Tjalling C. Koopmans Research Institute

How to reach the authors
Please direct all correspondence to the first author.

Gerrit de Geest
Utrecht University
Utrecht School of Economics
Vredenburg 138
3511 BG Utrecht
The Netherlands
Tel: 0031 (0)30 253 9800
Fax: 0031 (0)30 253 7373
Email: g.degeest@econ.uu.nl

Guiseppe Dari Mattiacci
George Mason University School of Law
3301 North Fairfax Drive,
Arlington, Virginia 22201
U.S.A.
E-mail: gdarimat@gmu.edu

Jacques Siegers
Utrecht University
Utrecht School of Economics
Vredenburg 138
3511 BG Utrecht
The Netherlands
E-mail: j.siegers@econ.uu.nl

The Intrinsic Inferiority of Efficiency Wages to Damages and Conditional Bonuses

Gerrit de Geest[a]
Giuseppe Dari Mattiacci[a b]
Jacques Siegers[a]

[a]Utrecht School of Economics
Utrecht University

[b]George Mason University of Law
U.S.A.

April 2004

Abstract
In this paper, we argue that, as an enforcement mechanism, efficiency wages are intrinsically inferior to damages and to conditional bonuses – an alternative positive sanction system overlooked in the labor economics literature, under which rents are only paid if monitoring has effectively taken place (and the employee is not found shirking). While all three alternatives succeed in incentivizing agents and satisfy the participation constraint of non-shirking employees, damages (and negative sanctions in general) do so at lower costs because they do not require the payment of any rents. Of the two positive sanction systems, conditional bonuses are less expensive than efficiency wages because the latter also pay rents when no monitoring has taken place and may allow employees who are found shirking to keep some rents.

Moreover, we find that monitoring levels are inefficiently low under efficiency wages. While efficiency wages (if they are completely non-retroactive) remove the employer's incentive to falsely sanction the employee, they solve this appropriation problem in a less rational way than some decoupling mechanisms that can be used under damages and conditional bonus regimes.

Therefore, it seems extremely unlikely that employers would ever opt for efficiency wages on such a massive scale that structural unemployment would result, as Shapiro and Stiglitz (1984) suggested.

1 Introduction

According to the efficiency wage theory of Shapiro and Stiglitz (1984), it can be rational for an employer to overpay his employees when monitoring is expensive. Above-market wages make the sanction of dismissal more severe. This increased sanction allows the employer to lower monitoring levels and to save on monitoring costs. These savings on monitoring costs can be higher than the rents paid to employees. But if many employers try to save on monitoring costs by increasing wages, the effect disappears since an employee who loses an 'overpaid' job can easily find another 'overpaid' job. In that case, the economic system may need structural unemployment in equilibrium to make dismissal unattractive and keep monitoring costs low according to Shapiro and Stiglitz.

Shapiro and Stiglitz' article was written in the 1980s, at a time when many economists started to believe that unemployment was structural, rather than cyclical. Their theory is still considered as a possible explanation of unemployment and is widely mentioned in textbooks on labor economics. In law and economics scholarship, however, the idea of sanctioning contract breach by withdrawing bonuses has hardly been explored.[1] Efficiency wages are just another technique to prevent breach of contract: in order to incentivize an employee to perform, he is overpaid. If he breaches the contract, he does not have to pay damages, but he loses a rent. The economic literature on contract remedies has focused exclusively on negative sanctions like specific performance, expectation damages, reliance damages, or penalty clauses. It has not compared the advantages of positive sanctions to those of negative sanctions yet.[2] The overpayment technique is also relatively unexplored in the economic literature on public law enforcement.[3]

In this paper we argue that, as an enforcement mechanism,[4] efficiency wages are a less likely choice than generally accepted in the labor economics literature. Efficiency wages are an

---

[1] One exception is Shavell (1997), who took into account the possibility for firms to pay efficiency wages to employees, in a discussion on the optimal level of corporate liability.

[2] Moreover, the economic literature on contract law has focused on cases in which there is a 100% chance that the breach will be discovered (as is the case in most commercial contracts; for instance, if a builder promises to build me a house and he does not, it is extremely likely that I will notice that). The efficiency-wage model focuses on cases in which monitoring is relatively expensive, so that the apprehension rate of contract breach is less than 100%. This is a typical feature of many employment contracts. If a researcher reads a newspaper instead of doing research for which he is paid, the university might not notice it. Similarly, if a manager of a firm enriches himself in a fraudulent way, it is not guaranteed that the shareholders will notice it on time.

[3] An exception is the economic literature on corruption in which a few recent papers can be found that consider bonus payments as an alternative solution to corruption. See Polinsky and Shavell (2001), and Cooter and Garoupa (2001).

[4] In this paper we do not consider other functions efficiency wages might have, such as the income-satiation (Rasmusen, 1992) and the gift-reciprocity (Akerlof, 1982) arguments.

intrinsically wasteful technique of incentivizing employees and will only be chosen by rational parties under rather unlikely conditions.

Efficiency wages are a form of positive sanctions; that is, they reward agents who have not been found shirking, rather than punishing agents who have been found shirking. Negative sanctions like damages,[5] however, incentivize as well but do so at lower costs for the employer: the employer has to pay only the market wage and no rents. Though the absence of rents will make the contract less rewarding for employees, they will accept a contract with negative sanctions as soon as they receive the market wage. That is, their participation constraint is fulfilled as soon as the wage equals the market wage; they should not be troubled by the sanctions for shirking because they decide not to shirk since the non-shirking constraint is fulfilled. In essence, an employer who pays efficiency wages tries to obtain something (non-shirking behavior) by paying rents, while he can obtain the same without paying rents (by threatening with damages).

We will also consider a third alternative, which we call *conditional bonuses*. This alternative seems to have been completely overlooked in the literature.[6] It is a positive sanction just like efficiency wages but, unlike efficiency wages, the rent is awarded only if monitoring takes place (and if the employee is not found shirking). This means that non-controlled employees receive a rent under efficiency wages but not under conditional bonuses. To illustrate, suppose that the employee's benefit of shirking is 2 and that the market wage is 100. If the employer wants to employ efficiency wages and to adopt a 10% monitoring level (i.e. a 10% chance that shirking is discovered), then the overpayment (that the employee loses if

---

[5] Since our aim is to compare efficiency wages with damages, we set in both cases the magnitude of the sanction at the lowest value that satisfies the non-shirking constraint (i.e. that is sufficient for deterring shirking). For instance, if the shirking benefit e is 2, and the apprehension rate (i.e. the monitoring level) x is 10%, then the overpayment (i.e. the wage minus the market wage) should be 20 under an efficiency wages regime, and the damages should be 20 under a damages regime. How a damage measure that equals xe should be translated into legal terms may vary. Legal concepts such as reliance damages, expectation damages, or punitive damages are defined by reference to the *harm* to the promisee and not by reference to the *benefit of breach* for the promisor. In the real world, the *harm* can have any value above the shirking benefit e (the harm cannot be lower than e because that would mean that it was not beneficial to have a contract at all, since the employee produces something that costs more than the value to the employer), but how much higher the harm is may vary. Depending on that ratio, the 'damages' in our model can be 'reliance damages', 'expectation damages', or anything below the reliance measure, in-between the reliance and the expectation measure, or above the expectation measure, in which case they might be called 'extra-compensatory' damages (though they will always be lower than 'punitive damages' which are defined in the law and economics literature as $x \cdot harm$, which is higher than xe since the harm is higher than e). From a technical legal viewpoint, it is even possible that what we call 'damages' will be considered a 'fine', like when the employer is the government and the employee a civil servant, and the 'breach' falls under criminal law rather than contract law.

[6] Shapiro and Stiglitz (1984) also briefly discuss *bonds* as an alternative sanctions mechanism, but bonds are analytically the equivalent of damages and are no bonus system. A bond is an amount of money from the employee that is posted on a blocked bank account before the contract is performed (ex ante). If the employee breaches the contract, this amount of money is paid to the employer. Damages are an amount of money that the employee pays to the employer after breaching the contract (ex post). They are analytically equivalent, at least for our purposes. Making a distinction between them would be useful only when the employee's wealth might change over time.

shirking is discovered) should be at least 20 (which brings the employee's wage to 120). Now, consider a conditional bonus system. Here, the employer pays a wage of 100, but announces that there will be a 10% chance that the employee will be monitored. If monitoring takes place, and the employee is not found shirking, then he will receive a bonus of 20. This creates exactly the same incentives for the employee: shirking is sanctioned by a 10% chance to lose (the opportunity to receive) 20. But the expected cost for the employer is 10 times lower (if employees do not shirk): a conditional bonus has an expected cost of only 2.

Conditional bonuses are a more effective and less wasteful instrument than efficiency wages because the latter is also paid to employees when no monitoring has taken place. Conditional bonuses, in contrast, are paid only when monitoring has effectively taken place (and the agent has not been found shirking). To put it differently, efficiency wages are bonuses that are also paid to shirking agents who are lucky enough not to be monitored. These bonuses are a waste of money from the principal's perspective (paying a bonus in case no control has taken place does not alter the marginal incentives not to shirk).

This paper is structured as follows. In section 2, we present our general framework (one-period model). The employer first chooses the monitoring level that minimizes his expected total enforcement expenses (i.e. the sum of monitoring costs and the expected overpayment expenses) for each sanction system. We show that monitoring levels are set at an optimal level under damages and conditional bonuses but at a higher than optimal level under efficiency wages (section 3).

Next, the employer chooses the sanction type that costs least to him while satisfying three constraints: the non-shirking constraint, the maximum sanction constraint, and the employee participation constraint. We demonstrate that rational employers will always prefer damages or conditional bonuses to efficiency wages (section 4). In the same section, we show that rational employers will prefer, in principle, to employ damages; only when the employee's wealth (i.e. the maximum sanction available under damages) is lower than a certain value will it be profitable to employ conditional bonuses. We will determine the conditions for that shift. In section 5, we analyze combined sanctions and find that only under certain conditions will employers use combined sanctions (damages and conditional bonuses) rather than damages. In section 6, we replace the one-period model with an *n*-period model. We demonstrate that the conclusions reached under the one-period model are reinforced in the *n*-period model, at least when efficiency wages are not retroactively withdrawn from an agent who is found shirking in a later period. If efficiency wages are not retroactive, then the agent who is found shirking keeps some rents, which further reduces the marginal incentives not to shirk.

Section 7 discusses the appropriation problem. Damages and conditional bonuses have an inherent risk of appropriation: the employer may benefit from falsely declaring that the employee shirked. This appropriation problem exists in case of efficiency wages as well if they are partly or fully retroactive (if the employer can claim back past rents or can avoid having to pay a rent for the present period, he has an incentive to falsely sanction the employee). While completely non-retroactive efficiency wages remove the employer's incentive to falsely sanction the employee, they solve this appropriation problem in a less rational way than some decoupling mechanisms that can be used under damages and conditional bonus regimes.

The concluding section 8 sums up under what conditions efficiency wages would ever be used, critically analyzes Shapiro and Stiglitz (1984), and discusses the implications of our findings for labor economics, the economic analysis of contract law, and anti-corruption policies.

2   The basic framework (one-period model)

An employee is hired to work for an employer during a specified period of time. The employee has to choose between shirking and non-shirking. For reasons of simplicity, we assume that this choice is discrete: the employee fully shirks or does not shirk at all. If he shirks, he obtains a *shirking benefit e* (an avoided effort cost). *e* is always positive: in the absence of a sanction, the employees prefers shirking. For mathematical simplicity, we assume that if the expected sanction on shirking equals the shirking benefit, the employee is not indifferent but prefers non-shirking.

We further assume that the harm caused by shirking is so high that the employer always prefers non-shirking[7] and always chooses a sanction-apprehension rate combination that guarantees non-shirking behavior (non-shirking constraint).

At the end of the period, the employer has the opportunity to monitor the employee. If monitoring takes place, shirking will always be discovered. The contract states the *monitoring level x*, that is, the chance that monitoring will take place at the end of the period (the contract is of course binding, so the employer cannot change the monitoring level ex post).[8] The choice of the monitoring level is not discrete but continuous: *x* can be any value between 0 and 1. To illustrate, if *x* = 1, this means that monitoring will always take place and that the apprehension

---

[7] Assuming that the optimal shirking level is zero is plausible for many (though not all) contracts. For instance, banks want to reduce the employee's fraud to 0. Similarly, states want judges never to be corrupt. The assumption that the optimal shirking level is 0 can be restated as the equivalent assumption that the marginal enforcement costs are always lower than the marginal net social costs of shirking – an assumption which is common in many economic models on law enforcement.

rate for shirking is 100%. If $x = 0.2$, there is a 20% chance that monitoring will take place and, consequently, an apprehension rate of 20%.

The *cost of a monitoring session* is constant and is denoted by m. The expected monitoring costs are mx. We also assume that the cost of a monitoring session is lower than the employer's wealth, $m < w^p$.

The contract also states the magnitude of the *sanction s* that is applied to shirking. The sanction can be the payment of damages (negative sanction), or the loss of an overpayment (positive sanction). Except for these damages or lost bonuses there are no other sanctions for breach. More specifically, we assume that there are *no turnover costs*,[9] and no transaction costs caused by the payment of bonuses or damages.

The employer sets the monitoring level x in such a way that his total enforcement costs (i.e. the sum of expected monitoring costs and expected bonus payments) are minimized. His choice, however, is limited by three constraints. First, the *non-shirking constraint*: the expected sanction must be high enough to prevent employees from shirking. Second, the *employee participation constraint*: the expected wage (taking bonuses and damages into account) of a non-shirking employee must at least equal the market wage (or price) p^m. Third, the *maximum sanction constraint*: the maximum sanction or bonus is not infinite. In the case of damages (negative sanction), we assume that the maximum feasible sanction is equal to the employee's wealth w^a. In the case of bonuses, we assume that the maximum feasible bonus is equal to the employer's wealth w^p.[10]

All parties are assumed to be rational, utility maximizing, and risk neutral.[11] All parties have perfect information on e, x, s, m and the market wage for labor (given the required effort

---

[8] By the same token, we assume that ex post renegotiation between the employee and employer is not possible.

[9] Introducing turnover costs (the transaction costs that are incurred by the parties for terminating the contract, and finding an alternative job or an alternative employee, respectively) would not qualitatively alter our conclusions. Turnover costs borne by the employee are additional negative sanctions. They do not change the analysis for damages since they reduce the assets available for paying damages but at the same time create an additional negative sanction for the employee. Hence, the total loss incurred by the shirking employee is likely to remain constant. Although turnover costs change the analysis of positive sanction systems in that they turn a pure positive sanction system (conditional bonuses or efficiency wages) into a combined sanction system (positive plus negative sanctions), they do not change the superiority of pure negative sanction systems over combined sanction systems and of combined sanction systems over pure positive sanction systems. Turnover costs borne by the employer do not change our analysis either since they will never be effectively borne if the non-shirking condition is fulfilled.

[10] The assumption that the maximum sanction corresponds to the maximum wealth of the party who has to pay or finance the sanction may seem to be extreme and unrealistic, but it is frequently made in the economic literature on law enforcement (Becker, 1968, at 193, Polinsky and Shavell, 2000, at 310) and it is useful for analytical purposes. Alternatively, w^a and w^p could be defined as the maximum budget of the agent and the principal, respectively, that is available for sanctions. This would not change the analysis and the conclusions with respect to the superiority of damages over conditional bonuses and conditional bonuses over efficiency wages.

[11] Introducing risk aversion would not substantially alter our results; it would just strengthen our conclusions with respect to the superiority of negative sanctions to positive sanctions. When agents are risk averse, negative sanctions may become more detrimental, while positive sanctions may become less beneficial in expected terms.

level) p^m.

Notation:

e =   the employee's shirking benefit (avoided effort), $e>0$;

x =   the probability of employer's monitoring (monitoring level); in theory $x=[0,1]$; however, in the analysis we will determine a feasibility range $x=[x^{min},1]$;

s =   the sanction; $0 \leq s \leq s^{max}$; $s^{max}=w^a$ (the employee's wealth) for damages, while $s^{max}=w^p$ (the employer's wealth) for conditional bonuses and efficiency wages;

m =   the cost of a monitoring session; $m<w^p$;

p^m =   the market wage.

2.1 *Non-shirking constraint under efficiency wages*

Under efficiency wages, the employee's wage is the market-clearing wage p^m plus a bonus s. He loses the bonus if monitoring takes place and he is found shirking.[12] The employee's expected net returns in the case of non-shirking are: p^m+s. The employee's expected net returns in the case of shirking are: $p^m+e+(1-x)s$. The employee will prefer non-shirking to shirking if: $p^m+s \geq p^m+e+(1-x)s$, which can be re-written as: $xs \geq e$ or $s \geq e/x$.

The lowest sanction that satisfies the non-shirking constraint at a given monitoring level is $s=e/x$.

2.2 *Non-shirking constraint under conditional bonuses*

Under conditional bonuses, the employee's wage is the market clearing wage p^m plus a bonus s that is only paid if monitoring effectively takes place and he is found non-shirking. The employee's expected net returns in the case of non-shirking are: p^m+xs. The employee's expected net returns in the case of shirking are: p^m+e. The employee will prefer non-shirking to shirking if: $p^m+xs \geq p^m+e$, which can be re-written as: $xs \geq e$ or $s \geq e/x$.

The lowest sanction that satisfies the non-shirking constraint at a given monitoring level is $s=e/x$.

2.3 *Non-shirking constraint under damages*

Under damages, the employee's wage is the market clearing wage p^m. He never receives a

---

[12] We assume in this one-period model that when the employee is caught shirking at the end of the period he will retroactively lose the overpayment of that period. Without such retroactivity, efficiency wages would not make sense, as they would provide no incentives (whether or not the employee is caught shirking, he can keep the overpayment). In the *n*-period model, partially retroactive or non-retroactive efficiency wages still create some

bonus. If monitoring takes place and he is found shirking, he has to pay a sanction s. The employee's expected net returns in the case of non-shirking are: p^m. The employee's expected net returns in the case of shirking are: $p^m + e - xs$. The employee will prefer non-shirking to shirking if: $p^m \geq p^m + e - xs$, which can be re-written as: $xs \geq e$ or $s \geq e/x$.

The lowest sanction that satisfies the non-shirking constraint at a given monitoring level is $s = e/x$.

2.4 Participation constraint

The employer may determine the terms of the contract (wage, monitoring level, sanction), but these terms must not be so uninteresting that no employee is willing to participate. Since the employer only wants to attract non-shirking employees, we have to account for their payoffs when analyzing whether the participation constraint is satisfied. If the employer pays the market wage, the non-shirking employee's expected payoffs are the following:

(i) Under efficiency wages: $p^m + s$.
(ii) Under conditional bonuses: $p^m + xs$.
(iii) Under damages: p^m (non-shirking employees never pay damages).

Since the participation constraint is satisfied if employees receive at least p^m (the market wage given the required effort level) and since x and s are always positive, the participation constraint is always fulfilled under any of these three contracts.

2.5 Maximum sanction constraint (creating a feasibility constraint for x)

The fact that sanctions (damages or bonuses) cannot be infinitely high restricts the range of x in the non-shirking constraint $s = e/x$. x can no longer be any value between 0 and 1, but has to be a value between the *minimum monitoring level* x^{min} and 1. The minimum monitoring level that satisfies the non-shirking constraint is found by the following formula:

$x^{min} = e/s^{max}$

Since s^{max} corresponds to the employer's wealth w^p in the case of bonuses and to the employee's wealth w^a in the case of damages, x^{min} may have a different value for positive sanctions (efficiency wages and conditional bonus) than for negative sanctions (damages).

2.6 A note on normalized efficiency wages

In theory, the employer could 'normalize' efficiency wages, by asking a prior payment from employees before entering into the overpaid position (a solution suggested by Becker and

---

incentives. See section 6 for further analysis.

Stigler, 1974, in a paper on preventing police corruption). Normalization would mean that the normal wage is set at $(p^m–s)$ and that a bonus s is added, which the employee can keep if he is not found shirking. Hence, a non-shirking employee always receives $(p^m–s)+s=p^m$, while a shirking employee receives the same if no monitoring takes place and $p^m–s$ if monitoring takes place. This shows that the payoffs are mathematically the same as under damages in a one-period model.[13] Since they are analytically identical there is no need to make a separate analysis of normalized efficiency wages. As a matter of fact – though the name suggests otherwise – 'normalized efficiency wages' are not efficiency wages but damages.[14]

3   Chosen monitoring levels (one-period model)

In this section, we will analyze what monitoring levels the employer chooses under each sanction, and compare them to the optimal monitoring level under that sanction.

Proposition 1. For each type of sanction, the optimal monitoring level is the lowest possible monitoring level x^{min}, i.e. the level that corresponds with the highest possible sanction. Under damages and conditional bonuses, the employer chooses x^{min}, while under efficiency wages the monitoring level will generally be higher than optimal.

From a social point of view, damages are purely redistributive: what the employee loses corresponds to what the employer gains.[15] This also applies to bonuses: what the employer pays corresponds to what the employee receives. Consequently, the only costs that remain to be minimized are the expected monitoring cost mx. Therefore, the optimal monitoring level is the

[13] In an *n*-period model, the payoffs are identical when the whole contract duration is taken into account, but the timing is different since a normalized efficiency wages regime makes the employee pay a high sum in advance, which he earns back over time.

[14] Under normalized conditional bonuses, the employer sets the normal wage at $(p^m–xs)$ and adds a bonus s, which the employee receives if monitoring takes place (probability x) and he is not caught shirking. So the expected wage of a non-shirking employee is $(p^m–xs)+xs=p^m$, while a shirking employee receives $p^m–xs$. For instance, if the market wage p^m is 100, the bonus s is 20 and the monitoring chance x is 10%, a non-shirking employee's wage would be 98 if no monitoring takes place and 118 if monitoring takes place. Normalized conditional bonuses are not mathematically identical to damages. They are inferior to a damages regime though from a risk perspective (they create artificial risks among employees) despite creating the same marginal incentives as damages (and non-normalized conditional bonuses) and leading to the same expected expenses for the employer while satisfying the employee's participation constraint as well.

[15] Our analysis does not take activity levels into account. For instance, if employers have to pay bonuses, they buy labor at a price that is above the real costs (if we assume that market prices and market wages reflect the real costs in competitive markets). Consequently, employers will buy less labor than optimal. As a result, the socially optimal monitoring level may be higher than the minimum: the lower the monitoring level, the lower the monitoring cost (which is a good thing for society) but also the higher the substitution effects (which is a bad thing). This trade-off between monitoring costs and substitution costs does not exist for damages. Under damages, the employer pays the market wage to non-shirking employees. So introducing activity levels into the model would make the case for damages even stronger.

lowest possible monitoring level x^{min}.

As argued in the former section, the lowest possible monitoring level is not equal to 0, due to the maximum sanction constraint. x^{min} is the monitoring level that corresponds with the highest possible sanction.[16]

Under conditional bonuses, the total enforcement costs, which the employer seeks to minimize, are the sum of the expected monitoring costs and the expected bonus payment costs, xs. As a consequence of the non-shirking constraint, $xs=e$, the expected bonus payment costs are always equal to e, which is a constant. Therefore, the employer's minimization problem is:

$$\min_x [mx + e]$$

The lowest value is given by x^{min}, which is the optimal monitoring level.

Under damages, the employer does not have to pay any bonuses. The damages are entirely paid by the employee and are a bonus that the employer receives. However, since the non-shirking constraint has to be fulfilled, the employee will never shirk and consequently never pay damages. Therefore, the costs that the employer seeks to minimize only consist of the expected monitoring costs.

$$\min_x [mx].$$

The chosen monitoring level will clearly be the lowest possible monitoring level x^{min}.

Under efficiency wages, the employer will minimize his total enforcement costs, which is the sum of the expected monitoring costs and the expected bonus payments:

$$\min_x [mx + s] = \min_x \left[mx + \frac{e}{x} \right].$$

The cost function is no longer linear, as under damages and conditional bonuses, but, to the contrary, is convex.[17] Let $x^\wedge$ denote the solution. The first order condition yields $x^\wedge = \sqrt{\dfrac{e}{m}}$ or $x^\wedge = \sqrt{\dfrac{x^{\min} w^p}{m}}$, as $x^{min}=e/w^p$. It can be easily shown that $x^\wedge = x^{min}$ only if the minimum monitoring level, x^{min}, is equal to 1, as we have assumed $m<w^p$. Only under these extreme conditions would efficiency wages lead to efficient monitoring; generally, they will instead lead to excessive monitoring levels.

---

[16] The lowest possible monitoring level x^{min}, may be different for negative and positive sanctions since $x^{min} = e / s^{max}$ and s^{max} equals the wealth of the employee w^a for negative sanctions and the wealth of the employer w^p for positive sanctions. However, in this section we compare the chosen monitoring level under each sanction with the optimal monitoring for that particular sanction. Therefore, it makes sense to use the general symbol x^{min} for all sanctions in this section.

[17] The first order condition is $m-ex^{-2}=0$. The second order condition is $2ex^{-3}>0$.

4   Chosen sanction (one-period model)

The employer will opt for the sanction system that yields the lowest total costs.

(i)   Under efficiency wages, the employer's total enforcement costs are *mx+e/x*.

(ii)   Under conditional bonuses, the total enforcement costs are *mx+e*.

(iii)   Under damages, the total enforcement costs are *mx*.

Proposition 2. Rational employers always prefer conditional bonuses to efficiency wages.

The employer will always prefer conditional bonuses to efficiency wages (unless the optimal monitoring level is equal to 1 in which there is no longer any analytical difference between both) for a simple reason. In spite of the fact that both mechanisms are subject to the same maximum sanction constraint, the employer's wealth, and hence to the same lower limit to the monitoring level, conditional bonuses have two advantages: first, the expected payment to the employer is always lower, as it takes place only if monitoring occurs ($e \leq e/x$, since $x \leq 1$); second, the expected monitoring costs *mx* are lower too, since supra we have demonstrated that monitoring levels under conditional bonuses will be lower than or equal to the monitoring level under efficiency wages. The total enforcement costs would be the same only if $x^{min} = 1$.[18]

Proposition 3. Rational employers always prefer damages to conditional bonuses, if the employee's wealth w^a is greater than the monitoring cost m.

We have already shown that under both sanctions, the employer will choose the minimal monitoring level x^{min}. However, the lowest possible monitoring level x^{min}, may be different for the two cases, since $x^{min} = e/s^{max}$ and s^{max} equals the wealth of the employee for damages and the wealth of the employer for conditional bonuses. In the former sections we used the same symbol x^{min} to denote both (which made sense since we compared the chosen monitoring level under each sanction with the optimal monitoring level for that particular sanction). In this section we will use two symbols to avoid confusion. x^{mind} will denote the minimum monitoring level under damages and x^{minb} will denote the minimum monitoring level under conditional bonuses. Under damages, the employer's total enforcement costs are: $mx^{mind}=me/w^a$. Under conditional bonuses, the employer's total enforcement costs are: $mx^{minb}+e=me/w^p+e$.

The employer will prefer damages to conditional bonuses (and this will be accepted by the other party since the participation constraint is always satisfied) if $me/w^a < me/w^p+e$. This expression can be written as:

$$w^a > m\frac{w^p}{m+w^p}.$$

Since $w^p/(m+w^p)$ is lower than 1, whenever the employee's wealth is greater than the cost of a monitoring session, a rational employer will choose damages. However, this is just a sufficient condition, as when the employer's wealth is very low, even an employee's wealth lower than the cost of a monitoring session justifies the use of damages.

5 Chosen monitoring levels and sanctions under combined sanctions (one-period model)

Let us now examine the chosen monitoring level under a combined sanction. The advantage of combining a positive with a negative sanction lies in the possibility to increase the maximum available sanction and hence to reduce the monitoring level. If combined sanctions are employed, the maximum feasible sanction is given by the sum of the employee's wealth (the maximum feasible sanction in the case of damages) and the employer's wealth (the maximum feasible sanction in the case of conditional bonuses and efficiency wages).

We have already demonstrated that conditional bonuses will always be preferred over efficiency wages; therefore, also in the case of combined sanctions, the employer will prefer to combine damages with conditional bonuses rather than with efficiency wages.

Proposition 4. Rational employers will prefer a combined damages-plus-conditional-bonuses sanction to damages only if the employee's wealth is lower than the cost of a monitoring session. They will employ damages otherwise.

Corollary. Rational employers will always prefer a combined damages-plus-conditional-bonuses sanction to pure conditional bonuses.

Let s^b denote the positive sanction (bonus) and s^d the negative sanction (damages). Under a combined damages-plus-conditional-bonuses sanction, the employer will minimize the sum of expected monitoring costs and expected bonus payments:

$$\min_x\left[mx + xs^b\right],$$

where x is the monitoring level given by $x=e/(s^d+s^b)$. Therefore, the former expression is equal to

---

[18] See proposition 1.

$$\min_{s^d,s^b}\left[\frac{e}{s^d+s^b}m+\frac{e}{s^d+s^b}s^b\right]=\min_{s^d,s^b}\left[e+e\frac{m-s^d}{s^d+s^b}\right].$$

With respect to damages, the function is clearly decreasing in s^d, so that the employer will choose the maximum value of $s^d=w^a$ for damages.

With respect to conditional bonuses, on the one hand, the function is also decreasing in s^b if $s^d<m$. Since the employer will always choose $s^d=w^a$, the condition becomes $w^a<m$. In this case, the value of s^b that minimizes the total enforcement cost is the maximum sanction also under conditional bonuses $s^b=w^p$. On the other hand, if $w^a>m$, the function is increasing in s^b, so that the value of s^d which minimizes the total enforcement costs is $s^d=0$. Therefore, rational employers will choose a combined damages-plus-conditional-bonuses sanction instead of damages only if the employee's wealth is lower than the cost of a monitoring session.

From this result it is also possible to conclude that, since the total enforcement cost is decreasing in s^d, the total enforcement cost if $s^d=0$ will be higher than the enforcement cost for any positive value of s^d. That means that a rational employer will always prefer a combined damages-plus-conditional-bonuses sanction to conditional bonus alone, which proves the corollary.[19] In synthesis, if combined sanctions are feasible, the employer will opt either for damages plus conditional bonuses or for damages.

6   *N*-period model

In the former sections, we analyzed a one-period model. Now we consider an *n*-period model in which the contract period is divided into *n* sub-periods, with possible monitoring sessions at the end of each sub-period (e.g. a one-year contract is divided into 12 one month periods, with a chance that monitoring takes place at the end of each month). Do our findings still hold?

We should preliminarily clarify what happens in an efficiency wages regime when an employee is found shirking with rents he received in the previous periods.[20] When efficiency wages are *fully retroactive,* an employee who is found shirking in a specific sub-period earns no rent for that sub-period and has to pay back the rents he received in all previous sub-periods. For example, if he is found shirking in the 8th period, he receives no rent for the 8th period and has to pay back the rents received in the first seven periods. In addition, he will not receive any

---

[19] Monitoring level will obviously be efficient.
[20] The retroactivity issue is of course not relevant for a damages regime in which employees never receive any bonuses, making the question whether previously received bonuses have to be given back pointless. The retroactivity issue is not relevant for conditional bonuses since an employee who is caught shirking cannot have received a bonus in a previous sub-period because bonuses are paid only after monitoring, and we assume that in cases of monitoring, shirking is always discovered. Retroactivity becomes an issue under efficiency wages because employees are also overpaid when no monitoring takes place, which makes it possible that shirkers receive some rents.

rents in future periods since he is dismissed. When efficiency wages are *partly retroactive*, an employee who is found shirking, does not receive the rent for the sub-period in which he was found shirking but is allowed to keep the overpayments received in previous sub-periods. For example, if he is found shirking in the 8th period, he will have to leave the firm, having received an above-market wage during 7 months and a market wage during one month. When efficiency wages are *completely non-retroactive*, an employee who is found shirking is dismissed, but can keep the overpayment related to the current and all previous periods. For example, an employee who is found shirking in the 8th period will have received an above-market wage during 8 periods. In that case, the sanction consists only in the foregone future rents.

If efficiency wages are *fully retroactive*, there is no substantial analytical difference between an N-period and a one-period. Sanctioning employees who are caught shirking in one of the 12 sub-periods by a loss of all rents is the equivalent of having no overpayment in the first 11 periods and one big bonus at the end of the contract (the only difference is that a shirking employee may be dismissed before the end of the contract, but that holds for damages, conditional bonuses and efficiency wages). Since there is no analytical difference, our conclusions with respect to the superiority of damages over conditional bonuses and of conditional bonuses over efficiency wages also hold here.

If efficiency wages are *partially retroactive*, then efficiency wages become even more ineffective since employees who are caught shirking keep some of their bonuses. From the principal's perspective, it is wasteful to pay some rents to employees who are found shirking. This further undermines the marginal incentives that efficiency wages are meant to create.

If efficiency wages are *completely non-retroactive*, they have two additional shortcomings compared to partially retroactive efficiency wages. First, they are even more wasteful since an agent who is found shirking in a sub-period still gets the bonus for that sub-period. Second, the incentives not to shirk are completely absent in the last sub-period (whether the agent shirks or not, he will keep the bonus for the last sub-period anyway).

7   A Note on the Appropriation Problem

Shapiro and Stiglitz (1984) briefly discussed 'bonds'[21] as an alternative to efficiency wages. A bond is an amount of money posted by the employee at the beginning of the contract and paid back at the end of the contract if the employee has not been found shirking. This is in essence a damages system: non-shirking employees get back their bond and will have received the market

---

[21] See also Shapiro and Stiglitz (1985).

wage during the contract period, while shirking employees have a chance to lose some money. Shapiro and Stiglitz (1984) remarked that bonds may not be feasible because of the appropriation problem: the employer can falsely pretend that the employee shirked and might consequently refuse to pay back (appropriately) the bond. Is it correct that efficiency wages are less vulnerable to appropriation by the employer than damages?[22]

The appropriation problem arises whenever the employer benefits from falsely sanctioning the employee. This is the case when the damages are paid to the employer but also when a falsely withheld bonus had to be financed by the employer. The latter holds not only for conditional bonuses but also for (partly or fully) retroactive efficiency wages.

When efficiency wages are completely non-retroactive, they do solve the appropriation problem. In that case, a sanctioned employee loses something (foregone future rents), but the sanctioning employer does not gain anything because the future rents that the first employee loses is paid to the newly hired employee, who receives the same high wage.

It is important to understand, however, that the fundamental mechanism behind this solution of the appropriation problem is *decoupling*.[23] Decoupling means that what the sanctioned employee loses (the rents of future periods) does not go to the employer, but instead goes to a third party: the newly hired employee.

Decoupling can perfectly solve the appropriation problem under the other sanction regimes as well. For instance, a contract could stipulate that the damages paid by the employee should not go to the employer but should be distributed among the other employees. Similarly, it could stipulate that the conditional bonuses that have not been paid in a period should be distributed among all employees who have not been monitored in that period. Alternatively, the employer could create a 'bonus fund' and divide a fixed amount among all non-shirking employees in

---

[22] Appropriation is a form of the more general concept of opportunism: the party who has to perform last refuses to perform or refuses to keep his promise. We discuss only one type of appropriation: falsely sanctioning a non-shirking employee. A second type of appropriation consists in giving incorrect information ex post on the real monitoring levels (understating the monitoring levels in order to claim a higher fine or overstating the monitoring levels in order to reduce the bonus). This problem can be very simply solved by writing the magnitude of the sanction in the contract, so that the effective monitoring level has no influence on the magnitude of the damages or bonus. A third type of appropriation consists in organizing less monitoring than announced in the contract to save monitoring expenses. This is a version of the more general enforcer's paradox: when agents think the enforcer will sufficiently control, they comply, but when the enforcer knows that the agents will comply, he has no reason to control. This enforcer's paradox equally holds for all enforcement regimes. A fourth type of appropriation occurs in conditional bonus regimes only. It consists in organizing less monitoring sessions than announced to save bonus expenses (monitoring less means that non-shirking employees receive bonuses less often). It does not exist under damages and efficiency wages regimes because if the non-shirking constraint is satisfied, no party has to pay anything after a monitoring session. Still, the solution for conditional bonuses is simple. Creating a bonus fund, the magnitude of which is contractually determined, such that non-paid bonuses are not earned by the employer but remain in the fund until they eventually go to other employees is a form of decoupling that may solve the problem.
[23] We have borrowed the concept of 'decoupling' from the economic literature on tort and litigation. See Polinsky and Che (1991) on 'decoupling liability'.

each period so that the bonus payment expenses of the employer do not decrease by falsely declaring that an employee 'shirked'.

The decoupling mechanism under non-retroactive efficiency wages, however, is less rational than the decoupling mechanisms that can be organized under damages or conditional bonuses. The reason is that non-retroactive efficiency wages decouple towards a third party (the newly hired employee), who is not a part of the original contractual relationship between the former employee and the employer. The decoupling mechanisms under damages and conditional bonuses that we described distribute to parties with whom the employer has a contractual relationship. Therefore, these benefits remain internalized. Rational contracting parties will try not to create positive externalities for third parties and will therefore prefer internal to external decoupling. Non-retroactive efficiency wages cannot be decoupled internally: if the rent that the shirking employee would have received in later periods goes to the other employees, then the newly hired employee would receive no rents and would have no incentives not to shirk.

8   Concluding Remarks

Efficiency wages are (a) rents that (b) are also paid when no monitoring takes place. Moreover (c), in an *n*-period model, non-fully retroactive efficiency wages allow employees who are caught shirking to keep the rents they received in previous periods in which they were not monitored. Efficiency wages are a more expensive incentive device than damages because of characteristic (a) and a more expensive incentive device than conditional bonuses because of characteristics (b) and (c).

Under what conditions, then, will efficiency wages ever be used? Since efficiency wages are a less effective incentive device than the two alternative devices we mentioned, they will be used only when it is impossible to use any of the two alternatives. This means that two conditions must be simultaneously fulfilled.

First, it must be impossible to work with damages. One reason why this could be the case is that the law forbids damages. This is what many contemporary employment regulation systems seem to do to some extent. Note, however, that when the market wage is sufficiently above the legal minimum wage, damages can mathematically be rewritten as normalized efficiency wages. For instance, if the market wage is 100, the damages 20, and the minimum wage is 80 or lower, a contract with a damages regime can be rewritten as a contract with a standard wage of 80 plus a bonus payment of 20 if no shirking has been discovered. We are not aware of any legal obstacles to such 'bonus' schedules. In section 2.6, we explained why such a normalized

efficiency wages regime is in essence just another name for a damages regime. No efficiency wages are paid since no rent is paid.

A second reason why working with damages may not be possible is the potential insolvency of the employee. In proposition 3, however, we have shown that these insolvency problems must be relevant before parties would switch to bonus systems: one necessary (though not sufficient) condition is that the cost of a monitoring session is higher than the highest damages the employee is able to pay. Even then, as shown in proposition 4, parties will work with combined sanctions rather than with a purely positive sanction system.

The second condition that must be fulfilled is that it should be impossible to work with conditional bonuses. Since appropriation problems can be solved by decoupling, we can think of only one major type of circumstance under which conditional bonuses cannot be used. That is, when a monitoring session can sometimes show that an agent shirked, but can never reveal that the agent did not shirk. A good example is corruption. A typical feature of corruption is that in principle the only parties who have information on the corrupt act are the two parties involved (the agent and a third party). Therefore, it may be possible to prove corruption only when that third party betrays the agent or when the principal coincidentally acquires the relevant information. As a consequence, an employer who wants to know whether an employee is corrupt may not know more after organizing a monitoring session. To put it different, the employer cannot go after evidence but has to wait until the evidence comes to him.

To see why under those circumstances efficiency wages do and conditional bonuses do not incentivize, consider that under conditional bonuses, there are three possible outcomes: (a) no monitoring takes place, in which case the employee receives no bonus; (b) monitoring takes place and the employee is found non-shirking, in which case he earns the bonus; and (c) monitoring takes place and the employee is found shirking, in which case he receives no bonus. If outcome (b) does not exist because it is impossible to prove that an agent is non-shirking, then conditional bonuses make no sense, since there is no difference between (a) and (c) and hence no incentive not to shirk. Under efficiency wages, however, employees receive a bonus under (a) and nothing under (c), so that employees have an incentive not to shirk.

Why did Shapiro and Stiglitz (1984) come to more positive conclusions regarding efficiency wages? The reason, in essence, is that they insufficiently considered the alternatives: damages and conditional bonuses. They briefly discussed a form of damages ('bonds') but concluded that they may not work because of the appropriation problem. They did not realize, however, that the appropriation problem affects all enforcement devices, unless the sanctioning party does not gain anything from sanctioning because of decoupling. Non-retroactive

efficiency wages decouple in an irrational way: they give what the sanctioned party loses to someone who is not a partner in the original contract.

Shapiro and Stiglitz (1984) also overlooked the possibility of working with conditional bonuses, but they seem to have analyzed situations in which conditional bonuses make no sense. In their basic model, apprehension rates are exogenous. The cases where conditional bonuses cannot work, largely coincide with the case where monitoring is exogenous. Exogenous monitoring means that the principal cannot increase the level of evidence. The most likely cases are when the evidence has to come from a third party or when detecting shirking is a matter of pure coincidence. A typical example is corruption, which we discussed above.

It is remarkable that in the real world, most forms of shirking for which the apprehension rate is exogenous are subject to criminal sanctions. Examples are corruption, insider trading, or fraud. This means that the types of shirking Shapiro and Stiglitz may have had in mind are, in the real world, deterred by negative sanctions with a built-in decoupling mechanism (criminal sanctions are subject to public enforcement, which means that it is not the victim who sanctions the agent, but a third party.)

We now turn to some implications of our findings for labor economics, economic analysis of contract law, and anti-corruption policies.

With respect to *labor economics*, our analysis suggests that the importance of efficiency wages as a mechanism to reduce monitoring costs has been overstated in the literature. Efficiency wages make sense only under very exceptional conditions. Therefore, it seems extremely implausible that employers will ever opt for efficiency wages on such a massive scale that structural unemployment would result.

The economic literature on *contract law* has until now focused on negative sanctions for breach (damages, specific performance, liquidated damages and penalty clauses) and has nearly completely overlooked positive sanctions. Our analysis suggests indeed that negative sanctions are in general superior to positive sanctions.

Finally, our analysis suggests that the overpayment of judges and other public officials is not a first-best method to conquer *corruption*. The first best method is to work with negative sanctions, like damages or criminal sanctions. In all legal systems, corruption is subject to criminal sanctions. As long as these criminal sanctions are not at their highest possible level (given other concerns like marginal deterrence), it seems irrational to work with efficiency wages, unless efficiency wages serve other functions like income-satiation or gift-exchange, which were beyond the scope of this paper.

References

Akerlof, George (1982), 'Labor Contracts as Partial Gift Exchange', **97** *Quarterly Journal of Economics*, 543-569.

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', **3** *Journal of Legal Studies*, 1-18.

Cooter, Robert D. and Garoupa, Nuno (2001), 'The Virtuous Circle of Distrust: A Mechanism to Deter Bribes and Other Cooperative Crimes', *The Berkeley Law & Economics Working Papers*, Vol. 2000, No. 2, Article 13.

Polinsky, A. Mitchell and Che, Yeon-Koo (1991), 'Decoupling Liability: Optimal Incentives for Care and Litigation', **22** *Rand Journal of Economics*, 562-570.

Polinsky, A. Mitchell and Shavell, Steven (2001), 'Corruption and Optimal Law Enforcement', **81** *Journal of Public Economics*, 1-24.

Polinsky, A. Mitchell, and Shavell, Steven (2000), 'Public Enforcement of Law', in Bouckaert, Boudewijn and de Geest, Gerrit (eds.), *Encyclopedia of Law and* Economics, Cheltenham, Edward Elgar, vol. V, 307-344.

Rasmusen, Eric (1992), 'An Income-Satiation Model of Efficiency Wages', **30** *Economic Inquiry*, 467-78.

Shapiro, Carl and Stiglitz, Joseph E. (1984), 'Equilibrium Unemployment as a Worker Discipline Device', **74** *American Economic Review*, 433-444.

Shapiro, Carl and Stiglitz, Joseph E. (1985), 'Can Unemployment be Involuntary: Reply', **75** *American Economic Review*, 1215-1217.

Shavell, Steven (1997), 'The Optimal Level of Corporate Liability Given the Limited Ability of Corporations to Penalize Their Employees', **17** *International Review of Law and Economics*, 203-213.