



# Employee Crime and the Monitoring Puzzle

## Citation

Dickens, William T., Lawrence F. Katz, Kevin Lang, and Lawrence H. Summers. 1989. Employee crime and the monitoring puzzle. *Journal of Labor Economics* 7(3): 331-347.

## Published Version

<http://dx.doi.org/10.1086/298211>

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:3645199>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# Employee Crime and the Monitoring Puzzle

William T. Dickens, *University of California, Berkeley, and  
National Bureau of Economic Research*

Lawrence F. Katz, *Harvard University and National Bureau  
of Economic Research*

Kevin Lang, *Boston University and National Bureau  
of Economic Research*

Lawrence H. Summers, *Harvard University and  
National Bureau of Economic Research*

The simplest economic theories of crime predict that profit-maximizing firms should follow strategies of minimal monitoring with large penalties for employee crime. We investigate possible reasons why firms actually spend considerable resources trying to detect employee malfeasance. We find that the most plausible explanations for firms' large outlays on monitoring of employees—legal restrictions on penalty clauses in contracts and the adverse impact of harsh punishment schemes on worker morale—are also consistent with the payment of premium (rent-generating) wages by cost-minimizing firms.

We have benefited from helpful discussions with George Akerlof, Jeremy Bulow, Louis Kaplow, Alan Krueger, Edward Lazear, Kevin M. Murphy, Steven Shavell, and Janet Yellen. We thank Alan Gerber for excellent research assistance. This research was supported by Lang's Olin Fellowship at the National Bureau of Economic Research, the Institute of Industrial Relations at Berkeley, and National Science Foundation grants SES-8409380, SES-8606139, SES-8553244, and SES-8809200.

[*Journal of Labor Economics*, 1989, vol. 7, no. 3]  
© 1989 by The University of Chicago. All rights reserved.  
0734-306X/89/0703-0005\$01.50

Employee theft is believed to transfer between \$15 and \$56 billion per year from businesses to their workers and to account for between 5% and 30% of business failures each year!<sup>1</sup> These numbers can be placed in perspective by noting that federal corporate income tax revenues in fiscal year 1987 were \$104 billion and that the financial cost of street crime is estimated at approximately \$4 billion (Arnold 1985).

Employee pilferage is especially conspicuous in some industries. It is believed that 80% of shipping losses in the freight shipping and airport cargo handling industries arise from employee theft (Willis 1986). Forty-two percent of all retail shrinkage, which averages 1.8% of sales, is due to employee theft according to a 1986–87 survey of 113 retailers (Arthur Young and Company, 1987). Hollinger and Clark (1983) find that about 30% of retail employees misuse discount privileges or directly steal merchandise from their employers, 27% of hospital employees take hospital supplies at least once a year, and about 9% of manufacturing workers falsify their time cards. It has also been estimated that unauthorized use of the federal long-distance telephone system by employees inflates the federal phone bill by \$89.5 million a year (Lipman and McGraw 1988).

While some employee theft is sanctioned as a form of compensation, a great deal of employee crime and shirking, including, most obviously, sabotage and vandalism, almost certainly does not represent an efficient transfer from firms to workers. An extreme example of inefficient employee “time theft” comes from the Peach Bottom nuclear power plant, which was fined \$1.35 million (*New York Times*, August 12, 1988) and is expected to be closed for over 20 months because of an inability to control widespread sleeping in the control room (Wald 1988). This closure has resulted in the owners “spending about \$14 million a month for replacement power, most of the cost borne by stockholders.”

To our knowledge, employee crime has received little attention from economists. This is unfortunate. Employee theft provides a natural testing ground for economic theories of principal-agent relationships generally and of crime specifically. In his seminal analysis of the economics of crime, Gary Becker (1968) pointed out that the cost of achieving any given degree of deterrence is minimized by combining an infinitesimal probability of detection and an arbitrarily large punishment. One might expect this strategy to better describe employers’ responses to employee malfeasance than

<sup>1</sup> An American Management Association study reported in Hollinger and Clark (1983, p. 3) estimated that in 1975 employee theft (employee pilferage, kickbacks and bribery, fraud, and embezzlement) cost businesses between \$17.5 and \$29 billion. These figures were converted to 1987 dollars by the Gross National Product deflator yielding a range of \$34 to \$56 billion. Shepard and Dutson (1988) conclude from a review of the existing literature that a conservative estimate of the annual economic losses of U.S. businesses from employee theft ranges from \$15 to \$25 billion. Evidence on the role of internal theft in business failures is discussed by Lipman and McGraw (1988).

it does the law enforcement behavior of governments, whose actions are constrained by notions of fundamental fairness and constitutional limitations that *punishments should be set to "fit the crime."* Indeed, Mirrlees (1974) showed that the "burn them in oil" solution arises naturally in certain principal-agent models. Moreover, employee crime is perhaps the most vivid, although not the most important, form of employee shirking. It can therefore shed light on efficiency-wage theories in which effort-elicitation considerations may lead firms to pay premium (rent-generating) wages (e.g., Shapiro and Stiglitz 1984).

This article offers some observations on employee crime, economic theories of crime, limits on bonding, and the efficiency-wage hypothesis. We begin by demonstrating that the simplest economic theories of crime predict that profit-maximizing firms should follow strategies analogous to Kolm's (1973) recommendation that an optimizing government should "hang tax evaders with probability zero." Finding overwhelming empirical evidence that firms expend considerable resources trying to detect employee malfeasance, we investigate a number of possible reasons why the simple model's predictions fail. It turns out that some of the plausible explanations for firms' large outlays on monitoring of employees also justify the payment of premium wages. A priori arguments that firms should not pay efficiency wages become much weaker once it is recognized that these firms expend significant resources on monitoring.<sup>2</sup>

Section I lays out the simplest formal model of employee crime. Section II discusses a number of respects in which this model is clearly inadequate as an explanation for observed personnel practices. Section III examines a number of possible reasons for the model's failure. Section IV discusses the implications of our analysis for the efficiency-wage hypothesis. Section V concludes this article.

### I. A Formal Model

We illustrate our basic argument by considering a firm's decision about monitoring intensity, wages, and bonding in a simple 1-period model. Consider a very simple form of employee shirking/"crime"—failing to supply the effort contracted for by the firm (or analogously stealing an equivalent amount of output). At the beginning of the period workers may post a bond  $b$ . During the period, they may or may not supply effort  $e$ , which takes on the value one if effort is supplied and zero otherwise. Workers, whose only source of income is employment, maximize a utility function— $U = C + (1 - e)$  where  $C$  represents consumption. The probability of a shirking worker being detected is given by  $p(m)$  where  $m$  represents outlays on monitoring. We assume that  $p' > 0$ ,  $p'' < 0$ , and that

<sup>2</sup> This point has been noted before by Eaton and White (1983) and Bulow and Summers (1986).

$p(0) = p^* > 0$ .<sup>3</sup> If workers are caught shirking, they forfeit their bond and do not receive the wage  $w$ . However, they are immediately employed elsewhere and earn the reservation wage,  $r$ ,  $r > 1$ . If they are not caught shirking, they receive the wage and their bond is returned.

It is apparent that it always pays the firm to make arrangements that induce its labor force to work rather than to shirk. The only possible equilibria are those in which  $e = 1$ . The firm's problem is

$$\max Y = R(L) - (w + m)L, \quad (1)$$

with respect to  $L$ ,  $w$ ,  $m$ , and  $b$ , and subject to the constraints that

$$w \geq r, \quad (1a)$$

$$w + b \geq 1 + [1 - p(m)](w + b) + p(m)r, \quad (1b)$$

$$m \geq 0, \quad (1c)$$

where  $Y$  represents the firm's income,  $L$  represents labor input and  $R(L)$  is revenue net of nonlabor costs. The first constraint implies that the firm must pay its workers at least the reservation wage. The second is the no-shirk condition requiring that given the wage, monitoring intensity, and posted bond, workers prefer to work rather than to shirk. The third constraint guarantees that monitoring expenditures are nonnegative.

Exhibiting the first-order conditions to the firm's problem is not very informative. It is straightforward to show that the solution to (1) is given by

$$m = 0, \quad (2a)$$

$$w = r, \quad (2b)$$

$$b \geq 1/p^*. \quad (2c)$$

This solution is clearly feasible. To see that it is optimal, one may note that (1a) requires that  $w$  be as great as  $r$ , (1c) insures that  $m$  cannot be negative, and that the bond  $b$  does not enter the firm's objective function (1).

Although we have analyzed a simple single-period model with risk-neutral workers, the basic result that firms should not spend resources on

<sup>3</sup> This guarantees that, even if the firm spends no money on monitoring, there is a positive probability that shirking workers will be detected. This simplifies the proof and seems reasonable since normal management coordination activities should yield a positive detection probability even without explicit monitoring expenditures.

monitoring workers when bonding is unrestricted is easily shown to be valid in multiperiod settings. Indeed, the possibility of basing wages on seniority actually eases the firm's effort-elicitation problem (Lazear 1981). The basic conclusions of this section also follow when workers are risk averse as long as inspection errors in which the firm falsely identifies innocent workers as shirkers are not a major problem. We address this latter issue below.

## II. The Pervasiveness of Monitoring and Employee Crime

Cursory inspection of almost any workplace indicates that firms expend substantial resources monitoring employees, and that much of this monitoring activity is aimed at deterring shirking and stealing by workers. Data from the 1983 Current Population Survey (CPS) indicate that, even under a fairly narrow definition of supervisory jobs, approximately 7% of workers in the U.S. nonagricultural, private sector are employed as supervisors or inspectors.<sup>4</sup> Even if a small fraction of their time were devoted to controlling worker misconduct, such expenditures would be quite substantial. Furthermore, a recent survey indicates that U.S. businesses spend approximately \$12 billion a year on security products, personnel, and services, and that a major concern in budgeting these expenditures is the control of employee crime (Whitehurst 1987). Arthur Young and Company (1986) report that a group of 168 large retailers spent an average of 0.42% of sales on security and loss prevention in 1985 and that employee theft was the largest component of the inventory losses these expenditures attempted to control. The Bank of America utilizes "20 people, a lot of computer software and about \$1 million a year" to keep tabs on the 3,500 employees in its credit card division (Dolan 1985, p. 15).

Furthermore, there is some evidence that the intensity of monitoring is increasing over time. In 1974, only 10% of Fortune 1000 companies used undercover security agents. By 1983, the incidence had increased to 50% (Lipman and McGraw 1988). Shepard and Duston (1988) report that employer demand for guard and investigative services is expected to increase rapidly through the early 1990s because of growing business concern over employee theft, drug abuse, and similar problems.

In gauging the extent of monitoring outlays, there is the obvious problem that supervisors produce a joint product: they both deter shirking and help to organize production. One striking example suggesting the presence of more than minimal monitoring to deter employee cheating is the typical company's treatment of employee expense accounts. Rather than spot checking these accounts and attaching large penalties to detected abuses,

<sup>4</sup> We classified as supervisors employees in the following three-digit 1980 Census of Population occupation codes: 35-36, 243, 303-306, 413-415, 433, 448, 456, 485, 489, 494, 497, 503, 553-558, 613, 633, 689-693, 796, 803, and 843.

firms inspect to at least some degree every request that is submitted. Furthermore, many firms hire outside accountants who periodically audit the company's financial records to prevent frauds and thefts committed by managers with the aid of staff accountants or bookkeepers. It is difficult to view these activities as costless byproducts of supervisory activities aimed at increasing the efficiency of production.

As noted in the introduction, despite this extensive monitoring the level of employee crime is considerable. Firms appear to be unwilling to adopt a level of deterrence sufficient to deter employee deviant behavior. An excellent example is provided by their response to frequent flyer programs.<sup>5</sup> When firms pay for their employees' travel, any discounts logically belong to the firm. Moreover, when employees receive rewards for traveling on particular airlines, they may make travel decisions which are not in the best interest of the firm. Experts estimate that firms could save up to \$3 billion annually by recovering their employees' travel awards. While one major oil company has three clerks who spend all their time claiming bonus awards for corporate use, it appears that most major companies have given up the effort to collect bonus miles because the costs outweigh the potential benefits. None have tried to ensure compliance by announcing a program of spot checks and severe penalties as would be suggested by our simple model.

It might be argued that what is labeled employee crime simply reflects an efficient transfer of resources from the firm to the worker. Wages may be reduced to offset anticipated employee cheating. While this is clearly true in some contexts (e.g., workers drinking the remaining wine from a company function), this is unlikely to be the whole story. The extra travel costs to companies of their employees arranging flights to capture frequent flyer miles almost certainly is greater than the value of the bonus trips earned by employees. A shirking manager may impose costs on the firm far greater than his personal gains.<sup>6</sup> The efficient-compensation argument is also not appropriate when workers steal supplies that they need to expend considerable resources to sell.

### III. Explaining Monitoring and Employee Crime

The pervasiveness of monitoring outlays and employee crime makes it apparent that some important consideration is omitted from our statement of the firm's problem. We first examine a modification of (1) in which it may be efficient for firms to expend resources monitoring workers although there are no constraints on bonding. Finding this to be an incomplete explanation, we then take up four possible considerations that may limit

<sup>5</sup> This discussion of frequent flyer programs is based on Brancatelli (1985).

<sup>6</sup> This point is emphasized by Klein (1980).

bonding: liquidity constraints, firm moral hazard, legal restrictions on the form of contracts, and social limits on enforceability.<sup>7</sup>

### False Positives and Worker Risk Aversion

Polinsky and Shavell (1979) have suggested that if there is a risk of false detection and citizens are risk averse, a strategy of low expenditures on detection and large penalties for alleged offenders will not maximize most standard social welfare functions even though that strategy conserves on law enforcement resources. In parallel fashion, one can argue that firms expend money monitoring workers to insure them against being falsely labeled as shirkers and having to forfeit their bonds. Positive monitoring outlays and smaller bonds may be optimal because of the high wage necessary to compensate workers for the risk of having a large bond erroneously forfeited.

We doubt that this consideration explains a substantial part of actual monitoring activities. As long as it is possible for firms to maintain some positive probability, no matter how small, of detecting shirking workers without risking false positives, firms should not spend resources on monitoring; they should instead require large bonds. Indeed, even if workers are risk averse and inspection errors are possible, Nalebuff and Scharfstein (1987) and Bolton (1987) show that as long as accurate (but not necessarily perfectly accurate) tests are available and large penalties are feasible, there is still no reason to spend more than an infinitesimal amount on monitoring.

Experience with improved monitoring technologies casts doubt on the hypothesis that they largely benefit workers by insuring them against unjust dismissal. More accurate monitoring by computers is so strenuously objected to by workers that there have been recent attempts to introduce legislation in Congress to restrict employer use of computer-based performance-monitoring technologies.

### Limits on Bonding

A natural way to modify the problem stated in equations (1) so that it has a plausible solution is to add a constraint on the size of bonds— $b \leq b^* < 1/p^*$ . This additional constraint will bind. The no-shirk condition (1b)

<sup>7</sup> There is a further traditional objection to the low detection probability/large penalty approach to deterrence. If maximum penalties are exacted for minor infractions, people have no incentive to commit minor rather than major crimes (Stigler 1970). If a thief is executed for taking \$10, he might as well take \$10,000. Marginal deterrence requires marginal punishments. We have difficulty seeing how creating marginal deterrence of more extreme forms of shirking could be an important consideration leading firms to closely monitor workers. If there are no limits on bonds, firms can set bonds large enough to deter any type of shirking without expending resources on monitoring. Firms could also create marginal deterrence by making the fraction of the bond forfeited a function of the seriousness of detected malfeasance.



directly implies that, in the presence of this constraint on the size of the bond, either the firm will pay wage premia or spend on monitoring or both. We observe spending on monitoring. If workers are heterogeneous in their shirk-and-theft propensities and these differences are not perfectly observed by employers, many firms are unlikely to find it profitable to increase wages and/or their monitoring intensities high enough to deter all worker malfeasance when bonds are limited. In this case, we are likely to observe the occurrence of shirking and employee crime. We would also see firms expending resources on screening workers to determine their trustworthiness. In fact, background checks, psychological tests, and polygraph tests are widely used in some industries to try to screen out “theft-prone” individuals. The problem then is to explain limits on bonds. Below we consider several possible causes for such limits.

### Liquidity Constraints

Perhaps the most direct explanation for the failure of workers to post bonds is that they lack sufficient cash and are liquidity constrained. Workers cannot post bonds with liquid assets that they do not have. A similar argument is often used to explain why societies rely on prisons rather than fines to punish criminals. While superficially plausible, this argument must confront an immediate question. Even if perfect bonding is not possible, why do firms not take at least some cash from newly hired workers or at least set initial wages to zero? If liquidity constraints were the only limitation on bonding, firms would ask workers for whatever up-front payments workers could make—at least up to the point where workers’ reservation utility constraint is exactly satisfied (Carmichael 1985). Most workers have at least some liquid assets. Yet explicit, up-front bonds and entrance fees appear to be quite rare in practice. Thus, liquidity constraints do not appear to provide an adequate explanation for limits on bonding.

### Firm Moral Hazard

A second standard reason why firms cannot rely fully on bonds is firm moral hazard. Once a bond is posted, a firm has a strong incentive to label a worker a shirker and to claim his bond. Similarly, when inspections can generate false positives, firms may be unable to commit to a low rate of inspections once a large bond is posted (Bolton 1987). Except for the unlikely event that third parties can be relied on to accurately determine whether a worker has shirked, workers will be willing to post bonds only if they are convinced that the firm will not take these bonds under false pretense. When workers are uncertain of the trustworthiness of firms, they are unlikely to be willing to post large bonds.

It is often suggested (e.g., Carmichael 1985) that the use of third parties can circumvent these moral-hazard difficulties. The likelihood of firm default can be reduced if the firm does not expect to gain anything from a

default. For example, baseball teams fine their players but give the proceeds to charity. But a moment's thought makes it clear that third-party payments do not really eliminate the firm's moral-hazard problem. Imagine that if a worker is deemed to have shirked, his bond will go to the Red Cross, to whom his employer is indifferent. The employer can nonetheless benefit by threatening to forfeit the worker's bond unless the worker works harder, accepts poorer conditions, or agrees to work for a lower wage in the future. The point is simple. Would the reader be willing to entrust us with \$10,000 even if we could not keep it but could only give it to charity? If we were well positioned to ask for a favor, it seems unlikely. Firm moral-hazard considerations will thus limit not only simple bonding arrangements but also arrangements that include third parties.<sup>8</sup>

It seems clear that moral-hazard considerations do place some constraint on the size of bonds that firms can ask their workers to post.<sup>9</sup> But we doubt that this constraint binds for the typical employer. If firm moral-hazard constraints often bound, we would expect large, established firms with significant reputations to pay lower wages (net of the bond) than small, less established firms. On the contrary, large firms appear to pay higher wages at all experience levels (Brown and Medoff, in press). Perhaps there are limits on the size of bonds that bind before reputation constraints are reached for many firms.

### Legal Restrictions on Contracts

There is the further issue of what types of contracts governments will enforce. An important aspect of American and English common law is the refusal of the courts to enforce contract provisions that are found to be "penalties." In particular, contract provisions calling for "liquidated damages" for breach of contract that exceed the loss caused by the breach are viewed as penalties and not enforced.<sup>10</sup> When the detection probability

<sup>8</sup> Note that the problems with third-party schemes are not avoided by using a workers' colleagues as the third party as suggested by the burgeoning literature on tournaments (e.g., Malcolmson 1984).

<sup>9</sup> Moreover, if firm moral hazard limits bonds, it also limits efficiency-wage payments since, with finite-horizon contracts, efficiency-wage earnings profiles involve a large lump-sum payment at the end of the contract (Akerlof and Katz, in press).

<sup>10</sup> Posner (1972, p. 59), in his discussion of contract law, writes: "The law permits parties to a contract to specify in advance the damages to be assessed in the event of a breach. Their specification will be enforced unless the court finds that the parties' intention was to impose a penalty rather than to estimate the actual damages caused by the breach." The current practice of courts with respect to liquidated damages is codified in the Restatement (Second) of Contracts (cited in Ferris [1982]): "A term fixing unreasonably large liquidated damages is unenforceable on grounds of public policy as a penalty." Farber (1980) notes that under this "penalty doctrine" courts will not award supercompensatory damages that exceed the harm caused by the breach.

is low, performance bonds far in excess of the actual damage done to the firm by any particular episode of detected worker shirking may be required to provide adequate deterrence. Contracts in which workers forfeit large, explicit bonds when caught shirking are unlikely to be enforced by the courts. This legal limitation on the size of penalties means that firms will need to expend resources on monitoring to raise detection probabilities.<sup>11</sup>

It is interesting to note that the termination of an employee for shirking is not typically viewed as a “penalty” by the courts. Firms have traditionally been able to fire employees at will even if lost seniority pay exceeds any damage caused by employee shirking. The termination of an employee known to shirk is viewed as a legitimate business decision since the employee may be expected to shirk again in the future. On the other hand, if a firm tried to deter shirking by making a worker explicitly post a large, up-front performance bond, the courts would be likely to treat the confiscation of the bond from a worker caught shirking as an unenforceable penalty.

A seemingly strong case for attributing the absence of bonds to legal limitations comes from an examination of the history of penal clauses in contracts. Penal bonds, in which the entire amount of the bond was due immediately upon a breach regardless of the actual damage caused by the breach, were commonly used to secure performance of contracts in early English legal history (Ferris 1982). Courts, however, often intervened to mitigate the harsh results where the breach did not cause any actual damage. The legal gymnastics that succeeded in nullifying the required forfeiture of a “pound of flesh” for contractual breach in *The Merchant of Venice* provides a vivid dramatization of the unwillingness of courts to enforce “inequitable” penal bonds even in this time period. Eventually, legislatures in both England and the United States prohibited the use of penal bonds.

Beyond the general prohibition of penal bonds, there are various specific prohibitions of practices that look like bonding in both the housing and labor markets. Despite the plausible efficiency case that could be made for a rich variety of contracts, most American jurisdictions have laws that prevent landlords from asking their tenants for more than 1 or 2 months’ rent as a security deposit. Explicit legal rules regarding the labor market prohibit the actions most like those that firms utilizing a strategy of low detection probabilities and large penalties would pursue. The most natural way for a firm to implement a strategy of bonding through deferred payments would be to have an unvested pension. Yet firms are required to vest pensions after a certain number of years. Actions to dismiss highly paid senior workers and replace them with younger ones are often subject to legal challenge in the United States (especially since the passage of the

<sup>11</sup> Farber (1980) argues analogously that the prohibition of supercompensatory damages in commercial transactions also leads to excessive levels of breach and excessive monitoring of contract compliance.

Age Discrimination Act of 1978) and are unthinkable in parts of Western Europe.

The frequent unwillingness of many societies to enforce contracts in which one party is empowered to “punish” another is somewhat mysterious. Even parties with very unequal strengths should be able to write efficient contracts. Limiting one type of provision hardly seems to do much to protect “weak” parties to contracts.<sup>12</sup>

The common law restriction of supercompensatory liquidated damages is often explained by suggesting that it is desirable to avoid incentives for “induced breach,” situations where one party to a contract tries to get another to breach it. The unenforceability of penalty clauses may also save on transaction costs by saving transactors the costs of “reading the fine print” of a contract to determine whether it contains a hidden penalty for breach. Furthermore, as Becker and Stigler (1974) argue, large penalties attached to dismissal will generate costly attempts by employees to litigate against efforts to fire them. Since parties are not fully charged for court services in resolving disputes, and since courts cannot abdicate completely from disputes involving liquidated damages provisions, the socially efficient outcome may be simply not to enforce such provisions.

These arguments raise the possibility that legal restrictions may help explain why firms devote resources to monitoring and rely on termination threats rather than cash penalties to deter inappropriate behavior by their workers.<sup>13</sup> Perhaps these laws reflect broader social sentiments that would constrain employer behavior even in the absence of formal legal restrictions.

### Social Limits on Bonding

While legal constraints on penal bonds may not fully explain the absence of bonds, they nevertheless reflect a widely shared opposition to such contractual arrangements. Individuals in charge of enforcing the use of punitive measures will be reluctant to punish workers if the punishment exceeds the crime. Firms are also reluctant to use certain punishment strategies because of their adverse effects on the attitudes and performance of nonshirking workers and on their public image. We suspect that these “enforcement” problems are the most empirically important considerations limiting the use of performance bonds.

Carr-Hill and Stern (1979), in an analysis of the criminal law which closely parallels our analysis of the employment relation, conclude that a

<sup>12</sup> On the other hand, restrictions against penalties have been justified under the doctrine of unconscionability as protecting parties who may have signed contracts under duress (Rea 1984).

<sup>13</sup> However, there are some problems with this view. If the “penalty clause constraint” bound frequently, we would expect to see a large caseload defining precisely its limits in labor contracts. In fact, outside of cases involving the termination of franchises, the penalty clause issue appears to have aroused a small caseload and little practical interest in labor-market settings.

positive theory of criminal law enforcement must include a cost of departing from the socially determined appropriate punishment. The same considerations of fairness that limit the punishments society imposes for criminal offenses also limit society's willingness to enforce contracts in which persons put themselves at the mercy of other private parties. These considerations are particularly important in limiting firms' ability to discipline, fine, and discharge workers.

In a classic discussion of the influence of criminal law on industrial penalties, Sanford Kadish (1964, p. 125) notes that "the criminal law and the process of disciplining employees for unsatisfactory conduct are peas from the same pod; that as a consequence each system gives rise to fundamental issues which are essentially similar." Kadish (p. 127) argues that "punishment must serve to accomplish its purposes at a cost to an individual which is not regarded as excessive" and that punishments that violate this condition will be viewed as "arbitrary and unfair." If punishments appear excessive relative to the offense, as is likely to be the case when firms follow a low monitoring intensity/high penalty strategy, these penalties will tend to be nullified in practice by persons charged with the administration of the discipline system and may create such a sense of injustice as to damage the firm's relations with its employees.

Akerlof (1982) follows dozens of personnel textbooks in stressing the importance of reciprocal good will (gift exchange) in maintaining employee morale and efficient production. Firms may find that asking workers to post bonds may signify a lack of trust and may negatively affect workers' morale and productivity. Akerlof's data on bill posters and everyday observation suggest that many workers do more than their jobs' stated requirements despite the apparent absence of pay or promotion incentives to do so.<sup>14</sup> If measures to prevent shirking by a few antagonize the many, these measures may be counterproductive. There are many things people will do if asked but refuse to do when threatened.<sup>15</sup>

There can be no serious question that employers who generate "good will" in the workplace are likely to enjoy greater productivity than those who do not. Letting a small minority of miscreant workers go seems to be a strategy that deters shirking or theft more effectively than alternative approaches that are likely to give more offense to workers who do not shirk.

<sup>14</sup> Titmuss (1971) provides a dramatic example of how partial reliance on pecuniary incentives can undermine "moral" behavior in the context of blood donations.

<sup>15</sup> A similar point is made by Carr-Hill and Stern (1979, pp. 294-95) when they note: "Someone who is subject to punishment he sees as unjust, may be so embittered that he is more likely to offend than if punishment had been lower. Thus offences may increase with punishment over a certain range."

#### IV. Efficiency Wages and Monitoring

The foregoing arguments give some insight into the structure of firms' compensation arrangements. Firms might find courts unwilling to enforce contracts involving large bonds, and large bonds might have counterproductive effects on worker morale. Under these conditions, firms will expend resources on monitoring and may find it profitable to pay workers premium wages. In fact, two experts on private security and employee theft conclude that "the most important control against employee theft is a good salary or wage" (Lipman and McGraw 1988, p. 58).

But this line of argument begs a fundamental question. If it is somehow efficient to give workers rents on the job for reasons of deterring shirking and crime or anything else, why do profit-maximizing firms not extract these rents up front from workers queued for such jobs? If the direct sale of jobs is somehow precluded, perhaps by legal restrictions, firms could surely find other ways of extracting some surplus from queuing workers—by requiring them to do menial work as a condition for attaining attractive jobs or by requiring them to purchase their own uniforms or training. Such rent-extraction strategies would cause the reservation-wage constraint (1c) to bind and so eliminate any involuntary unemployment or queuing even if explicit bonding or large punishments were impossible.

This is the essence of the challenge to efficiency-wage theories put forward by Carmichael (1985) and Murphy and Topel (1987). The argument surely proves too much, for it appears to be empirically invalid in at least some circumstances. For example, most economists believe that minimum wage laws give rise to at least some job rationing. This prediction is supported by data on job application rates analyzed by Holzer, Katz, and Krueger (1988). Arguments paralleling those of the efficiency-wage critics have been used to suggest that no queues should result from minimum wage laws because firms fully extract rents by adjusting working conditions, training, or fringe benefits (e.g., Wessels 1980). This does not seem to happen. Nor do ways seem to be found of dissipating all the rents and thereby eliminating the queues associated with rent-control laws, price controls, or certain dramatic instances of the payment of efficiency wages (e.g., the Ford case documented in Raff and Summers [1987]).

At least two different explanations can be given for why one does not observe all rents being eliminated, even where devices to do so are available.<sup>16</sup> First, firms may care about the ease with which they can recruit labor. Firms' ability to recruit labor depends upon the lifetime utility of the total compensation package associated with their job offers. Raising

<sup>16</sup> See Holzer, Katz, and Krueger (1988) for a more detailed analysis of the factors determining whether firms find it in their interests to eliminate ex ante employment rents.

lifetime earnings may increase the unobserved aspects of the quality of applicants (Weiss 1980) and increase the speed with which firms can fill vacancies (Lang 1988). Selling jobs may be unprofitable because it reduces the value of this package and makes it more difficult to recruit. When rent extraction is inefficient, in the sense that more is taken from workers than is gained by firms, it particularly may not pay because recruiting costs are likely to rise by more than the firm gains from rent-extraction activities.

Second, it may be that the effort-elicitation considerations that shape compensation arrangements have both forward- and backward-looking elements. If workers' performance on the job depends not just on a prospective comparison of the compensation packages associated with current working and shirking decisions but also, retrospectively, on what surplus has already been earned, then it will not be profitable for employers to extract all *ex ante* rents from workers. In fact, Kahneman, Knetsch, and Thaler (1986) find that a substantial majority of surveyed households view the auctioning off of a job to qualified candidates as "unfair," and Akerlof and Yellen (1987) provide a wide variety of evidence indicating that workers often respond to personnel practices perceived as unfair through the restriction of output.

Loyalty as it is usually understood has a retrospective, as well as prospective, element. Strategies such as "tit for tat," which seem to be successful in an evolutionary sense, involve conditioning behavior on one's partner's previous actions. Any such element in employees' effort decisions would militate against the collection of up-front payments or the use of other similar devices to extract surplus. Means of surplus extraction that cost employees much more than one dollar for each dollar transferred to the employer, such as requiring those who are about to hold responsible positions to do menial work, are likely to be particularly galling and particularly subversive to morale. In a similar vein, Okun (1981) persuasively describes how firms' compensation policies during recessions are shaped by the need to maintain the loyalty of their work force when the next boom comes.

It is instructive to contrast the labor market and certain product markets. Firms that rent bicycles, cars, or video cassettes wish to deter misuse and damage but are not usually thought to create rents with their pricing. Instead they rely on deposits. At least in part this is because denying future access to those who misuse products is likely to be difficult. On the other hand, landlords often price apartments below the market to insure loyalty on the part of their tenants rather than relying wholly on a huge security deposit. Landlords do this because the threat of denial of future access is credible and because a tenant's "attitude" is important.

These considerations hardly prove that firms pay efficiency wages, but they do demonstrate that *a priori* arguments ruling them out are inappropriate.

## V. Conclusions

Employee crime provides a natural testing ground for economic theories of principal-agent relationships. The compensation of top executives, as surveyed by Jensen and Murphy (1988), provides another one. Both suggest that compensation arrangements are driven to a substantial degree by considerations that lie outside of those emphasized in conventional treatments of the economics of agency and incentives. These fairness considerations directly impinge on the structure of compensation through firms' concerns with the impact of the perceived equity of their pay practices on worker productivity as well as on their stock of public goodwill. These considerations indirectly affect compensation arrangements through legal restrictions on the types of contracts that can be used to control employee cheating. We suspect that future research in this area should focus on more extensive comparisons of actual arrangements with those predicted by theory and with the elaboration of theory to cover the arrangements we actually observe.

## References

- Akerlof, George A. "Labor Contracts as Partial Gift Exchange." *Quarterly Journal of Economics* 87 (November 1982): 543-69.
- Akerlof, George A., and Katz, Lawrence F. "Workers' Trust Funds and the Logic of Wage Profiles." *Quarterly Journal of Economics* (in press).
- Akerlof, George A., and Yellen, Janet L. "The Fair Wage/Effort Hypothesis and Unemployment." Unpublished manuscript. Berkeley: University of California, 1987.
- Arnold, Gregory B. "Employee Theft: A \$40-Billion Crime." *Management World* 14 (November 1985): 27-28, 43.
- Arthur Young and Company. *An Ounce of Prevention*, 1985-86 ed. New York: National Mass Retailing Institute, 1986.
- Becker, Gary S. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (March/April 1968): 169-217.
- Becker, Gary S., and Stigler, George J. "Law Enforcement, Malfeasance, and the Compensation of Enforcers." *Journal of Legal Studies* 3 (January 1974): 1-18.
- Bolton, Patrick. "The Principle of Maximum Deterrence Revisited." Working Paper no. 8749. Berkeley: University of California, Department of Economics, August 1987.
- Brancatelli, Joe. "And The Winner Is." *Frequent Flyer Magazine* (November 1985), pp. 64-74.
- Brown, Charles, and Medoff, James L. "The Employer-Size Wage Effect." *Journal of Political Economy* (in press).
- Bulow, Jeremy, and Summers, Lawrence H. "A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination, and Keynesian Unemployment." *Journal of Labor Economics* 4 (July 1986): 376-414.



- Carmichael, H. Lorne. "Is Unemployment Involuntary? Comment." *American Economic Review* 75 (December 1985): 1213-14.
- Carr-Hill, R. A., and Stern, N. H. *Crime, the Police, and Criminal Statistics*. London: Academic Press, 1979.
- Dolan, Carrie. "Many Companies Now Base Workers' Raises on Their Productivity." *Wall Street Journal* (November 15, 1985).
- Eaton, C., and White, W. D. "The Economy of High Wages: An Agency Problem." *Economica* (1983), pp. 175-81.
- Farber, Daniel A. "Reassessing the Economic Efficiency of Compensatory Damages for Breach of Contract." *Virginia Law Review* 66 (1980): 1443-84.
- Ferris, Susan V. "Liquidated Damages Recovery under the Restatement (Second) of Contracts." *Cornell Law Review* 67 (1982): 862-79.
- Hollinger, Richard C., and Clark, John P. *Theft by Employees*. Lexington, Mass.: Lexington, 1983.
- Holzer, Harry; Katz, Lawrence F.; and Krueger, Alan B. "Job Queues and Wages." Working paper. Cambridge, Mass.: Harvard University, September 1988.
- Jensen, Michael C., and Murphy, Kevin J. "Performance Pay and Top Management Incentives." Harvard Business School Working Paper no. 88-059. Cambridge, Mass.: Harvard University, May 1988.
- Kahneman, Daniel; Knetsch, Jack L.; and Thaler, Richard. "Fairness as a Constraint on Profit Seeking: Entitlements in the Market." *American Economic Review* 76 (September 1986): 728-41.
- Kadish, Sanford. "The Criminal Law and Industrial Discipline as Sanctioning Systems: Some Comparative Observations." In *Labor Arbitration—Perspectives and Problems: Proceedings of the 17th Annual Meeting of the National Association of Arbitrators*. Washington, D.C.: Bureau of National Affairs Books, 1964.
- Klein, Benjamin. "Transaction Cost Determinants of 'Unfair' Contractual Arrangements." *American Economic Review* 70 (May 1980): 356-62.
- Kolm, Serge-Christophe. "A Note on Optimum Tax Evasion." *Journal of Public Economics* 2 (July 1973): 265-70.
- Lang, Kevin. "Persistent Wage Dispersion and Involuntary Unemployment." Working paper. Boston: Boston University, 1988.
- Lazear, Edward. "Agency, Earnings Profiles, Productivity, and Hours Restrictions." *American Economic Review* 71 (September 1981): 606-20.
- Lipman, Mark, and McGraw, W. R. "Employee Theft: A \$40 Billion Industry." *Annals of the American Academy of Political and Social Science* 498 (July 1988): 51-59.
- Malcolmson, James. "Work Incentives, Hierarchy, and Internal Labor Markets." *Journal of Political Economy* 92 (June 1984): 486-507.
- Mirrlees, James. "Note on Welfare Economics, Information, and Uncertainty." In *Essays on Economic Behavior under Uncertainty*, edited by M. Balch, D. McFadden, and S. Wu, pp. 243-58. Amsterdam: North Holland, 1974.
- Murphy, Kevin M., and Topel, Robert H. "Efficiency Wages Reconsidered:

- Theory and Evidence." Mimeographed. Chicago: University of Chicago, May 1987.
- Nalebuff, Barry, and Scharfstein, David. "Testing in Models of Asymmetric Information." *Review of Economic Studies* 54 (April 1987): 265-78.
- New York Times*. "\$1.25 Million Fine Is Sought for Napping at Pennsylvania Reactor." (August 12, 1988).
- Okun, Arthur M. *Prices and Quantities*. Washington, D.C.: Brookings, 1981.
- Polinsky, A. Mitchell, and Shavell, Steven. "The Optimal Tradeoff between the Probability and Magnitude of Fines." *American Economic Review* 69 (December 1979): 880-91.
- Posner, Richard. *Economic Analysis of Law*. Boston: Little, Brown, 1972.
- Raff, Daniel M. G., and Summers, Lawrence H. "Did Henry Ford Pay Efficiency Wages?" *Journal of Labor Economics* 5, no. 4, pt. 2 (October 1987): S57-S86.
- Rea, Samuel A. "Efficiency Implications of Penalties and Liquidated Damages." *Journal of Legal Studies* 13 (January 1984): 147-67.
- Shapiro, Carl, and Stiglitz, Joseph S. "Equilibrium Unemployment as a Worker Discipline Device." *American Economic Review* 74 (June 1984): 433-44.
- Shepard, Ira M., and Duston, Robert. *Thieves at Work: An Employer's Guide to Combating Employee Dishonesty*. Washington, D.C.: Bureau of National Affairs, 1988.
- Stigler, George J. "The Optimum Enforcement of Laws." *Journal of Political Economy* 78 (March/April 1970): 526-36.
- Titmuss, Richard. *The Gift Relationship: From Human Blood to Social Policy*. London: Allen & Unwin, 1971.
- Wald, Matthew L. "The Peach Bottom Syndrome." *New York Times* (March 27, 1988).
- Weiss, Andrew. "Job Queues and Layoffs in Models with Flexible Wages." *Journal of Political Economy* 88 (June 1980): 526-38.
- Wessels, Walter. "The Effect of Minimum Wages in the Presence of Fringe Benefits: An Expanded Model." *Economic Inquiry* 18 (April 1980): 293-313.
- Whitehurst, Susan. "Security Forecast, 1987." *Security* 24 (January 1987): 40-44.
- Willis, Rod, "White-Collar Crime." *Management Review* 75 (January 1986): 22-32.