

NBER WORKING PAPER SERIES

EMPLOYEE CRIME, MONITORING, AND
THE EFFICIENCY WAGE HYPOTHESIS

William T. Dickens

Lawrence F. Katz

Kevin Lang

Lawrence H. Summers

Working Paper No. 2356

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 1987

We have benefited from helpful discussions with George Akerlof, Jeremy Bulow, Alan Krueger, Edward Lazear, and Janet Yellen. This research was supported by Lang's Olin Fellowship at the NBER, the Institute of Industrial Relations at Berkeley and NSF grant numbers SES-8409380 and SES-8553244. The research reported here is part of the NBER's research program in Labor Studies. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

Employee Crime, Monitoring, and the Efficiency Wage Hypothesis

ABSTRACT

This paper offers some observations on employee crime, economic theories of crime, limits on bonding, and the efficiency wage hypothesis. We demonstrate that the simplest economic theories of crime predict that profit-maximizing firms should follow strategies of minimal monitoring and large penalties for employee crime. Finding overwhelming empirical evidence that firms expend considerable resources trying to detect employee malfeasance and do not impose extremely large penalties, we investigate a number of possible reasons why the simple model's predictions fail. It turns out that plausible explanations for firms large outlays on monitoring of employees also justify the payment of premium wages in some circumstances. There is no legitimate a priori argument that firms should not pay efficiency wages once it is recognized that they expend significant resources on monitoring.

William T. Dickens
Department of Economics
University of California
Berkeley, CA 94720

Lawrence Katz
Department of Economics
Harvard University
Cambridge, MA 02138

Kevin Lang
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215

Lawrence Summers
Department of Economics
Harvard University
Cambridge, MA 02138

Employee theft is believed to transfer between \$34 and \$56 billion per year from businesses to their workers.¹ Time theft, through dishonestly reported absences for illness, habitual late arrival and early departure from work, inordinately long lunch hours, and other forms of deliberate abuse of job time, is estimated to have cost employers approximately \$170 billion (or 4.5 hours a week per employee) in 1986.² These numbers can be placed in perspective by noting that federal corporate income tax revenues last year were \$83 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 percent of shipping losses in the freight shipping and airport cargo handling industries arise from employee theft (Willis, 1986). Hollinger and Clark (1983) find that about 30 percent of retail employees misuse discount privileges or directly steal merchandise from their employers and that 27 percent of hospital employees take hospital supplies at least once a year. While some employee theft is sanctioned as a form of compensation, a great deal of employee crime including, most obviously sabotage and vandalism, almost certainly does not represent an efficient transfer from firms to workers.

To our knowledge, employee crime has received little attention from

¹An 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. We used the GNP deflator to convert this estimate to 1987 dollars.

²This figure is from a study by Robert Half International based on interviews with top executives at 330 major companies (Security, June 1987, p. 11).

economists. This is unfortunate. Employee theft provides a natural testing ground for "economic" theories of crime. 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 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." 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 lead firms to pay premium wages (e.g. Calvo, 1979; Shapiro and Stiglitz, 1984).

This paper 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 and do not impose extremely large penalties, we investigate a number of possible reasons why the simple model's predictions fail. It turns out that plausible explanations for firms large outlays on monitoring of employees also justify the payment of premium wages in some circumstances. There is no legitimate a priori argument that firms should not pay efficiency wages once it is

recognized that they expend significant resources on monitoring.³

Section I lays out the simplest possible formal model of employee crime. Section II discusses a number of respects in which it is clearly inadequate empirically. Section III examines a number of possible reasons for the model's failure and their implications for the efficiency wage hypothesis. Section IV concludes by discussing some further implications of our analysis.

I. A Formal Model

We illustrate our basic argument by considering a firm's decision about monitoring intensity, wages, and bonding in a simple one period model. Consider a very simple form of employee shirking/"crime" -- failing to supply the effort contracted for by the firm. 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 1 if effort is supplied and 0 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 $p(0) = p^* > 0$.⁴ If workers are caught shirking, they forfeit their bond and do

³This point has been noted before by Eaton and White (1983) and Bulow and Summers (1986) though it seems to have escaped the attention of critics of efficiency wage models.

⁴This 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 by allowing us to avoid the use of "limit" arguments, and seems reasonable since normal management coordination activities should yield a positive detection probability even without explicit monitoring expenditures.

not receive the wage w . However, they are immediately employed elsewhere and earn the reservation wage, r . 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:

$$(1) \text{ Maximize } Y = R(L) - (w+m)L$$

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

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

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

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

where Y represents the firm's income, L represents labor input and $R(L)$ is revenue net of non-labor 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:

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

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

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

The solution given by (2) 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 monitoring workers when bonding is unrestricted is easily shown to be valid in multi-period 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 Type II errors in which the firm falsely identifies innocent workers as shirkers are not a 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. This is borne out by a recent survey that indicates 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 percent 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 3500 employees in its credit card division (Dolan, 1985). Most firms have auditors who scrutinize to at least some degree every expense account

request made by their executives. 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. These examples are particularly striking because it is difficult to see them as costless byproducts of supervisory activities directed at increasing the efficiency of production. While there are elements of ambiguity in any given situation, we regard it as a fair judgement that a sizeable fraction of the monitoring of employees by employers is directed at deterring worker malfeasance.

Data from the 1983 Current Population survey indicate that even under a fairly narrow definition of supervisory jobs approximately 7 percent of workers in the U.S. nonagricultural, private sector are employed as supervisors or inspectors.⁵ Even if a small fraction of their time is devoted to controlling worker misconduct, such expenditures would be quite substantial.

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.⁶ When firms pay for their employees' travel, any discounts logically belong to the firm. Moreover, when employees receive rewards for travelling on particular airlines, they may make travel decisions which are

⁵We classified as supervisors employee in the following three-digit 1980 Census of Population code occupations: 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.

⁶This discussion of frequent flyer programs is based on Brancatelli (1985).

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 labelled employee crime simply reflects an efficient transfer of resources from the firm to the worker. While this is clearly true in some contexts (e.g. workers drinking the remaining wine from a company function), this cannot be the whole story. It is instructive to consider a case where the "efficient compensation" argument seems relatively plausible -- absenteeism. If absentees were permitted in an efficient manner, we would expect the level of absenteeism to be countercyclical as firms permitted workers to take time off when there was less need for the employees' services. In fact, absenteeism is strongly procyclical (Leigh, 1985). While inconsistent with the "efficient compensation" model, this observation is consistent with efficiency wage formulations emphasizing the cost of job loss as an important determinant of worker behavior.⁷

⁷ While absenteeism may be easily observed by an employer, whether an employee has a legitimate reason is not. Hence the monitoring of illegitimate absenteeism is costly. An efficiency wage model generates tendencies towards a negative relation between absenteeism and the unemployment rate if the cost of job loss is positively related to the unemployment rate.

III. Explaining Monitoring and Employee Crime

The pervasiveness of monitoring outlays and employee crime makes it apparent that some consideration of fundamental importance is omitted from our statement of the firm's problem. We first examine two modifications of (1) in which it may be efficient for firms to expend resources monitoring workers although there are no constraints on bonding. Finding them to be incomplete, we then take up three possible considerations that may limit bonding: liquidity constraints, firm moral hazard, and social limits on enforceability.

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 will not maximize most standard social welfare functions even though it conserves on law enforcement resources.⁸ In parallel fashion, one can argue that firms expend money monitoring workers to insure them against being falsely labelled 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.

While this point is analytically valid, we doubt that it 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

⁸ If a crime is defined as engaging in a proscribed activity without a "legitimate" reason, then false positives occur both when those not engaging in a proscribed activity are punished and when those engaging in the activity for a "legitimate" reason are punished.

spend resources on monitoring and require large bonds. Casual observation suggests that it is extremely uncommon for innocent employees to be punished for stealing from their employers. Firms of at least moderate size often operate under a "just cause" standard requiring substantial direct evidence or "proof" to discipline or dismiss workers for engaging in malfeasance (Koven and Smith, 1985, p.10). When innocent employees are sacked, the reasons probably have more to do with a firm's desire to eliminate them, than with its being ignorant of their innocence. In this case, direct limits on bonding are still necessary to account for monitoring outlays.⁹

The reader who is skeptical at this point should ponder the following question. Would most workers be sorry if they were supervised less closely because lower supervision would result in a somewhat greater risk of their being falsely accused of malfeasance? Or would they instead welcome the increased opportunity to loaf? Recent experiences in many establishments with the introduction of new computer-based monitoring technologies suggest that workers do not welcome more accurate monitoring even if it reduces the likelihood of false positives.¹⁰ In fact, more accurate monitoring by computers is so strenuously objected to by workers that there have been recent attempts to introduce legislation in Congress to prohibit the use by employers of computer-based performance monitoring technologies. Worker risk aversion and the prevention of false detection may account for some

⁹ If effort were treated as continuous, this analysis would have to be modified. The dichotomous effort choice assumption made here is probably a reasonable one for forms of worker malfeasance such as employee theft or sabotage.

¹⁰ Howard (1985) presents numerous examples and a detailed discussion of the conflicts that arise with the introduction of computer-based performance monitoring systems.

monitoring, but it seems doubtful that it accounts for much of the monitoring expenditures that we observe.¹¹

The Prison Guard Problem

There is a second 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 dollars, he might as well take \$10,000. Where capital punishment is illegal, it becomes difficult to deter those imprisoned for life from attacking prison guards. 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. As long as more serious types of shirking like stealing are easier to detect than less serious varieties like loafing, any given bond will provide marginal deterrence against more severe offenses. Firms could also create marginal deterrence by making the fraction of the bond forfeited a function of the seriousness of detected malfeasance.

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)

¹¹ Arguments similar to those of Polinsky and Shavell have been made to explain positive monitoring outlays in the context of principal-agent models. Our analysis here suggests that important elements may be omitted from standard principal-agent formulations.

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.¹² 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 don't firms take at least some cash from newly hired workers or at least set initial wages to zero? Firms should ask workers for whatever upfront payments they can make at least until the point where workers' reservation utility constraint is exactly satisfied (Carmichael,

¹²Terris and Jones (1982) discuss the prevalence and effectiveness of the use of pre-employment screening devices to control employee theft.

1985). Most workers have at least some liquid assets. Yet explicit upfront bonds and entrance fees appear to be quite rare in practice.

Firm Moral Hazard

A second standard reason why firms cannot fully rely 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. Unless, as is unlikely in practice, third parties can be relied on to determine whether a worker has shirked, workers will only be willing to post bonds if they are convinced that the firm will not take them under false pretense. Workers should trust firms not to expropriate bonds falsely so long as the bond is smaller than the value to the firm of maintaining its reputation as an employer. 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 moral hazard problem. Imagine 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.¹³

It seems clear that moral hazard considerations do place some constraint on the size of bonds that firms can ask their workers to post. 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, 1985).¹⁴ Industries that pay high starting wages also pay high wages to senior workers (Dickens and Katz, 1987; Krueger and Summers, 1987). Perhaps there are limits on the size of bonds that bind before reputation constraints are reached for many firms.

Enforcement Problems

We suspect that enforcement problems are the most empirically important consideration limiting the use of performance bonds. Firms are reluctant to use certain punishment strategies because of their adverse effects on the attitudes and performance of non-shirking workers. Society appears unwilling to enforce the type of contracts that result when firms follow the strategy of combining low monitoring intensity with large penalties for eliciting effort. We discuss each of these considerations in turn.

Carr-Hill and Stern (1979), in an analysis of the criminal law which

¹³Note 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).

¹⁴Brown and Medoff also find that the tenure profile is steeper in large firms and establishments. This suggests that implicit bonds through upward sloping earnings profiles may be utilized to a greater extent by large employers.

closely parallels our analysis of the employment relation, conclude that a 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. If measures to prevent shirking by a few antagonize the many, these measures may be counterproductive.¹⁵ There are many things people will do if asked but not threatened.¹⁶

A major concern of personnel departments is fostering employee loyalty. Managers strive to develop corporate cultures which lead workers to internalize important organizational values. Every personnel text stresses the importance of treating workers fairly. It is hard to imagine a worse way of beginning a relationship with an employee than to require him to post so large a bond or to accept such menial work at the outset that he is indifferent to taking the job.

Beyond firms' reluctance to carry out severe punishments, there is the further issue of what types of contracts governments will enforce. As a general legal principle, courts will not enforce contract provisions calling for "liquidated damages" such that the penalties for breach exceed the damage done.¹⁷ When the detection probability is low, penalties far in excess of the actual damage done to the firm by worker malfeasance may be

¹⁵ Titmuss (1971) provides a dramatic example of how partial reliance on pecuniary incentives can undermine "moral" behavior in the context of blood donations.

¹⁶ A similar point is made by Carr-Hill and Stern (1979, pp. 294-5) 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."

¹⁷ 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."

required to provide adequate deterrence. More specifically explicit legal rules regarding the labor market prohibit the actions that look 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 Age Discrimination Act of 1978) and are unthinkable in parts of Western Europe. Despite the plausible efficiency case that could be made for a rich variety of contracts, most American jurisdictions have laws which prevent landlords from asking their tenants for more than one or two months rent as a security deposit. Is it likely that the authors of these laws would permit firms to ask workers to post large bonds?

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. The common law restriction of 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. 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.

IV. Conclusions

Pervasive employee crime and efforts to deter it are economic phenomena that require explanation. If bonding were costless and unfettered, firms would make workers post large bonds and devote no resources to monitoring worker malfeasance. Since firms do devote substantial resources to monitoring, there must be limits on bonding. Given these limits, there is no legitimate a priori argument for excluding the possibility that firms pay premium wages. In almost any plausible model, increases in the wages firms pay will permit reductions in their monitoring outlays. They may therefore be beneficial to firms.

We have focused on the question of why firms do not require their workers to post large enough bonds in order to obviate their need for more than infinitesimal monitoring outlays. The arguments in the preceding sections suggest that firms might find courts unwilling to enforce contracts involving large bonds, and that large bonds might be counterproductive because of their impact on worker morale. Under these conditions, firms may find it profitable to pay workers premium wages.

A challenge to the argument that firms pay workers premium wages might run as follows (Murphy and Topel, 1987). If some aspect of incentives leads firms to make jobs sufficiently attractive that workers queue, why don't firms find some way to extract the surplus from new hires? If explicit payments are ruled out, perhaps firms could expose workers to poor working

conditions or require them to do menial jobs. As long as there is some way of burdening workers that raises the firms' profits, firms should eliminate any queues outside their gates.

Despite its considerable logical appeal, this argument surely proves too much for it denies the possibility of rents in almost any situation. Consider the case of a minimum wage law. As long as job conditions and fringe benefits are unregulated, arguments paralleling those of Murphy and Topel would suggest that no involuntary unemployment should result. Similarly, their arguments would suggest that as long as any aspect of rental transactions is unregulated, rent control should interfere with efficiency but should not mean that tenants in uncontrolled apartments will envy those in controlled apartments. We doubt that most observers would accept these predictions. Presumably, this is because firms choose not to extract all possible surplus from workers for reasons related to the morale considerations discussed in the previous section.

Our analysis suggests that limits on the ability of firms and workers to enter contracts that would permit sufficiently large penalties to be exacted against shirking workers are the most important explanation for monitoring outlays. The two most plausible limits on bonding that we have identified - the potential negative impact of bonds on employee morale and the society's unwillingness to enforce contracts with explicit bonds -- are in a sense related. Both are connected to notions of fairness that lie outside of conventional treatments of the economics of agency and incentives. Understanding these constraints and their operation is an important subject for future research.

The allocative consequences of the payment of efficiency wages,

particularly the possibility of involuntary unemployment, have attracted a great deal of attention. Premium wages represent transfers and have only second order welfare effects in otherwise undistorted economies. In contrast, efforts to deter employee crime are pure waste relative to the unattainable first best outcome. Excessive monitoring is almost certainly a socially more important consequence of limitations on bonding than the payment of efficiency wages. Indeed it can be easily verified in the context of the model presented in the first section that, starting from an interior optimum in which firms are paying premium wages and engaging in costly monitoring, a marginal substitution of higher wages for lower monitoring expenditures represents a Pareto improvement. Such an intervention leaves profits and employment unchanged and raises the welfare of employees.¹⁸

Whether bonding limitations give rise to efficiency wages as well as monitoring outlays is an empirical question whose answer will vary with different circumstances. Rather than engaging in sterile theoretical arguments, future research should focus on evaluating the quantitative importance of supra-competitive wage payments caused by limitations on bonding.

¹⁸Shapiro and Stiglitz (1984) provide a further discussion of the welfare issues that arise when bonding is limited and monitoring is endogenous.

References

- Akerlof, George. "Labor Contracts As Partial Gift Exchange." Quarterly Journal of Economics 87 (November 1982): 543-69.
- Arnold, Gregory B. "Employee Theft: A \$40-Billion Crime." Management World 14 (November 1985): 27-28 and 43.
- Arthur Young & Company. An Once of Prevention, 1985-86 edition. 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 George J. Stigler. "Law Enforcement, Malfeasance, and the Compensation of Enforcers." Journal of Legal Studies 3 (January 1974): 1-18.
- Brancatelli, Joe. "And The Winner Is." Frequent Flyer Magazine (November 1985): 64-74.
- Brown, Charles and James L. Medoff. "The Employer-Size Wage Effect." Mimeo, Harvard University, November 1985.
- Bulow, Jeremy and Lawrence H. Summers. "A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination and Keynesian Unemployment." Journal of Labor Economics 4 (July 1986): 376-414.
- Calvo, Guillermo. "Quasi-Walrasian Theories of Unemployment." American Economic Review 69 (May 1979): 102-7.
- Carmichael, H. Lorne. "Is Unemployment Involuntary? Comment." American Economic Review 75 (December 1985): 1213-14.
- Carr-Hill, R.A. and N.H. Stern. 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.
- Dickens, William T. and Lawrence F. Katz. "Inter-Industry Wage Differences and Theories of Wage Determination." NBER Working Paper No. 2271, June 1987.
- Eaton, C. and W.D. White. "The Economy of High Wages: An Agency Problem." Economica (1983): 175-181.
- Hollinger, Richard C. and John P. Clark. Theft By Employees. Lexington, MA: Lexington Books, 1983.
- Howard, Robert. Brave New Workplace. New York: Penguin Books, 1985.

Kadish, Sanford. "The Criminal Law and Industrial Discipline As Sanctioning Systems: Some Comparative Observations." Labor Arbitration--Perspectives and Problems, Proceedings of the 17th Annual Meeting of the National Association of Arbitrators. Washington, D.C.: BNA Books, 1964.

Kolm, Serge-Christophe, "A Note on Optimum Tax Evasion," Journal of Public Economics 2 (July 1973): 265-70.

Koven, Adolph M. and Susan L. Smith. Just Cause: The Seven Tests. San Francisco: Coloracre Publications, 1985.

Krueger, Alan B and Lawrence H. Summers. "Reflections on the Inter-Industry Wage Structure," in K. Lang and J. Leonard, eds., Unemployment and the Structure of Labor Markets. New York: Basil Blackwell, 1987.

Lazear, Edward. "Agency, Earnings Profiles, Productivity, and Hours Restrictions." American Economic Review (September 1981): 606-20.

Leigh, J. Paul. "The Effects of Unemployment and the Business Cycle on Absenteeism." Journal of Economics and Business 37 (May 1985): 159-70.

Malcolmson, James. "Work Incentives, Hierarchy, and Internal Labor Markets." Journal of Political Economy 92 (June 1984): 486-507.

Murphy, Kevin M. and Robert H. Topel. "Efficiency Wages Reconsidered: Theory and Evidence." Mimeo, University of Chicago, May 1987.

Polinsky, A. Mitchell and Steven Shavell. "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 & Company, 1972.

Security. "Time Theft Taxes Business," June 1987, p. 11.

Shapiro, Carl and Joseph S. Stiglitz. "Equilibrium Unemployment As A Worker Discipline Device." American Economic Review 74 (June 1984): 433-44.

Stigler, George J. "The Optimum Enforcement of Laws." Journal of Political Economy 78 (March/April 1970).

Terris, William and John Jones. "Psychological Factors Related to Employees' Theft in the Convenience Store Industry." Psychological Reports 51 (December 1982): 1219-38.

Titmuss, Richard. The Gift Relationship: From Human Blood to Social Policy. London: George Allen & Unwin, 1971.

Whitehurst, Susan. "Security Forecast 1987." Security (January 1987): 40-44.

Willis, Rod. "White-Collar Crime." Management Review (January 1986): 22-32.