# Carrots versus Sticks

Gerrit De Geest [a, b, *], Giuseppe Dari-Mattiacci [c, d]

[a] *T. Koopmans Institute, Utrecht University, Janskerkhof 12, 3512 Utrecht, The Netherlands*
[b] *Ghent University, Faculty of Law, Universiteitstraat 4, 9000 Ghent, Belgium*
[c] *University of Amsterdam ACLE, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands*
[d] *Tinbergen Institute, Roetersstraat 31, 1018WB Amsterdam, The Netherlands*

**PRELIMINARY DRAFT – PLEASE DO NOT QUOTE OR DISTRIBUTE**

## Abstract

This paper proposes the building blocks towards a general theory on the optimal use of carrots and sticks in legal and social enforcement systems. Although in principle carrots and sticks are equivalent with respect to marginal incentives, the central theme of the paper is that they are nonequivalent in many other respects. Our central finding is that sticks are intrinsically superior to carrots because they are meant not to be applied: sticks incentivize simply by threatening, while carrots incentivize by actually rewarding. Consequently, sticks minimize sanctioning costs, risk bearing and distributional side-effects. We show that a perfectly informed principal will never use carrots. However, sticks are also inherently more dangerous tools in the hands of non-benevolent principals since enforcement under carrots is always Pareto efficient while enforcement under sticks is not even necessarily Kaldor-Hicks efficient. As society becomes more complex, labor more specialized and decisions more decentralized, carrots tend to be increasingly desirable.

*JEL classification*: H20, K00, K42, L14.
*Keywords*: negative sanctions, positive sanctions, activity level, bonus.

---

* *Corresponding author:* Phone +31-30-253.98.00, Fax +31-30-253.73.73. *E-mails*: g.degeest@econ.uu.nl (G. De Geest), gdarimat@uva.nl (G. Dari-Mattiacci).

The authors would like to thank Hugo Acciarri, Eugene Bardach, Arnoud Boot, Roger Bowles, Vincent Buskens, Lloyd Cohen, Bob Cooter, Ben Depoorter, Ehrling Eide, Sarah Hernandez, Oren Gazal, Fernando Gomez, Aristides Hatzis, Bruce Johnsen, Christine Kymn, Henrik Lando, Francesco Parisi, Dan Polsby, A. Mitch Polinsky, Jan Potters, Werner Raub, Carol Rose, Bernd Schäfer, Goran Skögh, Robert Scharff, Maarten Pieter Schinkel, Tom Ulen, and the participants in workshops at Utrecht University, George Mason University, University of Amsterdam, Ghent University, and the annual conferences of the Scandinavian Association for Law and Economics (Helsinki) and the Latin American and Caribbean Law and Economics Association (Berkeley).

## 1. Introduction

Incentives can be generated by either carrots – promises to reward, such as with prizes or bonuses – or sticks – threats to punish, such as with fines or damages. Carrots and sticks are *prima facie equivalent*, because any behavioral change induced by promising compliers a $100 reward can also be obtained by threatening violators with a $100 punishment. Parents may incentivize their children by promising toys if they behave well or by threatening to take away their toys if they misbehave. Yet, while both carrots and sticks seem to work indifferently, they are not chosen at random; some general patterns can be observed across legal systems. Incentives for careful driving are generally created by holding negligent drivers liable under tort law, rather than by rewarding careful drivers. Likewise, theft is discouraged by penalizing thieves and not by rewarding those who do not steal. Incentives to invent are instead created by rewarding successful inventors with patents or academic prizes rather than by punishing all of the others. While differences between carrots and sticks are often ascribed to psychological or behavioral aspects[1] – a perspective whose importance we do not deny –, we examine their properties within a narrowly-defined, rational-choice framework, in which their apparent equivalence is more poignant. The purpose of this paper is to make a step towards a general theory of the optimal use of carrots versus sticks.

Our central finding is that sticks are intrinsically superior to carrots because they are meant not to be applied. Sticks incentivize by threatening while carrots incentivize by actually rewarding; in an ideal world, fines are never paid, but prizes are. We exploit this simple finding to show that sticks yield less transaction costs, less risk for complying agents, and less distortions in the distribution of wealth. However, in the hands of non-benevolent principals, sticks are more dangerous than carrots. The reason is that carrots always lead to a Pareto improvement, while sticks can be employed to reach outcomes that may not even be Kaldor-Hicks improvements. Our analysis suggests that, as societies become more complex, labor more specialized, and decisions more decentralized, carrots tend to be used more often.

In the economic, legal, and psychological literature, there are numerous studies that implicitly or briefly deal with positive and negative sanctions. It is virtually impossible to list all such publications; thus, we will only comment on those who more directly relate to our analysis. The idea that carrots and sticks are equivalent with respect to marginal incentives is implicit in Coase

(1960), who showed that the same result can be obtained either by having polluters pay damages or by paying polluters not to pollute. Such an analytical symmetry between carrots and sticks is more explicitly mentioned in Becker (1968), who suggested that the more frequent use of sticks in legal systems may be due to historical coincidences, or difficulties in measuring benefits.

The first systematic analysis of the choice between positive and negative sanctions using an economic framework is Wittman (1984). In his seminal paper, Wittman argued that the choice between carrots and sticks is ultimately a matter of transaction costs. Compensation will less often need to be paid when inefficient behavior is sanctioned, rather than efficient behavior rewarded, because the former is usually less likely to occur than the latter, given the incentives created by the legal system. Hence, theft should be penalized, rather than non-theft subsidized, because theft is inefficient and thus less frequently occurring than the alternative. Moreover, compensation is easier to compute when the reference point (to which the actual behavior is compared) is the typical case, rather than a rare or even never occurring inefficient outcome.

The idea that the choice between carrots and sticks may affect long-run activity levels can be found in the literature on the use of taxes versus subsidies to control pollution (e.g., Polinsky, 1980), and in the literature on the duty to rescue (e.g., Landes and Posner, 1978, arguing that if non-rescuers were punished instead of rescuers rewarded, potential rescuers might lower their activity levels). This activity level argument was translated into a transaction costs argument by Wittman (1984): activity level distortions can be avoided either by society determining who should start the activity (which involves high information costs), or by subsidizing everyone who might start an activity (which involves high payment costs). For instance, the consumer should pay for taking apples, instead of the farmer paying the consumer for not taking apples, because the latter would either require the farmer to compensate all consumers who did not take apples, or society to know who should eat which apples.

In the economic literature on law enforcement, a number or papers explicitly consider the use of carrots as alternatives to sticks (Becker and Stigler, 1974, Cooter and Garoupa, 2000, Depoorter and De Mot, 2006). Note that in the criminological literature (and in some economic studies, e.g., Frey, 2004), the terms carrots and sticks are often used in a different sense that the one taken in the present article: working with 'carrots' stands for improving the attractiveness of the alternatives to a criminal career, for instance by increasing education levels, reducing unemployment, or improving

---

[1] See footnote 2 below and accompanying text.

social cohesion. Law and economics papers that explore the similarities and differences between carrots and stick regimes are Levmore (1985, comparing restitution and tort law), Levmore (1986, comparing carrots and sticks with respect to the duty to rescue), and Gordon (1982, comparing intellectual property law, tort law and restitution). The distinction between carrots and sticks has also received some attention in labor economics (e.g., Lazear, 1991, Aron and Olivella, 1994), in the economic literature on slavery (e.g., Fogel and Engerman, 1974, arguing that also carrots were used on southern plantations), and in the literature on international economic relations (e.g. Chang, 1997).

The different reactions of human beings (and animals) to negative and positive sanctions is also widely studied in psychology. Prospect theory (Kahneman and Tversky, 1979) suggests that negative incentives are more effective than positive incentives because of loss aversion. Still there is a general tendency among psychologists to be more favorable towards carrots than towards sticks (see Kohn, 1993). This emphasis on positive incentives is often attributed to the influence of behaviorism, the most popular school of thought in American psychology. The so-called Law of Effect (Thorndike, 1898) states that when behavior is followed by a positive consequence (which Skinner, ????, called 'positive reinforcement') it will be repeated. When behavior is followed by a negative consequence, the result may also be that the activity is avoided

This paper's contribution consists in offering a more general theory than previous attempts. In section 2, we discuss the definition of carrots and sticks, illustrate simple tests to decide whether a sanction is a carrot or a stick, and derive four main differences between carrots and sticks. In section 3, we present a simple framework of analysis and discuss the equivalence claim (incentives provided by carrots are exactly the same as incentives provided by sticks) and show its limitations. In particular we show that sticks are intrinsically superior to carrots. In section 4, we extend the analysis to include information and evidence problem and strategic behavior. We show that as societies become more complex, carrots also become more desirable. In section 5 we employ the framework of sections 3 and 4 to discuss some applications of our theory. In section 6, we conclude.

## 2. Definition and general characteristics of carrots and sticks

Can carrots and sticks be objectively defined, or are they just the same phenomenon seen from a different angle, similar to the difference between a half-full or half-empty glass of water? In this section we show that the difference between carrots and sticks is not just a matter of framing. We

provide an objective definition, from which we derive four characteristic differences between carrots and sticks.

*2.1. Defining carrots and sticks*

We use the following

***Definition****: carrots are transfers that increase the agent's wealth; sticks are transfers that decrease the agent's wealth compared to the status quo. Both are conditional on the agent's behavior.*

Carrots are positive sanctions; sticks are negative sanctions. A sanction is an (ex post) wealth transfer aimed at inducing the agent to (ex ante) comply with a rule. A carrot is a transfer that increases the agent's wealth compared to the status quo; a stick is a transfer that decreases the agent's wealth. To put it differently, a carrot is a wealth transfer *to* the agent, while a stick is a wealth transfer *from* the agent.

Since carrots and sticks are defined by referring to the status quo, whether the agent's wealth is increased or decreased is determined by comparing the ex post wealth with the wealth before the sanction was applied. Note that the idea that the existing (ex ante) wealth and not the expected future wealth should be taken as a reference point, also underlies one of the central concepts of modern welfare economics – the concept of Pareto optimality. Therefore, if one accepts that a Pareto improvement can be objectively defined, then it should be accepted that a carrot can be objectively defined as well.[2]

Of course, while receiving a stick is a direct cost to the agent, not receiving a carrot involves an opportunity cost as well. Still, this does not mean that a foregone carrot is a stick. For instance, there is a difference in terms of transaction costs between a wealth transfer to the agent that effectively takes place and a wealth transfer from the agent that does not take place. The former

[2] According to prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1981), whether something is seen as a gain or a loss depends on the 'framing', and this depends on the reference point, which is not an objective point but a subjective one. (ALT: According to prospect theory, individuals define a reference point, and this point determines whether an outcome is seen as a gain or a loss.) Prospect theory seems to imply that carrots and sticks cannot be objectively defined, but depend on subjective evaluations ('framing'). If the agent expects to get the carrot, he may feel penalized by not getting it. Although the agent is not worse off compared to the status quo, he may *feel* worse off by taking the carrot as the reference point. Again, the solution to come to on objective reference point is to take the status quo as reference, not the wealth that the agent expects to get at a point in the future.

involves transaction costs; the latter does not. Further in this paper we also show that there are major differences with respect to risk allocation and distributional consequences.

*2.2. Four fundamental differences between carrots and sticks.*

***Difference 1****: a carrot is applied upon compliance, a stick upon violation*. This follows from the two building stones of the central definition: the purpose of the transfer (incentivizing the agent to comply) and the positive or negative sign of the transfer (doing otherwise, i.e. applying a carrot upon violation or a stick upon compliance would create incentives to violate the rule, which is not the principal's aim).

***Difference 2****: carrots incentivize by effectively rewarding, while sticks incentivize only by threatening.* This follows immediately from the previous difference: when the goal – compliance – is obtained, the stick is not applied.

***Difference 3****: a carrot is used up upon compliance*, *while a stick is used up upon violation*. This follows from the first difference as well. When a reward has been paid to a complying agent, the same amount of money cannot be used again to incentivize the agent to comply with another rule (in other words, there is 'exhaustion' of the principal's budget). In contrast, when a threat is not executed, the same threat can generally be repeated.[3]

***Difference 4****: under carrots, non-monitored agents receive the same treatment as monitored violating agents; under sticks, non-monitored agents receive the same treatment as monitored complying agents.* This characteristic is a useful tool to determine whether a sanction is a stick or a carrot in practice (see section 2.6).[4]

---

[3] See Dari-Mattiacci and De Geest (2005), further elaborating the point that only sticks have a multiplication effect, and showing that its consequences for law enforcement.

[4] Besides, this characteristic can also be interpreted as how these regimes differently allocate the burden of proof. Sticks have a built-in presumption of innocence, because in the absence of any information on the agent's behavior, the agent is treated like a complier (he does not get a stick). Carrots, on the other hand, have a built-in presumption of 'guilt' since in the absence of such information the agent is treated like a violator (he does not get the carrot). Note that for annullable carrots and sticks the opposite is true: annullable bonuses implicitly presume that the agent complied, and annullable sticks implicitly presume that the agent violated the rule (see De Geest, Dari-Mattiacci and Siegers, 2006). This again shows that annullable carrots are analytically closer to sticks than to carrots, and annullable sticks closer to carrots than to sticks (see section 2.5).

### *2.3. Precompensated carrots and sticks*

Some confusion may arise because sanctions can be accompanied by side payments that compensate ex ante for the possible use of carrots or sticks. For instance, a child may first receive a toy and then be incentivized by the threat of having to give the toy back. Similarly, a seller may receive a higher contract price as a compensation for his future liability, or an employee may pay an implicit entry fee (in the form of a lower initial wage) to later get a contract with high bonuses.

We will use the general term of *precompensation* to refer to these unconditional (i.e. not conditional on the agent's behaviour) payments. A distinction should be made between *full* and *average* precompensation. Full precompensation equals the stick or carrot; average precompensation equals the *expected* stick or carrot. To illustrate, if the effort cost is e = \$2, the monitoring level p = 0.1 and the magnitude of the stick S = \$20, the agent is fully precompensated for the possible stick if he receives ex ante a fee of \$20 before being subjected to the rule. The agent is averagely precompensated if he ex ante receives \$2.

### *2.4. Can a stick be rewritten as an algebraically identical carrot?*

Some authors have made the point that carrots can be rewritten to algebraically identical sticks and vice versa (e.g., Lazear, 1991, Aron and Olivella, 1994, Cooter and Ulen, 2003). To illustrate, consider a construction contract with a price of \$100, but a penalty for late performance of \$90. This contract can be rewritten as one with a price of \$10 and a bonus for timely performance of \$90. In both cases the builder gets \$100 if he finishes in time and \$10 if he finishes late.

This certainly holds if monitoring is non-probabilistic (p=1).[5] In that case, if $w^{normal}$ denotes the normal price and $w^{precomp}$ the price that includes precompensation ($w^{precomp} = w^{normal} - s$ ???denote bonus as b???), then the payoffs in case of compliance are $w^{normal}$ under a stick regime and $w^{precomp} + b = w^{normal}$ under a carrot regime. The payoffs in case of violation are $w^{normal} - s$ under a stick regime and $w^{precomp} = w^{normal} - s$ under a carrot regime. However, if p<1, sticks and precompensated carrots are no longer mathematically identical.

**Proposition 1**. *A stick can be rewritten as an algebraically identical (precompensated) carrot and a*

[5] For simplicity we assume here that the precompensation and the carrot or stick are paid at the same moment. If precompensation is paid ex ante and carrots or sticks ex post, there may be differences with respect to problems of opportunism. If precompensation is paid at t3, there is no difference (it is a matter of framing, though from a scientific point of view the more simple concepts of stick and carrot look superior to the more

*stick as an algebraically identical (precompensated) carrot only if p=1. Even in this case, there remains a difference with respect to procedural costs.*[6]

(see also table 1 ???simplify and keep notation consistent by deleting normal wage???)
If p = 1, there is generally no analytical difference between carrots and sticks, in the sense that pure sticks and precompensated carrots have identical payoffs, and that pure carrots and precompensated sticks have identical payoffs as well. Still, there remains a difference between pure sticks and pure carrots, in the sense that the latter have a built-in compensation mechanism. To put it differently, pure sticks and carrots differ with respect to the amount of precompensation that is required to obtain or undo their intrinsic wealth effects. In addition, (pure and precompensated) sticks and (pure and precompensated) carrots remain different with respect to the number of times a sanctioning or rewarding procedure needs to be started. The fundamental reason is the fact that sticks are triggered by violation, while carrots are triggered by compliance.

If $p < 1$, sticks and precompensated carrots (just like carrots and precompensated sticks) no longer have identical payoffs because they treat agents who have not been monitored differently. The non-monitored agent's payoffs are $w^{normal}$ under a stick regime and $w^{precomp}$ under a carrot regime, which is clearly different. The payoffs for the monitored complying agent are $w^{normal}$ under a stick regime and $w^{precomp} + s$ under a carrot regime. These payoffs are identical only when precompensation would fulfill ($w^{precomp} = w^{normal} - s$), but if $p<1$, the appropriate amount of precompensation required to make the outcomes identical in expected terms is $ps$ (average precompensation). The same applies to the payoffs of the monitored violating agent which are $w^{normal} - s$ under a stick regime and $w^{precomp}$ under a carrot regime. Note that while under *average* precompensation the payoffs would be the same under both regimes *in expected terms* (the complying agent receives $w^{normal} = w^{precomp} + ps$, and the violating agent receives $w^{normal} - s = w^{precomp}$), there would still be differences in real terms with respect to transaction costs (the ex post transaction costs are lower under a stick regime since sticks need not be applied when the agent complies, see section 3.3) and risk allocation (a carrot regime creates risks for compliers, a stick regime risks for violators, see 3.4). For an application and further numerical illustration, see section 5.4.

---

complex concepts of precompensated carrot and precompensated stick).

[6] Note that the cases of probabilistic monitoring also include systems that are set up to be non-probabilistic but in which monitoring is less than perfect because evidence is less than perfect.

*2.5. Annullable carrots and sticks*

Another source of confusion lays in the existence of annullable carrots and sticks. Annullable sanctions are sanctions that are applied *unless* monitoring takes place and the agent is found to fulfil the condition for getting the (positive or negative) sanction. For instance, an annullable carrot is an amount (of $20 in the previous example) that the agent will receive, unless he has been monitored (with p = 0.1) and found violating. An annullable stick is a similar amount that the agent will have to pay, unless he has been monitored and found complying. In essence, *an annullable carrot is a threat to take back*, *an annullable stick a promise to give back*.

A real-life example of annullable carrot regimes is efficiency wage contracts (De Geest, Dari-Mattiacci and Siegers, 2006???). Under those contracts, employees are overpaid in order to give them an incentive not to shirk (since they lose the rents upon dismissal). Efficiency wages are annullable carrots because bonuses are paid to all employees, except for those who have been caught shirking (which is probabilistically monitored).

The difference between sanctions and their annullable variants consists in what happens when monitoring does not take place (???insert). At the surface, annullable carrots may look like carrots, but since they are algebraically identical to fully precompensated sticks they have all characteristics of pure sticks, except for some wealth effects and the danger of abusive rulemaking by non-benevolent principals. Hence, efficiency wages are analytically closer to sticks than to carrots.[7]

**Proposition 2**. *Annullable carrots are identical to fully precompensated sticks.*

| | Monitored Violators | Nonmonitored Agents | Monitored Compliers |
|---|---|---|---|
| Carrot | *0* | *0* | *+s* |
| Averagely Precompensated | *–ps* | *–ps* | *–ps + s* |

[7] Note that if the agent would pay entry fees to get efficiency wage contracts (as in Becker and Stigler, 1974), the result would be the precompensation of a precompensated stick, which is identical to a pure stick regime, except for the timing of the payments (i.e. the implicit bond posting at the start of the efficiency wages contract). To illustrate, assume that the market wage for a job with a certain effort level (which includes effort cost d which can be monitored at no cost and effort cost e which is costly to monitor and hence probabilistically monitored) is $100. Under an efficiency wage contract, the wage is set at $120, but if the agent is found shirking, the wage is only $100. This contract works with an annullable bonus: the agent gets a bonus of $20, unless he is found shirking. Now suppose that to get this contract, the agent pays an entrance fee of $20. Hence, if the agent complies or violates without being monitored, he gets a net wage of $120-$20=$100, while if he violates and is monitored, he gets a wage of $120-$20-$20=$80. This is algebraically identical to the payoffs in a simple stick contract.

| | | | |
|---|---|---|---|
| Carrot (precomp = *-ps*) | | | |
| **Fully Precompensated Carrot (precomp = *–s*)** | ***–s*** | ***–s*** | ***–s + s = 0*** |
| ***Annullable stick*** | ***–s*** | ***–s*** | ***0*** |
| Stick | *–s* | *0* | *0* |
| Averagely Precompensated Stick (precomp = *ps*) | *ps – s* | *ps* | *ps* |
| **Fully Precompensated Stick (precomp = *s*)** | ***s – s = 0*** | ***s*** | ***s*** |
| ***Annullable carrot*** | ***0*** | ***s*** | ***s*** |

*Table 1: sanctions, bonuses, and precompensation under the various regimes*

**Proposition 3**. *If p = 1, sticks are identical to annullable sticks, averagely precompensated carrots and fully precompensated carrots, and carrots are identical to annullable carrots, fully precompensated sticks and averagely precompensated sticks.*

This proposition can be proven by simply comparing the payoffs of table 2.

| | Monitored Violators | Monitored Compliers |
|---|---|---|
| Carrot | *0* | *s* |
| Averagely Precompensated Carrot (precomp = *–s*) | *–s* | *–s + s = 0* |
| Fully Precompensated Carrot (precomp = *–s*) | *–s* | *–s + s = 0* |
| ***Annullable stick*** | ***–s*** | ***0*** |
| Stick | *–s* | *0* |
| Averagely Precompensated Stick (precomp = *s*) | *s – s = 0* | *s* |
| Fully Precompensated Stick (precomp = *s*) | *s – s = 0* | *s* |
| ***Annullable carrot*** | ***0*** | ***s*** |

*Table 2: sanctions, bonuses, and precompensation under the various regimes (p=1)*

If monitoring takes place with certainty (p = 1), there is no difference between average and full precompensation. The differences between sticks and annullable sticks and between carrots and annullable carrots disappear as well, since these differences consisted only of what happens to non-

monitored agents. Note that (pure) carrots and sticks still have different distributional effects.

*2.6. Tests to determine whether a regime is a carrot or a stick regime*

We are now looking for objective tests to determine whether a regime is a sticks or a carrots scheme.

*(a) Transfer-to-whom test*. Since a carrot is a transfer to the agent and a stick a transfer from the agent, the most obvious test would be to analyze the direction of the transfer that effectively took place. The test may be unreliable, however, when the carrot or stick may be *set off* against another financial obligation. Suppose the agent is entitled to receive a carrot of $20, but she also owes the principal $100 tort damages. If both parties agree to set off, the agent pays $80 to the principal. Although the principal paid a carrot of $20, there was never any effective payment from the principal to the agent.

*(b) Normalcy test.* Aron and Olivella (1994) present a definition of bonuses and penalties in a context of contracts with non-probabilistic monitoring (in which case carrots can indeed be rewritten as algebraically identical sticks, see section 2.4). In their definition, bonuses are payments above the customary (i.e. more often observed) wage, and penalties are negative deviations from the customary wage. Their consideration is that, while the distinction between bonuses and penalties is an empty one from an economic point of view, people tend to take the normal outcome as their reference point.

While this definition is probably a valuable description of what people consider as bonuses and penalties, its weakness consists in the fact that it implicitly assumes that Wittman's normative argument – that the behavior of the majority should not be penalized or rewarded (see section ???) – is applied in all real-life situations.

*(c) Non-monitored agent test.* This test analyzes the risk allocation, and is based on the fact that carrots create risks for compliers, while sticks create risks for violators (section 3.4). The test consists in analyzing what would happen if the apprehension rate were lower than 1. If non-monitored agents get the same treatment as monitored violators, then we have a carrot. If non-monitored agents get the same treatment as monitored compliers, then we have a stick.

To illustrate, suppose that after being caught shirking a director is degraded to a lower position in the same firm. Is this a stick or just a withdrawn carrot? According to this test, the question is what the employee's position would have been if shirking had not been discovered. Since, in that case, the employee would still be director, the degrading should be considered as a stick. Similarly,

an economic boycott of, say, a merchant who violated some moral or religious norms should be seen as a stick rather than as a non-awarded carrot.

???Note that according to the non-monitored agent test, annullable bonuses like efficiency wages are sticks. Supra we have argued, however, that annullable bonuses are identical to fully precompensated sticks and hence closer to sticks than to carrots.???

## 3. Analysis under the basic framework

### *3.1. Basic model: benevolent principal with perfect information on the agents' compliance costs*

*Rule and sanction are chosen by benevolent principal*. We employ a simple framework, in which a principal (an employer, a parent, a state) exercises some authority deriving from the law or the use of force over a group of agents (employees, children, citizens) and decides on the rule that the agents have to follow and the sanction attached to this rule. Because of the principal's authority, the agent's participation constraint is always assumed to be fulfilled. We start by assuming that the principal is a *benevolent dictator*, who maximizes social wealth. In section 4.4, we will consider *non-benevolent dictators* and, in section 5.10, we will comment on *contractual principals*, thereby relaxing the previous assumption on the agent's participation constraint. The agents are unable to influence the principal's decisions through the use of threats, bribes, or signaling. (footnote: in this respect, our framework differs from e.g., Chang, see footnote ???).

*Probabilistic carrot or stick*. The principal monitors each agent with some positive probability ($0 < p \leq 1$).

*Carrots and sticks are monetary*. The principal applies a monetary (or monetary-equivalent) sanction, in the form of either a reward for compliance (a carrot) or a fine for violation (a stick).

*No ex post opportunism by the principal or the agent*. The agent is monitored with a probability $p$ as announced, the principal monitors in good faith, and the sanction is applied as announced, that is, the principal pays the rewards, and the agent pays the penalties. In section 4.5 we relax this assumption.

*Compliance cost*. The rule of conduct yields a positive compliance cost for each agent, i.e. effort cost $e$ (which is also the benefit of violation).

*Benefit of compliance*. Compliance does not generate a direct private benefit to the agent (the compliance cost is a net cost). Compliance may generate a benefit either for the principal or for

society at large. Since the agent is a member of society, she may receive a benefit from the other agents' compliance (for instance the agent may enjoy a clean beach when the other agents do not litter). Hence there are two ways for compensating the agent's effort costs: explicitly, through carrots, or implicitly, if they participate in the benefits of compliance.

*Discrete choice between compliance and violation*. For simplicity, we assume that the agent can either fully comply or fully violate. The agent will comply with the rule if the cost of compliance is less than or equal to the cost of violating the rule, given the probability of control and the magnitude and type of sanction.

*The principal has information on the agent's effort cost. All agents are identical.* There is only *one type of agent*, and the principal knows the agent's compliance cost. Since the principal knows the magnitude of the effort cost of the agent, the expected carrot or stick will be set to equal the compliance cost ($e = pS$ or $e = pC$). In an *extension* (*specification problems*, section 4???), we add a second type or agent, i.e. an agent who is not able to comply (as the compliance costs are infinitely high) and assume that the *principal's knowledge of the agent's type* is imperfect: the principal knows how many agents are able to comply, but does not know which ones.

*Rationality and risk neutrality*. Both the principal and the agent are rational, wealth-maximizing actors. The principal is risk neutral. The agent is risk neutral, except for in sections ??? in which we explicitly consider the consequences of risk aversion.

*No precompensation is paid*. ??? This assumption is made to simplify the comparison of carrots and sticks.(Note: see also footnotes???).

*Timing of the game*. At $t_0$, Nature assigns wealth $w^A$ to the agent and wealth $w^B$ to the principal. At $t_1$, the principal announces the rule, the type and magnitude of the sanction, and the probability of monitoring. At $t_2$, the agent complies or violates. At $t_3$, monitoring takes place with probability $p$, the sanctioning or rewarding procedure takes place, and sticks or carrots are paid.

*Wealth of the parties. Intragroup versus extragroup financing of carrots*. The principal has a limited budget available for paying carrots, equal to his wealth $w^B$. If the principal does not use his own resources, carrots can be financed either through intragroup taxes or through taxes collected outside the group.[8] For simplicity, we assume that under intragroup financing the taxes are paid by all

[8] Similarly, the term intragroup redistribution of sticks could be used when the sticks that have been paid are redistributed among all agents, and the term extragroup redistribution of sticks if they are redistributed among the group of outsiders. Under the basic framework that is consider in section 3, however, the distinction is meaningless since all agents comply.

agents, also those who comply with the rule. The alternative rule – letting only the violators pay – would imply that it would be impossible to finance the carrots when all agents comply.

### *3.2. Imperfect equivalence with respect to the agent's marginal incentives*

#### 3.2.1. Prima facie equivalence.

Carrots and sticks are *prima facie* equivalent mechanisms to induce an agent to comply with a rule of conduct, as summarized in the following proposition:

**Proposition 4 (incentive equivalence)**. *Any behavioral change that can be induced by a monetary stick can also be induced by a monetary carrot of the same magnitude and vice versa, if the agent is risk neutral, sufficiently solvent, and not saturated.*

In the basic setting presented in subsection 3.1, the agent's choice between compliance and violation is grounded in a simple comparison between the cost of compliance and the cost of violation. A carrot (probability times reward) reduces the cost of compliance, while a stick (probability times fine) increases the cost of violation. With risk neutral parties, if the carrot minus the effort cost (the cost of compliance) is greater than zero (the cost of violation) the agent will comply with the rule. Likewise, with sticks, the agent will comply if the effort cost is greater than the stick. It is easy to see that a stick is just a carrot brought from the left-hand side to the right-hand side of the equation, with a change in the sign. Given the probability of detection, a carrot and a stick of the same magnitude have exactly the same effect on the agent's behavior, providing for a strong equivalence argument that holds both in qualitative and in quantitative terms.

To illustrate, suppose, that the police know that an individual in their custody has valuable information on a crime. Suppose that the monetary equivalent of the cost of revealing that information is $900,000. They might offer a reward of $1,000,000 to obtain information; the same result could be reached by threatening him with a fine of $1,000,000 if he does not reveal valuable information.

This equivalence result is a fundamental implicit assumption behind the Coase Theorem, since re-allocating property rights also generates a switch from carrots to sticks or vice versa. Following Coase (1960), it does not matter whether the rancher has to pay for the pollution caused to neighboring property (a stick) or whether the rancher is paid for not polluting (a carrot); the final outcome will be the same. In other words, sticks for polluting ranchers and carrots for non-polluting

ranchers create the same incentive effect. The equivalence argument needs to be qualified on account of non-constant marginal utility of money and constraints on the maximum applicable sanction, but its qualitative implications remain valid.

### 3.2.2. Decreasing marginal utility of money (risk-aversion).

**Proposition 5**. *If an agent is risk averse, carrots need to be higher than sticks to induce the agent to do an effort with the same disutility. This holds also for non-probabilistic sanctions (p = 1).*

First we consider the case of probabilistic sanctions (p < 1). Risk aversion makes a probabilistic monetary carrot less attractive, and a probabilistic monetary stick more severe. To illustrate, consider a risk-adverse agent who has a 10% chance to receive a carrot or stick of $20. In net expected terms, the carrot may have a value of only $1.8 for the agent, while the stick may have a net value of $2.2. Hence, in absolute monetary terms, a carrot will need to be larger than a stick to have the same effect. The relative difference becomes larger as the carrot or stick become higher.

Second we consider non-probabilistic sanctions (p = 1). Risk-aversion implies that the marginal utility of money for the agent is decreasing. This implies that losing a sum provides a greater variation in utility than gaining the same sum. In this case, a carrot and a stick of the same magnitude do not create the same incentive effect, and the magnitude of the carrot will have to be increased so to generate the same incentives as the stick.[9] This difference comes down to the difference between willingness-to-accept (for carrots) and willingness-to-pay (for sticks), which in turn is a consequence of a wealth effect. The agent's wealth will be higher after a carrot scheme than after a stick scheme, because he gets compensated for his effort under the former and not under the latter.[10]

### 3.2.3. Incentives undermined by insolvency and saturation. Differences with respect to the maximum magnitude of the sanction, the minimum monitoring level, the maximum effort the agent can be induced to make.

---

[9] The same result holds in case of *loss aversion*. Note that in the inverse but unlikely case of increasing marginal utility, sticks will need to be higher than carrots.

[10] Hence if sticks are precompensated, the difference when p=1 will disappear. The difference when p<1, however, remains also when sticks are (averagely) precompensated.

**Proposition 6.** *The maximum monetary stick $S^{max}$ is determined by the wealth of the agent $w^A$, the maximum monetary carrot $C^{max}$ is determined by the wealth of the principal $w^P$.*

The maximal magnitude of a stick is determined by what the agent can pay, the agent's wealth $w^A$. A fine larger than his wealth would simply result in insolvency.[11] Likewise, the use of carrots is constrained by the principal's budget $w^P$. Note that if the principal has several agents, the maximum carrot per agent is $w^P$ divided by the number of agents.

The fact that $S^{max}$ will usually differ from $C^{max}$ is another source of nonequivalence between carrots and sticks. *Given some p, only one of them may be effective.* Of course, a possibly lower sanction can be compensated for by a higher probability of detection. But then the nonequivalence takes the form of *different minimum monitoring levels*.

**Proposition 7**. *Happy agents can be induced to higher efforts by sticks than by carrots. Unhappy agents can be induced by higher efforts by carrots than by sticks.*

For the sake of this (and the following) proposition, we express the required effort cost in terms of utility (instead of in monetary terms). We denote the maximum utility as $u^{max}$, and the actual happiness of the agent at $t_0$ as $u^{t0}$. We define a *happy agent* as an agent with $u^{t0} > u^{max}/2$, and an *unhappy agent* as one with $u < u^{max} / 2$. For simplicity,[12] we assume that the highest possible stick reduces the agent's utility to u = 0 and that highest possible carrot leads to u = $u^{max}$. In addition we assume that the agent's actual happiness has no effect on $e$.[13]

The idea here is that there is a maximum disutility the agent can be induced to incur and that this disutility is different for carrots and sticks.

Since the sanction will be equal to the effort cost divided by $p$, this proposition can also be restated as follows:

---

[11] Here we assume for simplicity that the agent's wealth $w^A$ is not reduced by the effort cost $e$, so that the agent's wealth after $t^2$ equals his initial wealth.

[12] In reality, legal systems put some limits on the maximum sticks that can be used, both with respect to physical punishments (e.g. limits on torture), as with respect to financial sanctions (e.g. limits on the assets that can be seized in case of bankruptcy, financial help for the poor). If the maximum stick corresponds to $u^{legalmin} > 0$, our definition of happy agents needs to be adapted to $u^{t0} - u^{legalmin} > (u^{max} - u^{legalmin})/2$ or $u^{t0} > (u^{max} + u^{legalmin})/2$. By the same token, the legal system may not be able to offer maximum happiness (given the limited total wealth in a society if utility depends on monetary goods, or given the impossibility to transfer nonmonetary goods, e.g., to give love to the agent). In that case, the definition of happy agents needs to be further adjusted.

[13] In case happiness would have a positive effect on e, the results need to be qualified.

**Corollary**. *Under a stick system, happy agents need to be monitored (as much as or) less than unhappy agents. Under a carrot system, unhappy agents need to be monitored (as much as or) less than happy agents.*

This follows from the fact that the maximum carrot applicable to happy agents is less than the maximum stick, thus the probability of monitoring must be larger. The same logic applies to unhappy agents.

**Proposition 8**. *Sticks can be undermined by the agent's insolvency, carrots by the agent's saturation.*

That sticks can be undermined by the agent's insolvency has been discussed extensively in the literature (e.g., Shavell, 1987, Dari-Mattiacci and De Geest, 2005). That there is an inverse problem for carrots – saturation – has not been previously discussed. Still, every dog trainer knows that a dog with a full belly may no longer respond to candy, while the same saturated dog will still respond to sticks.

Saturation could be defined as follows. An agent cannot be incentivized by a carrot because of saturation if the disutility of the required effort is higher than the utility that an infinitely high carrot could generate. By extension, one can speak of a saturation problem if the agent is not willing to exert the effort for the maximum carrot that the principal is able to offer. Full saturation can be defined as the situation in which an agent no longer responds to carrots, not even for an infinitely small effort cost, because the agent completely satisfied with his actual wealth. Full saturation implies that the marginal utility of money equals 0, or is even negative.

To put it different, if sticks do not work because of insolvency of the agent, a carrot may work and vice versa, if a carrot does not work because of saturation a stick may work.[14]

### *3.3. Differences with respect to transaction costs*

Transaction costs (or 'administrative costs') consist of monitoring costs, sanctioning procedure costs, and the costs of payment. Since in principle the monitoring levels will be equal under carrots and sticks (except for the qualifications made in subsection 3.2), the monitoring costs will be equal

[14] Note that in theory it could occur that an agent responds neither to sticks or carrot. For example, under some old (hard) bankruptcy regimes, some debtors were left without any assets, and faced an implicit tax of 100% on newly acquired financial resources, so that they no longer responded to financial incentives.

as well. Hence, in this subsection we focus on the other components of transaction costs. Sanctioning procedure costs may include the costs of collecting hard evidence (for instance, making a picture of the car that is driving too fast), collecting additional evidence (for instance, on the identity of the driver or on the driver's offence history), interpreting counter-evidence (for instance, on the driver being in an emergency situation), and determining the magnitude of the sanction or reward (if the magnitude depends on the severity of the offence). These sanctioning costs (and the costs of payment itself) depend on whether the carrots or sticks are paid. In case of carrots, these costs need to be made when the agent is found complying. In case of sticks, these costs need to be made when the agent is found shirking.

**Proposition 9**. *Carrots generate transaction costs in case of compliance, sticks in case of violation.*

This proposition follows directly from *characteristic difference 1* (see subsection 2.2): sticks are applied upon violation, a carrots upon compliance.

**Proposition 10**. *If the Principal is perfectly informed about the agent's effort costs (so that she sets the expected sanction high enough to make the agent comply), transaction costs are lower under sticks*.

If the principal knows the agent's effort costs, so that the chosen expected sanction is high enough to be effective, the rational agent will comply, and hence the transaction costs related to the sanctioning procedure and the payment of the sanction are avoided.

Note that the fundamental reason why sticks are superior in this respect is the fact that *compliance is the normal outcome*. More generally, absent serious informational shortcomings at the principal's side (see section 4.1), compliance can be presumed to occur more often than violation, because compliance is the purpose of the rule.

*3.4. Differences with respect to risks*

Risk becomes an important issue when agents are risk averse and monitoring is probabilistic ($p<1$), so that penalizing or rewarding becomes probabilistic as well. Also with respect to risks, sticks are superior within our basic framework, as summarized by the following propositions.

**Proposition 11**. *Sticks create risks for violators; carrots create risks for compliers.*

**Proposition 12**. *If the Principal is perfectly informed about the agent's effort costs (so that she sets the expected sanction high enough to make the agent comply), the risk costs are lowest under sticks.*

The proof is similar as for the previous proposition. Under sticks, there is a transfer of wealth from the agent in case of violation. Under carrots, there is a transfer of wealth to the agent in case of compliance. Hence, if sticks are probabilistic sticks, the risk of the stick is borne by the violators and the risk of the carrot is borne by the compliers. If the principal is perfectly informed about the agent's effort costs, the expected sanction will be set high enough: there will be only compliers. As a result, risk costs are lower under sticks.

Proposition 10 shows that carrots have an undesirable property: they create risks for the 'good' agents. In contrast, sticks create risks for the 'bad' agents, and since no agent has an incentive to violate, sticks create no risks at all – at least not in the framework we are considering here, in which the principal is perfectly informed about their agents' effort costs.

Note that when principals have to finance the carrots, a carrot system is more risky for them as well. Even if the principals would receive the penalties under a stick system, such a system would not create any risk at all since all agents comply (at least within the current framework).

### *3.5. Differences with respect to distributional side-effects*

The purpose of a penalty or reward is to incentivize agents to comply with a rule; its purpose is not to change the distribution of wealth in society. However, under some conditions the effect of a penalty or reward may be that some agents become poorer or richer than other agents. The effect may also be that some outsiders get poorer (by paying carrot taxes) of richer (by receiving benefits without compensating the agents' effort costs).

To illustrate, consider a duty to keep the beach clean that applies to all 1000 visitors of a beach. The rule requires an effort cost of $2 from each individual, but if all individuals do so, the rule also generates a benefit of $3 per individual. If sticks are used (with an expected value of at least $2), the result is that the individual wealth of all members of the group increases with 1$. While the agents do not receive explicit compensation for their effort costs, they receive in-kind compensation in the form of the benefit of the rule. In contrast, if carrots are used and these carrots are financed by outsiders (the taxpayers of that state, the principal herself…), then the anti-trash rule makes the agents richer at the expense of these outsiders. Now change the example and suppose that instead of requiring an effort from all agents, the best way to keep the beach clean is to ask an

effort of $1000 from one specialized professional. If a stick is used, that professional will be made substantially poorer. If a carrot of $1000 is used, the rule will not change his wealth.

Distributional distortions are socially costly for two reasons. First, assuming that we can conceive of an optimal distribution of wealth in society, carrots and sticks directly social welfare by altering such distribution. Second, they often distort ex ante incentives, just like taxes and the prospect of economic rents do.

For the sake of our analysis, we measure the degree of distortion as the sum of the (absolute values of the) overpayments (rents) or underpayments of producing agents, and overtaxation and undertaxation of outsiders. To determine what are rents and underpayments, we compare the compensation that producers receive with their true effort costs. The underlying idea – which is generally accepted in welfare economics – is that all surpluses should fall at the consumers' side, and not at the producers' side. Hence, the agents are perfectly compensated for their effort costs, and no agent or outsider is overtaxed. That is, those who do the work are perfectly compensated for their work, and those who receive the benefits bear the full costs of producing these benefits.

For instance, if two agents receive a compensation of $2 each for their effort, but the costs of effort are $1 for one agent and $3 for the other, than the distributional distortion is $2 (i.e. $1 rent and $1 undercompensation). In contrast, if an agent whose effort costs were already sufficiently compensated receives a $2 carrot financed by an outsider, then the outsider's wealth is distorted by $2 (overtaxation), and the agent's wealth by $2 (rent).

Note that in this subsection we do not consider distributional changes that are the side-effect of probabilistic carrots (as when only 10% of the agents get a carrot that is 10 times higher than their effort costs). While this probabilistic nature may make some agents – those who were lucky enough to get monitored – their distributional effects have indirectly been analyzed in the previous section on risks. ???Therefore we assume for simplicity throughout this subsection that p=1???. Since in the basic framework of the current section all agents are identical with respect to their effort cost and their benefit, the only distributional distortions that remain to be analyzed are the inter-group distortions, that is, distortions between the group of agents on the one hand and the group of outsiders on the other hand. (In section ???, we also consider intragroup distortions that may occur when agents vary with respect to effort cost or received in-kind benefit).

**Proposition 13 (intragroup financing of carrots)**. *If carrots are financed solely by the group of agents, then the complying agent's expected wealth is the same with carrots as with sticks.*

Since the principal knows e, all agents comply. Under a stick system all agents pay e. Under a carrot system all agents receive on average pC – e = 0. Therefore, if the carrots are fully financed by outsiders, each agent is on average *e* wealthier under carrots than under sticks. If the carrots are fully financed by the agents through a pro capita tax *t* equal to average carrot, there is no longer any difference between carrots and sticks since all agents pay on average e + t + pC = e.

Note that this holds irrespective of who benefits from the agents' compliance with the rule.

**Proposition 14**. *If the carrots are at least partly financed by extragroup taxes, agents are on average richer with carrots than with sticks.*

Note that this also holds irrespective of who benefits from the agents' compliance with the rule.

**Proposition 15 (Inter-group distortions).** *If the carrots are at least partly financed by extragroup taxes and if all benefits go to the agents (and agents are identical with respect to e and b) then sticks cause less inter-group distortions than carrots. If all benefits go to outsiders, then carrots cause less inter-group distortions than sticks.*

If all benefits go to the agents, and the agents have identical e and b, then sticks do not distort. Carrots, on the other hand cause inter-group distortions as soon as they are partly financed by outsiders (because these outside taxpayers are made worse off by the rule). If all benefits go to outsiders, sticks distort by making the agents poorer.

## 4. Extensions

### *4.1. The principal has imperfect information on the agents' compliance costs*

In this subsection we relax the assumption that the principal has perfect information on the agent's compliance costs. The general point of this subsection is that *if the principal is less informed* (which is the case as societies get *more specialized* with respect to *skills*) *it becomes more likely that carrots will be used*.

In section 4.1.1 we analyze a simple case in which some agents are unable to comply, and all agents who are able to comply have identical costs. In section 4.1.2 we briefly analyze the mirror image, i.e., the case in which some agents are unable to violate the rule. In sections 4.1.3 and 4.1.4

we analyze the case in which agents have varying compliance costs.

### 4.1.1 Some agents are unable to comply

Here we assume that there are two types of agents, one with a compliance cost *e* lower than the social benefit of the rule, and one with infinitely high compliance costs. In other words we assume that some agents are able to comply with the rule at reasonable costs, and some others are simply unable to do so. The principal knows the distribution and compliance costs of the types, but does not know the type of an individual agent.

For instance, suppose that in the example about revealing information about a crime of section 3.2 the police do not know which are the 100 people out of 60 millions citizens who have valuable information. All 100 potential compliers have identical compliance cost, which is assumed to be equal to $1 million. The state may create the proper incentives either by offering a $1 million reward to those who come up with information, or by giving all 60 million citizens a $1 million fine, except for those who come up with information. Note that if the state would know who those 100 informed people were, sanctions would be directed towards them only. In this case, the majority (even all of them) would comply, so that a stick would be the optimal sanction.

As a result of this lack of information, the group of non-compliers (those who do not offer information to the police) consists of two types of agents: those who possess information but do not offer it, and those who do not possess information because they are unable to comply with the rule. Since the principal is unable to distinguish between the two, they need to be treated equally, either by giving none of them a carrot, or by giving both of them a stick.

**Proposition 16**. *If the principal does not know which agents are able to comply, a stick regime may sanction agents who are unable to comply. The larger the proportion of agents who are unable-to-comply, the greater the transaction costs, risk costs, and distributional side-effects generated by sticks*

The larger the group of unable-to-comply agents, the more sticks lose their superiority with respect to transaction costs and risk costs. In our basic framework with full information on the agent's compliance costs, sticks generate no transaction costs (in the sense of sanctioning costs) and no risk costs because no agent violated the rule. Sticks were applied to 0% of the monitored agents, while carrots were applied to 100% of the monitored agents. In the framework we are now considering,

some agents do violate the rule because they are simply unable to do otherwise. As a result, the cost of administering a sanction decreases if carrots are applied when the majority violates the rule and sticks are applied when the majority obeys the rule. This normative proposition can be summarized as follows.

**Proposition 18 (Wittman's rule)**. *A stick should be used if the majority of agents comply with the rule, and a carrot should be used if the majority of the agents violate the rule (after introducing the sanction).*

This rule could also be summarized as follows: don't sanction or reward the majority. The rule was first translated into an economic framework by Wittman, although in essence it simply restates the old idea that customary behavior should not be sanctioned or rewarded (which can be found in e.g., Ellickson, 1973). Wittman's (1984) fundamental explanation was grounded in saving transaction costs. The less often sanctions have to be applied, the less often transaction costs will have to be incurred. The idea is that every applied sanction is in essence a transaction. To illustrate, should a teacher underline and deduct marks for wrong answers (sticks) or underline and give marks for correct answers (carrots)? If the average student has more than 50% answers correct, then underlining (and deducting points for) wrong answers takes less time.

It should be remarked that with only two types of agents, there is no difference between carrots and sticks with respect to distributional distortions. In the example on the information about a crime, compliers receive a $1,000,000 reward under a carrot system, but also incur an effort cost of $1,000,000 for coming up with the evidence (otherwise the reward would have been lower); violators (who do not come up with evidence) receive no reward but incur no costs either. Under a stick system, compliers incur an effort cost of $1,000,000 while violators incur a $1,000,000 fine. Formally, under a stick system the able agents pay e, while the unable pay $pS = e$. Under a carrot system the able agents receive $pC - e = 0$ while the unable receive 0.

**Proposition 19**. *A benevolent principal who is fully informed about the agent's compliance costs will only use sticks (in the absence of distributional side-effects).*

#### 4.1.2 Some agents are unable to violate

**Proposition 20**. *If the principal does not know whether the agent is able to violate, a carrot regime*

*may reward agents who are unable to violate. The higher the proportion of agents who are unable to violate, the greater the transaction costs, risk costs, and distributional side-effects generated by carrots.*

Many criminal prohibitions deal with acts that not everyone is able to commit, either because he is unable (e.g., hacking) or because he does not have the opportunity to do so (e.g., corruption and insider trading). In these cases the problem is not that some agents are unable to comply, but that some agents are unable to violate. While the former weakened the case for sticks, the latter weakens the case for carrots, because carrots may reward agents who have no merits. Paying a carrot for not hacking to someone who is not a computer expert makes no sense; it does not improve incentives, while it increases transaction costs, introduces risk and causes distributional distortions in a larger group of agents.

**4.1.3 Agents have varying compliance costs**

To illustrate, suppose that a sinking ship's load (with a value of $1,000,000) can be rescued at a cost that varies between $0 and $100,000, depending on the rescuer's skills. Denote the highest value of $e$ as $e_{max}$, and assume that $e$ is uniformly distributed between 0 and $e_{max}$, so that the average rescue cost $e_{av}$ is equal to $e_{max}/2$.

If carrots and sticks are *general*, i.e., if they have the same magnitude for all agents (irrespective of their effort costs) and they are high enough in incentivize all agents (i.e. they are equal to the maximum effort cost), then they lead to the same distributional distortion. Under general carrots, the overpayment is on average $e_{max}/2$ (all agents receive $e_{max}$, which is the lowest amount under which all agents comply, and incur on average a cost of $e_{max}/2$). Under general sticks, the agent does not receive anything, but loses between $0 and $100,000, and his underpayment is on average $e_{max}/2$. Note that if sticks are made specific, i.e. if the magnitude of the stick would be adapted to the effort cost (so that for instance an agent with an effort cost of $25,000 would be threatened by a $25,000 stick), the result would be the same. The fundamental reason is that sticks are not paid, so that their magnitude does not affect the distributional result.

If carrots are *specific*, however, so that for instance an agent with an effort cost of $25,000 would be promised a $25,000 carrot, the distributional distortion disappears (agents are no longer overpaid, and nor are they getting underpaid).

If we consider costs that are not uniformly distributed, the average compliance cost is not

necessarily equal to $\boldsymbol{e_{max}}/2$. If the average compliance cost is relatively low ($e_{av}<e_{max}/2$), general sticks generate less distributional distortions than carrots (as the average underpayment under sticks is $e_{av}<e_{max}/2$, while the average overpayment under carrots is $e_{max}- e_{av}$ which is higher than $e_{max}/2$). Conversely, general carrots will distort less when the average compliance cost is relatively high ($e_{av}>e_{max}/2$). These results of can be summarized as follows:

**Proposition 19**. *If agents have varying effort costs, general sticks distort less than general carrots if the average effort cost is relatively low. The opposite holds if the average effort cost is relatively high. Perfectly specific carrots do not cause distributional distortions, while specific sticks distort as much as general sticks.*

*4.2. Agents derive varying benefits from the rule*

In this subsection we relax the assumption that the principal has perfect information on the agent's benefits. The general point of this subsection is that *as agents differ more with respect to tastes it becomes more likely that carrots will be used*, if the principal has sufficient information on these benefits to introduce a benefits-based tax system.

We assume for simplicity that all benefits of the rule go to the agents, and not to outsiders.

**Proposition 13**. *The more unequal the individual rule-benefits are, the more likely it is that carrots distort distributions less than sticks. If benefits vary so that some agents have lower benefits than effort costs, then carrots and sticks have equal distortions unless carrots are financed through benefit-based taxes. In this case carrots distort less.*

To illustrate, consider a rule that obliges all beach visitors to spend some effort to keep the beach clean. Suppose that all 1,000 individuals have identical effort costs ($2). Now suppose that no one benefits from having a clean beach, except for one individual, who cares enormously about it ($3,000). Under a stick system, the wealth of that individual is significantly increased, while the wealth of all other agents is slightly reduced. Under a carrots system, these distributional changes can be reduced if the benefiting agent is the one who has to finance the carrot to all other individuals.

To prove this proposition, assume for simplicity that all agents have identical effort costs (and that all agents are able to comply and violate). The change in individual wealth is determined by the

benefit plus the received bonus payment minus effort costs and the individual tax. Under sticks, the wealth decreases if the benefit minus effort costs is negative. This implies that the more individuals have a benefit that is lower than their effort costs, the higher the number of individuals whose wealth is decreased by an overall wealth-increasing rule. Under a carrot system, if the bonus is financed equally by all individuals, the outcome is not changed. But if the tax is at least partly a function of the received individual benefits, the losers lose less (or even nothing at all) and the winners win less, and hence the existing distribution is less altered. Note that if $e$ and $b$ vary but $b/e$ is constant, then sticks do no longer have distributional distortions.

*4.3. The principal can only observe either compliance or violation (one-sided evidence)*

In our basic framework, if monitoring takes place, the principal finds out whether the agent complied with or violated the rule. For some rules, however, it may only be possible to prove violation, and not compliance or vice versa.[15] For instance, it may be possible to find evidence that an agent has been corrupt, but in case no such evidence is found this does not prove the agent was honest. Proving that the agent was honest during a certain period may be prohibitively costly (as it would require monitoring the agent 24h/24h during that period). Many other criminal acts have the same characteristic. For instance, it is not possible to prove at a reasonable cost that an agent did not commit a murder, or did not rob a bank. In such cases, the only sanctions that are effective are sticks, and annullable carrots. Pure carrots do not work: since the carrot is given upon compliance and this cannot be proven. We formulate the following more general proposition.

**Proposition 20** *If only compliance can be proven, then only carrots are effective. If only violation*

[15] Note that there is some relationship between one-sided evidence and lack of information on the agent's type (specification problems). Both are informational shortcomings: in the first case, the principal may have no information on whether the agent complied, in the second case the principal may have no information on whether the agent was able to comply. Both informational shortcomings can be closely related. Consider again the mafia example, but redraft the rule as 'all Italians are obliged to give information on the mafia if they have such information'. If an Italian comes up with information, he proves that he complies with the rule, because the fact that he gives information proves that he had such information. If an Italian does not come up with information, this does not prove that he violated the rule, because he might no have such information. Therefore, this rule – the way it is formulated – would be affected by one-sided evidence: only compliance can be proven. As a consequence, only carrots could work. In our original example the rule was "all Italians are obliged to give information on the mafia". This rule is not affected by one-sided evidence: compliance is proven by giving information, and violation is proven by not giving information. Both carrots and sticks are effective. While the one-sided evidence problem has disappeared, another type of information problem – lack of information on the agents' type – is introduced. Carrots may be the best rule for transaction cost reasons (assuming that only a minority has information on the mafia).

*can be proven, then only sticks are effective.*

To understand why, consider the following table.

| | Monitored violators | Non-monitored | Monitored compliers |
|---|---|---|---|
| Carrot | 0 | 0 | +C |
| Stick | -S | 0 | 0 |

If compliance can be proven, then the column of the 'monitored violators' no longer exists. By comparing the two remaining columns – the non-monitored agents and the monitored compliers – it become clear that a stick regime have no effect on the agents' incentives, because in both cases the agent remains at 0. A carrot regime does generate incentives to comply. The opposite holds when only violation can be proven. In summary, carrots are linked to compliance, and sticks to violation. Hence, if compliance can never be proven, carrots make no sense, and the same holds for sticks when violation can never be proven.[16]

*4.4. The principal is non-benevolent ex ante*

**Proposition 21** *Sticks involve a higher risk of abuse by a non-benevolent principal: if the Principal's decides on the rule, rules enforced by carrots are always Pareto improvements, while rules enforced by sticks are not necessarily Kaldor-Hicks improvements.*

**Proof.** Since the principal decides on the rule (including its positive or negative sanctions), we may reasonably assume that he is better off by the rule (though for a benevolent principal the benefit may derive from altruism). The agent is the one who decides on whether to comply with or violate the rule.

Under a carrot, if an agent violates the rule, he is not worse off compared to the situation before. As a matter of fact, violating the rule is choosing for the status quo. If he decides to comply with the rule, it must be because the expected carrot is larger than the effort cost, which implies that he is better-off compared to the status quo. As a result enforcement is a Pareto improvement.

The agent's situation is different for a stick. If the agent violates the rule, he is worse off

[16] Note that since the annullable variants of carrots and sticks are linked to the opposite behavior (carrots are annihilated upon violation, sticks are annihilated upon compliance), one of these will work in case of one-sided evidence (annullable carrots in case only violation can be proven, annullable sticks in case only compliance can be proven).

compared to the status quo since he will bear a negative sanction. If the agent complies with the rule, this means that compliance is more beneficial than violation (i.e. that the effort cost *e* is lower than the expected sanction *pS*), but it does not necessarily mean that compliance is an improvement over the status quo (since e may be larger than the agent's individual benefit from the compliance of other agents in society). Therefore, enforcement under sticks is not necessarily Pareto efficient, and even not necessarily Kaldor-Hicks efficient.

From proposition 21 it does not inevitably follow that enforcement under sticks yield socially inefficient results but the result points out a sticks' perilous ability to foster exploitation. Unsurprisingly, even when sticks are costly to apply, they are largely preferred to carrots in master-slave or criminal-victim relationships, where exit is generally not available.

Could we say that a rule with a carrot is an invitation to act, whilst a rule with a stick is an order to act? Is it more respectful to incentivize others by carrots than by sticks? This viewpoint is incorrect in the sense that both carrots and sticks create an amount of pressure on an individual. The viewpoint is correct in that carrots have a built-in participation constraint: the agent who chooses to violate remains at his status quo position. Carrots allow violating agents to preserve the status quo.

*4.5. The principal or the agents are opportunistic ex post*

A first type of ex post opportunism consists in the principal falsely pretending that the agent did not comply, so that either the agent has to pay a penalty to the principal, or that the principal does not have to pay a reward. Since the enrichment for the principal would be the same in both cases, the danger is similar.

**Proposition 22** *If an opportunistic principal receives or finances the payments, then the danger of opportunisms (consisting in unjustly to appropriate carrots or wrongfully to apply sticks) is equally high under carrots and sticks*

Another type of ex post opportunism consists in the principal not paying the carrot or the agent not paying the stick.

**Proposition 23**. *With sticks the agent has incentives to become a disappearing defendant; with carrots the principal has incentives to become a disappearing defendant*

In the economic analysis of tort law, *"disappearing defendant"* is a concept used to describe a situation where the injurer's incentives are undermined by his ability to escape the legal system ex post. The concept can be extended, not only to a contractual context (where the party that performs last can be the disappearing defendant) but also to a law enforcement context (as we are considering here). The disappearing defendant problem can be seen as a type of ex post opportunism, where the one who has to perform last behaves opportunistically.

Under a stick system, there is a danger that the agent will not pay the penalty. Under a carrots system, there is a danger that the principal will not pay the reward. The type of the sanction indeed determines the direction of the payment at $t^3$, and hence it also determines which party may have a chance to act as a disappearing defendant.

Note that both carrots and sticks can be undermined by the disappearing defendant problem. The extent to which the incentives are affected by this problem, however, depends on which party is least trustworthy. In that respect there is again an imperfect equivalence.

## 5. Applications

### *5.1. Sticks for negative externalities, carrots for positive externalities?*

In principle, the sign of the externality does not automatically affect the sign of the sanction, because any behavioral effect of a stick can also be obtained by a carrot (proposition 4). For instance, subsidies can also be used to prevent pollution. As a matter of fact, in case only a majority of plants can efficiently prevent pollution, working with carrots is even desirable to minimize transaction costs and risk costs (Wittman's rule, proposition ???, and proposition ???).

In reality, the choice is more complex because the type of sanction for preventing negative externalities also affects the incentive to start the potentially negative-externality-generating activity. Hence, a single sanction enforces two rules (to optimally start the activity and to optimally pollute). To put it different, two sanctions for two rules are set-off against each other, so that a single sanction remains. To illustrate, suppose that of 1,000 firms only 10 should start a potentially polluting activity, and that of those 10 only 3 should invest in technology that completely prevents pollution (though the principal does not know who should start and who should pollute). If only the decision to prevent pollution were taken into account, carrots should be paid to the pollution-preventing minority (to save transaction costs and risk costs). But with respect to the decision to

start the activity, sticks should be paid by the starting minority rather than carrots to the non-starting majority. If both rules sanction systems are set-off against each other, there is a choice between letting the 7 polluters pay a stick or paying a carrot to the 990 non-starter plus the 3 non-polluting starters. The stick regime minimizes transaction costs and risk costs, notwithstanding the fact that the polluters form a majority within the group of starters.

Since the group of non-starters will generally be the largest one, the result is that negative externalities should be treated with sticks. For similar reasons, positive externalities will generally be treated with carrots.

The results of this subsection are already in Wittman (1984) who argued that in general, sticks (liability) should deal with negative externalities and carrots (restitution) should deal with positive externalities (though the case of positive externalities is less clear because transaction costs and incentives for activity levels lead to different directions). Our findings on risk and distributional effects generally lead to the same conclusions as Wittman's transaction-cost-based conclusions.

### *5.2. Carrots for omissions (violations of obligations to do), sticks for actions (violation of obligations not to do)?*

In principle, both carrots and sticks work for both types of obligations. Obligations to do something can also be enforced by sticks (e.g., report information on crime), while obligations not to do something can also be enforced by carrots (e.g., an apple for children who do not complain). This follows from the incentive equivalence of carrots and sticks (proposition 4).

In many cases, however, there will be one-sided evidence. The general reason is that if an agent did something, this proves that he was able to do it, while if an agent did not do anything, this does not prove anything on his ability to do it. Hence, obligations to do something (like report information on crime) will have to take into account that some agents are unable to comply. If these agents are numerous, the case for using carrots becomes stronger (because sticks may sanction unable agents). Obligations not to do something (like hacking into a computer system) will have to take into account that some agents are unable to violate. If these agents are numerous, the case for using sticks becomes stronger (because carrots may reward unable agents). Therefore, carrots will more often be used for obligations to do something, than for obligations not to do something, and sticks will be used relatively more often for obligations not to do something.

*5.3. The incentives to reveal information, to self-report, to collect evidence and to litigate*

*Incentives to reveal information or to self-report* can be given by either carrots or sticks, as both generate the same marginal incentives. However, if the principal does not know which agents possess some information, it will nearly always be the case that only a minority of the group of possible information possessors effectively does possess the information. Hence, carrots will generally be superior with respect to transaction costs and risks. This is illustrated by the example on revealing information on a criminal act: it is better to give a $1 million carrot to 100 people than to give all 60 millions citizens a $1 million fine.

For self-reporting the answer is more complex because incentives to self-report may indirectly influence the incentives to commit the crime in the first place. The situation is similar to the one discussed in section 5.1: whether pollution is sanctioned or pollution abatement rewarded affects the decision to start the activity. This means that a single criminal sanction is used to create incentives for two different rules: not to commit a crime, and to self-report after having committed a crime. The carrot for the latter (additional years of freedom) is set-off against the stick for the former (additional years of imprisonment).

With respect to the incentives to litigate over the principal's sanctioning or rewarding decision and to collect information there is in principle no difference either, since the incentive is a function of the amount at stake which is equal for sanctions of an equal magnitude. If carrots are used, agents have an incentive to collect evidence that proves they are entitled to a carrot. If sticks are used, agents have an incentive to collect evidence that proves they are not subject to a stick. The only difference (which will be further discussed in section 5.6) consists of the fact that a non-monitored agent has an incentive to come up with evidence of his compliance under a carrots system, while he will not come up with evidence of his violation under a sticks system. The two things are different: failing to prove violation does not prove compliance and vice versa.

*5.4. Can legal restrictions on sticks be circumvented by rewriting them as carrots?*

The common law forbids penalty clauses, but does not forbid bonuses or rewards that are extra-compensatory. Can the penalty doctrine be evaded by rewriting a penalty scheme into a bonus scheme? Cooter and Ulen (2003, p. 252-253) argued that the law's refusal to enforce penalty clauses merely creates incentives to re-draft identical contracts with bonuses for performance instead of penalties for breach. Katz (2004) argued that the penalty doctrine may be evaded by

providing for an up-front payment that is characterized as the purchase of an option. He also made the point that the law indirectly limits bonus contracts by regulating deposits and performance bonds (if the prepayment is characterized as a deposit or performance bond, the law will generally limit the extent to which it can be made non-refundable).

Similarly, the law generally restricts penalties but does not restrict bonuses in employment contracts. Can these restrictions be circumvented by rewriting penalty schemes into bonus schemes, as Lazear (1991) suggested?

In section 2.4 it was shown that a stick can be rewritten as an algebraically identical (precompensated) carrot only if the stick is non-probabilistic ($p = 1$). Suppose that the price of a service is \$100, but to deter breach (which consists in delivering an inferior service) there is a penalty clause of \$90 in the contract (while the real damage is only \$10). Suppose that the inferior quality of the service can only be proven in 50% of the cases. Under a penalty regime, the seller receives \$100 in case of performance and on average \$55 (i.e., \$100 – 0.5(\$90)) in case of breach. Under a bonus regime that is made identical in expected terms, the base price is set at \$55 and a bonus of \$90 is promised in case a good quality service is delivered, which can also be proven in only 50% of the cases. Under this bonus contract, the seller receives on average \$100 (i.e., \$55 + 0.5 x \$90) in case of performance and exactly \$55 in case of breach. While the payoffs are equal in expected terms, the risk allocation is different (the penalty contract is risky for breachers, while the bonus contract is risky for performers), just as the transaction costs (which are higher under the bonus contract in case of performance). Note that the transaction cost difference holds even if sanctions are non-probabilistic.

### *5.5. Public good provision based on carrots or sticks?*

**Proposition 24**. *Sticks allow for contributions to a public good in the form of services, while carrots require the contribution to be in the form of a tax.*

Consider again the duty-to-keep-the-beach-clean example. In the first case, a stick system obliged all 1,000 visitors of the beach to do an effort of \$2, but they received an in-kind compensation of \$3 that consisted in enjoying a cleaner beach. Since keeping the beach clean can be seen as a public good, the example illustrated that sticks allow for public good provision without a tax system. In the case of the professional cleaner who received a carrot to do an effort of \$1000, the same public

good was provided through a tax system. Carrots require taxes for the simple fact that they need to be financed.

This again reinforces the statement that in older, less complex societies with less specialized markets sticks are more likely to be used. Sticks are an old and cheap technique for public good provision, but this cheapness comes at a price: distributional distortions when citizens start to vary with respect to preferences and effort costs. Sticks may still remain superior for general public goods, for which all members of society do an effort, and all members of society benefit, like the duty not to use violence, not to steal, not to damage the collective infrastructure, or the duty to apply hygienic norms.

Proposition 24 also implies that there is a correlation between the increased use of carrots and the increased public expenditure in modern societies. Both are driven by the same cause: increased specialization (in production and consumption), and increased individual differences.

### *5.6. Why probabilistic carrots are rare*

Consider the difference between the following two enforcement regimes. Parents A want their children to help old ladies cross the street and announce that when by coincidence ($p = 0.1$) they learn that a child did not do so it will be sanctioned. Parents B also want their children to help old ladies cross the street and announce that when by coincidence ($p = 0.1$) they learn that a child did so it will be rewarded. If a child notices that it has not been monitored in the stick regime of parents A, it will not tell its parents that it violated the rule. But if a child notices that it has not been monitored in the carrot regime of parents B, it will probably tell its parents that he helped an old lady cross the street. Under a carrots system, complying agents have an incentive to try to increase the monitoring level $p$, either by delivering the evidence to the principal, or by manipulating the environment so that the principal finds out himself. There is no similar incentive under a sticks system.

This may help to explain why probabilistic monitoring of the agent's output with carrots is rare. When carrot systems reward good behavior (input) in a probabilistic way, it is generally because the output is probabilistic (e.g., some luck is needed to invent something new for a which a patent can be granted).

### *5.7. Why combined carrots and sticks are rare*

If carrots and sticks are combined, the effect is even more powerful, and this could allow for even lower monitoring levels. Still, examples of combined carrots and sticks are rare. Because of

concerns for marginal deterrence, it will generally not be desirable to set sanctions at their theoretical maximum. In that case, using either a carrot or a stick has two major advantages over using a combination. Under the latter, transaction costs will always be higher, since a transfer is needed both in case of compliance and in case of violation. In addition, with respect to risks, combined sanctions create risks for compliers as well as violators, while single sanctions create only risks for one of these groups. Though the amount of the sanction or reward for which agents bear a risk will be lower under combined sanctions – on average it will be half of the amount of single sanctions – this effect is more than outweighed by the fact that the number of agents who bear a risk under a single sanction regime is less than half of the total number of agents (if in the single sanction regime the type of sanction is chosen by applying Wittman's rule, i.e., if it sanctions or rewards only a minority).

*5.8. Incentives in armies*

General conscription is a stick system. Men are obliged to contribute to a public good (safety) in the form of a service, under the threat of a criminal sanction. A professional army is a carrot system: volunteers are attracted to the army by a wage, which is financed by taxes. This simple fact illustrates two previously developed propositions. First (proposition 24), that sticks allow for contributions to a public good in the form of services, while carrots require the contribution to be in the form of a tax. Second (proposition 16???), that as societies become more specialized, and effort costs more variable, the case for working with carrots becomes stronger.

Once a soldier is in the army, he needs to be further incentivized. Military boot camps are notorious examples of stick-based enforcement systems. This can be explained by the fact that the optimal rules (like the rules on how to salute or how to march, like the fact that equipment has to be maintained, or that orders of higher-ranked officers have to be followed) are simple and clear. Still, armies use carrots as well. A first example is provided by medals for war heroes. These are not necessarily awarded to the soldiers who had the most dangerous tasks (like those who were in the line of fire), but rather to soldiers who undertook more dangerous tasks than they were asked to do. This indicates that there were specification problems because of the uniqueness of the situation, so that no specific orders could be formulated. A second example is promotion. Also here it is not prima facie clear which individuals are best suited for higher positions. By attaching carrots to the relatively scarce higher positions, individuals who believe they have the right capabilities are given an incentive to reveal themselves and exert the right efforts. Of course, penalties for those who did

not obtain a promotion would create the same incentives, but because only a minority is promoted, carrots are superior.

*5.9. Why coercive carrots are rare. The duty to rescue.*

Should rescuers be rewarded or non-rescuers sanctioned? Levmore (1986) found that legal systems that penalize non-rescue offer carrots (reimbursement) to rescuers as well. This balance between carrots and sticks is apparently in conflict with our analysis of section 5.7, which suggested that combining sticks and carrots is not desirable (except for when extremely powerful incentives need to be created that cannot be created by a single regime).

It is important, however, to notice that the stick is of a coercive kind, while the carrot is of an internalizing kind. The distinction between internalizing and coercive sanctions is a crucial one. Internalizing rules make the agent internalize the positive or negative effects he has caused, but leave the final decisions to do so to him (because he is believed to have more information about specific costs). Coercive sanctions, on the other hand, are set so high as to leave the agent no choice. It is easy to find examples of internalizing sticks (e.g., strict liability, pollution taxes), internalizing carrots (e.g. patents, market prices), or coercive sticks (e.g., criminal sanctions), but coercive carrots are rare. Extremely high rewards for helpful information on criminals (‘Wanted: Dead or Alive’’) may serve as a rare example. In market transactions, offers that are so high that they cannot be reasonably refused do exist but are rather exceptional.

The fundamental explanation lies in characteristic 2 (section 2.2): coercive sticks do not have to be paid (the threat is sufficient), while coercive carrots have to be paid. Hence, coercive carrots make sense only under exceptional circumstances, like when the majority of the agents is unable to comply with the rule (which is the case in the mafia information example).

In the duty-to-rescue case, judges have great difficulties in knowing the exact compliance cost (so that specific carrots risk to miss their effect in some cases), and although it is usually obvious that this cost is lower than the potential harm in case of non-rescue ($e_{max}$ < harm). Moreover, the costs are relatively low for most agents, so that the average effort cost will normally be relatively low compared to the highest possible cost at which the general stick or carrots will have to be set ($e_{av} < e_{max}/2$). As has been shown (proposition ???), it is desirable to use general sticks under those circumstances.

While the coercive function is fulfilled by sticks, adding carrots that are set at the average compliance cost may further reduce the distributional distortions. Note that a combination of a

coercive stick with an averagely internalizing carrot will generally be superior to a single general carrot regime.[17] Hence, Levmore's conclusion (1986) that legal systems that introduce sticks should also introduce carrots is consistent with the theoretical findings of this paper.

### *5.10. Breach of contract*

**Proposition 25**. *In complete contracts, sticks are superior. In relational contracts, carrots will often be superior.*

Should contract breachers be penalized or should non-breachers receive a bonus? In complete contracts, there are no specification problems ex post. There is no need to delegate authority to one of the parties, because all clauses are drafted ex ante. In addition, contract breach is typically verifiable, leaving again no reason to delegate the right-to-judge to one of the parties, and hence no danger of abuse of sticks by non-controlled non-benevolent principals.

Since a rationally drafted complete contract will never be breached, sticks seem superior with respect to transaction costs. Though one of the two components of these transactions costs – the costs of payment – is less important in contracts because the damages can often be deducted from the price, the second component – the costs of the sanctioning procedure – remains. In addition, probabilistic sticks remain superior to probabilistic carrots because of their risk property: sticks generate risks for violators, carrots for compliers. Since all agents comply with a complete contract, sticks are the optimal sanction type because of risk. Note that carrots involve more risks not only for agents but also for the principal who has to finance the carrots.

In relational contracts, where detailed terms need to be specified ex post by one of the parties and breaches are often non-verifiable, carrots become more attractive because of the danger of abuse of sticks by a non-benevolent principal. This may help to explain why legal systems usually restrict the use of penalties in employment contracts, but give employers the freedom to incentivize the employee with bonuses.

---

[17] A single general carrot regime will cause less distributional distortions only in the extreme cases when the difference between $e_{max}$ and $e_{av}$ is smaller than the standard deviation. Note also that if distributions are dense in the middle, combining high (coercive) sticks with averagely internalizing carrots becomes even more attractive than when the distribution is uniform.

### *5.11. The increased popularity of carrots in education*

**Proposition 26**. *In complex and dynamic societies, parenting will be increasingly based on carrots.*

Traditional parenting was mainly based on sticks. There seems to be a tendency in education to emphasize rewards (believing that carrots are better motivators). Some have suggested that this is an artifact of Skinner's popular behaviorism (Kohn, 1993). Modern psychologists generally emphasize the importance of positive sanctions (rewards, praise, …) in parenting and there is a tendency to recommend the use of carrots.

Our analysis suggests that the emphasis of contemporary psychologists on carrots may be more than a coincidental by-product of the popularity of behaviorism. It may simply be a consequence of a more complex society, in which it is less clear what to expect from children. Consider children that were raised at a traditional farm in pre-industrial times. For the parents it was clear what the children had to become farmers too. It was also clear what capabilities and attitudes good farming involved (exactly the same as the parents had, e.g., waking up early to milk the cows). Since there were no specification problems, sticks were the optimal type of sanction. It is instead less clear what makes a good doctor, attorney, merchant, or artist and it is not immediately evident what talent s a child possesses.

## 6. Conclusion

In this paper we have attempted to formulate a general theory for the use of carrots versus sticks. We have put some effort in distinguishing between factors that affect the choice between carrots and sticks and other irrelevant factors. The point with the greatest impact on the analysis is the simple observation that carrots are applied upon compliance, while sticks are applied upon violation. If enforcement is successful and agents comply with the rule, then sticks are never applied, while carrots need to be paid to all agents. From this observation it follows that sticks are preferable to carrots as they have a lighter impact of risk, distributional distortions, and transaction costs. We justify the increasing use of carrots in modern societies by reference to information and evidence problems.

A second point that clearly emerges from our analysis is that sticks are more dangerous than carrots if used by non-benevolent principals. While carrots have a built-in constraint and can only be used to enforce Pareto-optimal rules, sticks may result in situations that are not even Kaldor-

Hicks efficient.

**References**

Aron, Debra J. and Olivella, Pau (1994), Bonus and Penalty Schemes as Equilibrium Incentive Devices, with Application to Manufacturing Systems, 10 *Journal of Law, Economics, & Organization* 1-34.

Barzel, Yoram (1977), '???slavery???", 20 *Journal of Law and Economics*, 488-???.

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

Becker, Gary S. and Stigler, George (1974). Law Enforcement, Malfeasance, and Compensation of Enforcers. *Journal of Legal Studies*, *3*, 1-18.

Chang, Howard F. (1997), 'Carrots, Sticks and International Externalities', 17 *International Review of Law and Economics*, 309-324.

Coase, Ronald H. (1960), 'The Problem of Social Cost', 3 *Journal of Law and Economics*, 1-44.

Cooter, Robert D. and Garoupa, Nuno (2000). The Virtuous Circle of Distrust: A Mechanism to Deter Bribes and Other Cooperative Crimes. *The Berkeley Law & Economics Working Papers*, Vol. 2000 (2), Article 13.

Cooter, Robert D. and Ulen, Thomas (2003), *Law and Economics* (4th ed.), Boston, Addison-Wesley.

Dari Mattiacci and De Geest (2005) JP 4 technologies, JITE

Dari Mattiacci and De Geest on multiplication?

De Geest, Gerrit, Dari-Mattiacci, Giuseppe and Siegers, Jacques, 2006, Annullable Bonuses and Penalties, working paper, T. Koopmans Institute, Utrecht University.

Depoorter, Ben and De Mot, Jef (2006) Whistle Blowing: An Economic Analysis of the False Claims Act. *Supreme Court Economic Review*, *14*, 135-162.

Ellickson, Robert C. (1973), 'Alternatives to Zoning: Covenants, Nuisance Rules, and Fines as Land Use Controls', 40 *University of Chicago Law Review*, 681-781.

Ferster, Charles B. and B.F. Skinner (1957), *Schedules of Reinforcement*, New York: Appleton-Century-Crofts.

Fogel, Robert William and Engerman, Stanley L. (1974), *Time on the Cross: The Economics of American Negro Slavery*, 2 volumes, Boston: Little Brown.

Frey, Bruno S. 2004. *Dealing with Terrorism – Stick or Carrot?*, Cheltenham: Edward Elgar.

Gordon, Wendy J. 1982. Of Harms and Benefits: Torts, Restitution, and Intellectual Property. 21 Journal of Legal Studies, 449-???.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." 47 *Econometrica*, 263-291???or 363-391???

Katz, Avery Wiener, 2004, The Efficient Design of Option Contracts: Principles and Applications, Columbia Law School, Center for Law and Economic Studies, Working Paper No. 248 (http://ssrn.com/abstract=512146, version March 2004).

Kohn, Alfie. 1993. *Punished by Rewards: The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes*, New York: Houghton Mifflin.

Landes, William M. and Posner Richard A. (1978), 'Salvors, Finders, Good Samaritans, and Other Rescuers: An Economic Study of Law and Altruism', 7 *Journal of Legal Studies*, 83-128.

Lazear, Edward P. 1991. 'Labor Economics and the Psychology of Organizations' 5(2) Journal of Economics Perspectives 89-110.

Levmore, Saul (1985), "Explaining Restitution", 71, *Virginia Law Review*, 65-???.

Levmore, Saul (1986), "Waiting for Rescue: An Essay on the Evolution and Incentive Structure of the Law of Affirmative Obligations", 72, *Virginia Law Review*, 879-941.

Lipset, Seymour M. (1959), 'Some Social Prerequisite for Democracy: Economic Development and Political Legitimacy,' 53 *American Political Science Review*, 69-105.

Polinsky, A. Mitchell (1980), 'Notes on the Symmetry of Taxes and Subsidies in Pollution Control', 12 *Canadian Journal of Economics*, 75-82.

Shavell, Steven (1986), 'The Judgment Proof Problem', 6 *International Review of Law and Economics*, 45-58.

Summers, J., 1983, 'The Case of the Disappearing Defendant: An Economic Analysis', 132 *University of Pennsylvania Law Review*, 145-185.

Thorndike, Edward Lee (1898???)

Thorndike, Edward Lee (1913), *Educational Psychology*, New York: Columbia University Teachers College.

Tversky, Amos, and Daniel Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice." *Science* 211: 453–458.

Wittman, Donald A. (1984), 'Liability for Harm or Restitution of Benefit?', 13 *Journal of Legal Studies*, 57-80.

Wittman, Donald A. (2000), 'General Structure of the Law', in Bouckaert, Boudewijn and De Geest, Gerrit (eds.), *Encyclopedia of Law and Economics,* Vol. I, Cheltenham: Edward Elgar, 1072-1094.