# Prohibitions on Punishments in Private Contracts[☆]

Philip Bond

*University of Pennsylvania*

Andrew F. Newman

*Boston University and CEPR*

**Abstract**

In most contemporary economies loan contracts that mandate *exclusionary* penalties such as imprisonment or other non-pecuniary punishments for defaulting debtors are illegal, despite the fact that in some cases contracting parties might gain by being able to use them. A possible rationale for contracting restrictions of this type is that exclusion imposes negative externalities on individuals not party to the original loan contract. We explore the ability of such externalities to account for these restrictions. We contrast exclusion with enforceable collateral seizure, a widespread feature of developed financial systems. We also consider "behavioral" agents who underestimate their chances of being punished, and show that overconfidence of this type is a less compelling justification for restrictions on exclusionary punishments than often argued.

*Key words:* Exclusionary penalties; collateral; externalities; overconfidence.

*JEL codes:* C7, D6, G2, K1, O1.

[☆]Both authors thank the Institute for Advanced Study for hospitality and financial support. Bond thanks Deutsche Bank for financial support. We thank Benjamin Hermalin, Kathryn Spier, audience members at the Econometric Society Meetings, Vish Viswanathan (the editor) and especially two anonymous referees for helpful comments.

## 1. Introduction

A great deal of recent research in contract theory has focused on what has become known as "contractual incompleteness." By this, economists mean that certain contingencies are impossible to include in private contracts.[1] However, a second — and perhaps equally important — constraint on private contracting has received much less attention: the state typically restricts which kinds of contracts it will deem enforceable. A leading instance are the limitations on punishments that can be imposed on a party breaching a contract. For example, a contract cannot stipulate imprisonment, or bondage, or corporal punishment. Most jurisdictions place substantial limitations on the use of non-compete clauses in employment contracts; some (notably California) ban such clauses altogether.[2] Relatedly, courts rarely impose specific performance on a party who breaches a contract.

A key element of commonality among these various forms of punishment is that they are to some degree *exclusionary:* if permitted by the state and carried out by the contracting parties, they would effectively limit the punished individual's economic activities in the future.

The legal restrictions on exclusionary punishments stand in sharp contrast with the permissive attitude of the state toward transfers of pledged collateral. Indeed, many academics argue that enforceable collateral seizure is crucial for financial development, and hence for overall economic performance: see for instance de Soto (2000), Pande and Udry (2007), and Djankov *et al* (2008), and Levine (1997, 2006). However, the incentives that can be provided by collateral seizure are naturally limited,

[1] See, e.g., Hart (1995) for a survey.

[2] See, e.g., Garmaise (2006). Consistent with our paper's argument, Saxenian (1996) argues that enforcement of non-compete clauses in Massachusetts partly accounts for the decline of the Route 128 high-tech economy relative to its Silicon Valley, California counterpart; Samila and Sorenson (2009) present empirical evidence that allowing enforcement of such clauses impedes patent production and firm and job creation.

and consequently restrictions on exclusionary punishments potentially have large and important effects. For example, if loan contracts could threaten defaulting borrowers with imprisonment, credit constraints would be ameliorated and perhaps even eliminated.[3] From this perspective, limited punishments may even be construed to be at the root of many underdevelopment problems such as "poverty traps" (Banerjee and Newman, 1994), and their imposition by the state is therefore all the more puzzling. Put starkly, if effective enforcement of collateral seizure is so potentially important for economic performance, why not allow debt-bondage also?

In this paper we examine the extent to which restrictions on exclusionary contracts, together with the acceptance of collateral seizure, can be rationalized by a simple economic explanation, namely that exclusionary contracts impose negative externalities on other individuals not party to the original contract.

To fix ideas, consider the specific (and standard) case of a would-be entrepreneur endowed with an investment opportunity, but lacking funds. The entrepreneur can raise financing by promising some share of future output to a lender. However, if the amount promised (i.e., the interest rate) is too high, the entrepreneur's incentive to exert effort is low, and overall surplus is negatively impacted. Indeed, the entrepreneur's incentive efforts may be so diluted that expected output falls below the opportunity cost of capital, and financing is impossible.

One way for the entrepreneur to improve his access to credit would be to agree to accept an exclusionary punishment if he defaults. For example, he could sign a debt contract in which he must work for the lender for free if he defaults; or in which he is imprisoned. Indeed, historically contracts of this type were permitted, and widespread. The advantage of such a contract is that it gives the entrepreneur

[3] Under many circumstances, the bilateral welfare losses due to credit constraints and agency problems would disappear (or nearly so) if arbitrary punishments were possible. See, e.g., Mirrlees (1999) and Mookherjee and Png (1989).

a greater incentive to work hard, and so can be accompanied by a reduction in the interest rate.

The drawback of using exclusionary punishments to incentivize the agent is that with some probability the entrepreneur is unlucky and defaults in spite of working hard. While the entrepreneur internalizes the private cost of the exclusionary punishment (this, after all, is why the punishment provides incentives), he does not consider the full social cost. In particular, the entrepreneur does not consider the positive surplus that third-parties might gain from dealing with him, were they allowed do to so. Consequently, it is possible for the state to improve overall social surplus by restricting the use of exclusionary contracts.

Although simple, this externality-based explanation delivers predictions that are broadly consistent with the observed incidence of contracting constraints. First, the negative externality is larger when the growth rate of the economy is high, when uncertainty about the value of future economic interactions is high, and when the number of possible future economic interactions is large. These predictions provide an explanation for why debtor's prison was eliminated in the U.S. and western Europe at roughly the same time as industrialization occurred. Second, and as we discuss in Section 4, this explanation can account for the asymmetry between the treatment of exclusionary punishments and collateral seizure. Related, it is worth noting that restrictions are placed on the use of "tools of the trade" as collateral:[4] this is precisely a case in which collateral seizure has a significant exclusionary effect, and so this exception *supports* our argument. Third, and almost immediate, this explanation accounts for why the state itself continues to use exclusionary punishments such as

[4] See, for example, 522(f) of the U.S. Bankruptcy Code. We thank an anonymous referee for suggesting this point. Item 20 of the Magna Carta provides a second and older example: "a free man shall [not be fined] so heavily as to deprive him of his livelihood. In the same way, a merchant shall be spared his merchandise, and a husbandman the implements of his husbandry."

imprisonment even at the same time as it bans their use by private parties: the state is able to internalize the full effect of any externalities, while private parties do not.

We conclude the paper by briefly considering an alternative rationale for restrictions on exclusionary contracts, namely that individuals may underestimate the probability that the exclusionary punishment is actually imposed. This rationale is close to one that has been promulgated in the legal literature (McCormick, 1935). But it is also closely related to the main argument of the paper: instead of exclusionary contracts imposing an externality on third parties, they impose an externality on future selves. Nonetheless, we show that this rationale is less compelling than one might think. Specifically, if an individual overestimates his probability of success today,[5] he is likely to also overestimate the cost of being excluded in the future, and this effect mitigates the tendency to agree to exclusionary contracts today.

*Alternative explanations and related literature*

The main alternative to the externality-based arguments we consider is that restrictions on punishments stem from ethical concerns. This argument seems at best incomplete, because it overlooks the fact that, as noted, the state regularly uses imprisonment as a punishment for other offenses, even ones of a non-violent nature. Thus a debtor who consumes his loan instead of investing it and repaying his creditor cannot be imprisoned; while in a directly analogous setting a taxpayer who fails to pay his "debt" to the government may well suffer just such a punishment. If instead punishment restrictions arise from a need to control negative externalities, it is entirely consistent for the state to both restrict their use by private parties, and to allow itself to deploy them in some circumstances. Similarly, it is not clear why a

[5] For recent research on financing overconfident entrepreneurs, see, e.g., Gervais and Goldstein (2007), Hackbarth (2008), and Landier and Thesmar (2009).

greater ethical problem arises in enforcing a non-compete clause, for example, than in stripping an individual of valuable collateral. The latter sanction is, of course, entirely legal in almost all contemporary economies.

A second drawback to explaining punishment restrictions on ethical grounds is that widespread constraints on the use of non-monetary punishments are for the most part a historically recent phenomenon. Imprisonment for debt persisted into the nineteenth century in the United States.[6] The closely related institution of bound labor was not abolished until 1867 Antipeonage Act, i.e., *after* slavery had been abolished in the United States.[7] It is worth noting that advocates for the abolishment of debtor's prison deployed arguments based on overall social welfare and on new employment, very much in line with our paper. For example, Mann (2002, page 83) quotes the following passage from a 1754 Rhode Island pamphlet:

> it is best for Society, that his Creditors receive a Proportion of their Debts ... and his Person be sat at Liberty to seek new Employment; or that his Body be imprisoned for the Deficiency, until he pays the utmost Farthing, which is impossible?

Our paper is related to the small literature that has examined a much milder restriction on private contracts, namely the non-enforceability of penalty clauses for breach. Aghion and Bolton (1987), Chung (1992), and Spier and Whinston (1995) are leading examples, all of which seek to account for this restriction as stemming

[6]Mann (2002, page 79) summarizes the situation in United States shortly after independence in the following terms: "every colony north of the Potomac, with the possible exception of New Hampshire, permitted insolvent debtors to be bound in service to their creditors without their consent, typically for as long as seven years, the standard term for indentured servants." New York did not abolish imprisonment for indebtedness until 1831, and Pennsylvania did not do so until 1842 (Mann 2002, page 106).

[7]A concise legal history can be found in the Supreme Court case of Pollock vs Williams (322 U.S. 4).

from the fact that penalty clauses, while privately optimal, may be socially undesirable. In each case the social undesirability stems from the fact that penalty clauses can be used to deter entry into an industry. Also related is Diamond and Maskin (1979) who study the effects of penalty clauses on search and breach intensity in the labor market. Aghion and Hermalin (1990) and Spier (1992) have suggested that contractual constraints, among them constraints on punishments, exist as a way to prevent socially unproductive signaling.

Finally, in recent work Mookherjee and Lilienfeld-Toal (2008a, 2008b) consider issues that are related to our paper, though from a different perspective. They show that restricting the set of available punishments (they consider debt-bondage and collateral seizure) may help some agents via general equilibrium surplus-division effects.

Conceptually, our approach shares some elements of commonality with Camerer *et al* (2003), who seek to account for when apparently paternalistic policies (of which restrictions on non-monetary punishments is one example) can be justified on the grounds that they help boundedly rational agents more than they hurt rational agents. In common with Section 5 of our paper, Hynes (2004) points out that the implications of exactly what inefficiencies might stem from assuming that individuals are boundedly rational are less obvious than they might at first seem. Finally, Chwe (1990) characterizes conditions under which the use of physical violence to provide incentives is privately optimal; in contrast, our focus is on when such punishments are privately optimal but socially sub-optimal.

## 2. Model

On the one hand, exclusionary punishments incentivize an agent to work. On the other hand, they impose a cost on the agent and his future trading partners. To make the analysis as transparent as possible, we use the simplest model capable of capturing this trade-off.

There are two periods and three risk-neutral individuals — one agent, $A$, whose consumption must be non-negative (the entrepreneur in the example of the introduction) and two principals, $P_1$ and $P_2$ (e.g., lenders).

In each period $t \in \{1, 2\}$ the agent $A$ can contract with principal $P_t$. The value of the agent's time (i.e., his reservation utility) is $w$, which one can interpret either as the value of the agent's leisure, or as the amount he can produce in autarky. If he contracts with principal $P_t$, the principal supplies capital, with an opportunity cost that we normalize to 1. The agent either exerts high effort, i.e., "works," or low effort, i.e., "shirks." If the agent shirks, output is $H_t > 0$ with probability $p_t$, and output is 0 with probability $1 - p_t$. High effort raises the probability of output $H_t$ by $\Delta_t > 0$ to $q_t \equiv p_t + \Delta_t < 1$, but imposes a utility cost $B_t$ on the agent. Throughout, we assume that high effort is socially efficient, i.e., $\Delta_t H_t > B_t$.

A contract in period $t$ specifies principal $P_t$'s share of high output, $H_t - x_t$, and the agent's share, $x_t \geq 0$.[8] In addition to monetary incentives, principal $P_1$ has the option of taking some action that serves to deny the agent access to the period 2 labor market. As discussed above, imprisonment, non-compete clauses, debt-bondage, and (to a slightly lesser extent) corporal punishment all fall within this class. Accordingly, at date 1 the contract is a pair $(x_1, \pi)$, where $\pi$ is the probability that the agent is excluded in period 2 if the low output is realized.[9] An important assumption is that if the agent is excluded, there are frictions that impede principal $P_2$ from paying

[8] The agent cannot receive a negative amount: we assume he has no savings, and the only way to force him to borrow would be to threaten him with a non-pecuniary punishment — as in the contract under consideration.

[9] It is straightforward to show that neither the principal nor agent can gain from using a contract in which either the agent receives a strictly positive amount when output 0 is realized, or in which the agent is excluded with strictly positive probability when high output is realized.

principal $P_1$ to undo this exclusion.[10,11] For simplicity, we assume renegotiation of this type is simply impossible.

For simplicity, assume that the discount rate is zero, and that the agent cannot store any payments received in period 1 in order to ease the second period incentive problem. Let $U_2$ be the agent's expected utility in period 2, provided he is not excluded. The agent's period 1 incentive constraint is thus

$$q_1(x_1 + U_2) + (1 - q_1)(1 - \pi)U_2 - B_1 \geq p_1(x_1 + U_2) + (1 - p_1)(1 - \pi)U_2,$$

which reduces to

$$(x_1 + \pi U_2)\,\Delta_1 \geq B_1. \tag{IC$_1$}$$

Given the opportunity cost of capital, the principals' individual rationality constraints are

$$q_t\,(H_t - x_t) \geq 1 \tag{P$_t$-IR-W}$$

if the contract induces the agent to exert high effort, and

$$p_t\,(H_t - x_t) \geq 1 \tag{P$_t$-IR-S}$$

otherwise. In order to focus on the interesting case in which the contracting parties use exclusionary contracts, we assume throughout that there is no way to supply the agent with purely monetary incentives in period $t$ that both induce him to exert high

[10]For example, if principal $P_1$ is uncertain about how much surplus $P_2$ would derive from contracting with the agent, this will impede renegotiation (e.g., Spier 1994); the act of exclusion may simply be irrevocable, as in the case of corporal punishment, or state-controlled imprisonment (though this raises the question of whether one private party should be able to "buy" a second private party out of prison); or principal $P_1$ may decline to renegotiate for reputational reasons (this is most relevant if an agent released from exclusion is able to bargain so as to gain some utility from dealing with principal $P_2$).

[11]If instead principals $P_1$ and $P_2$ were able to contract to share the full social surplus available from an unexcluded agent ($S_2$, in the notation introduced below), with the agent receiving zero utility (as in exclusion), then a threat of exclusion would provide the agent with incentives without imposing any social cost. In this case, legal restrictions on exclusion would be unnecessary.

effort and satisfy principal $t$'s individual rationality constraint (P$_t$-IR-W):

$$q_t\left(H_t-\frac{B_t}{\Delta_t}\right)-1<0. \tag{1}$$

Finally, since some surplus is available at both dates, we need to specify how this surplus will be split. We adopt the standard randomized "take-it-or-leave-it" offers framework. That is, in period $t$ with probability $\theta_t$ the agent proposes a contract to principal $P_t$, who either accepts or rejects. Similarly, with probability $1-\theta_t$ the principal $P_t$ proposes a contract to the agent, who either accepts or rejects.

## 3. Contracts and tomorrow's trading partners

We start by illustrating the basic externality at work: when the agent and principal $P_1$ agree to use an exclusionary contract ($\pi > 0$) to incentivize the agent, they are imposing a cost on principal $P_2$ who no longer gets a share of the surplus in period 2. Because of this, if left unregulated the agent and principal $P_1$ will use exclusionary contracts more than is socially optimal.

If the agent is not excluded at the end of period 1, he is free to contract with principal $P_2$. Conditional on the agent not being excluded, we denote the total social surplus from the agent's relationship with principal $P_2$ by $S_2$, and the agent's utility by $U_2$. By assumption (1), the total social surplus attainable without an exclusionary contract is $S^{no-exc}\equiv\max\{p_1H_1,w+1\}+S_2$. The total social surplus if the period 1 contract induces the agent to work but imposes exclusion with probability $\pi$ is

$$S^{exc}(\pi)\equiv S^{no-exc}+\min\{\Delta_1H_1-B_1,q_1H_1-B_1-1-w\}-(1-q_1)\,\pi\,(S_2-1)\,.$$

That is, the threat of exclusion raises social surplus in period 1, but the imposition of exclusion then lowers social surplus in period 2.

The tendency of the agent and principal $P_1$ to overuse exclusion can be most clearly seen when the agent proposes the contract. In this case, principal $P_1$'s expected utility matches his outside option, 1.[12] Hence if the agent does not propose an exclusionary contract his utility is $U^{no-exc} = \max\{p_1 H_1 - 1, w\} + U_2$, while if he does propose an exclusionary contract (with the probability of exclusion equal to $\pi_A$; see Proposition 1 below) his utility is

$$U^{exc} = U^{no-exc} + \min\{\Delta_1 H_1 - B_1, q_1 H_1 - B_1 - 1 - w\} - (1 - q_1)\,\pi_A U_2.$$

Consequently

$$U^{exc} - U^{no-exc} = S^{exc}(\pi_A) - S^{no-exc} + (1 - q_1)\,\pi_A (S_2 - 1 - U_2)\,. \tag{2}$$

Equality (2) illustrates the source of contracting inefficiency. While the agent partially internalizes the social cost $S_2 - 1$ of exclusion, he ignores the portion of this cost that does not accrue to him, namely $S_2 - 1 - U_2$. As such, there are circumstances under which the agent proposes an exclusionary contract even though it destroys social welfare.[13] Moreover, and exactly as one would expect, if all the period 2 surplus accrues to the agent, i.e., if $\theta_2 = 1$ and hence $U_2 = S_2 - 1$, then the agent uses an exclusionary contract if and only if it is socially efficient to do so.

In order to more thoroughly characterize when social welfare destroying exclusion occurs, we must first determine parties' contractual choices in more detail. Throughout the paper, all proofs are relegated to the appendix.

**Proposition 1.** ***(Pareto frontier)***
*The exclusionary contract that maximizes the agent's utility subject to providing him with incentives and giving principal $P_1$ his outside option mandates an exclusion probability of $\pi_A \equiv \frac{1}{U_2}\left(\frac{B_1}{\Delta_1} - H_1 + \frac{1}{q_1}\right)$. The Pareto frontier attainable using exclusionary contracts is described by $U = U^{exc} - V/q_1$, where $U$ and $V \in [1, \bar{V}]$ are the agent's and principal $P_1$'s utilities, and principal $P_1$'s maximal attainable utility is*

[12] If this were not the case, the agent could increase his share of the success payoff $x$ without violating any incentive constraint.

[13] Conversely, if the agent's most preferred exclusionary contract increases social-welfare, the agent will propose it (in preference to a non-exclusionary contract).

$\bar{V} \equiv 1 + q_1 \left( U^{exc} - \frac{p_1 B_1}{\Delta_1} \right)$. *The exclusion probability increases in* $V$ *from* $\pi = \pi_A$ *at* $V = 1$ *to* $\pi = 1$ *at* $V = \bar{V}$.

From the expression for $\pi_A$ and equation (2), we obtain the following comparative statics:

**Corollary 1.** ***(Comparative statics)***
*The magnitude of the agent's bias in favor of the exclusionary contract is increasing in* $B_1/\Delta_1$, *and decreasing in* $\frac{U_2}{S_2-1}$, $H_1$ *and* $q_1$.

In particular, Corollary 1 formally demonstrates that the agent's bias towards exclusionary contracts is worse when the cost imposed upon him by exclusion, $U_2$, is only a small share of the total social cost of exclusion, $S_2 - 1$. Moreover, the agent's bias is worse when the moral hazard problem is more severe (as measured by $B_1/\Delta_1$); when there is sizeable probability of failure even when he works ($1 - q_1$ high); and when it is harder to supply adequate monetary incentives ($H_1$ low).

*Uncertainty, growth and mobility*

Historically, the introduction of restrictions on exclusionary punishments roughly coincides with industrialization. Consistent with this, our model predicts that the agent's bias in favor of socially inefficient exclusionary contracts is worse under three conditions commonly associated with industrialization — namely rapid growth; uncertainty; and occupational and/or geographic mobility. As such, one can view the introduction of contracting restrictions as an appropriate legislative response.

To see how our model delivers this prediction, consider the relation between the agent's loss of utility from period 2 exclusion, $U_2$, and the social cost of exclusion, $S_2 - 1$. The surplus generated by the agent contracting with principal $P_2$ is $S_2 - 1 - w$. Under our (standard) assumptions on how this surplus is divided, the agent receives $U_2 = w + \theta_2 (S_2 - 1 - w)$ when he is not excluded in period 2. Consequently, $\frac{U_2}{S_2-1} =$

$\theta_2 + (1-\theta_2)\frac{w}{S_2-1}$, and so an increase in total social surplus $S_2$ reduces the agent's share of the social cost of exclusion, worsening his bias.

Writing explicitly, $S_2 = E\left[\max\left\{p_2H_2, w+1\right\}\right]$, where the expectation corresponds to any uncertainty about principal $P_2$ as of period 1. So by Corollary 1:

**Corollary 2.** ***(Growth and uncertainty)***
*Either (1) an increase in the growth rate (higher expected $p_2H_2$), or (2) an increase in uncertainty in the value of $p_2H_2$, in the sense of second-order stochastic dominance, increases the agent's bias in favor of using an exclusionary contract.*

To capture an increase in mobility, suppose that instead of having the opportunity to deal only with principal $P_2$ in period 2, the agent can choose instead to deal with a third principal $P_2'$ who has a project with expected output $p_2'H_2'$. In this case the period 2 social surplus if the agent is not excluded is $E\left[\max\left\{p_2H_2, p_2'H_2', w+1\right\}\right] > E\left[\max\left\{p_2H_2, w+1\right\}\right]$. Consequently:

**Corollary 3.** ***(Increased mobility)***
*An increase in the agent's employment options at date 2 increases the agent's bias in favor of using an exclusionary contract.*

*Contract offers by principal $P_1$*

Returning to period 1, when principal $P_1$ makes the contract offer, from Proposition 1 he proposes a higher exclusion probability than the agent would. The reason is that the agent's incentive constraint ($\mathrm{IC}_1$) holds at equality, regardless of who makes the contract offer, since otherwise both the agent and principal could be made better off by lowering the exclusion probability. Since the principal's expected payoff is clearly higher when he makes the proposal, the agent's payment $x$ is lower in this case — and so the exclusion probability is in turn higher, for otherwise the incentive constraint would not hold. A related implication is that total social welfare is lower when the principal proposes an exclusionary contract than when the agent does.

However, the principal's use of higher exclusion probabilities also implies — perhaps surprisingly — that he is actually *less* likely to propose an exclusionary contract in the first place. To see this, consider parameter values under which the agent's gain from his most-preferred exclusionary contract is small. Now suppose principal $P_1$ makes the offer instead of the agent. Moving from the agent's most-preferred non-exclusionary contract to $P_1$'s most-preferred non-exclusionary contract gives $P_1$ more of the surplus, but the combined surplus itself is unaffected. On the other hand, moving from the agent's most preferred exclusionary contract to $P_1$'s most preferred exclusionary contract lowers the combined surplus because the exclusionary probability is increased. Hence, the increase in principal $P_1$'s utility (relative to when the agent makes the offer) is greater under non-exclusionary contracts, implying that he will offer these contracts in circumstances where the agent would not.

**Proposition 2.** ***(Agent's greater propensity to propose exclusionary punishments)***
*The agent proposes an exclusionary contract under (strictly) more circumstances that does principal $P_1$.*

## 4. Non-exclusionary punishments

Section 3 considers only exclusionary punishments. In practice, agents can also be punished in non-exclusionary ways. As discussed in the introduction, legal systems place far fewer constraints on the use of collateral to provide incentives. Indeed, the enforceability of collateral seizure is widely believed to be important for financial development and overall economic performance.

This sharp contrast between the legal treatment of different forms of punishment becomes even more surprising once one considers that collateral seizure may itself generate exclusionary-like effects (see, e.g., Ayotte 2007). Consider the case in which the agent has enough collateral to sign an incentive contract even in the second period. Then if the first principal seizes this collateral, an incentive contract is no longer possible in the second period. The agent is effectively "excluded" from working in an incentivized way.

In this section, we extend our basic model to allow for collateral seizure, and give two distinct rationales for the contrasting legal treatment of collateral seizure relative to more directly exclusionary punishments.

Formally, suppose the agent has collateral (a house, for example) with value $K$ to the agent. The collateral is non-divisible, and both principals $P_1$ and $P_2$ can impose randomized seizure of the house as a punishment for low output. Let $\chi_1$ and $\chi_2$ be the probabilities of seizure by the two principals respectively. For expositional ease, we focus on the case in which the agent has all the bargaining power in period 1 (this has little effect on the results). We also assume that $K$ is of no value to the principals. This assumption is made solely to make the punishment $K$ as closely comparable to the exclusionary punishment as possible, and relaxing this assumption would actually strengthen our results.

Our first observation is that it is always suboptimal for a jurisdiction to ban both collateral seizure and exclusionary punishments. While banning collateral seizure eliminates the externality in period 1, it does so at the cost of preventing collateral from providing incentives in period 2. In this sense, a total ban on collateral seizure is an instance of "cutting off the nose to spite the face."

**Proposition 3.** ***(No total ban on seizing collateral and exclusionary punishments)***
*It is a suboptimal policy to ban principals both from seizing the agent's collateral $K$, and from using exclusionary punishments.*

In our model, a lawmaker might still like to ban collateral seizure by principal $P_1$, while continuing to allow it for principal $P_2$. In practice, however, it may be difficult for a legal system to identify whether or not a particular principal is $P_1$ or $P_2$.

Proposition 3 leaves open the possibility of banning collateral seizure but allowing exclusionary punishments. Although the ban on collateral seizure removes the possibility of providing incentives in period 2, the tolerance of exclusionary punishments would allow incentives to be provided in period 1. Consequently, this legal regime might be attractive if collateral seizure imposed much greater social costs than exclusionary punishments.

However, our second argument in this section is that for many types of exclusionary punishments the ordering of social costs is just the reverse, i.e., exclusionary punishments are *more* socially costly than collateral seizure. Whenever this is the case, Proposition 3 implies that it never makes sense to place a blanket prohibition on collateral seizure.

The argument for why exclusionary punishments are more socially costly than collateral seizure is as follows. When principal $P_1$ provides the agent with incentives, he does so by threatening to impose a punishment (collateral seizure or exclusion) on the agent with some probability ($\chi_1$ and $\pi$ respectively). When the agent proposes the contract, and regardless of the punishment device used, he proposes $x_1 = H_1 - \frac{1}{q_1}$ so that principal $P_1$'s individual rationality constraint holds with equality. The probability with which the punishment is imposed after failure is then determined by the agent's incentive constraint, i.e.,

$$H_1 - \frac{1}{q_1} + (\text{cost of punishment to agent}) \times (\text{prob. of punishment}) = \frac{B_1}{\Delta_1}.$$

The equilibrium social cost can thus be written as

$$(1 - q_1) \left( \frac{\text{social cost of punishment}}{\text{cost of punishment to agent}} \right) \left( \frac{B_1}{\Delta_1} - \left( H_1 - \frac{1}{q_1} \right) \right).$$

Consequently, the social cost of providing incentives is minimized when the ratio of the agent's disutility from the punishment to its social cost is maximized.

Our next result, Proposition 4, establishes that in many situations this key ratio of the private cost of punishment to the agent to its social cost is greater for collateral seizure than for exclusion. The intuition is that, on the one hand, if the agent is excluded his private cost is approximately a fraction $\theta$ (his bargaining power) of the full social cost. But on the other hand, if the agent's collateral is seized, there are two effects. The direct consequence is that the agent loses his collateral. Clearly, he bears the full social cost of this loss. The indirect effect is that, as discussed, the loss of collateral eliminates the possibility of incentivizing the agent in period 2, and the agent suffers approximately a fraction $\theta$ of the consequent reduction in social surplus. Considering the two effects together, the agent bears more than a fraction $\theta$ of the social cost of collateral seizure.

These arguments are approximate because the actual division of social surplus depends on the outcome of the bargaining game, which is in turn affected by the agent's outside option of $w + K$ (or just $w$ if he has lost the collateral). As $w$ increases, the agent's share of lost social surplus increases beyond $\theta$, and does so at different rates for the two punishments. Consequently, it is possible for large values of $w$ to overturn the economic reasoning described above for why collateral seizure is a more efficient punishment. Formally:

**Proposition 4.** ***(Social loss from exclusionary punishments is worse)*** *The social loss of using the exclusionary punishment to incentivize the agent in period* 1 *is greater than the social loss of using collateral seizure whenever the agent's outside option $w$ is sufficiently low.*

Proposition 4 clearly relies on collateral seizure imposing a significant direct cost on the agent (namely, the loss of collateral), while the exclusionary contract hurts the agent only through restricted contracting opportunities in the future. This assumption is a good fit for exclusionary punishments such as non-compete clauses and debt-bondage. On the other hand, it applies less well to imprisonment (depending largely on the actual terms of imprisonment), and is a poor fit for corporal punishment. However, even for this last case collateral seizure may still be the more efficient punishment in cases where principals place some value on the agent's collateral (but take no direct pleasure from physically punishing the agent).

## 5. Overconfidence

An alternate motivation for restricting the use of punishment clauses in contracts is that individuals are poor at estimating the probability that they will be applied. For example, McCormick (1935, p. 601) writes:

> It is a characteristic of men, however, that they are likely to be beguiled by the "illusions of hope," and so feel so certain of their ability to carry out their engagements in future, that their confidence leads them to be willing to make extravagant promises and commitments as to what they are willing to suffer if they fail.

Many recent studies support this claim, by showing that individuals systematically overestimate their skill.[14]

We consider an agent who is overconfident, in the sense that he overestimates his success probability by $\varepsilon > 0$. That is, he believes his success probability is $q_t + \varepsilon$ if he works, and $p_t + \varepsilon$ if he shirks. Principals $P_1$ and $P_2$ continue to correctly perceive the success probabilities. To focus on the effects of overconfidence, we consider the case in which the agent has all the bargaining power in both periods 1 and 2.

In order to compensate principal $P_1$, in an incentive contract the agent must offer to pay him $1/q_1$ when he succeeds, while in a non-incentive contract he must offer to pay $1/p_1$. (Note that the agent is aware that he is optimistic relative to principal $P_1$.)[15] Write $\hat{U}_2$ for the overconfident agent's expectation of period 2 utility.

[14]To give just one (well-known) example, individuals systematically overestimate their driving ability — see Svenson (1981).

[15]Concretely, one might imagine the agent slowly increasing the amount he offers the principal until the principal accepts.

Consequently, an overconfident agent evaluates his expected utility from an incentive contract as

$$\begin{aligned}\hat{U}^{exc} &= (q_1+\varepsilon)\left(H_1-\frac{1}{q_1}\right)-B_1+(1-(1-(q_1+\varepsilon))\,\pi)\,\hat{U}_2 \\ &= U^{exc}+\varepsilon\left(H_1-\frac{1}{q_1}\right)+\varepsilon\pi\hat{U}_2+(1-(1-q_1)\,\pi)\left(\hat{U}_2-U_2\right),\end{aligned}$$

where $U^{exc}$ is the agent's "true" expected utility. Observe that the agent makes three mistakes in evaluating his expected utility, represented by the last three terms in the above equation: he overestimates his success probability, he underestimates the probability of exclusion, and he overestimates his expected utility when not excluded.

One immediate observation is that overconfidence lowers the exclusion probability proposed by the agent: the agent's perceived utility loss from exclusion, $\pi\hat{U}_2$, is determined by setting his incentive constraint ($\mathrm{IC}_1$) to equality, and since overconfidence raises $\hat{U}_2$, it lowers the exclusion probability $\pi$.[16]

We conclude with an even stronger result: there are circumstances in which an overconfident agent abstains from an exclusionary contract that would actually improve his welfare. To see this, note that an overconfident agent evaluates his expected utility from a non-exclusionary contract as

$$\hat{U}^{no-exc}=(p_1+\varepsilon)\left(H_1-\frac{1}{p_1}\right)+\hat{U}_2=U^{no-exc}+\varepsilon\left(H_1-\frac{1}{p_1}\right)+\left(\hat{U}_2-U_2\right).$$

Consequently, the agent evaluates the gain to using an incentive contract as

$$\hat{U}^{exc}-\hat{U}^{no-exc}=U^{exc}-U^{no-exc}+\varepsilon\pi\hat{U}_2-(1-q_1)\,\pi\left(\hat{U}_2-U_2\right)+\frac{\varepsilon\Delta_1}{q_1p_1},$$

where $\pi$ is the exclusion probability that the overconfident agent chooses if he wishes to implement high effort. The term $\varepsilon\pi\hat{U}_2$ relates to the fear of many observers,

[16] Observe that if we interpret overconfidence instead as an overestimation of the effects of one's own actions, so that $q_1$ is replaced by $q_1+\epsilon$ but $p_1$ is unchanged, then the increase in $U_2$ would be replaced by an increase in $\Delta_1$, leading to the same conclusion. Similarly, an analogue of Proposition 5 below holds under this alternative specification.

namely that overconfident agents overuse exclusionary punishments because they underestimate the probability of exclusion. However, the penultimate term captures a countervailing effect: overconfident agents also overestimate their utility from not being excluded, and this makes them overly hesitant to use exclusionary contracts. Finally, the last term reflects the fact that the overconfident agent thinks he is paying the principal too much in both the incentive and non-incentive contract.

Although the net impact of these three effects is hard to sign, it is important to note that there are circumstances in which the net effect is negative. In such cases, overconfidence leads the agent to avoid exclusionary contracts that would actually increase his objective expected utility. The existence of such cases suggests that overconfidence may be a less compelling justification for restricting exclusionary punishments than previously thought.

**Proposition 5.** ***(Overconfident agents may underuse exclusionary punishments)***
*There are circumstances under which an overconfident agent underestimates the value of an exclusionary contract, i.e.,* $\hat{U}^{exc} - \hat{U}^{no-exc} < U^{exc} - U^{no-exc}$.

As detailed in the proof of Proposition 5, the agent potentially underuses the exclusionary contract if the probability of period 1 failure $1 - q_1$, and hence exclusion is sufficiently high; and the degree of overconfidence $\varepsilon$ is small.[17]

## 6. Concluding remarks

Legal systems place severe constraints on what types of punishments contracting parties can agree to. Many forms of potentially incentive-improving punishments are banned outright (e.g., debt bondage), while others are tightly restricted (e.g.,

[17] Absent overconfidence, i.e., $\varepsilon = 0$, the agent uses exclusionary contracts in the socially optimal way: recall that in this section we assume that the agent has all the bargaining power in both periods.

non-compete clauses). Yet while restrictions on exclusionary punishments of this type receive little attention, a large and growing body of opinion posits that effective enforcement of collateral seizure is key to financial development, and economic growth more generally.

This paper argues that a simple contracting model can be used to make sense of this asymmetric treatment of different types of penalties. Exclusionary punishments reduce the future welfare not just of the individual actually excluded, but also of future trading partners. Contracting parties have no reason to internalize these externalities, and so regulation to restrict the use of exclusionary punishments may be welfare improving. In contrast, although collateral seizure potentially has exclusion-like effects, by reducing the incentives that can be provided in the future, banning collateral seizure merely ensures that collateral cannot be used to provide incentives at all. In this sense, banning collateral would be an instance of "cutting off the nose to spite the face." Moreover, the fraction of the social welfare loss that is not internalized by contracting parties is generally smaller in the case of collateral seizure than explicit exclusion.

To some extent, one can also justify restrictions on exclusionary punishments as existing to protect overconfident individuals. However, our analysis suggests that a caveat is worth noting: there are circumstances in which overconfident agents *under*use rather than overuse exclusionary punishments. Moreover, to the extent to which bans on exclusionary punishments are motivated by protecting overconfident individuals, our arguments of Section 4 are still relevant in explaining why collateral seizure is not similarly restricted.

Inevitably our analysis omits some important issues. Two deserve particular mention. First, in many situations an individual who "fails" (e.g., a defaulting debtor, or a poorly performing employee, etc.) suffers a loss of reputation. This lost reputation has exclusionary-like effects, in that it makes it harder for the individual to contract in the future. Consequently, our basic argument suggests that under some circumstances, individuals may be *too incentivized* from the standpoint of social efficiency — somewhat contrary to conventional wisdom.[18] Relative to the explicit forms of exclusion that are our main focus in the paper, exclusion via lost reputation seems harder to legislate against. That said, an important exception in many countries is the ceiling on the number of years that credit bureaus are allowed to report personal bankruptcy filings (see, e.g., Elul and Gottardi 2008 and Musto 2004).

Second, we have ignored the extent to which the outputs of different individuals are correlated. Consider an economy populated by a large number of agents, with projects that are highly correlated. If exclusionary contracts are used in such an economy, the supply of agents will be substantially reduced when projects produce low output. As a consequence, the bargaining power of any non-excluded agent will be raised. This effect means that, in equilibrium, at least some agents will refrain from writing socially inefficient exclusionary contracts in period 1, since given their high bargaining power in period 2 they suffer most of the cost. We leave a fuller analysis of the equilibrium outcomes of such an economy for future research.

Aghion, P. and P. Bolton, 1987. Contracts as a barrier to entry. *American Economic Review* 77, 388-401.

Aghion, P. and B. Hermalin, 1990. Legal restrictions on private contracts can enhance efficiency. *Journal of Law, Economics and Organization* 6, 381-409.

Ayotte, K., 2007. Bankruptcy and entrepreneurship: the value of a fresh start. *Journal of Law, Economics and Organization* 23, 161-185.

[18] A complete modeling of this case lies beyond the scope of the current paper. The main complication is that modeling reputation requires the introduction of an adverse-selection problem (in addition to moral hazard). Elul and Gottardi (2008) develop a model along these lines, and show that it is often optimal to mandate the "forgetting" of failure. The assumptions of their model imply that principals (lenders in their model) always make zero profits, so the motivation for this result is different from that in our paper.

Banerjee, A. V. and A. F. Newman 1994. Poverty, incentives, and development. *American Economic Review* 84, 211-215.

Camerer, C. , S. Issacharoff, G. Lowenstein, T. ODonoghue, and M. Rabin, 2003. Regulation for conservatives: Behavioral economics and the case for asymmetric paternalism. *University of Pennsylvania Law Review* 151, 1211-1254.

Chung, T. Y., 1992. On the social optimality of liquidated damage clauses: An economic analysis. *Journal of Law, Economics, and Organization* 8, 280-305.

Chwe, M. S. Y., 1990. Why were workers whipped? Pain in a principal-agent model. *Economic Journal* 100, 1109-1121.

de Soto, H., 2000. The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else. Basic Books.

Diamond, P. A. and E. Maskin, 1979. An equilibrium analysis of search and breach of contract, i: Steady states. *Bell Journal of Economics* 10, 282-316.

Djankov, S., O. Hart, C. McLiesh, and A. Shleifer, 2008. Debt enforcement around the world. *Journal of Political Economy* 116, 1105-1149.

Elul R. , and P. Gottardi, 2008. Bankruptcy: Is it enough to forgive or must we also forget? *Federal Reserve Bank of Philadelphia Working Paper.*

Garmaise, M., 2006. Ties that truly bind: Non-competition agreements, executive compensation and firm investment. *University of California, Los Angeles Working Paper.*

Gervais, S. and I. Goldstein, 2007. The positive effects of biased self-perceptions in firms. *Review of Finance* 11, 453-496.

Hackbarth, D., 2008. Managerial traits and capital structure decisions. *Journal of Financial and Quantitative Analysis* 43, 843-882.

Hart, O., 1995. Firms, Contracts and Financial Structure. Oxford University Press.

Hynes, R., 2004. Overoptimism and overborrowing. *Brigham Young University Law Review*, 127-168.

Landier A. and D. Thesmar, 2009. Financial contracting with optimistic entrepreneurs. *Review of Financial Studies* 22, 117-150.

Levine, R., 1997. Financial development and economic growth: views and agenda. *Journal of Economic Literature* 35, 688-726.

Levine, R., 2006. Finance and growth: Theory and evidence. In: Aghion, P. and S. Durlauf (Eds.), Handbook of Economic Growth, volume 1, Elsevier North-Holland, New York, pp. 865-934.

Mann, B. H., 2002. Republic of Debtors: Bankruptcy in the Age of American Independence. Harvard University.

McCormick, C. T., 1935. Handbook on the Law of Damages. West.

Mirrlees, J. A., 1999. The theory of moral hazard and unobservable behaviour: Part I. *Review of Economic Studies* 66, 3-21.

Mookherjee, D. and U. Lilienfeld-Toal, 2008a. A general equilibrium analysis of personal bankruptcy law. *Boston University Working Paper.*

Mookherjee, D. and U. Lilienfeld-Toal, 2008b. The political economy of debt bondage. *Boston University Working Paper.*

Mookherjee, D. and I. Png, 1989. Optimal auditing, insurance and redistribution. *Quarterly Journal of Economics* 104, 399-415.

Musto, D. K., 2004. What happens when information leaves a market? Evidence from post-bankruptcy consumers. *Journal of Business* 77, 725-749.

Pande, R. and C. Udry, 2007. Institutions and development: A view from below. In: Blundell, R., W. Newey, and T. Persson (Eds.), Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress, volume 2. Cambridge University Press.

Samila, S. and Sorenson, 0., 2009. Non-compete covenants: incentives to innovate or impediments to growth. *Brock University Working Paper.*

Saxenian, A., 1996. Regional Advantage: Culture and Competition in Silicon Valley and Route 128. Harvard University Press.

Spier, K. E., 1992. Incomplete contracts and signalling. *RAND Journal of Economics* 23, 432-443.

Spier, K. E., 1994. Settlement bargaining and the design of damage awards. *Journal of Law, Economics and Organization* 10, 84-95.

Spier, K. E., and M. D. Whinston, 1995. On the efficiency of privately stipulated damages for breach of contract: entry, reliance, and renegotiation. *RAND Journal of Economics* 26, 180-202.

Svenson, O., 1981. Are we all less risky and more skillful than our fellow drivers? *Acta Pschologica* 47, 143-148.

## 7. Mathematical proofs

**Proof of Proposition 1:** If the agent proposes an exclusionary contract with enough incentives to work, he sets $x_1$ as high as possible while satisfying principal $P_1$'s individual rationality constraint, ($\mathrm{P}_t$-IR-W), i.e.,

$$x_1 = H_1 - \frac{1}{q_1},$$

and the probability of exclusion as low as possible while satisfying his own incentive constraint ($\mathrm{IC}_1$), i.e.,

$$\pi U_2 + x_1 = \frac{B_1}{\Delta_1}.$$

So the exclusion probability $\pi_A$ solves

$$\pi_A U_2 = \frac{B_1}{\Delta_1} - H_1 + \frac{1}{q_1},$$

and is strictly positive by assumption (1). For the Pareto frontier, observe that to increase principal $P_1$'s utility by 1 requires a reduction in $x_1$ of $\frac{1}{q_1}$, and hence an increase in the exclusion probability $\pi$ of $\frac{1}{q_1 U_2}$ in order to maintain incentives. Hence the combined utility of the agent and principal $P_1$ is changed by $-\frac{1-q_1}{q_1 U_2} U_2$.

Since $P_1$'s utility is increased by 1, the agent's utility is changed by $-1 - \frac{1-q_1}{q_1} = -\frac{1}{q_1}$. Finally, to calculate $\bar{V}$, the principal's maximal obtainable utility from an exclusionary contract, observe that the exclusion probability is bounded above by 1. Since this is $1 - \pi_A$ above $\pi_A$, the corresponding value of $x_1$ is $H_1 - \frac{1}{q_1} - U_2 (1 - \pi_A) = \frac{B_1}{\Delta_1} - U_2$ (this is positive since, by (1), $q_1 \frac{B_1}{\Delta_1} > q_1 H_1 - 1$, which certainly exceeds $U_2$). The corresponding value of the agent's utility is simply $U = q_1 \left( \left( \frac{B_1}{\Delta_1} - U_2 \right) + U_2 \right) - B_1 = \frac{p_1 B_1}{\Delta_1}$. Hence principal $P_1$'s maximal obtainable utility is $\bar{V} = 1 + q_1 (U^{exc} - U) = 1 + q_1 \left( U^{exc} - \frac{p_1 B_1}{\Delta_1} \right)$. ∎

**Proof of Proposition of 2:** From Proposition 1, principal $P_1$'s utility from proposing an exclusionary contract is $V^{exc} \equiv \min \left\{ \bar{V}, 1 + q \left( U^{exc} - (w + U_2) \right) \right\}$, depending on whether or not the agent's individual rationality constraint binds. In comparison, $P_1$'s utility without an exclusionary contract is from a non-exclusionary contract is $V^{no-exc} \equiv \max \{ p_1 H_1 - w, 1 \} = 1 + U^{no-exc} - (w + U_2)$. So

$$\begin{aligned} V^{exc} - V^{no-exc} &\leq q \left( U^{exc} - (w + U_2) \right) - \left( U^{no-exc} - (w + U_2) \right) \\ &= q \left( U^{exc} - U^{no-exc} \right) - (1 - q) \left( U^{no-exc} - (w + U_2) \right). \end{aligned}$$

The result follows since $U^{no-exc} \geq w + U_2$. ∎

**Proof of Proposition 3:** The proofs of Propositions 3 and 4 make use of the following notation. Let $K + S_2^K$ be the expected social surplus in period 2 when the agent has collateral. Let $K + U_2^K$ be the agent's expected period 2 utility when he has the collateral. Similarly, let $K + S_2^{A,K}$ and $K + U_2^{A,K}$ be period 2 social surplus and agent utility when the agent makes the offer in period 2; and $K + S_2^{P,K}$ and $K + U_2^{P,K}$ be period 2 social surplus and agent utility when the principal $P_2$ makes the offer in period 2. Finally, let $S_2^{-K}$ etc. denote the corresponding quantities when the agent does not have the collateral in period 2.

We consider the case in which collateral seizure is allowed, but exclusionary contracts are banned. We calculate total social welfare in this case, and show that it exceeds total social welfare if both types of punishment are banned.

If the agent does not propose collateral seizure in one period, he will not do so in the other; and so neither will principal $P_2$.[19] In this case, a ban on collateral seizure has no impact. If instead the agent proposes collateral seizure in one period, he will do so in both. In this case, total social surplus is

$$\begin{aligned} & q_1 H_1 - B_1 + K + S_2^K - (1-q_1)\,\chi_1\left(K + S_2^K - S_2^{-K}\right) \\ = \; & q_1 H_1 - B_1 + K + S_2^K \\ & - (1-q_1)\,\chi_1\left(K + U_2^K - U_2^{-K}\right) \\ & - (1-q_1)\,\chi_1\left(\left(S_2^K - S_2^{-K}\right) - \left(U_2^K - U_2^{-K}\right)\right). \end{aligned} \tag{3}$$

Here, we have decomposed the expected social loss of imposing the punishment, i.e. $(1-q_1)\,\chi_1\left(K + S_2^K - S_2^{-K}\right)$, into the cost inflicted on the agent, $(1-q_1)\,\chi_1\left(K + U_2^K - U_2^{-K}\right)$, and the remainder. The cost experienced by the agent must be less than the agent's expected gain from using the incentive contract, which equals

$$\begin{aligned} & q_1 H_1 - B_1 - 1 - \max\{p_1 H_1 - 1, w\} \\ = \; & \min\{\Delta_1 H_1 - B_1, q_1 H_1 - B_1 - 1 - w\}. \end{aligned}$$

So expression (3) exceeds

$$\begin{aligned} & q_1 H_1 - B_1 + K + S_2^K \\ & - \min\{\Delta_1 H_1 - B_1, q_1 H_1 - B_1 - 1 - w\} - \left(S_2^K - S_2^{-K}\right) \\ = \; & \max\{p_1 H_1, 1 + w\} + K + S_2^{-K}, \end{aligned}$$

which is exactly the social surplus available if both possible punishments are banned.
■

[19] We are assuming that in period 1 the agent always proposes the contract. However, if the agent does not propose collateral seizure in period 1, neither would principal $P_1$.

**Proof of Proposition 4:** We use the notation defined at the start of the proof of Proposition 3. When the punishment threatened by principal $P_1$ is exclusion in period 2,

$$\frac{\text{cost of punishment to agent}}{\text{social cost of punishment}} = \frac{U_2^K}{S_2^K - 1}, \tag{4}$$

while when the punishment threatened by principal $P_1$ is collateral seizure,

$$\frac{\text{cost of punishment to agent}}{\text{social cost of punishment}} = \frac{K + U_2^K - U_2^{-K}}{K + S_2^K - S_2^{-K}}. \tag{5}$$

As a preliminary, note that the result is very straightforward if $S_2^K = S_2^{-K}$. In this case, collateral plays no role in the period 2 contracts, and so $U_2^K = U_2^{-K}$. So expression (5) equals 1, while expression (4) is (weakly) lower since $U_2^K \leq S_2^K - 1$.

The bulk of the proof consists of showing that if $S_2^K > S_2^{-K} > 1 + w$ and $w = 0$ then expression (5) strictly exceeds expression (4). By continuity, the same is true for all $w$ sufficiently small. We will show that

$$\frac{U_2^K - U_2^{-K}}{S_2^K - S_2^{-K}} > \frac{U_2^K}{S_2^K - 1} \tag{6}$$

and

$$\frac{x + U_2^K - U_2^{-K}}{x + S_2^K - S_2^{-K}} \text{ is weakly increasing in } s. \tag{7}$$

To establish (6), note that by straightforward algebra the inequality is equivalent to

$$\frac{U_2^K}{S_2^K - 1} > \frac{U_2^{-K}}{S_2^{-K} - 1}, \tag{8}$$

and hence, in turn, to

$$\frac{\theta_2 U_2^{A,K} + (1 - \theta\theta_2)\, U_2^{P,K}}{S_2^K - 1} > \frac{\theta_2 U_2^{A,-K} + (1 - \theta_2)\, U_2^{P,-K}}{S_2^{-K} - 1}.$$

When the agent makes the offer he is able to capture the entire social surplus other than principal $P_2$'s outside option, 1: hence $U_2^{A,K} = S_2^{A,K} - 1$ and $U_2^{A,-K} = S_2^{A,-K} - 1 = S_2^{-K} - 1$ (since without collateral, incentive contracts are impossible and social surplus is independent of who makes the offer). When the agent has lost his collateral, and the principal $P_2$ makes the offer, he can capture the entire social surplus, so $U_2^{P,-K} = 0$ (recall that $w = 0$). Hence (6) is in turn equivalent to

$$\theta_2 \frac{S_2^{A,K} - 1}{S_2^K - 1} + (1 - \theta_2) \frac{U_2^{P,K}}{S_2^K - 1} > \theta_2.$$

This inequality holds since $S_2^{A,K} \geq S_2^K$ and $U_2^{P,K} \geq 0$, with at least one these relations strict. To see this, observe that since $S_2^K > S_2^{-K}$ at least one of the parties must propose an incentive contract. It is straightforward to show that whenever principal $P_2$ proposes an incentive contract then the agent does also. If the agent proposes an incentive contract but principal $P_2$ does not, then $S_2^{A,K} > S_2^{-K} = S_2^{P,K}$, and so $S_2^{A,K} > S_2^K$. Finally, if both the agent and principal $P_2$ propose incentive contracts, then the agent's proposed contract uses less collateral seizure, and so $S_2^{A,K} \geq S_2^{P,K}$. Moreover, the inequality is strict unless $U_2^{P,K} = U_2^{A,K}$. So either $S_2^{A,K} > S_2^K$ or $U_2^{P,K} > 0$.

To establish (7), note that it is equivalent to

$$U_2^K - U_2^{-K} \leq S_2^K - S_2^{-K},$$

which (using the above expressions for $U_2^{A,K}$ etc.) is in turn equivalent to

$$\theta_2 \left( \left( S_2^{A,K} - 1 \right) - \left( S_2^{-K} - 1 \right) \right) + (1 - \theta_2) U_2^{P,K} \leq S_2^K - S_2^{-K},$$

i.e.,

$$(1 - \theta_2) U_2^{P,K} \leq (1 - \theta_2) \left( S_2^{P,K} - S_2^{-K} \right),$$

i.e.,

$$S_2^{-K} \leq S_2^{P,K} - U_2^{P,K}.$$

The right hand side is what the principal $P_2$ gets when he makes an offer and the agent has collateral. This must weakly exceed what the principal could get from instead ignoring the presence of collateral, namely $S_2^{-K}$ (since $w=0$, principal $P_2$ is able to capture the entire social surplus). Hence (7) holds.

Finally, we deal with the remaining case of $S_2^K > S_2^{-K} = 1+w$. In this case, principal $P_2$ and the agent do not contract at all without collateral, and so $U_2^{-K} = w$ and expression (5) equals $\frac{K-w+U_2^K}{K-w+S_2^K-1}$. Since $U_2^K \leq S_2^K - 1$ this is weakly greater than expression (4) for $w \leq K$. ∎

**Proof of Proposition 5:** From the main text,

$$\frac{\left(\hat{U}^{exc} - \hat{U}^{no-exc}\right) - \left(U^{exc} - U^{no-exc}\right)}{\varepsilon} = \pi \hat{U}_2 \left(1 - (1-q_1)\left(\frac{1}{\varepsilon}\left(1 - \frac{U_2}{\hat{U}_2}\right)\right)\right) + \frac{\Delta_1}{q_1 p_1}.$$

Since the agent makes the offer, $\pi \hat{U}_2 = \pi_A U_2$, where $\pi_A$ is as defined in Proposition 1. Substituting in,

$$\frac{\left(\hat{U}^{exc} - \hat{U}^{no-exc}\right) - \left(U^{exc} - U^{no-exc}\right)}{\varepsilon} = \left(\frac{B_1}{\Delta_1} - H_1 + \frac{1}{q_1}\right)\left(1 - (1-q_1)\left(\frac{1}{\varepsilon}\left(1 - \frac{U_2}{\hat{U}_2}\right)\right)\right) + \frac{\Delta_1}{q_1 p_1}.$$

To exhibit circumstances in which this is negative, i.e., the agent underuses exclusionary punishments, consider the case in which $p_2 H_2 > 1$ so that $U_2 = p_2 H_2 - 1$ and $\hat{U}_2 = (p_2 + \varepsilon)\left(H_2 - \frac{1}{p_2}\right) = \frac{p_2+\varepsilon}{p_2} U_2$. Then

$$\frac{\left(\hat{U}^{exc} - \hat{U}^{no-exc}\right) - \left(U^{exc} - U^{no-exc}\right)}{\varepsilon} = \left(\frac{B_1}{\Delta_1} - H_1 + \frac{1}{q_1}\right)\left(1 - \frac{1-q_1}{p_2+\varepsilon}\right) + \frac{\Delta_1}{q_1 p_1}.$$

This is potentially negative if $p_2 + \varepsilon < 1 - q_1$. In particular, it is negative if $p_2 + \varepsilon$ is sufficiently small, with $H_2$ sufficiently large to ensure that $p_2 H_2 > 1$. ∎