{aligned} \tag{A.3}$$

Note that $g_N = 0$ (no deterrence) cannot obtain in equilibrium if σ is finite, as we have noted in the text. Note further that $\overline{g_N} = 1$ (full deterrence) is an equilibrium iff $\gamma(1) \geq 0$, which in turn is satisfied iff $\sigma + \tau \leq h - \frac{1}{2}$. An equilibrium with partial deterrence is characterized by the values of $g_N \in (0,1)$ that are solutions of (A.1). To find those values we need to study the properties of the polynomial in the numerators of the right-hand side of (A.3).

We do so in the Online Appendix, Section (OA.1), where we show that this polynomial has either no, one, or two zeros, corresponding to the various cases in the proposition. Since $\gamma(0) > 0$, if the polynomial has no zero in $(0,1)$ then we must have $\gamma(g_N) > 0$ for all $g_N \in [0,1]$ and hence the only equilibrium is in full deterrence. If the polynomial has one zero, then $\gamma(g_N)$ must cross 0 from above at a unique interior value $\underline{g_N}$, which is then the unique partial deterrence equilibrium. Finally, if the polynomial has two zeros, then $\gamma(g_N)$ crosses 0 twice, the first time from above at $\underline{g_N}$ and the second time from below at $\hat{g_N}$, so that there are two candidate partial deterrence equilibria and a full deterrence equilibrium, $\overline{g_N} = 1$. Since, as we show, the solution $\hat{g_N}$ is dynamically unstable, there are two stable equilibria, $\underline{g_N} \in (0,1)$ and $\overline{g_N} = 1$. □

A.2 Proof of Proposition 2

By differentiating equation (9) we obtain:

$$\begin{aligned} \frac{dL}{d\sigma} \quad &= \quad \frac{\partial L}{\partial \underline{g_N}} \frac{\partial \underline{g_N}}{\partial \sigma} + \frac{\partial L}{\partial \sigma} \\ &= \quad -\frac{\sigma+\tau}{1-\underline{g_N}+\sigma+\tau}\left(h - \underline{g_N}\right)\frac{\partial \underline{g_N}}{\partial \sigma} - \underline{g_N}\left(h - \frac{1}{2}\underline{g_N}\right)\frac{(\sigma+\tau)\frac{\partial \underline{g_N}}{\partial \sigma}+1-\underline{g_N}}{\left(1-\underline{g_N}+\sigma+\tau\right)^2} \end{aligned} \tag{A.4}$$

The first addendum in (A.4) captures the effect of σ on the social cost of crimes committed in N, where $\frac{\sigma+\tau}{1-\underline{g_N}+\sigma+\tau}$ is the fraction of dishonest workers in N and $-\left(h - \underline{g_N}\right)\frac{\partial \underline{g_N}}{\partial \sigma}$

is the derivative of this cost with respect to σ due to the induced change in deterrence. The second addendum captures the effect on L of the induced variation in the proportion of dishonest workers in N and in C (due both to the direct effect of changing σ and the indirect effect of varying $\underline{g_N}$): $\frac{(\sigma+\tau)\frac{\partial \underline{g_N}}{\partial \sigma}+1-\underline{g_N}}{\left(1-\underline{g_N}+\sigma+\tau\right)^2}$ is in fact the derivative of the share of dishonest workers in N and $\underline{g_N}\left(h-\frac{1}{2}\underline{g_N}\right)$ is the difference between the net social loss in N and in C.

If $h<\frac{1}{2}$, we have $\sigma+\tau>0>h-\frac{1}{2}$, and hence we are either in case $1\,(c)$ or in case $2\,(b)$ of Proposition 1, yielding a unique interior equilibrium $\underline{g_N}<1$ for all values of σ and τ. Thus, full deterrence cannot be implemented in this case. We can then show[47] that when $h<\frac{1}{2}$ the share of dishonest workers in N increases in σ and the net social loss is greater in N than it is in C (since the deterrence in N in this case is always such that $\underline{g_N}<2h$). From these properties it follows that the derivative in (A.4) is strictly negative for $\underline{g_N}\geq h$ and hence it is optimal to set σ such that $\underline{g_N}<h$. We can then show that, when τ is sufficiently close to 0, the optimal σ is strictly positive since the optimal level of deterrence is bounded away from 0.

If instead $h\geq\frac{1}{2}+\tau$, then full deterrence is implementable and results in a net social loss equal to $L=0$, which is optimal unless partial deterrence is an equilibrium and results in $L<0$. We show in this case that a necessary condition for the optimality of partial deterrence is that there exists an equilibrium with partial deterrence at which $L=0$ (Fact OA.7). To find this equilibrium, we analyze the expression of $\gamma\left(g_N\right)|_{L=0}$ to show[48] that, if and only if h is lower than a threshold $\hat{h}$, which decreases in r, there exists a $\underline{g_N}\in(0,1)$ such that $\gamma\left(g_N\right)|_{L=0}=0$, ensuring that we have an interior equilibrium and, at such equilibrium, $L=0$. We then show that at this equilibrium we have $\frac{dL}{d\sigma}>0$, and hence deterrence can be further lowered to reduce the social loss below 0.

Therefore, in the region $h\geq\max\left\{\hat{h},\frac{1}{2}+\tau\right\}$ full deterrence is implementable and the net social loss with partial deterrence is always greater. Hence full deterrence and an infinite criminal record, $\sigma=0$, are optimal. In contrast, when $\frac{1}{2}\leq h<\max\left\{\hat{h},\frac{1}{2}+\tau\right\}$ the optimum features partial deterrence, either because full deterrence cannot be implemented ($h<\frac{1}{2}+\tau$) or because it can be implemented but a lower net social loss $L<0$ can be attained in a partial deterrence equilibrium, with a strictly positive value of σ, ($\frac{1}{2}+\tau\leq h<\hat{h}$). We also show that in this region the optimum for some parameter values features over deterrence and under deterrence for others. □

A.3 Proof of Proposition 3

The condition for the implementability of full deterrence with nonmonetary sanctions is $\left(V_N-V_I\right)|_{g_N=1}\geq 1$, which can be easily compared with the analogous condition with criminal records to identify the parameter values under which the condition holds in the two cases. Similarly, if we compare the equilibrium level of deterrence under non-monetary sanctions with that under criminal records using (8) and (11), we obtain

[47] See Facts (OA.6) and (OA.5) in the Online Appendix for details of the proofs.

[48] See the Online Appendix Section (OA.2).

the conditions under which the first one is greater than the second one. The last claim is established by studying the partial derivatives of $V_N - V_I$ with respect to θ and λ. Details are in Online Appendix Section OA.3. $\square$

A.4 Proof of Proposition 4

Before addressing the optimal policy question we need to establish the equilibrium levels of deterrence as a function of the expected length of imprisonment $\frac{1}{\zeta}$. As shown in Proposition (OA.1) in the Online Appendix, the equilibrium level of deterrence is unique and weakly decreases in ζ, so that $\zeta \leq \pi - \tau$ implements a unique equilibrium with full deterrence, $\zeta > \pi - \tau$ implements a unique equilibrium with partial deterrence decreasing in ζ, with $\zeta = \infty$ implementing no deterrence. It is easy to see that, conditional on being implementable ($\pi \geq \tau$), full deterrence is preferable to no deterrence if the harm from crimes is greater than the average benefit from crimes, $h \geq \frac{1}{2}$, and no deterrence is preferable in the alternative scenario, $h < \frac{1}{2}$. We need then to show that an interior equilibrium with partial deterrence never improves on these corner policies. Note first that full deterrence yields a loss equal to $L = 0$, while no deterrence yields a loss equal to $L = h - \frac{1}{2}$. The claim is established by showing that: (i) when $h < \frac{1}{2}$, there is no simultaneous solution to $L_I = h - \frac{1}{2}$ and $\gamma_I = 0$ for $\underline{g_N} \in (0, 1)$, for all parameter values; (ii) when instead $h \geq \frac{1}{2}$, there is no simultaneous solution to $L_I = 0$ and $\gamma_I = 0$ for $\underline{g_N} \in (0, 1)$ for all parameter values. Details are in Online Appendix Section OA.4. $\square$

A.5 Steady-state distribution in the model with nonmonetary sanctions and criminal records

In the model of Section 4.3, the wages in states N, I and C are as follows:

$$\begin{aligned}
w_N &= \pi - \tfrac{rp_N}{1-r+rp_N}\,(1-g_N)\,h \\
w_I &= (1-\theta)\,\pi - (1-\lambda)\,h \\
w_C &= \pi - (1-g_C)\,h
\end{aligned}$$

where $\frac{rp_N}{1-r+rp_N}$ is the fraction of workers in N who are dishonest, while in states I and C all workers are dishonest. The variable p_N denotes the share of dishonest workers who are in state N (relative to the total population of dishonest workers) and is derived below. The values of the wages reflect the fact that there is deterrence in N and C, but not in I where, however, productivity and expected harm from crime are reduced due to incapacitation. The equations determining these levels of deterrence are:

$$\begin{aligned}
g_N &= V_N - V_I \\
g_C &= V_C - V_I
\end{aligned} \tag{A.5}$$

As before, the lifetime expected utility in each state reflects the flow payoff in that state and the probability of transitioning to another state:

$$\begin{array}{rcl} V_N & = & \mathrm{E}_{\nu\sim\exp[1-g_N+\tau]}\left[\int_0^\nu \left(w_N + \frac{1-g_N^2}{2}\right) dt\right] + \frac{1-g_N}{1-g_N+\tau} V_I \\ V_I & = & \mathrm{E}_{\nu\sim\exp[\zeta+\tau]}\left[\int_0^\nu \left(w_I + \frac{(1-\lambda)}{2}\right) dt\right] + \frac{\zeta}{\zeta+\tau} V_C \\ V_C & = & \mathrm{E}_{\nu\sim\exp[1-g_C+\sigma+\tau]}\left[\int_0^\nu \left(w_C + \frac{1-g_C^2}{2}\right) dt\right] \\ & & + \frac{\sigma}{1-g_C+\sigma+\tau} V_N + \frac{1-g_C}{1-g_C+\sigma+\tau} V_I \end{array} \tag{A.6}$$

The steady-state share of dishonest workers in each state is equal to the stationary distribution $p = \{p_N, p_I, p_C\}$ of the continuous-time Markov chain with states $\{N, I, C\}$ and transition rate matrix

$$Q = \begin{bmatrix} -(1-g_N) & 1-g_N & 0 \\ \tau & -(\tau+\zeta) & \zeta \\ \sigma+\tau & 1-g_C & -(1-g_C+\sigma+\tau) \end{bmatrix}$$

Solving $pQ = 0$ subject to $p_N + p_I + p_C = 1$, we obtain:

$$\begin{array}{rcl} p_N & = & \frac{\tau(1-g_C+\sigma+\tau)+\zeta(\sigma+\tau)}{(1-g_N+\tau)(1-g_C+\sigma+\tau)+\zeta(1-g_N+\sigma+\tau)} \\ p_I & = & \frac{(1-g_N)(1-g_C+\sigma+\tau)}{(1-g_N+\tau)(1-g_C+\sigma+\tau)+\zeta(1-g_N+\sigma+\tau)} \\ p_C & = & \frac{\zeta(1-g_N)}{(1-g_N+\tau)(1-g_C+\sigma+\tau)+\zeta(1-g_N+\sigma+\tau)} \end{array}$$

Finally, solving the model yields the following value functions in the 3 states:

$$\begin{array}{rcl} V_N & = & \frac{1}{\tau}\frac{(\tau(1-g_C+\sigma+\tau)+\zeta(\sigma+\tau))\left(w_N+\frac{1-g_N^2}{2}\right)+(1-g_N)\zeta\left(w_C+\frac{1-g_C^2}{2}\right)+(1-g_N)(1-g_C+\sigma+\tau)\left(w_I+\frac{\lambda}{2}\right)}{(1-g_N+\tau)(1-g_C+\sigma+\tau)+\zeta(1-g_N+\sigma+\tau)} \\ V_I & = & \frac{1}{\tau}\frac{\zeta\sigma\left(w_N+\frac{1-g_N^2}{2}\right)+\zeta(1-g_N+\tau)\left(w_C+\frac{1-g_C^2}{2}\right)+(1-g_N+\tau)(1-g_C+\sigma+\tau)\left(w_I+\frac{\lambda}{2}\right)}{(1-g_N+\tau)(1-g_C+\sigma+\tau)+\zeta(1-g_N+\sigma+\tau)} \\ V_C & = & \frac{1}{\tau}\frac{\sigma(\zeta+\tau)\left(w_N+\frac{1-g_N^2}{2}\right)+(\zeta+\tau)(1-g_N+\tau)\left(w_C+\frac{1-g_C^2}{2}\right)+((1-g_N+\tau)(1-g_C)+(1-g_N)\sigma)\left(w_I+\frac{\lambda}{2}\right)}{(1-g_N+\tau)(1-g_C+\sigma+\tau)+\zeta(1-g_N+\sigma+\tau)} \end{array}$$

Note that as $\zeta \to \infty$, we have $V_I = V_C$ and $g_C = 0$, and both V_C and V_N converge to the values we found in the basic model.

A.6 Numerical analysis of equilibria

Using the equations derived in Appendix Section A.5, we analyze the equilibria for three numerical specifications of the model with Wolfram Mathematica 13.0. We consider the following parameter values: $\tau = 0.2$, $r = 0.6$, $\pi = 1$, $\lambda = 0.3$, and $\theta = 0.01$. Across specifications we vary the value of the social loss h and find the optimal policy in three scenarios: when only the criminal record is used — that is, we look for the optimal σ — when only nonmonetary sanctions are available — that is, we look for the optimal ζ — and when a nonmonetary sanction is followed by a period during which the worker has a criminal record — that is, we look for the optimal combination of ζ and σ. For the first two scenarios we calculate the levels of σ and ζ, respectively, that minimize the social loss. In the mixed policy scenario, we approximate the optimal policy using contour plots to find values of σ and ζ that yield lower levels of the social loss than at the optima found in the other two scenarios. Therefore, in this third scenario, the social loss could in fact be lower for some other values of σ and ζ. Yet, given our purpose — that is, to show that a mix of the two policy instruments combined can yield a lower net social loss than any of the two instruments taken in isolation — this is sufficient.

For each numerical specification, we first report the optimal length of the criminal record and the resulting equilibrium level of deterrence, next we do the same for non-monetary sanctions, and finally we identify a pair of values of σ and ζ such that using a combination of the two instruments at this level achieves a lower social loss than in both previous scenarios.

Tables A.1 to A.3 report the value of the optimal policy and the associated equilibrium values for deterrence, wages, steady-state distribution, and net social loss. Additional plots are provided in the Online Appendix. Figures OA.10 to OA.12 show the contour plots for loss function under the mixed policy (left-hand side graph) and the levels of deterrence under the approximately optimal policy identified in the contour plot (right-hand side graph). Finally, the graphs in Figure OA.13 show the patterns of the social loss for various values of the expungement policy σ when only criminal records are used. The loss under only nonmonetary sanctions follows similar patterns.

A.7 Numerical analysis tables

	Loss	$\underline{g_N}$	$\underline{g_I}$	$\underline{g_C}$	$\frac{1}{\varsigma}$	$\frac{1}{\sigma}$
Mixed case	-0.0812031	0.200576	0	0.0111422	0.1	2.
Only nonmonetary sanction	-0.0705881	0.168225	0	n.a.	1.37801	n.a.
Only criminal record	-0.0809172	0.202801	n.a.	0	n.a.	2.38323

(a) Main outcomes

	p_N	p_I	p_C	w_N	w_I	w_C
Mixed case	0.439858	0.0809317	0.47921	0.856998	0.675	0.555014
Only nomonetary sanction	0.526717	0.473283	0	0.834798	0.675	n.a.
Only criminal record	0.437324	0	0.562676	0.857893	n.a.	0.55

(b) Additional outcomes

Table A.1: The optimal policy with $h = 0.45$

	Loss	$\underline{g_N}$	$\underline{g_I}$	$\underline{g_C}$	$\frac{1}{\varsigma}$	$\frac{1}{\sigma}$
Mixed case	-0.02902	0.258353	0	0.0997493	1.	1
Only nonmonetary sanction	-0.0215738	0.252091	0	n.a.	2.56191	n.a.
Only criminal record	-0.0233253	0.256456	n.a.	0	n.a.	2.74943

(a) Main outcomes

	p_N	p_I	p_C	w_N	w_I	w_C
Mixed case	0.413348	0.397425	0.189227	0.8524	0.626	0.53187
Only nomonetary sanction	0.441126	0.558874	0	0.845134	0.626	n.a.
Only criminal record	0.431217	0	0.568783	0.848138	n.a.	0.48

(b) Additional outcomes

Table A.2: The optimal policy with $h = 0.52$

	Loss	$\underline{g_N}$	$\underline{g_I}$	$\underline{g_C}$	$\frac{1}{\varsigma}$	$\frac{1}{\sigma}$
Mixed case	0.0235794	0.348049	0	0.162501	2.	1
Only nonmonetary sanction	0.0294893	0.358622	0	n.a.	4.9224	n.a.
Only criminal record	0.0382238	0.34446	n.a.	0	n.a.	3.6598

(a) Main outcomes

	p_N	p_I	p_C	w_N	w_I	w_C
Mixed case	0.3785	0.499037	0.122463	0.85834	0.57	0.497501
Only nomonetary sanction	0.385966	0.614034	0	0.858897	0.57	n.a.
Only criminal record	0.419249	0	0.580751	0.848146	n.a.	0.4

(b) Additional outcomes

Table A.3: The optimal policy with $h = 0.6$

Online Appendix

OA.1 Details of the proof of Proposition 1

The proof proceeds in several steps. We start with some preliminary results. Let $\overline{\gamma}(g_N)$ denote the numerator of $\gamma(g_N)$, as defined in (A.3). Note that the denominator of $\gamma(g_N)$ is strictly positive for all $g_N \in [0,1]$. To find the solutions of (A.1) it suffices then to consider the zeros of $\overline{\gamma}(g_N)$. Note that $\overline{\gamma}(g_N)$ is a polynomial of degree 3, hence $\overline{\gamma}(g_N) = 0$ can have at most 3 real solutions. We also see from A.3 that $\overline{\gamma}(g_N)$ is a continuously differentiable function of $g_N \in [0,1]$. We have:

Fact OA.1. *$\overline{\gamma}(g_N)$ is a strictly convex function, and hence, $\overline{\gamma}(g_N) = 0$ has at most 2 solutions for $g_N \in [0,1]$.*

Proof. We have:

$$\overline{\gamma}''(g_N) = 3(1-r)(1-g_N) + 2(1-r)(\sigma+\tau) + \sigma + \tau > 0$$

for $g_N \in [0,1]$. □

Moreover, it is easy to verify that:

Fact OA.2. *$\overline{\gamma}(0) > 0$ for all values of σ, τ and h, while $\overline{\gamma}(1) \geq 0$ iff $\sigma + \tau \leq h - \frac{1}{2}$.*

Note that $\overline{\gamma}(1) = 0$ when $\sigma + \tau = h - \frac{1}{2}$. Evaluating then the first derivative of $\overline{\gamma}(g_N)$ at $g_N = 1$ when $\sigma + \tau = h - \frac{1}{2}$, we have:

$$\overline{\gamma}'(1)|_{\overline{\gamma}(1)=0} = \left(h - \frac{1}{2}\right)\left(\frac{1}{2} - h(1-r)\right)$$

This expression allows to establish the following:

Fact OA.3. *If $h > \frac{1}{2(1-r)}$, when $\sigma + \tau = h - \frac{1}{2}$ (which implies $h > \frac{1}{2}$) we have $\overline{\gamma}'(1)|_{\overline{\gamma}(1)=0} < 0$. If instead $\frac{1}{2} < h < \frac{1}{2(1-r)}$, we have $\overline{\gamma}'(1)|_{\overline{\gamma}(1)=0} > 0$. If $h = \frac{1}{2(1-r)}$, $\overline{\gamma}'(1)|_{\overline{\gamma}(1)=0} = 0$.*

Finally, we establish some properties of the original function $\gamma(g_N)$:

Fact OA.4. *If $h > \frac{1}{2}$, we have $\frac{\partial \gamma(g_N)}{\partial(\sigma+\tau)} < 0$ and $\lim_{\sigma+\tau\to 0} \gamma(g_N) > 0$ for all $g_N \in [0,1]$.*

Proof. We find:

$$\frac{\partial \gamma(g_N)}{\partial(\sigma+\tau)} = -\frac{g_N\left(h - \frac{g_N}{2}\right)}{(1-g_N+\sigma+\tau)^2} - \frac{(1-r)(1-g_N)h}{((1-r)(1-g_N)+\sigma+\tau)^2}$$

which is negative for all values of $g_N \in [0,1]$ if $h \geq \frac{1}{2}$. We also have:

$$\lim_{\sigma+\tau\to 0} \gamma(g_N) = \frac{h - g_N\left(1 - \frac{g_N}{2}\right)}{1 - g_N}$$

which is positive for all values of $g_N \in [0,1]$ if $h > \frac{1}{2}$. □

We prove now the main claims in Proposition 1, proceeding in two steps. We first characterize the possible equilibrium configurations and then identify which equilibria are stable. Given that, by Facts OA.1 and OA.2, $\overline{\gamma}(g_N)$ is strictly convex and is positive at $g_N = 0$, there can only be 3 situations: (1) $\overline{\gamma}(g_N)$ is always positive (in which case there is no equilibrium with partial deterrence but there is an equilibrium with full deterrence); (2) $\overline{\gamma}(g_N)$ crosses 0 once and is thereafter negative (in which case the only equilibrium has partial deterrence); or (3) $\overline{\gamma}(g_N)$ crosses 0 twice (in which case we have two equilibria with partial deterrence as well as one with full deterrence). More precisely, we have:

Proof of claim 1. If $h < \frac{1}{2(1-r)}$, combining Facts OA.2 and OA.3 we know that, when $\sigma + \tau = h - \frac{1}{2}$, $\overline{\gamma}(g_N)$ crosses 0 from below at $g_N = 1$ (that is, $\overline{\gamma}'(1)|_{\overline{\gamma}(1)=0} > 0$). Since $\overline{\gamma}(0) > 0$ there must then be 2 solutions of $\overline{\gamma}(g_N) = 0$: one at $g_N = 1$ and one at $g_N \in (0,1)$. We also know from Fact OA.4 that $\gamma(g_N)$ increases when $\sigma + \tau$ decreases and that $\lim_{\sigma+\tau\to 0} \gamma(g_N) > 0$ for all $g_N \in [0,1]$ when $h > \frac{1}{2}$. Therefore, there must be a threshold $\xi \in (0, h - \frac{1}{2})$ such that if $\sigma + \tau < \xi$ then $\gamma(g_N) > 0$ for all $g_N \in [0,1]$, while if $\xi < \sigma + \tau < h - \frac{1}{2}$, $\gamma(g_N)$ must cross 0 twice. When $\xi = \sigma + \tau < h - \frac{1}{2}$, $\gamma(g_N)$ is tangent to the horizontal axis. Leveraging on these observations, we have:

(a) If $\sigma + \tau < \xi < h - \frac{1}{2}$, then $\gamma(g_N) > 0$ for all $g_N \in [0,1]$. Thus there is no solution to $\gamma(g_N) = 0$ for $g_N \in [0,1]$; hence, the only equilibrium is full deterrence, $\overline{g_N} = 1$. See Figure OA.1.

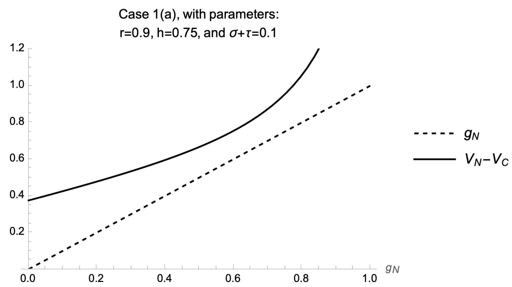


Figure OA.1: Case $1(a)$ with $h < \frac{1}{2(1-r)}$ and $\sigma + \tau < \xi$

(b) If $\xi < \sigma + \tau < h - \frac{1}{2}$, the equation $\gamma(g_N) = 0$ has two solutions for $g_N \in [0,1]$, labeled $\underline{g_N}$ and $\hat{g}_N$, with $\hat{g}_N \geq \underline{g_N}$. Moreover, $\overline{\gamma}(1) > 0$. Therefore, there is an (unstable, as will be shown below) equilibrium, $\hat{g}_N \in \left[\underline{g_N}, \overline{g_N}\right]$, which solves $\gamma(g_N) = 0$, and two (stable) equilibria: the first one solves $\gamma(g_N) = 0$, and features partial deterrence, $\underline{g_N} \in (0,1)$, the second one features full deterrence, $\overline{g_N} = 1$. Moreover, from the continuity of $\overline{\gamma}(g_N)$ and the facts that $\overline{\gamma}(0) > 0$ and $\overline{\gamma}(1) > 0$, it follows that we must have $\gamma'\left(\underline{g_N}\right) < 0$ and $\gamma'(\hat{g}_N) > 0$ and hence also $\frac{d\underline{g_N}}{d(\sigma+\tau)} = -\left.\frac{\frac{\partial\gamma}{\partial(\sigma+\tau)}}{\frac{\partial\gamma}{\partial g_N}}\right|_{g_N=\underline{g_N}} < 0$ and $\frac{d\hat{g}_N}{d(\sigma+\tau)} =$

$-\left.\frac{\frac{\partial\gamma}{\partial(\sigma+\tau)}}{\frac{\partial\gamma}{\partial g_N}}\right|_{g_N=\hat{g}_N} > 0$. Thus, as the sum of $\sigma+\tau$ decreases towards ξ, the two roots $\underline{g_N}$ and $\hat{g}_N$ approach each other. See Figure OA.2. Finally, note that when $\xi = \sigma+\tau < h-\frac{1}{2}$, $\gamma\left(\underline{g_N}\right)$ is tangent to 0 and hence $\underline{g_N} = \hat{g_N}$; when instead $\sigma+\tau = h-\frac{1}{2}$, we have $\gamma(1) = 1$ and hence $\overline{g_N} = 1 = \hat{g_N}$.

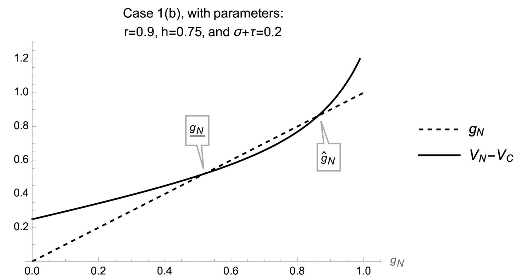


Figure OA.2: Case 1(b) with $h < \frac{1}{2(1-r)}$ and $\xi \le \sigma+\tau \le h-\frac{1}{2}$

(c) If $\sigma+\tau > h-\frac{1}{2}$, we have $\overline{\gamma}(1) < 0$. Again by continuity as well as the fact that $\overline{\gamma}(0) > 0$ we know that $\overline{\gamma}(g_N)$ can only have one solution, $\underline{g_N} \in (0,1)$. Hence, the only equilibrium is $\underline{g_N} \in (0,1)$. Also in this case, given the strict convexity of $\overline{\gamma}(g_N)$, we must have $\gamma'\left(\underline{g_N}\right) < 0$. See Figure OA.3.

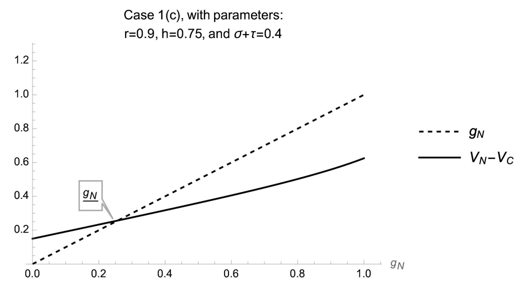


Figure OA.3: Case 1(c) with $h < \frac{1}{2(1-r)}$ and $\sigma+\tau > h-\frac{1}{2}$

Proof of claim 2. If $h \ge \frac{1}{2(1-r)}$:

(a) If $\sigma+\tau \le h-\frac{1}{2}$, we have $\overline{\gamma}(0) > 0$ and $\overline{\gamma}(1) \ge 0$ (Fact OA.2). Therefore, either $\overline{\gamma}(g_N)$ remains above 0 or it crosses 0 twice. Let us start by considering the case $\sigma+\tau = h-\frac{1}{2}$, at which point $\overline{\gamma}(1) = 0$. By the same argument as in the proof of claim 1(b) above, it follows that $\overline{\gamma}(g_N)$ crosses 0 from above at $g_N = 1$ when $h > \frac{1}{2(1-r)}$ and also that $\overline{\gamma}(g_N)$ is tangent to 0 when $h = \frac{1}{2(1-r)}$. Since $\overline{\gamma}(g_N)$ is strictly convex, this implies that $\overline{\gamma}(g_N) > 0$ (and hence $\gamma(g_N) > 0$), for all $g_N \in [0,1)$; therefore the unique equilibrium is full deterrence, $\overline{g_N} = 1$. Furthermore, note that $h \ge \frac{1}{2(1-r)}$ implies $h > \frac{1}{2}$ and hence by Fact OA.4 we know that $\gamma(g_N)$ increases when $\sigma+\tau$ decreases. Hence $\gamma(g_N)$ remains entirely above 0 if $\sigma+\tau < h-\frac{1}{2}$. In this case, there is no solution to $\gamma(g_N) = 0$ for $g_N \in [0,1]$ and hence, as stated in claim 2(a)

of Proposition 1, the only equilibrium is again full deterrence, $\overline{g_N} = 1$. See Figure OA.4.

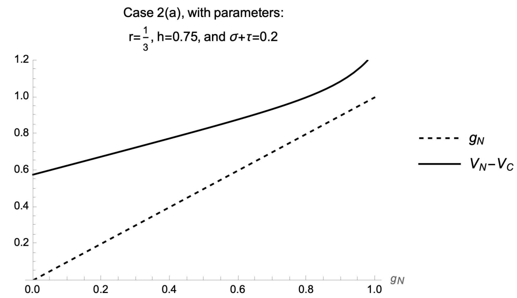


Figure OA.4: Case $2(a)$ with $h \geq \frac{1}{2(1-r)}$ and $\sigma + \tau \leq h - \frac{1}{2}$

(b) If $\sigma + \tau > h - \frac{1}{2}$, we have $\overline{\gamma}(0) > 0$ and $\overline{\gamma}(1) < 0$ (by Fact OA.2). By the Intermediate Value Theorem and the convexity of $\overline{\gamma}(g_N)$ it follows that there is a unique internal solution to $\overline{\gamma}(g_N) = 0$ and such solution obtains at a level of partial deterrence $\underline{g_N} \in (0, 1)$. By the same argument as above, in this case too we must have $\gamma'\left(\underline{g_N}\right) < 0$. See Figure OA.5.

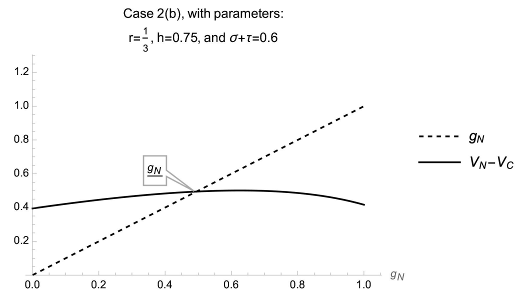


Figure OA.5: Case $2(b)$ with $h \geq \frac{1}{2(1-r)}$ and $\sigma + \tau > h - \frac{1}{2}$

We consider now the local stability of the three possible equilibria identified above, using a standard cobweb plot with coordinates g_N and $v(g_N) \equiv V_N - V_C$ as in Figure OA.6.

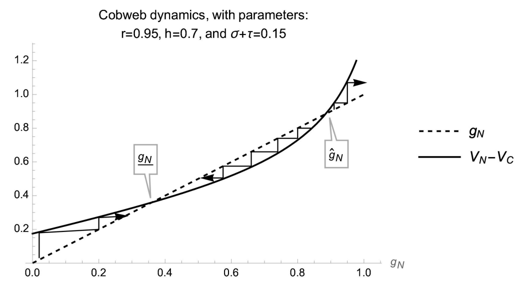


Figure OA.6: Cobweb Dynamics

The full-deterrence equilibrium $\overline{g_N} = 1$ is stable when $v(1) > 1$, that is, when $\sigma + \tau < h - \frac{1}{2}$, as it is easy to see from Figure OA.6. If $v(1) = 1$, that is, if $\sigma + \tau = h - \frac{1}{2}$, we have $\overline{g_N} = \hat{g}_N$ and the equilibrium is unstable, as we will see below. Next, an interior equilibrium is stable iff:

$$-1 < \left.\frac{\partial v}{\partial g_N}\right|_{\gamma(g_N)=0} < 1 \tag{OA.1}$$

Note that, using (8), we obtain:

$$\begin{aligned}\left.\frac{\partial v}{\partial g_N}\right|_{\gamma(g_N)=0} &= \frac{\left(\frac{\partial w_N}{\partial g_N} - g_N\right)(1-g_N+\sigma+\tau)+w_N-w_C-\frac{g_N^2}{2}}{(1-g_N+\sigma+\tau)^2} \\ &= \frac{\frac{\partial w_N}{\partial g_N}+V_N-V_C-g_N}{1-g_N+\sigma+\tau} \\ &= \frac{\frac{\partial w_N}{\partial g_N}}{1-g_N+\sigma+\tau}\end{aligned}$$

where in the second line we used the equilibrium property $\gamma(g_N) = V_N - V_C - g_N = 0$. We also have:

$$\frac{\partial w_N}{\partial g_N} = \left(\frac{\sigma+\tau}{(1-r)(1-g_N)+\sigma+\tau}\right)^2 rh > 0 \tag{OA.2}$$

which shows that the left inequality in (OA.1) is always satisfied. The right inequality in (OA.1) must be satisfied at $\underline{g_N}$ because the property $\gamma'\left(\underline{g_N}\right) < 0$ — which we have shown above in the proofs of claims 1(b), 1(c) and 2(b) — implies $\frac{\partial v}{\partial g_N} < 1$. In turn, this last inequality can be equivalently stated as

$$w'\left(\underline{g_N}\right) \quad < 1 - g_N + \sigma + \tau \tag{OA.3}$$

Conversely, note that $\gamma'(\hat{g}_N) > 0$ — shown above in the proofs of claims 1(b), 1(c) and 2(b) — is equivalent to $\frac{\partial v}{\partial g_N} > 1$, which in turn implies that $\hat{g}_N$ is unstable.

To conclude the proof, we establish how $\underline{g_N}$ and $\hat{g}_N$ change with respect to (infinitesimal) variations of $\sigma + \tau$ and r. Note that by the Implicit Function Theorem $\underline{g_N}$ is a continuously differentiable function of $\sigma+\tau$ for $\sigma+\tau \in (\xi, \infty)$ if $h < \frac{1}{2(1-r)}$ or $\sigma + \tau \in \left(h - \frac{1}{2}, \infty\right)$ if $h \geq \frac{1}{2(1-r)}$. Moreover, using (8), we have:

$$\begin{aligned}\left.\frac{\partial \gamma}{\partial(\sigma+\tau)}\right|_{\gamma=0} &= \frac{\partial}{\partial(\sigma+\tau)}\left(V_N - V_C - \underline{g_N}\right) \\ &= \frac{\partial}{\partial(\sigma+\tau)}\left(\frac{w_N - w_C - \frac{\underline{g_N}^2}{2}}{1-\underline{g_N}+\sigma+\tau}\right) \\ &= \frac{\frac{\partial w_N}{\partial(\sigma+\tau)}\left(1-\underline{g_N}+\sigma+\tau\right)-w_N+w_C+\frac{\underline{g_N}^2}{2}}{\left(1-\underline{g_N}+\sigma+\tau\right)^2} \\ &= \frac{\frac{\partial w_N}{\partial(\sigma+\tau)}-\underline{g_N}}{1-\underline{g_N}+\sigma+\tau}\end{aligned} \tag{OA.4}$$

where in the last line we used $V_N - V_C = \underline{g_N}$. Similarly:

$$\begin{aligned}\left.\frac{\partial\gamma}{\partial g_N}\right|_{\gamma=0} &= \frac{\partial}{\partial g_N}\left(\frac{w_N - w_C - \frac{\underline{g_N}^2}{2}}{1-\underline{g_N}+\sigma+\tau}\right) - 1 \\ &= \frac{\left(\frac{\partial w_N}{\partial g_N} - \underline{g_N}\right)\left(1-\underline{g_N}+\sigma+\tau\right)+w_N-w_C-\frac{\underline{g_N}^2}{2}}{\left(1-g_N+\sigma+\tau\right)^2} - 1 \\ &= \frac{\frac{\partial w_N}{\partial g_N}}{1-\underline{g_N}+\sigma+\tau} - 1\end{aligned} \tag{OA.5}$$

so that:

$$\frac{\partial \underline{g_N}}{\partial\left(\sigma+\tau\right)} = -\frac{\frac{\partial\gamma}{\partial(\sigma+\tau)}}{\frac{\partial\gamma}{\partial g_N}} = \frac{\frac{\partial w_N}{\partial(\sigma+\tau)} - \underline{g_N}}{1-\underline{g_N}+\sigma+\tau-\frac{\partial w_N}{\partial g_N}} < 0 \tag{OA.6}$$

where the numerator is negative because:

$$\frac{\partial w_N}{\partial\left(\sigma+\tau\right)} = -\left(\frac{1-\underline{g_N}}{\left(1-r\right)\left(1-\underline{g_N}\right)+\sigma+\tau}\right)^2 r\left(1-r\right)h < 0 \tag{OA.7}$$

and the denominator is positive because of the stability condition (OA.3). Hence, (OA.6) establishes the claim in the proposition.[49] Note that the same argument implies that the unstable equilibrium $\hat{g}_N$ is instead increasing in $(\sigma+\tau)$. The next claim, $\lim_{\sigma+\tau\to\infty}\underline{g_N} = 0$, follows from the fact that $\underline{g_N} = V_N - V_C$ and $\lim_{\sigma+\tau\to\infty}\left(V_N - V_C\right) = 0$.

The final claim in the proposition concerns the effects of changes in r. Note that:

$$\begin{aligned}\left.\frac{\partial\gamma}{\partial r}\right|_{\gamma=0} &= \frac{\partial}{\partial r}\left(V_N - V_C - \underline{g_N}\right) \\ &= \frac{\partial}{\partial r}\left(\frac{w_N - w_C - \frac{\underline{g_N}^2}{2}}{1-g_N+\sigma+\tau}\right) \\ &= \frac{\frac{\partial w_N}{\partial r}}{1-\underline{g_N}+\sigma+\tau} \quad < \quad 0\end{aligned}$$

because:

$$\frac{\partial w_N}{\partial r} = -\frac{1-g_N+\sigma+\tau}{\left(\left(1-r\right)\left(1-\underline{g_N}\right)+\sigma+\tau\right)^2}\left(\sigma+\tau\right)\left(1-g_N\right)h < 0$$

Therefore, we have:

$$\frac{d\underline{g_N}}{dr} = -\frac{\frac{\partial\gamma}{\partial r}}{\frac{\partial\gamma}{\partial g_N}} < 0 \tag{OA.8}$$

where we used the property, established above and used in OA.7, that the denominator is strictly negative. □

OA.2 Details of the proof of Proposition 2

We start by providing two preliminary results.

[49] Note that in Fact OA.4 we proved a different but compatible result for all g_N (thus, not only for $g_N = \underline{g_N}$) when $h \geq \frac{1}{2}$.

Fact OA.5. *If $h < \frac{1}{2}$, we have $\underline{g_N} < 2h$ for all $\sigma \geq 0$ and $\tau \geq 0$.*

Proof. The equilibrium equation (A.3), when we take the limit $\sigma + \tau \to 0$, reduces to $\frac{\underline{g_N}^2}{2} - \underline{g_N} + h = 0$ and the only solution of this equation for $\underline{g_N} \in [0, 1]$ is $1 - \sqrt{1 - 2h}$. Furthermore, when $h < \frac{1}{2}$ this solution is a real number and satisfies $1 - \sqrt{1 - 2h} < 2h < 1$. Since $\underline{g_N}$ is decreasing in σ and τ — as we showed earlier in expression (OA.6) — it follows that the maximum level of deterrence that can be attained when $h < 1/2$ is:

$$\lim_{\sigma+\tau \to 0} \underline{g_N} = 1 - \sqrt{1 - 2h}$$

and so $\underline{g_N} < 2h$ for all σ and τ. □

Fact OA.6. *If $h < \frac{1}{2}$, the equilibrium share of dishonest workers in N, $\frac{\sigma+\tau}{1-\underline{g_N}+\sigma+\tau}$, is increasing in σ for all $\sigma \geq 0$ and $\tau > 0$.*

Proof. The derivative with respect to σ of the equilibrium share of dishonest workers in N is:[50]

$$\frac{(\sigma + \tau) \frac{\partial \underline{g_N}}{\partial \sigma} + 1 - \underline{g_N}}{\left(1 - \underline{g_N} + \sigma + \tau\right)^2}$$

for all $\sigma \geq 0$ and $\tau > 0$. Since the denominator is positive, the above derivative is positive iff:

$$-\frac{\partial \underline{g_N}}{\partial \sigma} < \frac{1 - \underline{g_N}}{\sigma + \tau}$$

Using $\frac{\partial \underline{g_N}}{\partial \sigma} = -\frac{\frac{\partial \gamma}{\partial \sigma}}{\frac{\partial \gamma}{\partial \underline{g_N}}}$ and rearranging, we obtain:

$$\frac{\partial \gamma}{\partial \sigma} (\sigma + \tau) - \frac{\partial \gamma}{\partial \underline{g_N}} \left(1 - \underline{g_N}\right) > 0$$

Substituting the expressions for $\frac{\partial \gamma}{\partial \sigma}$ and $\frac{\partial \gamma}{\partial \underline{g_N}}$ derived in (OA.4) and (OA.5), the above inequality can be rewritten as:

$$1 - \underline{g_N} - \frac{1}{1 - \underline{g_N} + \sigma + \tau} \left(\frac{\partial w_N}{\partial \underline{g_N}} \left(1 - \underline{g_N}\right) - \frac{\partial w_N}{\partial \sigma} (\sigma + \tau) + \underline{g_N} (\sigma + \tau) \right) > 0$$

Note that, using (OA.2) and (OA.7) for $\frac{\partial w_N}{\partial \underline{g_N}}$ and $\frac{\partial w_N}{\partial \sigma}$ we have:

$$\left(1 - \underline{g_N}\right) \frac{\partial w_N}{\partial \underline{g_N}} - (\sigma + \tau) \frac{\partial w_N}{\partial \sigma} = \frac{\left(1 - \underline{g_N}\right) (\sigma + \tau) hr}{\left(1 - \underline{g_N}\right) (1 - r) + \sigma + \tau}$$

Substituting this in the above inequality we get:

$$1 - \underline{g_N} - \frac{1}{1 - \underline{g_N} + \sigma + \tau} \left(\frac{(\sigma + \tau) \left(1 - \underline{g_N}\right) hr}{(1 - r) \left(1 - \underline{g_N}\right) + \sigma + \tau} \right) - \frac{\sigma + \tau}{1 - \underline{g_N} + \sigma + \tau} \underline{g_N} > 0$$

[50] The expression for $\frac{\partial \underline{g_N}}{\partial \sigma}$ is analogous to (OA.6) derived in the proof of Proposition 1 in the Online Appendix.

Observe that:

$$-\frac{(\sigma+\tau)\left(1-\underline{g_N}\right)hr}{(1-r)\left(1-\underline{g_N}\right)+\sigma+\tau} = w_N - w_C - h = \underline{g_N}\left(1-\underline{g_N}+\sigma+\tau\right) + \frac{\underline{g_N}^2}{2} - h$$

where the first equality follows from equations (1) and (2) and the second equality follows from the equilibrium condition (8). Using this to rewrite the inequality above yields:

$$1-\underline{g_N}+\frac{1}{1-\underline{g_N}+\sigma+\tau}\left(\underline{g_N}\left(1-\underline{g_N}+\sigma+\tau\right)+\frac{\underline{g_N}^2}{2}-h\right)-\frac{\sigma+\tau}{1-\underline{g_N}+\sigma+\tau}\underline{g_N}>0$$

which in turn is equivalent to the following:

$$\sigma+\tau > \frac{h-\frac{\underline{g_N}^2}{2}}{1-\underline{g_N}}-1 \tag{OA.9}$$

To complete the proof note that the RHS of (OA.9) is negative because, if $h<\frac{1}{2}$, we have:

$$\underline{g_N}\left(1-\frac{\underline{g_N}}{2}\right) \leq \frac{1}{2} < 1-h$$

This completes the proof. □

We proceed by proving each claim in the proposition in turn.

Proof of claim 1. When $h<\frac{1}{2}$, we have $\sigma+\tau>0>h-\frac{1}{2}$, and hence, we are either in case $1\,(c)$ or in case $2\,(b)$ of Proposition 1, yielding a unique interior equilibrium $\underline{g_N}<1$ for all values of σ and τ. Thus, full deterrence cannot be implemented in this case. In (OA.6) above we derived an expression for $\frac{d\underline{g_N}}{d(\sigma+\tau)}$ which is well defined for all σ and τ. It then follows that $\frac{dL}{d\sigma}$ in (A.4) is well defined and the optimal policy is obtained when $\frac{dL}{d\sigma}=0$, if $\sigma>0$, or $\frac{dL}{d\sigma}\geq 0$, if $\sigma=0$. To prove under-deterrence note that the second addendum in (A.4) is negative since $h-\frac{1}{2}\underline{g_N}>0$ (by Fact OA.5) and $\frac{(\sigma+\tau)\frac{\partial \underline{g_N}}{\partial\sigma}+1-\underline{g_N}}{\left(1-\underline{g_N}+\sigma+\tau\right)^2}>0$ (by Fact OA.6). Thus, to have $\frac{dL}{d\sigma}\geq 0$, the first addendum in (A.4) must be positive, and hence, we must have that, at the optimum, $h-\underline{g_N}>0$. To show that if τ is close enough to 0 the optimal expungement rate is strictly positive, note that $\lim_{\tau\to 0}\frac{dL}{d\sigma}\Big|_{\sigma=0}=-\frac{\underline{g_N}}{1-\underline{g_N}}\left(h-\frac{1}{2}\underline{g_N}\right)<0$, hence $\sigma=0$ cannot be an optimum in this case, as it is profitable to increase σ above 0 to attain a lower social loss. □

Proof of claim 2. In this case, since $h\geq\frac{1}{2}+\tau$, full deterrence can be implemented by setting $\sigma=0$ (Proposition 1) and yields a social loss equal to $L=0$. We begin by establishing some additional instrumental results below.

Fact OA.7. *A necessary condition for partial deterrence to be optimal when $h\geq\frac{1}{2}+\tau$ is the existence of a partial deterrence equilibrium where the social loss is*

$L = 0$ *and the expungement policy is a strictly positive value satisfying:*

$$\sigma = \frac{h - \frac{1}{2}}{\frac{\underline{g_N}}{2} - \left(h - \frac{1}{2}\right)} - \tau \tag{OA.10}$$

which is positive iff $\underline{g_N} > 2h - 1$. *If such an equilibrium exists, it is unique.*

Proof. Note that[51] $\lim_{\sigma\to\infty} L = h - \frac{1}{2} > 0$, where the inequality is strict since we assumed $h \geq \frac{1}{2} + \tau$ and $\tau > 0$. Note also that the net social loss L is a continuous function of σ as long as we focus on the stable partial-deterrence equilibrium $\underline{g_N}$ (since we show below that such equilibrium is unique, when it exists). Since for partial deterrence to be optimal the equilibrium value of the loss must be $L < 0$, for this to happen L must cross 0 as σ decreases towards $\xi - \tau$, if $h < \frac{1}{2(1-r)}$ (claim 1 in Proposition 1), or towards $h - \frac{1}{2} - \tau$, if $h \geq \frac{1}{2(1-r)}$ (claim 2 in Proposition 1). As noticed above, for the social loss to be zero at an equilibrium with partial deterrence, there must be a value of the expungement policy $\sigma \geq 0$ satisfying (OA.10). The level of deterrence $\underline{g_N}$ appearing in the RHS of (OA.10) must then satisfy the equilibrium equation $\gamma\left(\underline{g_N}\right) = 0$ for this value to be attained in equilibrium and is then also a function of σ. Note that if $\sigma+\tau \leq h-\frac{1}{2}$, equation (OA.10) is satisfied only if the denominator is greater or equal to 1, which implies $\underline{g_N} \geq 2h + 1 > 1$, but this cannot hold true. Therefore, to have $L = 0$ at an equilibrium with partial deterrence, we must have $\sigma + \tau > h - \frac{1}{2}$. Hence, we must be either in case 1 (c) or in case 2 (b) of Proposition 1 so that the equilibrium, if it exists, is unique and features partial deterrence. Recall that full deterrence can still be implemented when $h \geq \frac{1}{2} + \tau$ by setting $\sigma = 0$. □

Fact OA.8. *Let:*

$$\begin{aligned} \chi(g_N) &\equiv \textit{Numerator}\left[\gamma(g_N)|_{L=0}\right] \\ &= g_N^3(1-r) - g_N^2(1-r)(2h+1) \\ &\quad + g_N(2h-4hr+r) - (2h-r)(2h-1) \end{aligned}$$

and

$$\begin{aligned} g_N^{max} &\equiv \arg\max \chi(g_N) \\ &= \tfrac{2h+1}{3} - \tfrac{1}{3}\sqrt{\tfrac{4(1-h)^2(1-r)+3(2h-1)}{1-r}} \end{aligned}$$

There exists a level of deterrence $g_N \in (0,1)$ *satisfying the equilibrium condition* $\gamma(g_N) = 0$ *and resulting in a loss* $L = 0$ *iff* $\chi(g_N^{max}) \geq 0$ *and* $g_N^{max} > 0$.

Proof. Let us evaluate the equilibrium expression $\gamma(g_N)$ when $L = 0$, using (OA.10) to replace $\sigma + \tau$ with $\frac{h-\frac{1}{2}}{\frac{g_N}{2}-\left(h-\frac{1}{2}\right)}$. It is immediate to verify that the numerator of the resulting function, which we denote $\chi(g_N)$, is as in the statement of Fact OA.8. We are interested in the existence of admissible solutions to

[51] This can be readily seen from equation (9), recalling that $\lim_{\sigma\to\infty} \underline{g_N} = 0$, as established at the end of the proof of Proposition 1.

$\chi(g_N) = 0$, which are values of $g_N \in [0,1]$ that satisfy $\gamma(g_N) = 0$ and yield $L = 0$. Note that we have:

$$\chi''(g_N) = 2(3g_N - 1 - 2h)(1-r) < 0 \quad \Longleftrightarrow \quad g_N < \tfrac{2h+1}{3}$$

and $\chi'(g_N) = 0$ yields two solutions. One solution is greater than 1 and hence not admissible.[52] The other solution is $g_N^{\max}$. Note that $g_N^{\max}$ is such that $g_N^{\max} < \frac{2h+1}{3} < 1$ and hence $\chi''(g_N^{\max}) < 0$, so that $g_N^{\max}$ identifies the unique maximum of $\chi(g_N)$.[53] Further, we have that $\chi(1) < 0$ (since $h > \frac{1}{2}$) and $\chi(0) < 0$, because $h > \frac{1}{2}$ implies $2h > r$. Therefore, there can only be two situations. The first one where $g_N^{\max} \leq 0$, in which case it follows that $\chi(g_N)$ is decreasing in g_N for all $g_N \in (0,1)$. Thus, recalling that $\chi(0) < 0$, it must be the case that $\chi(g_N) < 0$ for all $g_N \in (0,1)$ and, therefore, we have no solution to $\chi(g_N) = 0$, and hence to $\gamma(g_N) = 0$. Alternatively, we have $g_N^{\max} > 0$ in which case a solution to $\chi(g_N) = 0$ exists iff $\chi(g_N^{\max}) \geq 0$ by the continuity of $\chi(.)$ and hence by the Intermediate Value Theorem. $\square$

In the previous and the two following facts we focus our attention on the existence of a value of $g_N \in (0,1)$ that satisfies $\gamma(g_N) = 0$ and (OA.10), postponing till the end of the proof the verification that the associated value of σ, specified in (OA.10), is positive. Only when this last property has been verified we can say that such a g_N constitutes an equilibrium with partial deterrence where $L = 0$. Note that $\chi(g_N)$ is a continuous function not only of g_N but also of h and r. We show next the following:

Fact OA.9. *For every r, there exists a level of h, denoted $\hat{h} \in \left(\frac{1}{2}, 1\right)$, such that $\chi\left(g_N^{max}; \hat{h}, r\right) = 0$ and, if $h < \hat{h}$ we have $\chi(g_N^{max}) > 0$, while if $h > \hat{h}$ we have $\chi(g_N^{max}) < 0$.*

Note that, since neither $\chi(g_N)$ nor $g_N^{\max}$ are a function of σ or τ, then $\hat{h}$ will be only a function of r. If $\chi\left(g_N^{\max}; \hat{h}, r\right) = 0$, $g_N^{\max}$ lies in $(0,1)$ (see Fact OA.8) and satisfies the equilibrium condition $\gamma(g_N) = 0$ and $L = 0$. To prove Fact OA.9, we establish first a preliminary result.

Fact OA.10. *We have $\frac{d\chi(g_N^{max}; h, r)}{dh} < 0$ for all $h \in \left(\frac{1}{2}, 1\right)$ and $r \in (0,1)$ such that $g_N^{max} > 0$.*

Proof. By the Envelope Theorem, we have:

$$\begin{aligned}
\frac{d\chi(g_N^{\max})}{dh} &= \left.\frac{\partial\chi(g_N)}{\partial h}\right|_{g_N = g_N^{\max}} \\
&= -2(g_N^{\max})(g_N^{\max}(1-r) + 2r - 1) - 2(4h - 1 - r)
\end{aligned} \tag{OA.11}$$

[52] This solution is $g_N = \frac{2h+1}{3} + \frac{1}{3}\sqrt{\frac{4(1-h)^2(1-r)+3(2h-1)}{1-r}} > 1$. Recall that in claim 2 we are assuming $h \geq \frac{1}{2}$.

[53] Note that g_N^{max} is a candidate equilibrium level of deterrence with $L = 0$ only when $\chi(g_N^{\max}) = 0$.

A sufficient condition for the RHS of (OA.11) to be negative is that the following inequality holds:

$$-2x\left(x\left(1-r\right)+2r-1\right)-2\left(4h-1-r\right)<0 \tag{OA.12}$$

for all values of $x \in (0,1)$, $h \in \left(\frac{1}{2},1\right)$ and $r \in (0,1)$.[54] Note that (OA.12) is strictly concave in x. The FOC for the global maximum of the LHS of (OA.12) yields:

$$x^{*}=\frac{1-2r}{2\left(1-r\right)} \tag{OA.13}$$

Note that $\frac{1-2r}{2(1-r)}<1$. Consider first the case where $r<\frac{1}{2}$: hence $x^{*}>0$. Replacing (OA.13) into (OA.12) we get:

$$-8h+2+\frac{1}{2\left(1-r\right)}<0$$

for all $h \in \left(\frac{1}{2},1\right)$ and $r \in \left(0,\frac{1}{2}\right)$. Hence in this case (OA.11) always holds. Consider next the case where $r \geq \frac{1}{2}$ so that $x^{*} \leq 0$. Note that now, since the LHS of (OA.12) is strictly concave in x, reaches its maximum at a negative value of x, and is negative at $x=0$, it must also be negative for all positive values of x. This is true for all $h \in \left(\frac{1}{2},1\right)$ and $r \in \left[\frac{1}{2},1\right)$. Therefore, the RHS of (OA.11) is negative for all $h \in \left(\frac{1}{2},1\right)$ and $r \in (0,1)$. □

We can now turn to the proof of Fact OA.9:

Proof of Fact OA.9. We proceed case by case, as follows.

- If $r \in \left(0,\frac{2}{3}\right]$, it is easy to verify that $g_N^{\max}>0$ for all $h \in \left(\frac{1}{2},1\right)$. Thus, by Fact OA.10, we have $\frac{d\chi(g_N^{\max})}{dh}<0$ for all $h \in \left(\frac{1}{2},1\right)$ and $r \in (0,\frac{2}{3}]$. In this case the result in Fact OA.9 follows from the following observations:
  - $\lim_{h\to\frac{1}{2}} g_N^{\max}=\frac{1}{3}$ and $\lim_{h\to\frac{1}{2}}\chi\left(g_N^{\max}\right)>0$;
  - $\lim_{h\to 1} g_N^{\max}=1-\frac{1}{\sqrt{3(1-r)}}$ and $\lim_{h\to 1}\chi\left(g_N^{\max}\right)<0$,
- Consider next the case where $r \in \left(\frac{2}{3},1\right)$. Differentiating $g_N^{\max}$ with respect to h we obtain:

$$\frac{\partial g_N^{\max}}{\partial h}=\frac{2}{3}+\frac{2-8h-8\left(1-h\right)r}{6\sqrt{\left(1-r\right)\left(4\left(1-h\right)^{2}\left(1-r\right)+3\left(2h-1\right)\right)}}<0 \tag{OA.14}$$

To see why the above inequality holds, note that the expression on the RHS

[54] We disregard here the fact that $g_N^{\max}$ depends on h, r and let it move independently in its range.

of (OA.14) is increasing in h (when $r > \frac{1}{4}$):

$$\frac{\partial^2 g_N^{\max}}{\partial h^2} = \frac{(4r-1)(1-r)}{\sqrt{(1-r)^3\left(4(1-h)^2(1-r)+3(2h-1)\right)^3}} > 0$$

Therefore, to establish the inequality in (OA.14) it suffices to show that it is negative for $h \to 1$, which is always true for $r \in \left(\frac{2}{3}, 1\right)$:

$$\lim_{h\to 1} \frac{\partial g_N^{\max}}{\partial h} = \frac{2}{3} - \frac{1}{\sqrt{3(1-r)}} < 0$$

Furthermore, as established above, $\lim_{h\to\frac{1}{2}} g_N^{\max} = \frac{1}{3}$ and $\lim_{h\to 1} g_N^{\max} = 1 - \frac{1}{\sqrt{3(1-r)}}$, which is strictly negative when $r > \frac{2}{3}$. Hence if $r \in \left(\frac{2}{3}, 1\right)$, $g_N^{\max}$ is monotonically decreasing in h, starting from a positive value when $h = \frac{1}{2}$ and ending at a negative one when $h = 1$. Therefore, there is a level of h, labeled $\overline{h}$, such that $g_N^{\max} \leq 0$ for $h \geq \overline{h}$, and $g_N^{\max} > 0$ for $h < \overline{h}$. In this case:

- $\lim_{h\to\frac{1}{2}} \chi\left(g_N^{\max}\right) > 0$, by the same argument as when $r \leq \frac{2}{3}$;
- if $h = \overline{h}$ we have $\chi\left(g_N^{\max}\right) = \chi(0) < 0$.

The function $\chi\left(g_N^{\max}\right)$ is continuous in h — because both $\chi(g_N; h, r)$ and $g_N^{\max}$ are continuous in h — $g_N^{\max} > 0$ for $h < \overline{h}$ and by Fact OA.10 we have $\frac{d\chi(g_N^{\max})}{dh} < 0$ for all $h \in \left(\frac{1}{2}, \overline{h}\right)$. By the Intermediate Value Theorem there must then exist a value $\hat{h} \in \left(\frac{1}{2}, \overline{h}\right)$ such that for all $h < \hat{h}$ we have $\chi\left(g_N^{\max}\right) > 0$. Since $g_N^{\max} > 0$ for all $h < \overline{h}$, by Fact OA.8 there exists an admissible solution $g_N \in (0,1)$ to $\chi(g_N) = 0$. Note that for $h \in \left(\hat{h}, \overline{h}\right)$ we have $\chi\left(g_N^{\max}\right) < 0$ while for $h \in \left(\overline{h}, 1\right)$ $g_N^{\max} < 0$. Hence, by Fact OA.8 it follows that in both cases there is no admissible solution to $\chi(g_N) = 0$. □

Finally, it remains to verify that the value of $g_N \in (0,1)$ such that $\gamma(g_N) = 0$ and $L = 0$ whose existence we established above can be implemented as a partial-deterrence equilibrium with an admissible expungement policy $\sigma > 0$. As shown in Fact OA.7, this property holds iff $g_N > 2h-1$. It is easy to see that since $h < \hat{h}$ implies $\chi\left(g_N^{\max}\right) > 0$, we must have 2 values of $g_N \in (0,1)$ such that $\chi(g_N) = 0$, each implemented by a different value of σ, one to the right and one to the left of $g_N^{\max}$. Therefore, to show that at least one of these values of σ is feasible, it is enough to show that $g_N^{\max} > 2h-1$ for all $h < \hat{h}$. In that case, at least the value of g_N such that $\chi(g_N) = 0$ is to the right of $g_N^{\max}$ and also satisfies the condition $g_N > 2h-1$. The next fact proves that the condition above holds.

Fact OA.11. *If $h < \hat{h}$, then $g_N^{max} > 2h-1$.*

Proof. Recall that, as shown in the proof of Fact OA.9 above, $\lim_{h\to\frac{1}{2}} g_N^{\max} = \frac{1}{3}$. Hence for values of h higher but close to $\frac{1}{2}$ the condition stated in the fact is satisfied. Next, note that solving the equation[55] $g_N^{\max} = 2h - 1$ for $h \in \left(\frac{1}{2}, 1\right)$, we obtain a unique solution, denoted:

$$h^* \equiv \frac{1}{2} + \frac{\sqrt{1 + (1-r)^2} - 1}{1-r} \in \left(\frac{1}{2}, 1\right)$$

Since we established above that for h higher but close to $\frac{1}{2}$ we have $g_N^{\max} > 2h - 1$ and there is only one value of h such that $g_N^{\max} = 2h - 1$ it follows that for all $h < h^*$ we have $g_N^{\max} > 2h - 1$. To establish the claim in the fact, it suffices then to show that $h^* > \hat{h}$, as this ensures that $g_N^{\max} > 2h - 1$ for $h \le \hat{h}$. In turn, since by Fact OA.10 $\chi\left(g_N^{\max}\right)$ is monotonically decreasing in h, to show that $h^* > \hat{h}$ it suffices to show that $\chi\left(g_N^{\max}; h^*, r\right) < \chi\left(g_N^{\max}; \hat{h}, r\right)$ for all $r \in (0, 1)$. Recall that, by construction, $\chi\left(g_N^{\max}; \hat{h}, r\right) = 0$ for all r. Replacing $g_N^{\max}$ with the expression in Fact OA.8, we obtain:

$$\begin{aligned} \chi\left(g_N^{\max}; h^*, r\right) \quad = \quad & \tfrac{2}{(3(1-r))^3}\left(14z - 13(1-r)\right) r^3 \\ & + \tfrac{2}{(3(1-r))^3}\left(-3\left(1-r\right) - \left(84 + 4y\right) z + 5\left(1-r\right) y\right) r^2 \\ & + \tfrac{2}{(3(1-r))^3}\left(3\left(1-r\right) + \left(14 + 10y\right) z - 14\left(1-r\right) y\right) r \\ & - \tfrac{2}{(3(1-r))^3}\left(99\left(1-r\right) + 56z + \left(11\left(1-r\right) - 6z\right) y\right) \end{aligned}$$

where:

$$y \equiv \sqrt{r\left(5r - 14\right) + 11 - \left(6 - 4r\right)\sqrt{1 + (1-r)^2}}$$

and:

$$z \equiv \sqrt{1 + (1-r)^2}$$

Note that $\chi\left(g_N^{\max}; h^*, r\right)$ is only a function of r and by analyzing this function we see that $\chi\left(g_N^{\max}; h^*, r\right) < 0$ for all $r \in [0, 1]$. $\square$

We have thus established the existence, for all $h < \hat{h}$, of a policy σ that implements a partial-deterrence equilibrium $\underline{g_N}$ with $L = 0$. It is then easy to verify that, at this equilibrium, $\frac{dL}{d\sigma} < 0$. Hence, there exists another policy $\sigma' > \sigma$ for which a partial-deterrence equilibrium exists with a net social loss from crimes $L < 0$. Finally, we have:

Fact OA.12. $\frac{d\hat{h}}{dr} = -\frac{\frac{\partial \chi\left(g_N^{max}\right)}{\partial r}}{\frac{\partial \chi\left(g_N^{max}\right)}{\partial h}} < 0$

Proof. Recall that $\hat{h}$ is defined implicitly as a solution of the equation $\chi\left(g_N^{\max}; \hat{h}, r\right) =$

[55] After substituting the expression of $g_N^{\max}$ in Fact OA.8.

0. Note that:

$$\begin{aligned} \frac{d\chi(g_N^{\max})}{dr} &= \left.\frac{\partial\chi(g_N)}{\partial r}\right|_{g_N=g_N^{\max}} \\ &= -2\left(1-g_N^{\max}\right)^2\left(g_N^{\max}+1-2h\right) \\ &< 0 \end{aligned} \tag{OA.15}$$

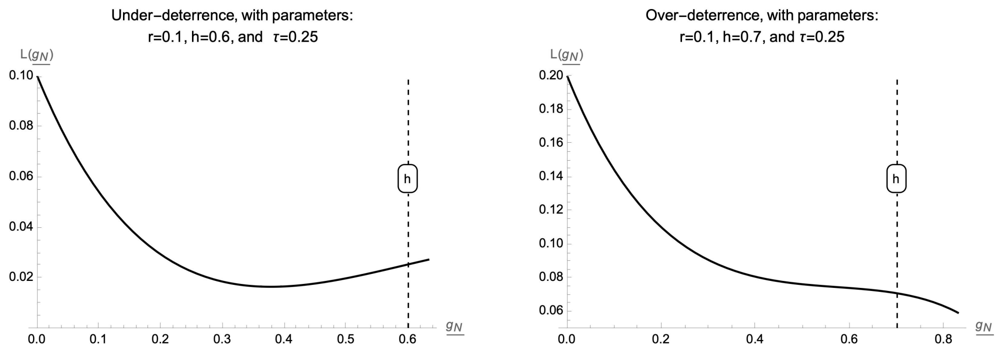
where the last term is positive because of the inequality in Fact OA.11, where we showed that $h^* > \hat{h}$ and hence that $g_N^{\max} > 2h-1$ when $h = \hat{h}$. The statement then follows from Fact OA.10. □

This completes the proof of claim 2. □

Proof of claim 3. In this case we have $\frac{1}{2} \leq h < \max\left\{\hat{h}, \frac{1}{2}+\tau\right\}$. Hence, either full deterrence cannot be implemented ($h < \frac{1}{2}+\tau$), and hence partial deterrence is optimal and can be implemented at the optimum by some $\sigma^* \geq 0$, or full deterrence is implementable but a lower social loss ($L < 0$) can be attained with a strictly positive expungement rate $\sigma^* > 0$ ($\frac{1}{2}+\tau \leq h < \hat{h}$). To prove the rest of the claim, it suffices to find a set of parameter values such that under-deterrence is optimal and a different set of parameter values such that over-deterrence is optimal. We do so by providing the results for two different numerical specifications. In both specifications the parameters are such that $h < \frac{1}{2}+\tau$ and $h > \frac{1}{2(1-r)}$, so that we are in case 2 (b) of Proposition 1, where partial deterrence is the unique equilibrium (full deterrence is not implementable). Note also that in both specifications the maximum level of deterrence that can be implemented is constrained because τ is sufficiently large.

The first specification has parameters $r = 0.1$, $h = 0.6$, and $\tau = 0.25$. Figure OA.7a reports the resulting value of the net social loss $L\left(\underline{g_N}\right)$, as a function of the equilibrium level of deterrence $\underline{g_N}$(rather than of σ)[56]. We see this function is convex and the maximum level of deterrence that can be implemented by setting $\sigma = 0$ is $\underline{g_N} = 0.63$, as evidenced by the range of attainable values of g_N on the horizontal axis. The loss is minimized by a level of deterrence equal to $\underline{g_N} = 0.38$, which is substantially less than the harm, $h = 0.6$ and hence implies under-deterrence compared to the first best.

[56] Since there is a 1-to-1 correspondence between $\underline{g_N}$ and σ this transformation is innocuous.



(a) Under-deterrence

(b) Over-deterrence

Figure OA.7: Optimal partial deterrence when $\frac{1}{2} \leq h < \max\left\{\hat{h}, \frac{1}{2} + \tau\right\}$

The second specification features the same values of r and τ but a higher value for the social harm from crimes, $h = 0.7$. The value of the net social loss for this specification is shown in Figure OA.7b, which we see is now a decreasing function of the equilibrium level of deterrence $\underline{g_N}$. The maximum level of deterrence that can be implemented in this case by setting $\sigma = 0$ is $\underline{g_N} = 0.83$. This value also minimizes the loss and we see it is greater than $h = 0.7$, so we have over-deterrence compared to the first best. □

OA.3 Details of the proof of Proposition 3

Proof of claim 1. Setting $(V_N - V_I)|_{g_N=1} = 1$ we obtain: $\zeta+\tau = (1-\lambda)\left(h - \frac{1}{2}\right)+\theta\pi$. Hence, if $(1-\lambda)\left(h-\frac{1}{2}\right) + \theta\pi - \tau \geq 0$ full deterrence can be implemented with nonmonetary sanctions. The threshold value of τ below which full deterrence can be implemented in equilibrium under nonmonetary sanctions is thus: $\tau = (1-\lambda)\left(h-\frac{1}{2}\right) + \theta\pi$. The corresponding equation in the case of criminal records is instead: $\tau = h - \frac{1}{2}$. Hence the threshold is higher with nonmonetary sanctions as compared to criminal records — that is, full deterrence is easier to implement — iff $\theta\pi > \lambda\left(h - \frac{1}{2}\right)$ and lower otherwise.

Proof of claim 2: If there is an equilibrium with partial deterrence (and so an internal solution for g_N) under incapacitation and/or under stigma, then the level of deterrence is higher with incapacitation as compared to stigma for any given $\zeta = \sigma$, if $\frac{w_N - w_I + \frac{1-(1-\lambda)-g_N^2}{2}}{1-g_N+\zeta+\tau} > \frac{w_N - w_C - \frac{g_N^2}{2}}{1-g_N+\zeta+\tau}$ for any g_N, that is, iff $w_I - w_C < \frac{1-(1-\lambda)}{2}$, which in turn holds iff $\theta\pi > \lambda\left(h-\frac{1}{2}\right)$.

Proof of claim 3: The claim can be easily verified by calculating the first partial derivatives of $V_N - V_I$ with respect to λ and θ. □

OA.4 Details of the proof of Proposition 4

Before proving the result, we need to characterize first the properties of an equilibrium under imprisonment. We do so in the following proposition.

Proposition OA.1. *When $\theta = \lambda = 1$, for any fraction $r \in (0,1)$ of dishonest workers, we have:*

1. *If $\zeta + \tau \leq \pi$, there is a unique equilibrium with full deterrence, $\overline{g_N} = 1$;*

2. *If $\zeta + \tau > \pi$, there is a unique equilibrium with partial deterrence, $\underline{g_N} \in (0,1)$. This equilibrium exhibits the property $\frac{\partial \underline{g_N}}{\partial(\zeta+\tau)} < 0$ and converges to no deterrence as $\zeta + \tau$ goes to infinity: $\lim_{\zeta+\tau\to\infty} \underline{g_N} = 0$.*

Proof. When $\theta = \lambda = 1$, the expression of the value functions derived in Section 4.1 simplifies as follows:

$$\begin{array}{rcl} V_N & = & \frac{1}{\tau}\frac{\zeta+\tau}{1-g_N+\zeta+\tau}\left(w_N + \frac{1-g_N^2}{2}\right) \\ V_I & = & \frac{1}{\tau}\frac{\zeta}{1-g_N+\zeta+\tau}\left(w_N + \frac{1-g_N^2}{2}\right) \end{array}$$

where w_N satisfies (10). Similarly, the equation determining the equilibrium level of partial deterrence simplifies to:

$$g_N = V_N - V_I = \frac{w_N + \frac{1-g_N^2}{2}}{1 - g_N + \zeta + \tau} \tag{OA.16}$$

This equation can be rewritten as:

$$\gamma_I(g_N) \equiv V_N - V_I - g_N = 0$$

Using (OA.16) and substituting the expressions of V_N, V_I, we obtain:

$$\begin{array}{rcl} \gamma_I(g_N) & = & \frac{((1-r)(1-g_N)+\zeta+\tau)\left(\pi+\frac{1-g_N^2}{2}\right)}{(1-g_N+\zeta+\tau)((1-r)(1-g_N)+\zeta+\tau)} \\ & & -\frac{r(\zeta+\tau)(1-g_N)h}{(1-g_N+\zeta+\tau)((1-r)(1-g_N)+\zeta+\tau)} \\ & & -\frac{(1-g_N+\zeta+\tau)((1-r)(1-g_N)+\zeta+\tau)g_N}{(1-g_N+\zeta+\tau)((1-r)(1-g_N)+\zeta+\tau)} \end{array}$$

Note that there are only two differences between $\gamma_I(g_N)$ and the corresponding expression $\gamma(g_N)$ with a criminal record: (1) the rate of expungement, σ, is replaced by the rate of release from prison, ζ, and (2) the term $\left(h - \frac{g_N^2}{2}\right)$ becomes $\left(\pi + \frac{1-g_N^2}{2}\right)$. Let $\overline{\gamma}_I(g_N)$ denote the numerator of $\gamma_I(g_N)$ and note that it is a polynomial of degree 3. As before, $\overline{\gamma}_I(g_N)$ is a continuously differentiable and strictly convex function of $g_N \in [0,1]$. Hence the equation $\overline{\gamma}_I(g_N) = 0$ has at most 2 solutions for $g_N \in [0,1]$. It is then easy to prove that the following two properties. The first is analogous to Fact OA.2 used in the proof of Proposition 1, while the second qualifies Fact OA.3, also used in that proof.

- $\overline{\gamma}_I(0) > 0$ for all values of ζ, τ, and π; while $\overline{\gamma}_I(1) \geq 0$ iff $\zeta + \tau \leq \pi$;

- $\left.\overline{\gamma}_I'(1)\right|_{\overline{\gamma}_I(1)=0} = 2\pi(rh-\pi) < 0.$

Then, following the same approach as in the proof of Proposition 1 yields the results. $\square$

Therefore, the equilibrium under imprisonment is always unique. The threshold value of τ below which full deterrence can be implemented (for some $\zeta \geq 0$) is now given by π, instead of $h - \frac{1}{2}$. Note that $\pi \geq 1 > h - \frac{1}{2}$. Hence, full deterrence is always easier to implement with imprisonment than with a criminal record.[57] We then see from (12) that the social loss in I is now equal to the productivity π, and is thus always greater than the one in C with criminal records, because $\pi > h - \frac{1}{2}$. The overall loss, however, also depends on the value of deterrence $\underline{g_N}$ attained in equilibrium, which, in turn, depends on the length of imprisonment, given by $\frac{1}{\zeta}$. The optimal level of ζ that balances these effects is given in Proposition 4, which we can now prove.

Note that $\overline{g_N} = 1$ is implementable with $\zeta = 0$ because $\pi > \tau$ (Proposition OA.1), and the social loss with full deterrence is $L_I = 0$. Moreover, we have $\lim_{\zeta\to\infty} L_I = h - \frac{1}{2}$ with no deterrence, since $\lim_{\zeta\to\infty} \underline{g_N} = 0$. Hence the choice between full and no deterrence trivially depends on whether h is above or below $\frac{1}{2}$. What remains to be proven is that there are no other levels of partial deterrence that achieve a lower social loss in each of these two cases.

Proof of claim 1. We consider the case $h < \frac{1}{2}$. We need to show that any equilibrium level of partial deterrence $\underline{g_N} \in (0,1)$ yields a social loss greater than $h - \frac{1}{2}$. Since L_I is continuous and the social loss with full deterrence is equal to $0 > h - \frac{1}{2}$, to show that the social loss cannot be less than $h - \frac{1}{2}$ it suffices to show that there is no simultaneous solution to $L_I = h - \frac{1}{2}$ and $\gamma_I = 0$ for $\underline{g_N} \in (0,1)$ and any value of the parameters. The release rate $\zeta\left(\underline{g_N}\right)$ that implements the equilibrium $\underline{g_N}$ is obtained by solving $\gamma_I = 0$ and is then given by:

$$\zeta\left(\underline{g_N}\right) = 1 + 2\pi - 2hr - 2\underline{g_N}(2 - r - hr + 2\tau) + \underline{g_N}^2(3 - 2r) \pm \sqrt{A} \quad \text{(OA.17)}$$

where:

$$\begin{aligned} A \;=\; & \left(1 - \underline{g_N^2} + 2\pi\right)^2 + 4\left(1 - \underline{g_N}\right)^2\left(h - \underline{g_N}\right)^2 r^2 \\ & -4\left(1 - \underline{g_N}\right) r\left(\left(1 - \underline{g_N}\right)\left(\underline{g_N^2} - \underline{g_N}(3h-1) + h\right) + 2\pi\left(h + \underline{g_N}\right)\right) \end{aligned} \quad \text{(OA.18)}$$

It is tedious but straightforward to show that $A > 0$ and that the "$-$" solution is negative while the "+" solution is positive and decreasing in $\underline{g_N}$. Replacing the latter value into the expression of the net loss L_I we obtain:

$$L_I^{\zeta}\left(\underline{g_N}\right) = \frac{1}{4r}\left(1 - \underline{g_N^2} + 2\pi - 2\left(1 - \underline{g_N}\right)(1-h)r - \sqrt{A}\right) \quad \text{(OA.19)}$$

We need to show that the equation $L_I^{\zeta}\left(\underline{g_N}\right) = h - \frac{1}{2}$ has no solution for $\underline{g_N} \in (0,1)$ and that this is true for all h, r, and π. We will do so by contradiction. Simple

[57] The condition stated in point 1. of Proposition 3 is always satisfied in the case of imprisonment.

manipulations show that a necessary and sufficient[58] condition for $L_I^\zeta\left(\underline{g_N}\right)=h-\frac{1}{2}$ is:

$$\left(1-\underline{g_N}^2+2\pi-2\left(h-\underline{g_N}\left(1-h\right)\right)r\right)^2=A$$

Expanding, dividing by $4\underline{g_N}r$, and rearranging yields:

$$\underline{g_N}^3\left(1-r\right)-2\underline{g_N}^2\left(1+h\right)\left(1-r\right)-\underline{g_N}\left(2\pi+1-2h\left(4-3r\right)\right)-4h\left(\pi-hr\right)-6h+4\pi+2=0$$

Solving for π yields:

$$\pi^*=\frac{1}{2}\left(\frac{\left(1-\underline{g_N}\right)\left(6h-2-\underline{g_N}\left(2h+1-\underline{g_N}\right)\right)-r\left(2h-\underline{g_N}\right)\left(2h-\left(2-\underline{g_N}\right)\underline{g_N}\right)}{2\left(1-h\right)-\underline{g_N}}\right)$$

which is decreasing in r[59] and is then maximal for $r=0$. Therefore:

$$\pi^*\le\pi^*_{r=0}=\frac{1}{2}\left(1-\underline{g_N}\right)\left(\frac{2h\left(3-\underline{g_N}\right)-2-\underline{g_N}\left(1-\underline{g_N}\right)}{2\left(1-h\right)-\underline{g_N}}\right)$$

which is increasing in h and hence maximal for $h=\frac{1}{2}$. Therefore:

$$\pi^*\le\pi^*_{r=0,h=\frac{1}{2}}=\frac{1}{2}\left(1-\underline{g_N}\right)^2\le\frac{1}{2}$$

Since $\pi\ge 1$, the equation $L_I^\zeta\left(\underline{g_N}\right)=h-\frac{1}{2}$ has no solution. As a result, we cannot have $L_I^\zeta\left(\underline{g_N}\right)<h-\frac{1}{2}$ for $\underline{g_N}\in(0,1)$ when $h<\frac{1}{2}$, which proves the claim.

Proof of claim 2. We consider next the case $h\ge\frac{1}{2}$. We need to show that any equilibrium level of partial deterrence $\underline{g_N}\in(0,1)$ yields a positive social loss. Given continuity and $\lim_{\underline{g_N}\to 0}L_I=h-\frac{1}{2}>0$, it suffices to show that there is no simultaneous solution to $L_I=0$ and $\gamma_I=0$ for $\underline{g_N}\in(0,1)$ for any value of the parameters. We proceed as before and use expressions (OA.18) and (OA.19). We need to show that the equation $L_I^\zeta\left(\underline{g_N}\right)=0$ has no solution for $\underline{g_N}\in(0,1)$ and for all h, r, and π. We will do so by contradiction. Simple manipulations show that a necessary and sufficient[60] condition for $L_I^\zeta\left(\underline{g_N}\right)=0$ is:

$$\left(1-\underline{g_N}^2+2\pi-2\left(1-\underline{g_N}\right)\left(1-h\right)r\right)^2=A$$

Expanding, dividing by $4\left(1-\underline{g_N}\right)r$, and rearranging yields:

$$\underline{g_N}^3\left(1-r\right)-\underline{g_N}^2\left(2h+1\right)\left(1-r\right)-\underline{g_N}\left(2\pi-\left(4h-1\right)\left(1-r\right)\right)-\left(2h-1\right)\left(2\pi+1-r\right)=0$$

[58] This is because $1-\underline{g_N}^2+2\pi-2\left(h-\underline{g_N}\left(1-h\right)\right)r>0$.

[59] This is because $\min_h\left[\left(2h-\underline{g_N}\right)\left(2h-\left(2-\underline{g_N}\right)\underline{g_N}\right)\right]=\frac{\left(1-\underline{g_N}\right)^2\underline{g_N}^2}{4}>0$

[60] This is because $1-\underline{g_N}^2+2\pi-2\left(1-\underline{g_N}\right)\left(1-h\right)r>0$.

Solving for h yields:

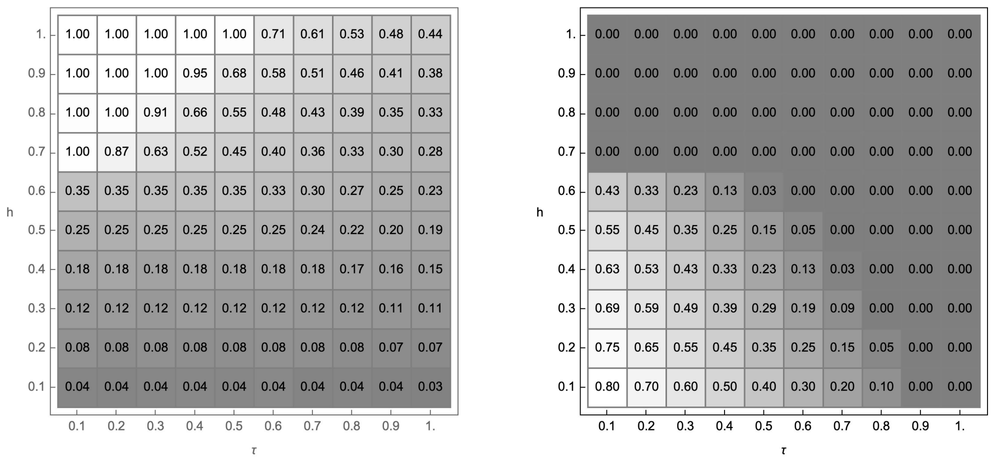
$$h^* = \frac{1}{2}\left(1 + \underline{g_N} - \frac{4\underline{g_N}\pi}{2\pi + \left(1 - \left(2 - \underline{g_N}\right)\underline{g_N}\right)(1 - r)}\right)$$

which is decreasing in r, and so is maximal for $r = 0$, and hence

$$h^* < \frac{1}{2}\left(1 - \underline{g_N}\frac{2\pi - \left(1 - \underline{g_N}\right)^2}{2\pi + \left(1 - \underline{g_N}\right)^2}\right) < \frac{1}{2}$$

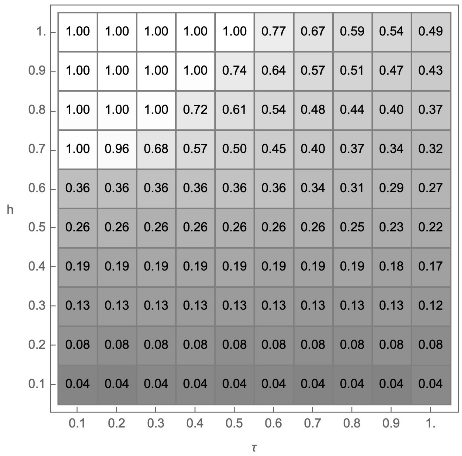
which cannot hold because $h \geq \frac{1}{2}$. Therefore, we cannot have $L_I^\zeta\left(\underline{g_N}\right) < 0$ for $\underline{g_N} \in (0, 1)$, which proves the claim. $\square$

OA.5 Optimal expungement policy for higher values of r

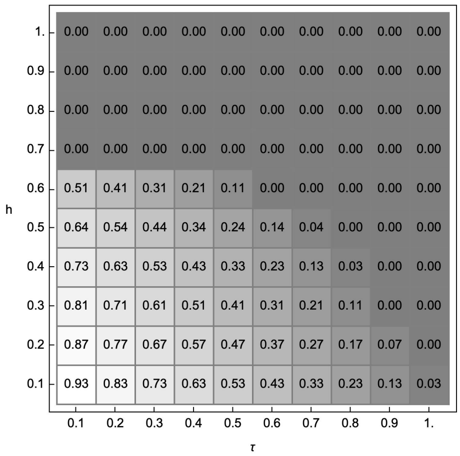


(a) Optimal deterrence (b) Optimal expungement rate

Figure OA.8: The optimal policy (with $r = 0.5$)



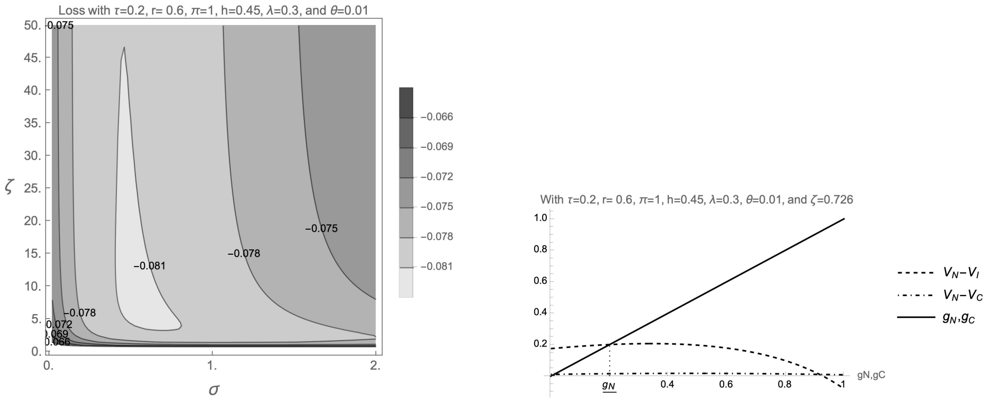
(a) Optimal deterrence



(b) Optimal expungement rate

Figure OA.9: The optimal policy (with $r = 0.75$)

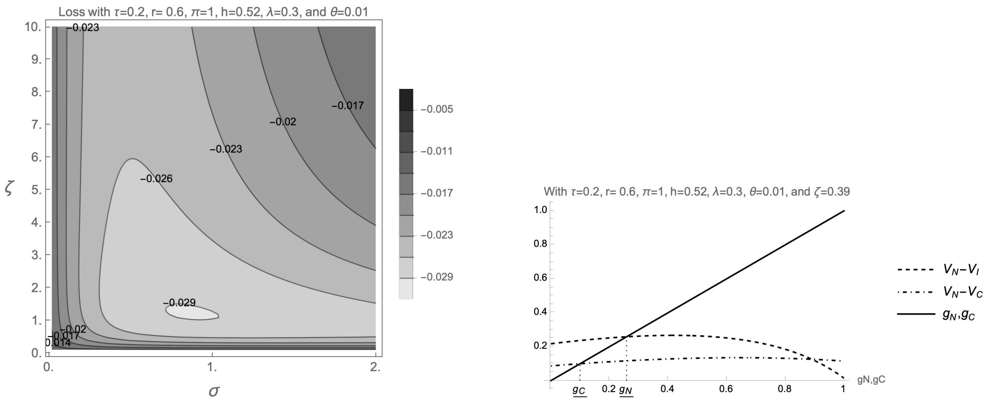
OA.6 Numerical analysis plots



(a) Social loss

(b) Equilibrium deterrence ($\zeta = 10$ and $\sigma = 0.5$)

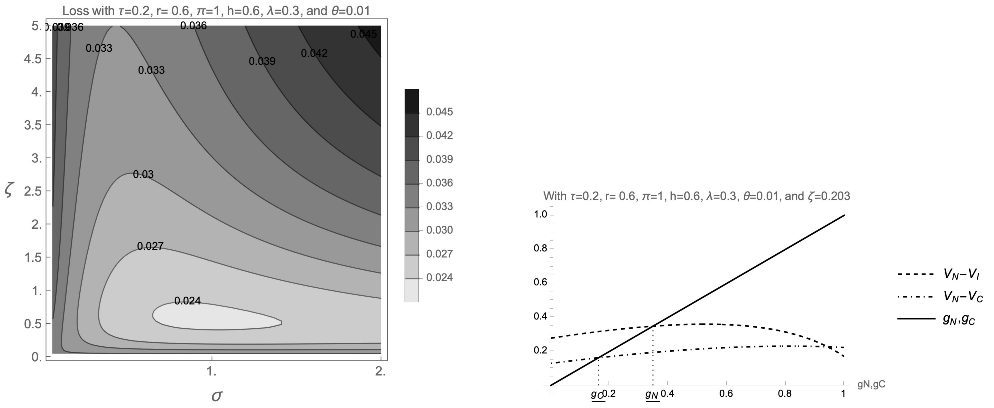
Figure OA.10: Welfare-improving mixed policy ($h = 0.45$)



(a) Social loss

(b) Equilibrium deterrence ($\zeta = \sigma = 1$)

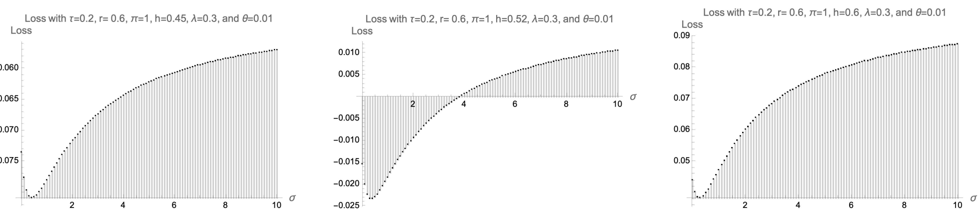
Figure OA.11: Welfare-improving mixed policy ($h = 0.52$)



(a) Social loss

(b) Equilibrium deterrence ($\zeta = 0.5$ and $\sigma = 1$)

Figure OA.12: Welfare-improving mixed policy ($h = 0.6$)



(a) Social loss with $h = 0.45$

(b) Social loss with $h = 0.52$

(c) Social loss with $h = 0.6$

Figure OA.13: The optimal policy for the criminal record