

UC Berkeley

Berkeley Program in Law and Economics, Working Paper Series

Title

The Conspirator Dilemma: Introducing the “Trojan Horse” Enforcement Strategy

Permalink

<https://escholarship.org/uc/item/98z174x0>

Author

Yadlin, Omri

Publication Date

2002-04-02

School of Law, Boalt Hall

U.C. Berkeley Law and Economics

Working Paper Series

Working Paper 2001 - 2

***The Conspirator Dilemma:
Introducing the “Trojan Horse” Enforcement Strategy***

Omri Yadlin

Tel Aviv University, Law School

School of Law
Center for the Study of Law and Society
University of California
Berkeley, California 94720

The Conspirator Dilemma:
Introducing the “Trojan Horse” Enforcement Strategy

Omri Yadlin*

I. INTRODUCTION

Scholars debate the optimal division of labor between public and private law enforcement. Pioneers in this area of scholarship were *Becker & Stigler* who called for the privatization of law enforcement.¹ According to their proposal, private entities would collect fines from outlaws on a *first to file collects the fine* basis. In response, *Landes & Posner* pointed at certain failures the market for private law enforcement suffers from.² They also questioned the novelty of *Becker & Stigler’s* proposal, noting that, in practice, the law enforcement enterprise already occupies bounty hunters, class action representatives and other legal entrepreneurs. On that basis, *Landes and Posner* concluded that there is no reason to doubt the efficiency of the current separation of responsibilities between the public and the private sectors of law enforcement.

This paper offers a different kind of enforcement mechanism, one that relies neither on public agencies or victims, nor on bounty hunters or similar private agents. The best enforcers of the law, goes my argument, are the most informed parties, namely, the law-breakers themselves. Thus, the mechanism I offer here entrusts the first felon filing a civil suit with a right to collect a portion of the fine from his accomplices; the

* Tel Aviv University, Law School. I have benefited from discussions with Malcolm Feeley, Ariel Porat, Alon Harel, Dan Rubinfeld, Jesse Fried, Omri Ben-Shahar, Sanford kadish, Ehud Kamar, Eric Talley, Ron Shapira. While developing my paper I learned that Bob Cooter and Nuno Garupa were independently developing a similar idea with different emphasis. While my paper addresses conspiracies in general, Cooter and Garupa focus on corruption of officials. We have exchanged drafts and benefited from discussions with each other.

¹ Becker & Stigler, Law Enforcement, Malfeasance, and Compensation of Enforcers, 3 J. Legal Stud. 1 (1974).

² Landes & Posner, The Private Enforcement of Law, 4 J. legal stud. 1 (1975).

rest of the fine will go to the government. Whereas public enforcers, in the course of investigating a case, may subject suspects to a prisoner dilemma by offering each one of them immunity for testifying against the others, the mechanism I offer here positions lawbreakers in the “conspirator dilemma” at the very minute they enter the illegal scheme. Each felon faces the risk that her partner would defect, and thus each partner’s best strategy must be to act as a *Trojan Horse* and collect the fine first. Hence, unlike public and private enforcement agencies, that typically start investigating a crime only after its completion, the *Trojan Horse* conspirator would sue his partners at the preliminary stages of the crime, often before any harm was done.³ The mechanism I offer, therefore, decreases the level of trust among criminals and increases their cost of cooperation. Thus it may serve as an effective deterrence apparatus against such conspiracies.

As shown in this paper, in some contexts the *Trojan Horse* mechanism can actually replace public enforcement. Thus, the *Trojan Horse* mechanism might allow us to decriminalize certain offences. In most cases, however, the *Trojan Horse* mechanism would only supplement the traditional enforcement mechanisms. In such cases, in addition to collecting the fine, a felon who wins a *Trojan Horse* suit against his partner would be immune from government’s prosecution.⁴ Such a suit would not

³ In a way, my point is motivated by the same vision that inspired Kaplow and Shavell’s paper on “Optimal Enforcement with Self-Reporting of Behavior” 102 J. Pol. Econ. 583 (1994), namely, that the law should be designed to induce wrongdoers to expose their own wrongdoing. However, the mechanisms differ in some very substantial ways. First, whereas self-reporting mechanisms are designed to deal with no-fault liability, the *Trojan Horse* mechanism mainly addresses intentional violations of the law. Second, self-reporting is an ex post mechanism in the sense that it induces lawbreakers to report themselves after they realize they have broke the law. The goal of the *Trojan Horse* mechanism, on the other hand, is to talk conspirators out of their plan ex ante.

⁴ This paper raises of course difficult legal questions. For example, if the two mechanisms, the public and the private one we offer here, operate concurrently, there is a problem of double jeopardy – arguably, the defendant would be immune from public enforcement even if he loses the case. This is clearly not my intention. One way to deal with this problem is to define the proceedings between the conspirators as civil litigation. Alternatively, *Trojan Horses* would operate like bounty hunters, i.e., they would only be allowed to file a complaint with a regulator, and the regulator will decide whether

stop the government from prosecuting the defendant or from running after a *Trojan Horse* who failed to prove his case.⁵

The following two examples demonstrate the way the Trojan Horse mechanism operates:

First Example: Price Fixing

Consider first the case of price fixing arrangements. Currently, public agencies and private entities work in tandem to enforce the antitrust laws. But both of these enforcement mechanisms are seldom privy to information about the formation and execution of anti-competitive agreements. Hence, some price-fixing arrangements escape sanctions altogether; others are detected only long after initiation, often after the cartel breaks apart.

Compared with any public or private enforcer, cartel members possess far more information about the anti-competitive nature of their conduct. The *first-felon-to-file-collects-the-fine* rule takes advantage of this comparative advantage, and situates cartel members in the Conspirator Dilemma. Since each member of the cartel faces the risk of being sued by the other, the best strategy for each member must be to defect first and file a suit as soon as sufficient evidence to support the case is gathered. Hence, the *Trojan Horse* mechanism raises significantly the cost of entering such an anti-competitive agreement.

Second Example: Illegal Employment

Another good example is the import and the employment of illegal immigrants. A large segment of the workforce in many industrial countries consists of illegal immigrants. Typically, the laws of these countries subject illegal immigrants and their employers to significant sanctions. The fact that these fines do not deter employers from hiring illegal immigrants suggests that public enforcement is ineffective. One explanation for this enforcement failure is that enforcement agencies are uninformed. Another explanation is that enforcement agencies are captured by interest groups that persuade the public enforcement agency to keep a high profile in the media but a low presence in the sweatshops.

Arguably, private agencies, in line with *Becker and Stigler's* proposal, are less influenced by interest groups and more informed than public agencies, and therefore, have the potential of somewhat tightening the enforcement of these laws. But the best enforcers of these laws are probably the illegal immigrants themselves, who know their status better than any public or private enforcer does. Under my suggestion, therefore, illegal immigrants will be authorized to collect fines from their employers and/or from the

to take over the case to the court. I believe the latter option misses upon some of the most important advantages of the mechanism I offer.

⁵ To be more precise, the rule I suggest is that the plaintiff's immunity will be limited to those offences the defendant was convicted for. To determine whether the conviction covers a certain criminal conduct or not we may use similar criteria to those used under the double jeopardy doctrine.

manpower agent who matched them with their employer. If they win their case, the court would order the defendant to deposit the fine at the country's gate and the immigrant would collect it on the way out of the country.

Again, the main purpose of the *Trojan Horse* mechanism is to deter, *ex ante*, the violation of the law. Thus, it is expected that the threat of an employee suit would force employers to hire less illegal immigrants. A side benefit of this mechanism, however, is that for those illegal immigrants who were hired by a local employer, this is a much more elegant and humane enforcement mechanism than the current regulation of illegal employment. Not only does it encourage illegal immigrants to end their employment relationships soon after being hired, but it also motivates them to leave the country voluntarily.

After introducing the mechanism we are in a position to discuss its details:

In the second chapter of the paper I use a simple model to compare the operation of the Trojan Horse mechanism in finite schemes like a one-shot bribery agreement or a series of n-shot tax evasions, with its administration in infinite continuous schemes like price fixing or illegal employment.⁶

The third chapter focuses on asymmetric schemes like illegal employment. Whereas cartel members, typically, derive similar rents from price fixing and are on equal footing to meet a fine, the relationship between illegal immigrants and their employers is less symmetric.⁷ Typically, employers' ill-gotten profits are higher than those of their illegal-employees and while employers can typically meet the fine, the

⁶ Another common property cartels and illegal employment is that in both schemes it is often the case that victims of these schemes are often not aware of the fact that they were victimized. Unlike robbery victims, consumers may not know they suffer losses due to a price fixing scheme and unemployed citizens may not know that the reason for their unemployment is the fact that their ex-employer replaced them with low-cost illegal immigrants. One problem with law enforcement in these areas, therefore, is that no one complains. The *Trojan Horse* mechanism provides an incentive for someone, in this case an offender, to complain.

⁷ Another important distinction between the two schemes is that, unlike illegal employment, price fixing is an unstable conspiracy. Not only are cartel members motivated to secretly provide discounts, more importantly from our perspective, cartel members benefit from their co-conspirator's defeat. Whereas illegal employees are made better off with their employer's prosperity, cartel members are made better off when their co-conspirators go bankrupt or, even better, to jail. For unstable conspiracies like price-fixing, therefore, it might be sufficient to declare them unenforceable and to offer amnesty to any cartel member that snitches on his colleagues. Allowing cartel members to sue each other is only one further step to reinforce instability. Illegal employment, on the other hand, is a stable law violation, and therefore, amnesty would not adequately deter employers and employees from

illegal employees, for all practical matters, are judgment proof. In the third chapter I modify the model developed in the second chapter to address this asymmetry.

II. Optimal Trojan Horse Programs

In this chapter I use a simple model to demonstrate the way the Trojan Horse mechanism operates in three different contexts. In the first scenario, two (or more) conspirators engage in only one incident of lawbreaking. In the second context, two conspirators engage in a series of n transgressions. Thirdly, I examine the operation of the Trojan Horse mechanism in infinite, serial conspiracies.

The one factor I will use to compare between the efficacy of enforcement systems is the minimum fine that is required, in each system, to deter wrongdoing. Assuming risk neutrality, and given the fact that financial resources of lawbreakers are restricted, criminals who cannot meet a certain fine will not be deterred by a mechanism whose effectiveness depends on such a fine. Thus, the lower the fine that obtains deterrence, the more criminal activity the law deters.

1st. The One-Shot Conspiracy Game

Consider the case of two agents, say a homeowner name Roe and a contractor name Doe, who conspire to misreport their business to the tax authorities. Roe and Doe expect to derive from their illegal interaction a gain of $\$G$ each. The fine the public enforcement agency imposes on parties to such a conspiracy is $\$F$ and the probability the plot will be detected and successfully prosecuted is P . I account for the fact that Roe and Doe engage in a one-shot scheme by assuming that they capture the ill-gotten gain (G) and produce evidence that might incriminate them (with

colluding. The *Trojan Horse* mechanism inserts instability into stable conspiracies, an instability that may revolutionize the battles against such crimes.

probability P), only after the scheme concludes. Assuming Roe and Doe are risk neutral, they would not engage in such a scheme if:

$$(1) (1-P)G + P(G-F) < 0$$

The minimum fine that satisfies this condition is:

$$(2) F^* = G/P$$

Now suppose we replace the public enforcement agency with *Trojan Horses*, namely, lawmakers rely solely on Roe and Doe to enforce the laws they enact. Again, I assume that Roe and Doe gather sufficient evidence to support such a suit only after they have completed the tax evasion plot. At that time, the probability a *Trojan Horse's* suit will prevail is given by Q and the cost of filing such a suit by C .⁸ In case the *Trojan Horse* prevails in trial the court orders the defendant to pay a fine ($\$F$), a portion of which $\$R$, where $R=F$, is allotted to the plaintiff and the rest goes to the state.

The 2*2 table below describes the dilemma the two conspirators face immediately after they have executed their commitments under the agreement:

Table A: One-Shot Conspiracy Game with a Trojan Horse Mechanism

Roe's payoff	Roe Cooperates	Roe Sues
Doe's payoff		
Doe Cooperates	G	$Q(G+R) + (1-Q)(G) - C$ $Q(G-F) + (1-Q)G$
Doe Sues	$Q(G-F) + (1-Q)G$ $Q(G+R) + (1-Q)(G) - C$	$[Q(G-F) + (1-Q)G + Q(G+R) + (1-Q)(G) - C]/2$ $[Q(G-F) + (1-Q)G + Q(G+R) + (1-Q)(G) - C]/2$

Clearly, it is Pareto optimal for Roe and Doe to evade the tax laws and to refrain from suing each other. However, under the following condition the Trojan Horse

mechanism induces the two of them to race to the courthouse as soon as the illegal scheme has ended:

$$(3) \quad Q(G+R)+(1-Q)G-C > G$$

Condition (3) can be reduced to $QR > C$

Knowing in advance, i.e., before commencing the scheme, that eventually the two will race to the courthouse, the two will engage in the scheme only if the following “disparticipation” condition is met:

$$(4) \quad [Q(G-F)+(1-Q)G+Q(G+R)+(1-Q)(G)-C]/2 < 0$$

Condition (4) can be reduced to $Q(F-R) > 2G-C$

Solving for the minimum fine (F^*) that would deter participation, we find that:

$$(5) \quad F^* = 2G/Q \text{ and } R^* = C/Q \text{ if } C < 2G$$

and

$$(6) \quad F^* = C/Q \text{ and } R^* = C/Q \text{ if } C = 2G$$

We may conclude, therefore, that if the costs (C) of filing a suit are large enough, more accurately if $C=2G$, the Trojan Horse mechanism should allow the Trojan Horse plaintiff to collect the whole fine the defendant is bound to pay. However, assuming the gains conspirators derive from an illegal scheme are significantly larger than the cost of filing a Trojan Horse suit, the condition $C < 2G$ does not seem to be very restrictive. In such cases, the portion of the fine that should be allotted to a Trojan Horse increases with the filing cost and decreases with the probability of winning a case. In the extreme case, when cost of filing are zero, the law may induce co-conspirators to snitch on one another by offering an infinitesimal portion of the fine.⁹

⁸For an analysis of the significance of these costs see, *infra*...

⁹ Arguably, if public enforcement is also in play, it might be sufficient to offer amnesty to the Trojan Horse. This is basically the anti trust amnesty plan, *supra* n... but notice that a price fixing arrangement is not a one-shot game.

Another conclusion we may derive from the model above is that the minimum fine, a fine that would talk Roe and Doe out of their plan to evade the tax laws, is not negligible. In fact, comparing public enforcement with the Trojan Horse enforcement system, we find that the two will be effective to the same degree only if the probability the *Trojan Horse* would prevail in trial is twice the probability the tax authorities would detect the evasion. The intuition behind this result is quite simple: when the public enforcement mechanism is in play, the two conspirators are forced to pay the fine to the government. When *Trojan Horses* are the only enforcers of the law, on the other hand, one conspirator gets off the hook and keeps his ill-gotten gains to himself and the fine, therefore, should be higher.

Although this comparison seems to favor the public agencies, it does not suggest that *Trojan Horses* are inferior enforcers. Conspirators, by definition, always detect, and have superior access to evidence concerning the crime they commit. Thus, it might be reasonable to assume that the probability Q is more than twice the probability P , in which case Trojan Horses are more effective than public agencies.

Another consideration we should take into account is that often there are more than two parties to a conspiracy. As shown hereto, the effectiveness of the Trojan horse mechanism is improving with the number of conspirators:

Assume exactly the same game but with N conspirators, where $N > 2$. Each agent faces the following dilemma: If all agents cooperate, namely, if they all refrain from suing, each one of the accomplices will gain G . However, an agent that takes the *Trojan Horse* position gets:

$$(7) U_i = Q(G+R) + (1-Q)G - C, R = (N-1)F$$

Thus, our agent will sue if:

$$(8) Q(G+R) + (1-Q)G - C > G, R = (N-1)F$$

Namely, if:

$$(9) R > C/Q, R = (N-1)F$$

Thus, we know that if condition (9) is met, all the agents would rush to court after committing the crime. Assuming all N conspirators are on equal footing to sue or to be sued, the expected gains from such a scheme to agent i is:

$$U_i = [Q(G+R) + (1-Q)G - C]/N + [Q(G-F) + (1-Q)G](N-1)/N$$

To ensure that the N agents do not conspire, their expected return must be negative:

$$(10) [Q(G+R) + (1-Q)G - C]/N + [Q(G-F) + (1-Q)G](N-1)/N < 0$$

Condition (10) can be reduced to:

$$(11) QF(N-1) - QR > NG - C$$

Solving for the minimum fine that would meet conditions (9) and (11), we find that:

$$(12) F^* = NG/(N-1)Q \text{ and } R^* = C/Q \text{ for any } C < NG$$

and

$$(13) F^* = C/(N-1)Q \text{ and } R^* = C/Q \text{ for any } C = NG$$

Based on the above conditions we may conclude that the larger the number of conspirators the more effective is the Trojan Horse enforcement mechanism.¹⁰ In other words, with the Trojan Horse mechanism in the background we should expect to find less conspiracies and those conspiracies that will take place would probably involve less accomplices.

¹⁰ In addition, conditions 12 and 13 imply that as the number of conspirators increase, the smaller the portion of the fine that should be allotted to the Trojan Horse, and the larger the portion of the fine that the government will collect.

B. *The Finite n-Shot Conspiracy*

Consider now, again, the case of the two conspirators, only now they embark on a series of n incidents of lawbreaking, n tax evasions, n bribes or any other series of n incidents of lawbreaking. Whereas in the one shot game the parties derive no benefit before the end of the scheme, an n -shot scheme is divided to n separate incidents, each incident enriches each conspirator by $\$g$. Another, more important distinction between n -shot and one-shot schemes is that whereas parties to a one-shot scheme can sue only after the scheme concludes, each shot in the n -shot game provides conspirators with the evidence necessary to file a suit against their associates. Just as in the one-shot game, however, the conspirators cannot sue before the first incident. To allow comparisons between the one-shot and the n -shot games I assume $g=G/n$. Assuming zero interest rates, the total gain is the same in both schemes. The cost of filing and the probability of winning the case remain $\$C$ and Q respectively.

Looking at the position of the two players after the n -th shot, it should look very similar to their position after the one shot game. Hence, the two of them will race to the court after conducting the last tax evasion if condition (3) above is met, namely, if:

$$(14) R > C/Q$$

The disparticipation condition in the serial case, however, is different from that of the one-shot conspiracy. This time, we should not ask ourselves whether the players will embark on the series of tax evasions that provides them a gain of $\$G$, but rather whether they would embark on the last tax evasion which yields only $\$g$. The following table presents their payoffs when they face the dilemma of whether to

cooperate with the last tax evasion and sue thereafter or whether they should sue beforehand:

Table B: The Finite n-Shot Game (at the n-1 Stage)

A's payoff B's payoff	A Cooperates	A Sues
B Cooperates	$[Q(G-F)+(1-Q)G+Q(G+R)+(1-Q)(G-C)]/2$ $[Q(G-F)+(1-Q)G+Q(G+R)+(1-Q)G-C]/2$	$Q(G(n-1)/n+R)+(1-Q)(G(n-1)/n)-C$ $Q(G(n-1)/n-F)+(1-Q)G(n-1)/n$
B Sues	$Q(G(n-1)/n-F)+(1-Q)G(n-1)/n$ $Q(G(n-1)/n+R)+(1-Q)(G(n-1)/n)-C$	$[Q(G(n-1)/n-F)+(1-Q)G(n-1)/n+Q(G(n-1)/n+R)+(1-Q)(G(n-1)/n)-C]/2$ $[Q(G(n-1)/n-F)+(1-Q)G(n-1)/n+Q(G(n-1)/n+R)+(1-Q)(G(n-1)/n)-C]/2$

The disparticipation condition in the last stage of the game, therefore, is:

$$(15) \quad Q(G(n-1)/n+R)+(1-Q)(G(n-1)/n)-C > [Q(G-F)+(1-Q)G+Q(G+R)+(1-Q)(G-C)]/2$$

This condition can be reduced to:

$$(16) \quad FQ+RQ > 2g-C$$

Solving for the minimum fine that would meet these conditions, we find that:

$$(17) \quad F^*=R^*=2(g-C)/Q \text{ if } C < 2g$$

and

$$(18) \quad F^*=R^*=C/Q \text{ if } C = 2g$$

Given that the above set of conditions is met, the parties would figure out that $n-1$ is actually the last tax evasion they embark on. Thus, they would face the same dilemma at the $n-2$ stage. By backward deduction we reach the conclusion that under this set of conditions, the two parties would race to court immediately after the first incident.

Our ambition, however, is to deter lawbreaking all together and thus we should search for the condition under which the two parties would not embark even on the first incident. This is actually the one shot game we presented earlier, only the gain from this shot is not G , but rather g :

(19) $F^*=2g/Q$ and $R^*=C/Q$ if $C<2g$

and

(20) $F^*=R^*=C/Q$ if $C=2g$

It is easy to see that these two conditions satisfy conditions (17)-(18) above, and thus they bind.

C. Infinite Conspiracy Games

In the finite game discussed above the parties knew from start that a day will come when the scheme will end and both will face the existential dilemma of whether “to sue or not to sue.” In the infinite game, on the other hand, they know that once they start they will always face the same dilemma: to sue or to go on with the scheme.¹¹ As shown below, this distinction has significant consequences.

For simplicity I will assume the following payoffs: In each stage of the crime the parties derive a gain of \$g. To demonstrate, g can be understood as the one-day gain each cartel member derives from price fixing. Interest rate is r and for consistency I assume that $g/r=G$. Each incident of lawbreaking produces sufficient evidence to allow conspirators to prevail in trial with probability Q. Filing such a Trojan Horse claim costs \$C. In case the Trojan Horse plaintiff wins the case, defendant pays a fine \$F, and the plaintiff collects a portion of this fine -- \$R. Thus, the payoff table the parties observe at each stage looks as follows:

Table C: Infinite Conspiracy Game with a Trojan Horse Mechanism

A's payoff	A Cooperates	A Sues
B's payoff		

¹¹ Even if both conspirators have no intentions to conspire together in the future, we would still treat it as an infinite game if they do intend to conspire with others. See discussion *infra*...

B Cooperates	$g/r=G$	$g/r=G$	QR-C
B Sues	QR-C	-QF	$[QR-C-QF]/2$

In equilibrium, both parties would rush to court soon after the first violation if:

$$(21) QR-C > G$$

Assuming condition (21) is met, both conspirators know, before engaging in the first violation, that the next morning both would race to the court. Hence, they would not engage in the first incident of lawbreaking if:

$$(22) [QR-C-QF]/2 + g < 0$$

Equation (22) can be reduced to:

$$(23) QF - QR > 2g - C$$

Searching for the minimum fine that would satisfy these conditions we find that:

$$(24) F^* = (G+2g)/Q \text{ and } R^* = (G+C)/Q \text{ if } C < 2g$$

and

$$(25) F^* = R^* = (G+C)/Q \text{ if } C = 2g$$

D. Analysis

The tables below summarize our findings.

Summary Table for Minimum Fines

	$C < 2g$	$2g = C < 2G$	$C = 2G$
<i>One-Shot Game</i>	$2G/Q$	$2G/Q$	C/Q
<i>N-Shot Game</i>	$2g/Q$	C/Q	C/Q
<i>Infinite Game</i>	$(G+2g)/Q$	$(G+C)/Q$	$(G+C)/Q$

Summary Table for Minimum Rewards

	$C < 2g$	$2g = C < 2G$	$C = 2G$
<i>One Shot Game</i>	C/Q	C/Q	C/Q
<i>N-Shot Game</i>	C/Q	C/Q	C/Q
<i>Infinite Game</i>	$(G+C)/Q$	$(G+C)/Q$	$(G+C)/Q$

In the following sections I discuss few of the conclusions that can be drawn from these tables:

D.1. The Role of Filing Cost

One intuitive conclusion that can be drawn from the first table is that the higher the filing costs the less effective is the *Trojan Horse* enforcement mechanism. The costs of filing a Trojan Horse suit are typically significant and stem from few sources.

First, assuming the immunity the law assigns to the *Trojan Horse* plaintiff does not extend beyond the crime in question, and for obvious reasons it should not, a conspirator with a “past” faces the risk that bringing a suit will expose him to prosecution for other crimes.

Second, filing a Trojan Horse suit impairs the plaintiff’s reputation in the outlaw community and portrays him as an unreliable partner. Hence, the *Trojan Horse* strategy is very costly for professional criminals who have a “respectable” past and a shining future in the outlaw community. Consider, for example, the case of a contractor who occasionally offers bribes to public officials. Being a repeated player in the bribery business, filing a suit would cut the stream of payments he expects to receive in the future because public officials would refrain from conspire with him.

One implication of this finding, therefore, is that the Trojan Horse mechanism is less effective against conspiracies between professional criminals. Another implication is that in response to the introduction of the Trojan Horse mechanism, we should expect criminals to try to raise filing costs. For example, one venue conspirators are expected to take in response to the *Trojan Horse* threat is violence. Arguably, therefore, although the Trojan Horse mechanism would probably increase entry barriers to the market for crime and reduce the level of crime in society, it might actually increase the level of violence among criminals.

D.2. The Effect of Divisibility

Another interesting conclusion we may draw from the first table is that n-shot schemes are more susceptible to the Trojan Horse mechanism than one-shot schemes. To put it more generally, the more divisible is the scheme, the more effective is the Trojan horse mechanism. From the criminals' point of view, this conclusion suggests a way for lowering the effectiveness of Trojan Horses: they should try to design their scheme less divisible. This tactic, however, is often quite difficult. Even if conspirators may be able to postpone capturing the ill-gotten dividends to the endgame, it seems almost impossible, or at least very expensive, to refrain from producing incriminating evidence along the way.

D.3. The Effect of Infinity

Comparing n-shot finite schemes with infinite schemes, we find that the minimum fine required to deter infinite schemes is significantly higher. We may conclude, therefore, that this enforcement mechanism is less effective against infinite schemes. Criminals, therefore, in response to the introduction of the Trojan Horse mechanism, are expected to devise their schemes with no finite horizon. Hence,

although the Trojan Horse mechanism erects entry barriers to the crime market, at the same time it might strengthen the Mafia.

D.4. Trojan Horses v. Bounty Hunters: Main Distinctions

As a matter of first impression, it seems that the *Trojan Horse* enforcement mechanism is only a unique case of any bounty hunter program. Arguably, there is nothing in *Becker & Stigler's* private enforcement model that denies co-conspirators the right to snitch and collect the fine. Similarly, many bounty hunter programs do not rule out the possibility of rewarding a squealing co-conspirator. But *Becker and Stigler's* model, as well as the prevailing bounty hunter programs do not provide adequate incentives for co-conspirators to come forward, as they fail to take into account the *Trojan Horse's* significant filing costs and, more importantly, his opportunity costs.¹²

Unlike bounty hunters, a *Trojan Horse* plaintiff waives the benefits of the illegal scheme. Hence, for the *Trojan Horse* to kick in, the law must ensure *Trojan Horses* a much larger reward than the reward it provides to any other private enforcer. In fact, as shown in the two tables above, it is often the case that the Trojan Horse mechanism is effective only if we allow Trojan Horses to collect the whole fine the defendant pays. As shown above, ensuring a very large reward to the plaintiff-conspirator is particularly essential in infinite schemes.

In practice, however, even those bounty hunter programs that do allow lawbreakers to squeal and collect a portion of the fine do not provide any beneficial treatment to co-conspirators. On the contrary, relatively to a law-abiding squealer,

¹² Another important distinction between the prevailing bounty hunter programs and the mechanism this paper endorses is in the government's role. Whereas in this paper the government stays out of the loop, the prevailing bounty hunter programs allow regulators the discretion as to whether, and to what extent to reward a snitch. Thus, although some laws empower regulators to reward co-conspirators for snitching, typically these regulators refuse to do so. For an excellent review, see M.J. Ferziger and

co-conspirators are typically awarded a lesser portion of fine.¹³ Under these circumstances it seems unlikely that conspirators would snitch on their partners voluntarily, i.e., before the police is on their tail, only if, like in the competitors' cartel for example, they derive indirect benefit from their co-conspirators' losses.

D.5. Trojan Horses and Amnesty Programs: Main Distinctions

State witness programs provide amnesty to conspirators who snitch on their partners. Hence, these programs, more than bounty hunter programs, resemble the Trojan Horse mechanism. Especially relevant here are programs like the Antitrust Division's Amnesty Program,¹⁴ that waives the prosecutorial discretion and guarantees immunity to the first cartel member who comes forward.

One important, qualitative distinction between the Trojan Horse mechanism and amnesty programs is that whereas amnesty programs rely on the operation of a public enforcement agency in the background, Trojan Horses can enforce the law on their own. Thus, the Trojan Horse mechanism allows us to decriminalize certain offences.

Assuming public agencies are always in the background, it can still be shown that amnesty programs do not deter certain offences, which are deterred by the Trojan Horse mechanism. Amnesty programs are quite effective in finite games. In fact, as shown earlier, if we assume very small filing costs, amnesty seems to provide sufficient incentives to squeal and there is no need to add to the immunity any

D.G. Currell, Snitching for Dollars: The Economics and Public Policy of Federal Civil Bounty Programs, 1999 Univ. Ill. L. Rev. 111 (1999).

¹³ For an excellent review, see M.J. Ferziger and D.G. Currell, Snitching for Dollars: The Economics and Public Policy of Federal Civil Bounty Programs, 1999 Univ. Ill. L. Rev. 111 (1999).

¹⁴ But see the anti-trust amnesty program mentioned earlier, *supra n. 8* – which provides amnesty but no financial reward.

financial reward. Still, assuming filing costs (C) are larger than the expected fine the government imposes (PF), amnesty programs are not effective.

More importantly, amnesty programs are much less effective against infinite schemes. To see why, consider the case of an infinite scheme in which each conspirator anticipates a stream of revenues with a net present value of G . Conspirators also know that there is a probability P they will be detected by the public authorities and forced to pay a fine with a net present value of F . For each conspirator, therefore, the net present value of the scheme is $G - PF$. As long as the net present value of the scheme is negative, public enforcement provide sufficient deterrence and there is no essential need for supplementing it with private enforcers. The interesting case, therefore, is when the net present value of the scheme is positive, namely when $F < G/P$. In these cases, it seems an amnesty program cannot provide sufficient motivation for conspirators to snitch one against the other. Even if filing costs were zero, as long as the net present value of the scheme is positive, it is worthwhile to continue engaging in the infinite scheme and it does not pay to squeal. Only a positive reward, like the reward this paper suggests to provide to a Trojan Horse plaintiff, has the potential of tipping the scales, persuading participants in a profitable scheme to incriminate their partners.

D.6. Opportunism

One of the most troubling aspects of the *Trojan Horse* mechanism is that it might provoke opportunistic behavior. Consider, for example, the following hypothetical: the general belief among producers in a certain market is that an agreement to set quality standards does not violate the antitrust laws. Although such an agreement would clearly benefit these producers, free riding predicaments entails

that no one of them is willing to sustain the costs of initiating negotiations for such an agreement. Now suppose that one of these competitors suddenly learns about the illegality of such a standardization agreement. Assuming this informed party believes that the other producers are uninformed, the *Trojan Horse* enforcement mechanism might lure that informed party to call upon his naive competitors and invite them to enter into such an illegal scheme. Many would find this effect troubling even if the opportunistic *Trojan Horse* would probably file the suit at the very early stage of the scheme before any harm is done. Most troubling is the fact that the law puts together a mechanism that encourages otherwise law-abiding citizens to join or even initiate an illegal scheme. Arguably, they would do so for the sole purpose of filing a suit and collect the fine from their “naive” accomplices.

One comforting answer to this objection is that the risk of such opportunism is probably not significant. The models I have used above missed this point because they assumed conspirators act contemporaneously. But in real life, any attempt to seduce others to join an illegal scheme subjects the seducer to the risk that the supposedly naive party would file the suit first. In our price fixing example, the knowledgeable party can rarely be sure that all his competitors are in fact ignorant about the illegality of the scheme. Given the fact that the seducer must make the first illegal move, the seducer faces the risk of being sued by one of his competitors, for the mere attempt to fix prices, before he finds his way to the court house.¹⁵

¹⁵This comforting answer is not satisfactory in asymmetric cases like the illegal employment example we have discussed in the previous chapter. With no money in his pocket, an immigrant can search for illegal employment with no risk of being sued. Notice, however, that in the context of illegal employment we may assume that the deep pockets are sophisticated and unlikely to be entrapped.

Another comforting answer to this concern is that opportunism might yield some socially beneficial by-products. One side effect of the fact the Trojan Horse mechanism may motivate (otherwise) law-abiding agents to entrap others is that it encourages agents to search for legal information and to publicize this information. Currently, the search for legal information is encouraged by the sanctions the law imposes on lawbreakers. With the *Trojan Horse* mechanism in the background, potential lawbreakers may also rip profits from legal information. Some parties would search for such information for the purpose of entrapping other agents and others for the sake of protecting themselves from the formers. The market for legal information, therefore, would certainly benefit from this mechanism.

If despite these two comforting answers we are still interested in restricting such opportunism, we may adopt a variant on the *first felon to file collects the fine* rule, a variant that would mitigate this problem. One radical solution would be to deny standing to anyone who knowingly conspires to violate the law. According to this solution, only innocent parties, who realize they engage in an illegal conduct only after its commencement, would be allowed to bring such a suit. The flaw in this solution is that although it might operate effectively to deter opportunistic behavior, at the same time it undermines the efficacy of the *Trojan Horse* mechanism. For example, it would allow the briber and the bribee to commit not to sue one another simply by declaring, in their bribery agreement, the illegality of their interaction.

Hence, to mitigate opportunism and, at the same time, to sustain the efficacy of this mechanism we must ignore the plaintiff's motivation and focus on the defendant's state of mind. One logical refinement would be to subject *Trojan Horses* to the same limitation any police informer is subject to under the entrapment

doctrine.¹⁶ Thus, the right to collect the fine would be contingent on the plaintiff's ability to demonstrate that the defendant would have engaged in the illegal scheme even without the plaintiff's involvement. Although this refinement does not exterminate opportunism, it does reduce the probability that the legal mechanism erected for the purpose of diminishing illegal schemes would in fact produce such schemes.

D.7.. The Problem of False Claims

Another concern the Trojan Horse mechanism provokes is that felons would file false complaints. The significant prize the mechanism offers – a portion of the fine and immunity, might induce felons to file false claims against innocent parties. Although I admit the problem of false claims is indeed troubling, I believe the risk of a false *Trojan Horse* claim is much lower than the risk of a deceptive testimony by a state witness or a false complaint of a bounty hunter. Whereas state witnesses and bounty hunters face almost no risk by filing a false claim, deceptive Trojan horses are subject to significant risks. Along side the huge benefits the *Trojan Horse* mechanism provides to successful plaintiffs, the sanctions imposed on a failing *Trojan Horse* are significant as well. In particular, unlike a state witness, a failing *Trojan Horse* is not immune from public prosecution and the evidence he produces as a plaintiff is likely to be used against him when the public enforcement agencies go after him. Hence, assuming the risk of failure is higher when the claim is false, it seems the risk of false *Trojan Horse* claims is small, at least relatively to the risk of a false testimony by a state witness or a bounty hunter. On the contrary, the main problem I see with the

¹⁶ Model Penal Code, Section 2.13

Trojan Horse mechanism is not that false claims will be filed but rather that only the very clear-cut cases will be prosecuted by this mechanism.

III. Financial Inadequacy

Any enforcement mechanism that is based on economic sanctions alone seems to be ineffective when utilized against judgment-proof defendants. Consider again the case of illegal employment. Economic sanctions may deter employers but, clearly, the employees' empty pockets protect them from economic sanctions. It seems that the only way to discipline judgment proof lawbreakers, like illegal workers, is through jail sentencing and deportation. Our experience, however, suggests that opportunity-costs for judgment proof felons are often so high that they would engage in the illegal conduct even if they face a very high risk of jail sentencing and/or deportation. Moreover, subjecting only judgment proof defendants to physical sanctions like jail sentences, or deportation in the case of illegal immigrants, is a very costly move for the government, mainly because of the human rights sentiments it arouses.

This brings us back to the advantages of the *Trojan Horse* mechanism. The argument that economic measures do not deter illegal employees and other judgment proof felons fails to recognize the fact that economic sanctions *can* work against judgment-proof lawbreakers if they are offered as carrots rather than sticks. When one party to an illegal scheme is judgment proof and the other is financially adequate to meet the fine, the *Trojan Horse* strategy operates as a stick against the deep pocket party and as a carrot for the judgment proof party. In the illegal employment example, the *Trojan Horse* strategy provides adequate deterrence against employees, just as it does against employers, only against the former it utilizes carrots instead of

sticks – it encourages illegal employees to sue and collect the fine from their employers, thereby ending the employment relationship. Similarly, the fact that they can collect the fine only on the way out of the country encourages illegal immigrants to leave the country voluntarily. The following model demonstrates few of these contentions.

(a) An Infinite Asymmetric Conspiracy Model

Consider an asymmetric game, like illegal employment, where conspirators have different level of financial resources and they engage in the illegal conduct for the long run with no endgame in the horizon. For simplicity, I will assume that one party has no financial resources to pay a fine and the other is a deep pocket. I will further assume, as often is the case, that the empty pocket criminal derives $\$m$ from each incident of lawbreaking and the deep pocket derives $\$g$ where $g > m$. Interest rate is still r , such that $m/r = M$ and $g/r = G$. If the empty pocket sues and proves his case, the deep pocket would pay a fine of $\$F$, a portion of which, $\$R$ ($R = F$), is allotted to the empty pocket and the rest to the government. The deep pocket conspirator may also sue but only for the purpose of protecting himself from a counter suit by the empty pocket conspirator. Filing costs for the deep pocket and for the empty pocket are $\$D$ and $\$E$,¹⁷ respectively. The probability of winning a case remains Q for both conspirators. Under this setting, the infinite game looks as follows:

Table D: The Infinite Asymmetric Conspiracy Game

Deep Pocket's payoff	Deep Pocket Cooperates	Deep Pocket Sues
Empty Pocket's Payoff		
Empty Pocket Cooperates	$m/r = M$	0
	$g/r = G$	$-D$

¹⁷ Despite the fact that empty-pocket conspirators cannot meet the fine, they may still be able to incur the filing costs even if filing costs are higher than the fines. As long as the net present value of filing a Trojan Horse complaint is positive, empty pocket conspirators should be able to find a financier for such a suit.

Empty Pocket	-QF	-[D+QF]/2
Sues	QR-E	[QR-E]/2

The Empty pocket would sue if:

$$(26) \quad QR-E > M$$

The deep pocket would sue only if he knows the empty pocket is about to sue and only if:

$$(27) \quad QF > D$$

We are left to make sure that the parties would not want to engage even in one violation of the law. The empty pocket would engage in one incident of lawbreaking and sue if:

$$(28) \quad m + QR - E > 0 \text{ and } QF > D$$

or if:

$$(29) \quad m + (QR - E)/2 > 0 \text{ and } QF < D$$

Since we assume $QR - E > M$ (condition 26), condition (28) and (29) always are satisfied. Hence, we may conclude that the empty pocket will always be interested in engaging in one incident of lawbreaking for the purpose of filing a suit if:

$$(30) \quad QR > M + E$$

Turning to the deep pocket, we know that he would engage in one violation only if:

$$(31) \quad g - QF > 0 \text{ and } D > QF$$

or if:

$$(32) \quad g - (D + QF)/2 > 0 \text{ and } D < QF$$

We may conclude, therefore, that the deep pocket would not engage in the first lawbreaking incident if:

$$(33) \quad g < QF \text{ and } g < D$$

Solving for the minimum fine that would meet conditions 30 and 33 we find that:

$$(34) \quad F^* = R^* = (M+E)/Q \text{ and } D > g$$

(b) the Efficacy of Trojan Horses in Asymmetric Cases

One lesson condition 34 teaches us is that the minimum fine that would deter asymmetric schemes is a function of the gains the empty-pocket conspirator anticipates to receive rather than those of the deep pocket. This result is in sharp contrast with the fines that public enforcement agencies should impose in such cases, which are a function of employer's gains. Thus, if filing-costs for empty pocket conspirators are small enough, the Trojan Horse mechanism would be more effective than public agencies even if the probability of success for public enforcement are as high as that of Trojan Horses. For similar reasons, comparing condition 34 with the minimum fine for infinite symmetric schemes, we find that the minimum fine in asymmetric schemes is lower. We may conclude, therefore, that Trojan Horses mechanisms are more effective in asymmetric schemes.

(c) The Problem of No Pressure to File

Notice, however, that while setting the fine and the reward according to condition 34 above ensures that employers would be deterred from hiring illegal immigrants, it does not subject employers and illegal immigrants in the conspirator dilemma. As long as filing costs for employers are lower than his expected fine ($D > FQ$), an employer who mistakenly hired illegal immigrants is better off continuing the employment relationship even if he knows the employee is likely to sue. Knowing this, an illegal employee that has been hired might prefer working for a

while and collect the fine later. He is not under any pressure to file a suit as early as possible.

The fact that *Trojan Horses* are not under a strong pressure to file fast may slow down the operation of this mechanism and may reduce its effectiveness. One obvious solution for this problem is to set the fine to mitigate it, such that $F > D/Q$. But we may also think of alternative competitive pressures.

First, and most obviously, we may preserve the threat of the traditional enforcement mechanisms. Similarly, the law could employ *Becker and Stigler's* mechanism and entrust the right to collect the fine with the first citizen to file a suit against the employer. Such potential competition from public or private enforcement agents would put illegal employees on notice that if other enforcers sue their employer before they do, they would lose his right to collect the fine from their employer. This threat would encourage illegal employees to pursue the carrot and to bring an end to their illegal employment at the very start of the employment relationship.

Second, the state may promise the financially adequate conspirator a reward for putting his empty pocket co-conspirators behind bars.¹⁸ More specifically in our example, an employer would be rewarded for filing a suit for the imprisonment or deportation of his illegal employee. The employer's right to collect the reward would be conditioned, of course, upon filing the suit before the illegal employee does so.

Lastly we may think of a more elegant solution, one that would force the *Trojan Horse* employee to gallop faster and still preserves the spirit of the Trojan

¹⁸ The fact that the government rewards a criminal should not bother us much. Currently, the government provides immunity and even promises financial rewards to suspects who spill the beans and turn against their collaborators. The *Trojan Horse* plaintiff renders a more beneficial service to society as he turns in criminals even before the police was on their tails, often even prior to committing the crime. It is only logical, therefore, that the price society should pay for this service is higher than the price society pays to a state witness.

horse mechanism. Conceptually, we could subject illegal employees to competition one against the other by entrusting the first illegal employee to come forward with the right to file a class action suit against his employer on behalf of all the defendant's illegal employees. The larger the piece of the pie the law places in the hands of the class representative and the smaller the portion of the passive class members, the stronger is the pressure-to-sue on the employees.

(d) Opportunism

As a matter of first impression it seems that the risk of opportunism is much higher in asymmetric cases. In the symmetric case, we have seen that an opportunistic agent who tries to seduce others to join a scheme for the mere purpose of suing them the morning after, exposes himself to a significant risk of being sued first. In the asymmetric case, on the other hand, the judgment proof agent may try to approach a deep pocket with almost no risk. Whereas employers of illegal immigrants face a significant risk of being sued by their employees, Condition 30 above teaches us that illegal immigrants face no such risk and, therefore, they would try to get a job for the mere purpose of suing their employers.

Hence, in the context of asymmetric schemes it might be particularly important to subject empty pocket Trojan Horses to the entrapment doctrine, namely that they will prevail in trial only if they prove that the defendant would have engaged in the illegal scheme even without their solicitation. In certain contexts, like that of illegal immigrants, it seems that this burden of proof is not too onerous. For example, illegal immigrants who were recruited for the job in their homeland would probably find it easy to prove that their employer would have hired an illegal immigrant even without their solicitation.

On the positive side of asymmetric cases we know that the likelihood of opportunism on the part of the deep pocket conspirators is nil because they have nothing to gain from seducing and suing judgment proof accomplices. This effect is important because typically the deep pocket parties are also the more informed and are in a better position to seduce the empty pockets. Employers, for example, are typically more familiar with their legal risks and rights than illegal immigrants who may learn about their right to sue their employer only after they immigrate and start working. Moreover, an attempt on the part of a potential illegal immigrant to seduce employers involves significant costs – traveling costs for example – and it is unlikely that someone would try to incur these costs only for the hope that he'll find an uninformed employer that would fall into the trap.

IV. Summary

This paper offered a new enforcement mechanism against conspiracies, one that relies solely on the conspirators themselves. According to my proposal, the law should position conspirators in the conspirator dilemma by entrusting each one of them to sue the others. The first conspirator to file will collect the fine, or a portion of the fine the defendant pays. Using a simple model, I have analyzed the conditions under which the *Trojan Horse* mechanism is effective, and compared its efficacy with that of alternative enforcement mechanism like public agencies, bounty hunters and amnesty programs.

One important conclusion of this paper is that the Trojan Horse mechanism is particularly effective in asymmetric cases, where one party to the conspiracy is judgment proof and the other is a deep pocket.

In this context, of asymmetric conspiracies, it is also easier to justify the Trojan Horse mechanism on moral grounds as it enforces the law on empty pocket agents by offering them carrots instead of sticks. Whereas in the bribery example, a moral argument can be brought against rewarding someone for engaging in a bribery agreement,¹⁹ very few people would argue that there is something immoral about illegal immigration.²⁰ Immigration law is not based on moral grounds but on society's preference of its own citizens over foreigners. In fact, many supporters of the criminalization of illegal employment, prostitution or drug-use would even concede that, from a moral perspective the violators of these laws are actually the victims. In such cases, therefore, any legal system that would put an end to these activities, without criminal sanctions will be morally superior. The *Trojan Horse* mechanism provides such a solution: it allows the prostitute to sue her pimp, the illegal employee to sue her employer and the drug user to sue his pusher. In such cases, therefore, cases in which the morally accepted view is that the lawbreaker is the actual victim, the *Trojan Horse* is more morally commended and politically conceivable than the traditional law-enforcement mechanisms.

¹⁹ This argument is not accurate also: the reward is not given for participation in the crime but rather for putting an end to the scheme.

²⁰ Arguably, the fact that many crimes cannot be justified on moral grounds might be typical to asymmetric schemes but clearly not exclusively. The fact that a certain conduct is not immoral does not mean the state has no ground for stopping it, but it does suggest that it is preferable to fight against such crimes by civil means rather than utilizing criminal law.

Another important conclusion of this paper is that the Trojan Horse mechanism is more effective than amnesty programs. Arguably, both the Trojan Horse mechanism and amnesty programs can be criticized on moral grounds, as both mechanisms encourage betrayal and squealing.²¹ But the Trojan Horse mechanism enjoys one significant advantage in this regard: a conspirator who offers to testify against his partners is typically offered an unconditional immunity. A Trojan Horse who files a suit, on the other hand, always faces the risk of losing the case and being exposed to prosecution by the public agency.

²¹ Arguably, the whole enterprise of law enforcement is designed to encourage squealing. In this regard it is quite hypocritical to criticize only the Trojan horse mechanism on this ground.