Does Post-Accident Drug Testing Reduce Injuries? Evidence from a Large Retail Chain

Alison D. Morantz, *Stanford Law School*, and Alexandre Mas, *University of California*

This study examines the effects on occupational injury claims of a recently implemented post-accident drug testing (PADT) program in a large retail chain. We find that claims have fallen significantly in affected districts, suggesting that PADT programs can reduce injury claims, even in workplaces that already utilize other forms of drug testing. Our results also suggest that some types of employees—such as full-time workers, male workers, and higher-tenure workers—are particularly responsive. Finally, we find some "circumstantial evidence" that a portion of the observed decline could be caused by employees' reduced willingness to report workplace accidents. (*JEL* D21, H11, H51, H73, H75, I18, I38, J32, J33, J38, J81, J88, K00, K13, K31, K32, L51, M50, M52)

1. Introduction

Drug testing in the workplace has remained controversial ever since its application to many federal government employees and widespread adoption by corporate America in the mid-1980s.[1] Proponents contend that the direct

Send correspondence to: Alison D. Morantz, Stanford Law School, Crown Quadrangle, 559 Nathan Abbott Way, Stanford, CA 94305-8610; E-mail: amorantz@law.stanford.edu or Alexandre Mas, U.C. Berkeley, S545 Haas School of Business, University of California, Berkeley, CA 94720-1900; E-mail: amas@haas.berkeley.edu.

1. President Ronald Reagan's Executive Order 12,564, the Drug Free Workplace Program, issued on September 15, 1986, ordered all federal agencies to develop programs to identify illegal drug users. The directive was targeted at employees in "sensitive"

American Law and Economics Review
doi:10.1093/aler/ahn012
Advance Access publication August 23, 2008

and indirect costs of employee drug use on the job are enormous, costing businesses at least $100 billion per year in lost productivity (Bensinger, 1988; Todd, 1987). Detractors not only have questioned the validity of such empirical claims (ACLU, 1999), but also have challenged the legality of such programs on the grounds that they infringe on the right to privacy (Todd, 1987; Westphal, 1987; Comer, 1994),[2] constitute an unreasonable search and seizure under the Fourth Amendment (Joseph, 1987; Mell, 1987; Todd, 1987),[3] violate Fifth or Fourteenth Amendment due process rights (Todd, 1987; Mell, 1987),[4] contravene the employment contract (Todd, 1987),[5] or are conducted in a tortious manner (Todd, 1987). With respect to public

positions. During the subsequent five-year period, many federal agencies—including the Department of Defense, the Department of Transportation, and NASA—passed regulations designed to comply with the directive. Subsequent legislation and regulations expanded the directive to include firms that accepted federal contracts, or that operated nuclear reactors (Ackerman, 1991). The growth of drug-testing in the private sector was roughly contemporaneous. See, e.g., Hartwell *et al.* (1996) (documenting growing trend toward workplace drug-testing programs in the private sector, and dating their proliferation from the mid-1980s); Morrow (1989) (citing recent evidence on growing popularity of workplace drug testing); Barnum and Gleason (1994) (citing statistics that 85 percent of surveyed corporations used some sort of drug testing in 1993, and that one in five American workers was tested in 1992). The growth of private-sector programs was probably spurred, at least in part, by President Reagan's Commission on Organized Crime, which in March of 1986 encouraged private employers to "consider the appropriateness" of a drug-testing programs. (Daily Labor Report, Bureau of National Affairs 43, at A-12 (March 5, 1986)).

2. See, e.g., National Treasury Employees Union v. Von Raab, 649 F.Supp. 380 (E.D. La. 1986), modified, 808 F.2d 1051 (1981) (holding that drug testing program implemented by U.S. Customs Service interfered with employees' right to privacy), Jennings v. Minco Technology Labs, Inc., 765 S.W.2d 497 (Tex. App.-Austin 1989, writ denied) (raising unsuccessful privacy-based challenge).

3. See, e.g., Capua v. City of Plainfield, 643 F.Supp. 1507 (D. N.J. 1986) (holding that urinalysis drug testing program violates Fourth Amendment rights against search and seizure), Lovvorn v. City of Chattanooga, 647 F.Supp. 875 (E.D. Tenn. 1986) (same), National Treasury Employees Union v. Von Raab, 649 F.Supp. 380 (E.D. La. 1986), modified, 808 F.2d 1051 (1981).

4. See, e.g., Jones v. McKenzie, 28 F.Supp. 1500 (D.D.C. 1986).

5. See, e.g., Black v. Kroger Co., 527 S.W.2d 794 (Tex. Civ. App. 1975) (false imprisonment-based challenge), Armstrong v. Morgan, 545 S.W.2d 45, 46–47 (Tex. Civ. App. 1976) (negligence-based challenge), Hogan v. Forsyth Country Club Co., 79 N.C. App. 483, 495–96, 340 S.E.2d 116, 123–24, disc. review denied, 317 N.C. 334, 346 S.E.2d 140 (1986) (intentional infliction of emotional distress-based challenge), Houston Belt & Terminal Railway Co. v. Wherry, 548 S.W.2d 743 (Tex. Civ. App. 1976), appeal dismissed, 434 U.S. 962 (1977) (defamation-based challenge).

sector employees, most constitutional challenges were laid to rest in 1989 with the Supreme Court's decision in *National Treasury Employees Union v. Von Raab*, which rejected a Fourth Amendment challenge to the random drug testing of certain Customs Service employees.[6] Since then, the primary legal battleground has shifted to state legislatures, many of which have enacted laws regulating the use of drug testing in the workplace (de Bernardo and Pedro, 2006).

There are several important distinctions among workplace drug-testing programs. Most importantly, some programs screen prospective employees at the prehiring stage, while others test incumbent workers. Preemployment testing programs are the more prevalent of the two types.[7]

Incumbent worker programs commonly include at least one of four testing triggers. Under "reasonable cause" testing, an individual worker may be tested if her behavior reasonably gives rise to the suspicion of drug use. "Comprehensive" testing involves the periodic, scheduled testing of all employees, such as during routine physical exams. "Random" testing involves testing all employees (or particular groups of workers) on an unannounced and variable schedule (Hartwell *et al.*, 1996). Finally, "post-accident" drug testing (PADT) subjects any employee who reports a workplace accident (and sometimes co-workers who were directly involved) to a drug test at the time the report is made, regardless of whether the reporting worker's conduct precipitated the incident.

Drug-testing programs also differ with respect to the method of testing applied. The collection of urine samples under direct supervision has raised particularly weighty privacy concerns. Less invasive methods, such as

6. 109 S.Ct. 1384 (1989). The affected employees in *Von Raab* were those seeking promotions to positions involved in drug interdiction or involving the use of a firearm. See also Betts (1990) and Walstatter (2001) (arguing that since many state workers' compensation and unemployment insurance systems deny benefits on the basis of a positive drug test, or refusal to submit to being tested, such tests, even when implemented in private sector, should be required to pass constitutional muster under the Fourth Amendment).

7. Gust and Walsh (1989) found that of 145,300 private-sector establishments surveyed in 1988, 85,200 tested job applicants, as compared to only 63,500 that tested current employees. Hartwell *et al.* (1996) noted that testing of new applicants appears to be more common than testing of current employees. This disparity likely is due to the fact that many of the legal protections that drug testing arguably violates—such as contractual just cause protection or protection against terminations that violate public policy—do not apply to prospective employees, who by definition are not yet covered by an employment contract.

alcohol breathalyzers and "cheek swab" saliva tests, are generally seen as less objectionable.[8]

From a public policy perspective, PADT programs raise unique risks. After all, incumbent employees can do little to avoid comprehensive or random testing, and the most one can do to evade reasonable-cause testing is to refrain from behaviors likely to raise a reasonable suspicion of drug use. In the case of PADT, however, an employee may respond by not only by taking more care on the job, but also by declining to report any accidents that do occur. It is even conceivable that workers who do not use drugs, but who dislike the prospect of undergoing a drug test, may underreport injuries simply to avoid being subjected to the test. (For example, some employees may perceive the tests as inherently unpleasant or intrusive, or may fear the risk of false positives or possible disclosure of private medical information.) The greater the psychic costs to an employee of taking the test, the greater the likelihood that some injuries—particularly those that are relatively slight or easily hidden—will go unreported.[9] Unlike other forms of employee drug testing, therefore, PADT programs pose a potential risk of underreporting.[10]

Drug-testing programs, and state laws regulating their use, have been implemented in two waves since the mid-1980s. The earliest programs, many of which date from the 1980s, screened job applicants (Coombs and West, 1991). Tests targeted at incumbent employees began to proliferate in the late 1980s and early 1990s. State laws regulating drug-testing programs have evolved in similar fashion, with the earliest laws generally focusing on

8. See Caruso v. Ward, 133 Misc.2d 544, 549, 506 N.Y.S.2d 789, 793 (N.Y. Sup. Ct. 1986) (containing dicta to this effect), cited in Todd (1987) (discussing particularly severe privacy concerns raised by urinalysis, since direct observation is the "only sure way to guarantee the integrity of the urine sample").

9. Crant and Bateman (1989) discuss how psychological factors such as anxiety and perceived privacy violation may affect employees' willingness to participate in drug testing program. White (2003) lists possible privacy concerns that may lead employees subject to PADT to underreport accidents.

10. MacDonald and Wells (1994) note that PADT programs may lead to underreporting. All forms of drug testing encourage drug users to devise new and clever ways to pass the test. Companies marketing such products as "The Urinator," "Urine Luck," and "Absolute Detox" have sprung into being, offering a dizzying array of products to help drug users evade detection by their employers. An online company called "Cleartest," for example, offers a variety of products designed to "help you pass a drug test" on its website. *See* http://www.cleartest.com/products (visited March 9, 2008).

preemployment programs, and those passed since the 1990s tending to cover a broader array of program types (Bureau of Labor Statistics, 1990–2008).

Although detailed data are scarce, the use of PADT beyond the transportation sector seems to be a relatively recent phenomenon. While many large firms instituted preemployment, reasonable cause, and/or random testing by the mid-1990s, most PADT programs outside transportation are of more recent vintage.[11] Similarly, most laws regulating the use of PADT date from the 1990s. As of 2004, twenty-two US states, two municipalities, and Puerto Rico had passed laws regulating the use of PADT in the private sector (de Bernardo and Pedro, 2006).

Another important facet of legislative activity in the PADT policy arena has been the setting of standards regarding when, if ever, the government may deny benefits to workers who test positive for drugs or alcohol. In many states, a worker who tests positive for drugs or alcohol following a work-related accident is deemed ineligible for workers' compensation and/or unemployment benefits (de Bernardo and Pedro, 2006).[12] Important details such as whether the refusal to take the test is deemed tantamount to a positive result, and whether the employee must rebut the presumption that drug use caused the injury, have become fertile areas of legislative activity and judicial scrutiny.[13]

---

11. We were unable to find detailed data on the prevalence of private-sector PADT outside transportation. However, the fact that a study of different types of programs conducted in 1996 did not even mention PADT as a major form of employee testing—instead focusing exclusively on random, comprehensive, and reasonable cause programs—suggests that the use of PADT beyond transportation is a relatively new phenomenon (Hartwell *et al.*, 1996).

12. This issue is not pertinent to this study insofar as no states encompassed in this study have a per se workers' compensation exclusion for workers who test positive from PADT. Some states included in the study have no statute at all. Among those that do have a pertinent statute, the employee is only rendered ineligible for benefits if his or her injury was "caused by" the use of a controlled substance. Since in practice it is the employer's burden to prove this causation wrong, and the Company as a matter of policy does not challenge employees' entitlement to benefits even if they test positive for illegal drugs, differences in law across states should not affect our results.

13. Walstatter (2001) summarizes the main areas of legislative and/or judicial activity. In 2002, for example, the Ohio Supreme Court overturned legislation requiring any employee who tested positive for drugs or alcohol, and wished to receive workers' compensation benefits, to rebut the presumption that it was the proximate cause of the injury (see *State ex rel. Ohio AFL-CIO v. Ohio Bur. of Workers' Comp.*).

Although drug testing is used widely in the private sector, it remains more prevalent among some companies than others. The most important predictor of drug testing is company size. Although a 1996 study by the American Management Association found that 80 percent of firms used some form of drug testing, such programs were virtually universal among the largest Fortune 500 companies (Hartwell *et al.*, 1996; McManis, 1999). Its prevalence also varies by industrial sector. A 1993 survey found that manufacturing; communications, utilities, and transportation; and mining and construction were the industries most likely to drug test, with estimated rates of prevalence ranging 60–72 percent. At the other extreme, estimated prevalence in the service industry, and in finance, insurance and real estate, was less than 30 percent. Wholesale and retail trade fell in between these extremes, with slightly over half (53.7 percent) of all firms estimated to conduct drug and alcohol testing (Hartwell *et al.*, 1996).

Advocates of workplace drug testing generally have premised their arguments on two empirical claims: that drug use lowers employee productivity, and that it increases occupational accidents. The first claim has been relatively well substantiated, with the majority of studies finding significant negative correlations between employee drug use and individual performance, and/or positive correlations between drug-testing programs and labor productivity.[14] Similarly, many controlled laboratory studies have found that the ingestion of drugs impairs an array of cognitive and psychomotor

14. See, e.g., Elmuti (1993) (finding that drug testing program in a Midwestern manufacturing plant improved objective measures of employee efficiency, productivity, and attendance); Zwerling *et al.* (1990) (finding statistically significant positive association between positive preemployment drug test results among postal employees and relative risk of turnover, absenteeism, and discipline); Normand *et al.* (1990) (finding that job applicants who tested positive for the use of illicit drugs, once hired, had higher rates of absenteeism and involuntary turnover); McDaniel (1989) (linking preemployment drug use of military personnel to likelihood of subsequent discharge); Blank and Fenton (1989) (correlating drug use among male navy recruits to lower retention rates). See also French *et al.* (2001) (finding chronic drug use to be significantly negatively related to the likelihood of employment for both genders, and to labor force participation for males, although similar effects were not found among light or casual users). But see Kaestner (1991) (finding that increased use of marijuana and cocaine is associated with higher wages in a cross-sectional two-stage least squares model); Kaestner (1994) (finding large, positive, and statistically significant effects of illicit drug use on wages in a cross-sectional model, but no statistically significant effects in a longitudinal fixed-effects model).

skills likely to affect on-the-job productivity (Chait *et al.*, 1985; Herning *et al.*, 1989; Murray, 1989; Yesavage *et al.*, 1985).[15]

Although the link between drug use and productivity has been fairly well explored, the hypothesized salutary effect of employee drug testing on occupational safety has received relatively little scholarly attention. To the best of our knowledge, only four prior studies have examined the effect of corporate drug-testing programs on occupational accidents. Of these, two were descriptive case studies that did not statistically evaluate programmatic effectiveness.[16] Although the remaining two studies used statistical methodologies to analyze company-level data, and both found that at least certain forms of drug testing lowered the frequency of injuries, each suffered from important empirical limitations.[17] Finally, in a related vein, two economists have examined the association between drug-testing *laws* and changes in

15. See also Kelly *et al.* (1990) (finding that even small amounts of alcohol affect performance and social behavior relevant to workplace); Jobs *et al.* (1990) (linking moderate alcohol consumption to extreme changes in business decision making).

16. The first case study, authored by an official of Southern Pacific Transportation Company, presents data from before and after implementation of a drug-testing program in 1984. The summary statistics presented indicate that the implementation of the program coincided with a decline in personal injuries. However, no statistical analysis was conducted, and even the summary statistics are incomplete (for example, the oldest data presented date to the year immediately prior to implementation, so one cannot reject the possibility of a preexisting long-term secular decline). Therefore, one cannot meaningfully evaluate the effect of the program (Taggart, 1989). Although a second study was conducted in 1987 to evaluate the effectiveness of drug testing at Southern Electric International, Inc., no statistical analysis was conducted, or indeed could have been, given the small sample sizes. *See* Sheridan and Winkler (1989). A third study of a corporate drug-testing program at Utah Power and Light Company did include some basic statistical analysis; however, the authors did not analyze the effect of the program as such, but rather the differences in key outcomes (absenteeism, vehicular accidents, medical benefit costs, and turnover) between employees with positive and negative drug tests. *See* Crouch *et al.* (1989). Moreover, the study was based on a sample of only twelve drug users, all of whom had been previously identified through the company's for-cause testing program, and eight of whom were tested precisely because they were involved in an accident, thereby precluding any meaningful causal inferences regarding the effect of drug use on accident rates. *See* American Civil Liberties Union (1999).

17. The first study, focusing on the construction industry, found that companies with a drug-testing program experienced at 51 percent decline in injury incident rates. However, the study suffered from two important drawbacks. First, since the only companies included were those willing to participate—and the company is also the smallest unit of observation analyzed—the sample may exhibit selection bias. Second, the study did not distinguish among different types of drug-testing programs, but simply treated all firms with at least one program of any kind as part of the treatment group. *See* Gerber and Yacoubian Jr.

occupational safety. A study by Kesselring and Pittman (2002) attempting to link state drug-testing laws to occupational injury and illness rates found that the state legal environment had no significant effect, although the study design suffered from important methodological shortcomings.[18] A more narrowly focused study by Mirielle Jacobson (2003) concluded that laws mandating drug testing of truck drivers led to a 9–10 percent reduction in truck accident fatalities.[19] A handful of studies addressing the related question of whether prospective or current employees' drug use predicts their subsequent likelihood of sustaining occupational accidents have

(2001). The second study found, rather paradoxically, that although the use of drug-testing programs did *not* reduce the frequency of occupational injuries as compared to firms using no drug testing at all, the implementation of PADT reduced companies' occupational injury and illness rates compared to (1) the same companies' pretesting period, and (2) companies that used only preemployment testing. The study also suffered from important limitations. First, the results were gleaned from questionnaires sent out to a large number of businesses, only a small number of which agreed to participate. This selection bias concern is exacerbated by the fact that the effective size of the sample was very small. (The study examined annual changes over a four-year period in 48 facilities, only 12 of which used any form of drug testing). Moreover, although the study analyzed OSHA recordable accident and illness rates, it relied only on company self-reports, and did not distinguish among the two types of OSHA recordable incidents, which vary markedly by severity (Feinauer and Havlovic, 1993).

18. Although provocative, the study suffers from important shortcomings. For example, only forty states were included in the study, since the remaining ten did not report the pertinent data. Moreover, although the results hinge crucially on the correct categorization of legal regimes, the three-part typology used is relatively crude. The categorization of each state as "restrictive," "supportive," or "neutral" does not differentiate among types of laws (for example, those affecting prospective employees and those affecting incumbent employees), and arbitrarily categorizes those states whose laws took effect in 1991 as having no laws at all (ignoring the possibility that workers and/or employers changed their behavior in anticipation of the new laws). The model also excludes important variables likely to affect state-level injury rates, such as detailed industry codes (only nine industry classifications are used, although each encompasses heterogeneous subindustries whose prevalence varies across states); workers' compensation laws; the presence of laws restricting workplace alcohol testing; and whether the state belongs to the federal OSHA enforcement regime.

19. Jacobson's study exploited the fact that thirteen states passed drug-testing laws in the late 1980s, several years before the passage of the federal Department of Transportation regulations (Jacobson, 2003). See also Mehay and Pacula (1999) (finding that a drug-testing program implemented by the military in 1981, combined with a "zero tolerance" dismissal policy, achieved its intended effect of significantly lowering drug use among military personnel).

yielded equally conflicting results.[20] Viewed as a whole, then, prior scholarship leaves many important questions unanswered about the relationship between drug testing and occupational safety.[21]

Our goal is to contribute to this relatively sparse body of empirical work by examining the effect of a recently implemented PADT program in a large Fortune 100 corporation (“the Company”), whose anonymity we have agreed to preserve for the purposes of this study. Like most other large corporations, the Company already uses both preemployment and “reasonable cause” forms of testing. The PADT program thus overlays and augments these programs.

Specifically, we examine the effect of the PADT policy on three types of accidents: minor ones that require no more than short-term first aid (“first aid” claims); those that incur medical costs through workers’ compensation but cause no loss of work (“medical only” claims); and accidents that are serious enough to cause a loss of work and thus the payment of workers’ compensation indemnity benefits (“indemnity” claims).[22] Our data therefore enable us not only to examine whether the PADT program affects the frequency of reported occupational accidents, but also to differentiate relatively serious claims from minor ones. The data also enable us to examine heterogeneity in employee responsiveness to PADT implementation depending on seniority, full-time status, gender, and age.[23] In addition to the data on claims

20. Compare, e.g., Normand *et al*. (1990) (finding no significant association between preemployment drug test results and frequency of subsequent occupational injuries and accidents) and Hoffman and Larison (1999) (finding that neither marijuana use nor cocaine use affected likelihood of survey respondents’ sustaining accident in past year) with Holcom *et al*. (1993) (finding substance abuse among municipal employees to be significantly correlated with accidents in high-risk jobs, although much of correlation apparently was plausibly explained individual demographic and personal background characteristics). Zwerling *et al*. (1990) found statistically significant positive association between positive preemployment drug test results and postal employees’ relative risk of accidents and injuries. Kaestner and Grossman (1998) found illicit drug use to be significantly correlated with workplace accidents among young adult males, although not among young adult females.

21. The National Academy of Sciences, which in 1994 analyzed the empirical literature on workplace drug testing, echoed this assessment. The report concluded that the then-existing evidence regarding the effect of illicit drug use on occupational safety outcomes was inconclusive (Normand *et al*., 1994).

22. All incidents in our sample fall into one of these three categories.

23. See, e.g., National Institute of Occupational Safety and Health (2004) Figures 2–71 and 2–72 (NIOSH Publ. No. 2004–146) (charting differences in number and rate of

frequency, the Company has also provided us with the results of PADT drug tests that have been administered to Company employees. The latter data allow us to explore which employee characteristics, if any, are predictive of a positive drug test result.

The remainder of this article unfolds in four parts. In Section 2, we describe the Company, the mechanics of its PADT program, and the data upon which our study is based. Section 3 describes the study design and our empirical methodology. Section 4 presents the results of the empirical analysis. Section 5, the concluding section, summarizes the key results and their relevance to the broader policy debate on post-accident drug testing (PADT).

2. Description of the Company, the PADT Program, and the Data

Although confidentiality restrictions preclude us from disclosing the Company's identity, its organizational characteristics make it an advantageous environment in which to analyze the effect of PADT. A major player in the retail sector, the Company's annual revenues place it in the Fortune 100. It operates over one thousand facilities nationwide and employs well over one hundred thousand employees (more than 75 percent of whom work part-time). Widely geographically dispersed, the Company maintains operations in more than twenty US states, with multiple districts in each state and multiple establishments in each district. Since organized labor has a significant penetration in the Company's workforce, the returns to tenure are significant (albeit declining in recent years). Since the Company's constituent facilities conduct their operations in a uniform and routinized manner, establishment-level variation in the Company's business practices (along with their attendant safety risks) are unlikely to bias our results.

In recent years, the Company's overall rate of occupational injuries has been fairly typical of firms with similar characteristics. From 2001 to 2006, for example, the Company's rate of total Occupation Safety & Health Administration (OSHA)-recordable injuries (per one hundred full-time worker equivalents) ranged from 7.0 to 8.4, and its rate of more serious

nonfatal injuries by sex and age); Kopstein and Gfroerer (1990) (detailing differences in prevalence of drug use by demographic characteristics); Frone (2006) (finding that "men used illicit drugs and were impaired by their use at work more often than women.")

OSHA-recordable injuries (i.e., those involving a loss or restriction of work) ranged from 3.3 to 4.2. These rates placed it between the fiftieth and seventy-fifth percentiles for the retail trade sector as a whole in each of these years, as well as for the subset of retail firms with similar industrial characteristics.[24]

The Company's mandatory accident reporting policy, which has been in effect since 1990, requires each employee to immediately report each work-related injury or illness to a supervisor, or risk "disciplinary action, up to and including termination." Since the early 1990s, the company also has conducted both preemployment and "reasonable cause" drug-testing programs across all of its establishments. The PADT program, the most recent addition to the company's drug-testing policy, converts each qualifying incident report into an automatic trigger for a drug test. Since the phase-in of the PADT component in 2004, each time a claim is reported, the claimant automatically receives a "cheek swab" drug test. Rather than implementing the policy immediately company-wide, however, the Company began by implementing it in two pilot divisions ("Division One" and "Division Two") that jointly span three US states.[25] Within each division, the program was progressively phased in across seventeen individual districts.[26] Division One

24. Specifically, after calculating the rate of total and more serious OSHA-recordable injuries for the Company using our data, we compared each of these rates to the fiftieth percentile (median) and seventy-fifth percentile rates published by the Bureau of Labor Statistics (BLS) (at http://www.bls.gov/iif/oshsum.htm) for three categories: (1) the entire "retail trade" sector; (2) for the years 2001 and 2002, the three-digit SIC code to which the Company belongs; and (3) for the years 2003 through 2006, the five-digit NAICS code to which the Company belongs. Since the BLS data are stratified by establishment employment size, in each case, we compared the Company's rates to those of companies within the same establishment employment size stratum, defined by the number of full-time worker equivalents employed at the establishment. (A full-time worker equivalent is defined as 200,000 hours worked per year.) Company officials explained that the injury rates calculated from the Company's database may somewhat overstate the rates reported to BLS, because the OSHA injury logs that form the basis of the BLS reports are filled out contemporaneously with the initial accident report. If an injury appears to be minor at the time it is reported, it may not be recorded as a "restricted or lost workday" injury in the OSHA log, or indeed, may not be recorded at all. If the injury worsens over time, however, it will still be flagged retrospectively as an "OSHA-reportable" claim in the Company's internal claims database (from which the data used in our study was extracted). Therefore, it is possible that the Company's OSHA-reportable accident rates are even closer to the median for its respective industrial grouping.

25. In fact, the implementing divisions spanned five US states, but districts in two of these states did not implement the PADT program.

26. A typical district encompasses approximately fifteen facilities.

districts put the policy into practice between late 2004 and mid-2005, while districts in Division Two followed suit between mid-2005 and early 2006. The Company made no other changes in its accident reporting policies or protocols coincident with the roll-out of the PADT program.[27]

For the purposes of our empirical analysis, one particular aspect of the program merits special scrutiny. Formally, the Company's PADT policy encompasses only those injuries that result from a traumatic on-the-job event. Injuries whose onset is cumulative or gradual—such as carpal tunnel syndrome, progressive hearing loss, mental disorders, dermatitis, respiratory diseases, and so forth—are expressly excluded. In theory, since cumulative injuries are not drug tested, workers should have no incentive to underreport them.

In reality, however, there are two reasons to expect that cumulative injuries might respond to PADT implementation. First, Company officials with whom we spoke expressed doubt about whether employees were aware of their technical exclusion from the policy. They also observed that many store supervisors have been uncertain about whether to treat particular injuries as cumulative, and have been routinely instructed, "When in doubt, test."[28] Interestingly, the company's internal records also provide support for the view that employees reporting cumulative injuries do, in fact, risk being tested.[29] Second, of all occupational injuries, those defined as "cumulative" are probably, by far, the easiest to hide. In light of these considerations, our empirical analysis explores the possibility that cumulative injuries also responded to PADT implementation, notwithstanding their formal exclusion from the ambit of the program.

---

27. Although the Company assured us that it made no changes in its policy, the renewed emphasis on mandatory accident reporting may have made the policy more prominent and visible to employees. If so, the enhanced salience of the policy presumably would have *lowered* the prevalence of underreporting in the wake of PADT implementation.

28. Telephonic interview with Director of Loss Prevention of Division 1, October 3, 2006.

29. The Company's records reveal that one employee in a treatment division received first-aid treatment for a cumulative injury, was given a PADT, tested positive for marijuana use, and was terminated the following week. Since only 108 employees to date have been terminated for drug use through the PADT program, and cumulative injuries comprise only about 5 percent (or less) of all claims, it is possible that a sizable number of cumulative injuries have been subjected to drug testing.

The PADT program applies to two classes of employees: those who suffer a work-related injury that requires medical attention; and those who are "actively involved" in a "qualifying incident" that causes death or injury to third parties or $250 of property damage. Any employee who reports a qualifying incident is automatically required to submit a cheek saliva swab in a supervisor's presence immediately after the report is made. The test can detect the presence of all major drug metabolites, but not alcohol. The policy further specifies that any employee testing positive "will be subject to disciplinary action up to and including immediate termination of employment." Although in theory this language grants the Company discretion to allow an employee to continue working at the Company, in practice, a positive drug test has always resulted in immediate termination.

The data provided by the Company consists of three files. The first file, the "claims file," contains detailed information on all reported claims, including workers' compensation claims, in the two pilot divisions (Divisions One and Two) as well as two divisions used as comparators (Divisions Three and Four) since 2002. It specifies the date, nature, and treatment of each reported incident, as well as basic demographic information on each claimant. The second file, the "hours file," specifies the number of total hours worked in each company facility during each of thirteen "reporting periods," which approximate, but differ slightly from, calendar months.[30] Although the data was not designed for research purposes, we found it to be of high quality, with relatively few implausibly outlying values.[31] The third file, the "test

30. The Company's fiscal year consists of thirteen 28-day periods, each of which begins on a Sunday and ends on a Saturday. Company years are not necessarily coterminous with Gregorian years, nor are company periods coterminous with Gregorian months. Approximately once every four years, the Company inserts a "leap week" into period 13, to make up for days that the Company calendar has progressively "lost" relative to the Gregorian calendar.

31. We made the following adjustments to our dataset due to missing and/or incomplete data: (1) In one of the four divisions under analysis, first-aid reports were not recorded until the company time period corresponding to late July and early August of 2002. To avoid biasing our results, we dropped all claims from previous periods. (2) Because the data fields for average weekly wage and occupation were frequently missing and/or incorrectly coded, we omitted these variables in our analysis. (3) Nineteen claims were recorded in the company's data system without information on in which facility the injury took place. Because without facility information we could not match these claims to their respective districts, we deleted them. (4) Eighteen claims were recorded with invalid facility information. Since the company informed us that there was

result file," contains information on all employees who tested positive for use of illegal drugs (and were subsequently discharged) after an accident. By linking the "test result file" with the "claims file," we were able to determine the results of each administered PADT test.

In merging the three files, we aggregated both claims and hours to the district level. We chose to use the district (rather than the facility) as the unit of analysis because the PADT program was phased in at the district level, and therefore conducting our analysis at the district level most accurately reflects the respective sizes of the "treatment" and "control" groups observed in our study.[32] In effect, then, each cell in the final dataset corresponded to data for a given district during a particular reporting period.[33] Among the fields examined for each district-period were: total incidents, total hours worked, total incidents broken down by claimant characteristics (sex, tenure, age, and full-time versus part-time status), and total incidents broken down by claim type ("first aid," "medical only," and "indemnity"). Finally, the data on employees who tested positive for illegal drug use includes information on sex, age, tenure, full- versus part-time status, and the type of drug detected. The time period analyzed in the study spans five calendar years (July 2002 through July 2007).

no reliable way to recover the correct information, we deleted them. (5) There were five instances in which members of "skeleton crews" undergoing preliminary training before a new facility officially opened were injured and filed workers' compensation claims. Because these preparatory training periods are qualitatively different from normal facility operations, we deleted these five claims.

32. As a robustness check, we reestimated all of the models presented in this paper using the facility (rather than the district) as the unit of analysis. As expected, the magnitude of all of the estimates was nearly identical, but the standard errors (and *P*-values) were generally smaller, since analyzing the data at the facility level expands the effective size of the sample. (These results are available upon request.) Since the treatment actually varies at the district rather than the facility level, we believe that the estimates presented here—all of which use the district as the unit of analysis—more accurately reflect the true strength of our statistical findings.

33. In the second phase of our analysis, which explores the heterogeneity of the treatment effect by worker characteristics, each cell reflects the total number of claims in a given district during a given period for a given worker characteristic (sex, tenure, age, and full-time/part-time status).

3. Study Design and Empirical Methodology

The present study is, to the best of our knowledge, the first to analyze the effects of a PADT program in a large US corporation. It is also the first to attempt to isolate the impact of a PADT policy in a company that, like most large employers, already uses the more traditional preemployment and "reasonable cause" approaches to detecting and deterring drug use.

PADT programs are implicitly premised on two assumptions: that testing for illicit drug use will deter employees from using drugs; and that lowering drug use among employees will, in turn, lessen the frequency of on-the-job accidents, thereby reducing workers' compensation claims. In reality, however, implementation of PADT implicitly confronts employees with at least three distinct, albeit interrelated, decisions: (1) whether or not to take illegal drugs; (2) whether to report an occupational injury if one occurs;[34] and (3) whether to take extra care in an effort to avoid sustaining an accident.

Assuming that each individual can rationally compare the costs and benefits of each alternative,[35] his or her decision along each dimension is likely to depend on both programmatic and personal factors. Key programmatic factors include the relative penalties for using drugs and for failing to report an accident, the relative intrusiveness of the drug testing method used, and the level of medical care and benefits available through the workers' compensation system. Among those personal factors likely to enter an employee's decision calculus are: whether he or she is a user of illegal drugs, how averse he or she is to undergoing a drug test, the relative likelihood of his or her

34. It would not be surprising to find evidence of an "underreporting" effect in this context, since earlier empirical scholarship has found that the frequency of workers' compensation claims is sensitive to its costs and benefits. See, e.g., Butler and Worrall (1983) (finding that changes in workers' compensation laws substantially affected the number of accident claims filed).

35. Whether drug addiction deprives individuals of the capacity to engage in cost–benefit analysis is a matter of ongoing empirical controversy. On one hand, there is empirical support for the proposition that drug users are responsive to changes in the price of addictive substances and/or penalties of usage (Becker and Murphy, 1988; Elster and Skog, 1999; Heyman, 1996). On the other hand, some medical studies suggest that drugs that cause impairment may diminish insight, so that a user may be significantly impaired in his or her objective functioning without being aware of that fact. *See, e.g.,* National Highway Traffic Safety Administration (2003), Sect. 5 (and studies cited therein). For the purposes of our analysis, we assume that at least some employees who use drugs may respond to the program by taking more care.

sustaining an accident conditional on taking different kinds of precautions, the costs of alternative avenues for treating occupational injuries (such as health care plans[36]), and how highly he or she values the job.

Although we cannot directly observe any of the latter personal characteristics, the Company's full-time employees—constituting about a fifth of its workforce—are likely, *ceteris paribus*, to value their jobs more highly than their part-time peers, since job loss will trigger a larger shock (in absolute terms) to their stream of income. One might also expect higher-tenure employees to value their jobs more highly, since their seniority puts them at a higher point on the wage–tenure profile, and the extra wages and benefits that they have accrued at the Company would be lost if they changed jobs.[37]

Although lower drug use, underreporting, and greater care could each trigger a decline in the frequency of workers' compensation claims, their respective real-world consequences differ sharply. If PADT encourages drug users to kick their habit (or exit the workplace), its beneficial effects on addicts in particular, and on occupational safety in general, could be profound and far-reaching. Even if the policy's primary effect is to encourage drug users to take more care on the job, it could still improve the overall level of occupational safety.[38] On the other hand, if most employees respond simply by hiding their injuries, then the policy's net impact on employees' health and safety might be far less salutary.[39]

36. Technically, of course, all occupational injuries are supposed to be treated through the workers' compensation system, not through ordinary health care coverage. In many instances, however, employees enjoy considerable *de facto* discretion over whether to characterize an injury as work-related, and therefore the possibility of "claim migration" between the workers' compensation and health care systems is a commonly voiced concern.

37. Moreover, unionized jobs with relatively steep wage–tenure profiles—like that which the Company offers to incumbent employees—are becoming increasingly scarce across the industry.

38. Any such beneficial effect on occupational safety, however, could be counterbalanced by lower productivity if "taking care" to avoid an accident lessened the quantity or quality of time spent performing other tasks.

39. Rather than deterring legitimate workers' compensation claims, it could be that drug testing simply reduces the prevalence of fraudulent claims, especially among illicit drug users. If so, then PADT—even if it does not deter a single worker from using drugs—could be socially desirable. To date, there is no consensus on the prevalence of fraudulent workers' compensation claims. Some observers have claimed that 20 percent or more of all claims are fraudulent, while others have contended that the true figure is

With such policy concerns in mind, we examine four distinct yet interrelated issues. First, we explore the threshold question of whether the implementation of the PADT program is associated with a significant decline in the frequency of claims over time. Second, we parse the data more finely to determine whether the observed effects (if any) vary by accident type and/or worker characteristics. Third, we probe whether any employee characteristics are correlated with the likelihood of a positive drug test. Finally, we consider the possibility that the observed trends are caused, at least in part, by underreporting.

Several methodological issues require clarification. First, the PADT policy was progressively phased in during a 15-month period in districts located in Divisions One and Two. Unobservable characteristics of these "treatment" districts—such as idiosyncratic aspects of management practices, or employee culture, within its constituent facilities—conceivably could drive cross-district differences in reported claim rates. To mitigate such concerns, we use fixed-effects models in all specifications, exploiting the panel nature of our data. We also include period (time) dummies in all specifications. In so doing, we hope to control not only for any general time trends in claiming behavior, but also for omitted variables that may differ between districts but remain constant over time.

Another important preliminary issue was how to model the distribution of claims. Since claim frequency—our dependent variable—is a form of "count" data, it raises special methodological concerns. Count data are not normally distributed, but exhibit a rightward skew, and therefore ordinary-least-squares estimation is technically inappropriate. A Poisson distribution may be used instead to analyze count data if the mean and variance of data are the same. However, our data show signs of overdispersion, i.e., we can reject the null hypothesis that the conditional variance is equal to or less than the conditional mean of the distribution. Therefore, we use a negative

no greater than 1 or 2 percent. Compare, e.g., Fricker (1997) (suggesting that fraud is rare) and Leigh *et al.* (2000) with Texas Department of Insurance (1998) (suggesting that fraud may be as high as 30 percent). Unfortunately, we have no data available on the frequency of fraudulent claims at the Company before and after PADT implementation. Given such uncertainties, we cannot draw any conclusions about how a decline in the reporting of claims (if any) is likely to affect social welfare. Therefore, we use the term "underreporting" in a value-neutral sense to denote any observed fall in claims relative to the preimplementation baseline.

binomial model, the approach conventionally used when analyzing count data with these characteristics.[40]

A final threshold methodological question was which districts to use as the comparison or "control" group when analyzing the effects of the PADT program. Our goal was to compare each PADT-implementing district with other districts that had no PADT program, but otherwise were as similar as possible. In addition to the small handful of districts within Divisions One and Two that never implemented the program,[41] Company officials recommended that we use two other divisions ("Division Three" and "Division Four"), in the same census region as Divisions One and Two, as control districts. Although located in different states, Divisions Three and Four drew from a similar labor pool and experienced similar regional economic trends during the periods examined. Therefore, we used Divisions Three and Four (sometimes in combination with the nonimplementing districts in Divisions One and Two) as comparison groups throughout our empirical analysis.

As table 1 reveals, there are observable disparities between our treatment and (entire) control groups with respect to the "baseline" frequencies of claims. The mean frequency of indemnity claims is somewhat higher, and the mean frequency of medical-only workers' compensation claims is somewhat lower, in the control divisions than in the treatment divisions. Most striking is the disparity in the baseline frequency of "first aid" reports across regions. Company officials informed us that such cross-state variations in baseline rates are common, and are believed primarily to reflect cross-state differences in workers' compensation regimes (and secondarily, more subtle differences in customary practices and/or management practices

40. Fixed-effects negative binomial models are estimated using conditional maximum likelihood, i.e., conditioning on each district's total number of claims across all periods (Σ_t y_{td}). In using a conditional fixed-effects negative binomial model, we implicitly allow the conditional variance of the distribution to be proportional to a scaling parameter, rather than constraining it to be the same as the population mean.

41. All of the implementing districts in Divisions One are located in the same state, and all of the implementing districts in Division Two are located in three adjacent states. However, Division Two also includes a small handful of nonimplementing districts in noncontiguous states. In some versions of our models, we include these nonimplementing districts as control districts, along with Divisions Three and Four. However, since these nonimplementing districts might be demographically dissimilar to the other districts in Division Two, we also (as a robustness check) run the same models using only Divisions Three and Four as controls.

Table 1. Sample characteristics for a large retail chain, 2002–2007

	Treatment districts		Control districts
	Before post-accident drug testing program implementation	**After post-accident drug testing program implementation**	
Mean number of noncumulative claims/reports per 100,000 hours worked			
Total workers' compensation claim rate	4.85 [2.0]	4.40 [1.7]	4.13 [2.5]
Indemnity workers' compensation claim rate	1.26 [0.9]	1.17 [0.8]	1.44 [1.1]
Medical-only workers' compensation claim rate	3.60 [1.7]	3.23 [1.5]	2.69 [2.1]
First-aid report rate	12.14 [4.2]	9.63 [4.3]	5.54 [5.5]
Total number of claims/reports	18,141	11,686	32,609
From noncumulative injuries (percent of total)	17,180 (94.7%)	11,316 (96.8%)	31,081 (95.3%)
From cumulative injuries (percent of total)	961 (5.3%)	370 (3.2%)	1,528 (4.7%)
Total number of hours worked (in 100,000s)	1,013	804	3,126
Total number of company divisions	2	2	
Total number of company districts	17	27	
Total number of company time periods analyzed (equivalent number of years)	65 (5)	65 (5)	

Notes: Standard deviations are presented in brackets. All rates are calculated per 100,000 hours worked. The unit of observation over which means were calculated is district × "company time period." ("Company time periods"—approximately thirteen of which comprise a Gregorian year—are 28 days long apiece.) All summary statistics presented above were drawn from a "cut" of the data that excluded data from facilities that closed before any treatment districts implemented the PADT program, and included data from the small handful of nonimplementing districts in the treatment divisions. The statistics presented under the heading "mean number of noncumulative claims/reports per 100,000 hours worked" exclude cumulative injury claims, while the statistics presented under the heading "total number of claims/reports" include them.

across regions). Prior research finding that differences in state workers' compensation laws affect the "base" frequency of claims—especially less serious claims—lends credence to this explanation (Boden and Ruser, 2003; Barkume and Ruser, 2001; Butler and Worall, 1985; Card and McCall, 1996; Ehrenberg, 1989; Krueger, 1990; Krueger and Burton, 1990; Meyer *et al.*, 1995; Ruser, 1991; Thompson, 1981; Thomason *et al.*, 2001; Viscusi and Moore, 1987). Since our focus is on changes over time, and all of our models include district-level fixed effects as well as time dummies, such cross-district differences in "base rates" should not affect our ability to identify the effects (if any) of PADT implementation.

Having resolved these preliminary issues, we conduct the empirical analysis in four sequential stages. First, we estimate a set of models to determine whether there was a significant difference in the frequency of claims before and after implementation of the PADT program. Specifically, we estimate fixed-effects negative binomial models of the following form:

$$E[y_{td}|\alpha_d] = \exp(\ln(\alpha_d) + \Phi \mathit{POST\text{-}IMPLEMENTATION}_{td} + \eta_t + \beta \ln(\text{hours}_{td})), \quad (1)$$

where t denotes the period; d denotes the district; y_{td} is the number of injuries reported in district d during period p; and α_d denotes the district-level fixed effect. The dummy variable $\text{POST-IMPLEMENTATION}_{\text{td}}$ takes on the value of "1" if district d has implemented PADT by time period t, and "0" otherwise. We also include period dummies (η_t) to account for any general time patterns in injuries over the sample period, as well as the logarithm of total people-hours worked in each district in each period ("hours_{td}").[42] The parameter of interest in this baseline specification is Φ, the average post-treatment effect of PADT across districts. We estimate Model (1) for all claims, as well as separately for claims of different severity (indemnity workers' compensation, "medical only" workers' compensation, and "first aid" claims).

In addition to estimating Model (1), we also perform two different types of falsification tests to probe the robustness of our findings. Our goal is

42. We could treat hours as an exposure variable by constraining the coefficient β to be 1. However, we chose not to impose this restriction because of the possibility that the elasticity of injuries with respect to hours worked is different than 1 (for example, if injuries are more likely to occur during overtime hours).

to determine whether the average post-treatment effect derived from our model, Φ, is a likely outcome of the distribution of estimates produced by a false "placebo" treatment. If so, this would cast doubt on the validity of our results.

Our first robustness check is designed to probe whether pre-PADT trending of injury rates, rather than PADT implementation itself, could plausibly be driving our results. For example, if districts with recent declines in injury rates were selected for immediate PADT implementation, then any observed negative correlation between PADT phase-in and injury rates could be (at least partly) spurious. Although Company officials assured us that such was not the case—and district-level injury rates had nothing to do with the sequencing of district-level implementation—we perform a simple robustness check to ensure that the data bear this out. Specifically, we estimate a composite model containing not only the "true" implementation dummy, but also a second "fake" implementation dummy which turns on five periods before the true implementation date.[43] If inclusion of the fake dummy eliminates the magnitude or significance of the estimated value of Φ, the "true" estimated treatment effect, this would suggest that pre-PADT trending of incident rates is driving our results.

As a second robustness check, we apply a randomization inference technique to confirm that our model can correctly differentiate between true "treatment" and "control" districts in our sample. Following Donohue and Ho (2007), we randomly select subsets of twenty-two districts (the true number of treatment districts) to serve as "pseudo-treatment" districts, and assign each a randomly chosen PADT implementation date. We assume (as with our original model) that once a district implements the PADT policy, it continues to do so for the remainder of the sample period. In other words, each iteration consists of three steps: (1) randomly selecting a new set of pseudo-treatment districts, (2) randomly assigning them implementation dates, and (3) re-estimating Equation (1). By repeating the above simulation one thousand times, we derive the (nonparametric) randomization distribution of the

43. Since employees in each district were typically informed of the PADT policy 1–3 months before implementation, district-level injury rates 5 months prior to implementation should be free from "anticipation effects," yet still close enough in time to pick up any pre-PADT trending in injury rates. As a robustness check, however, we also estimated models with "placebo" implementation dummies that phased in two, three, four, and six periods (respectively) prior to the true PADT implementation date.

treatment coefficients, and compare them to the estimated treatment effects presented in table 2. We perform two slightly different variations of this technique: (1) one in which the pseudo-treatment districts are drawn from the entire population of districts in our sample; and (2) one in which the pseudo-treatment districts are drawn only from the two "control" divisions (i.e., Divisions Three and Four). If this procedure yielded many values of the randomization distribution at least as large as the observed treatment effect derived from estimating Equation (1), this would similarly cast doubt on the specificity of our model and the validity of our findings.

Although Model (1) represents an important first cut into the data, it may mask important variations across time during the post-implementation period. For example, even if there are transitional effects immediately after implementation, such effects could disappear within a few months. To gain a more nuanced picture, we estimate another set of models to capture the evolution of injury rates in treatment districts before and after implementation. Specifically, for each injury category, we estimate a fixed-effects negative binomial model in which we allow the effect on claims from belonging to the treatment group to vary by each individual period, τ, before and after implementation. The model takes the following form:

$$E[y_{td}|\alpha_d] = \exp(\ln(\alpha_d) + \phi_\tau \cdot \mathrm{TREATMENT_d} + \eta_t + \beta \ln(\mathrm{hours}_{td})). \quad (2)$$

As before, the parameter α_d represents the district fixed-effect, and we include both time dummies (η_t) and the log of total people-hours worked ("$\mathrm{hours_{td}}$"). The dummy variable $\mathrm{TREATMENT_d}$, which takes on the value of "1" if the district ever implemented PADT and "0" otherwise, indicates whether the district is in the treatment or control group. The parameters of interest, ϕ_τ, represent the factor change in reported injuries in all treatment districts at τ periods before or after the respective implementation date of each district. Since we track claims in the treatment group for a year prior to PADT implementation and 24 months after PADT implementation, the subscript τ ranges from -12 to $+24$.[44] Importantly, the parameter ϕ_τ

44. We include the immediate preimplementation period, defined as twelve periods prior to PADT phase-in, in our analysis to account for the possibility of preimplementation trends in reporting (caused by anticipation effects and/or other unrelated factors). As a robustness check, we varied the number of periods included in the preimplementation

Table 2. The effect of post-accident drug testing on claim frequency for a large retail chain, 2002–2007

	Fixed-effects negative binomial models									
			Indemnity workers' compensation claims							
	(1) Total workers' compensation claims		(2) Excluding cumulative injuries		(3) Including cumulative injuries		(4) Medical-only workers' compensation claims		(5) First-aid reports	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
Post-implementation	−0.123**	−0.229**	−0.0635	−0.162^	−0.171**	−0.199*	−0.153**	−0.272**	−0.198**	−0.272**
	(0.029)	(0.048)	(0.052)	(0.087)	(0.050)	(0.085)	(0.034)	(0.057)	(0.025)	(0.041)
"Placebo" post-implementation		0.130**		0.121		0.0336		0.147**		0.0916*
		(0.048)		(0.087)		(0.084)		(0.056)		(0.040)
Log of hours	1.125**	1.102**	1.130**	1.108**	1.057**	1.051**	1.093**	1.068**	0.952**	0.935**
	(0.13)	(0.13)	(0.23)	(0.23)	(0.22)	(0.22)	(0.16)	(0.16)	(0.12)	(0.12)
Observations	2,795	2,795	2,730	2,730	2,730	2730	2795	2795	2860	2860
Districts	43	43	42	42	42	42	43	43	44	44

Notes: Standard errors are presented in parentheses. Levels of significance are as follows: **1%, *5%, ^ 10%. The dependent variable in the above models is the number of injury claims/reports. The post-implementation dummy takes a value of 1 for all periods following implementation in treatment districts, and 0 otherwise. The unit of observation is district × company time period. All models include district-level fixed-effects and time-period dummies. The number of observations and districts fluctuates slightly across the models because one very small district reported no indemnity claims, and another very small district reported no indemnity claims or medical-only claims, during our entire sample period, and therefore were automatically dropped from the applicable models. Specification (a) of each model examines the effect of the implementation of the PADT program on the frequency of claims/reports; specification (b) of each model also contains a "placebo" post-implementation dummy which takes a value of 1 five periods prior to the true implementation date. All models were estimated on a cut of the data, which excluded data from facilities that closed before any treatment districts implemented the PADT program; included data from nonimplementing districts in the treatment divisions; and (with the exception of Models 3(a) and 3(b)) excluded claims associated with cumulative injuries.

Besides the specifications presented above, we conducted the following additional robustness checks:

(1) We included claims and hours data associated with stores that closed before any of the treatment districts implemented the PADT program. This robustness check produced parameters exhibiting significance levels identical to those displayed above, with identical signs on statistically significant parameters.

(2) We excluded claims and hours data associated with the five nonimplementing districts within the treatment divisions. This robustness check produced parameters exhibiting significance levels identical to those displayed above, with identical signs on statistically significant parameters.

(3) We included claims associated with cumulative injuries in all models (in addition to Models 3(a) & 3(b)). The estimated parameters in all such models exhibited significance levels identical to or more significant than those presented above, with identical signs on statistically significant parameters.

(4) We re-ran all models using facility × time period as the unit of observation. All estimated parameters exhibited significance levels identical to or more significant than those displayed above, with identical signs on statistically significant parameters.

(5) In addition to the specifications (b) of each model presented above, in which the "placebo" implementation dummy comes into effect five periods before implementation, we estimated models in which the "placebo" implementation dummy came into effect six, four, three, and two periods, respectively, before the true implementation date. In the first three such cases, the significance levels of the post-implementation dummy were identical to those displayed above. In the fourth case, however (in which the "placebo" dummy comes into effect two periods beforehand), the *P*-value on the post-implementation dummy became significant at the 5% level in Model 1(b), and the post-implementation dummy became significant at the 10% level in Model 4(b). Although the coefficient on the "placebo" post-implementation dummy was significant in some specifications and insignificant in others, regardless of which "placebo" dummy was included, it uniformly carried a positive sign.

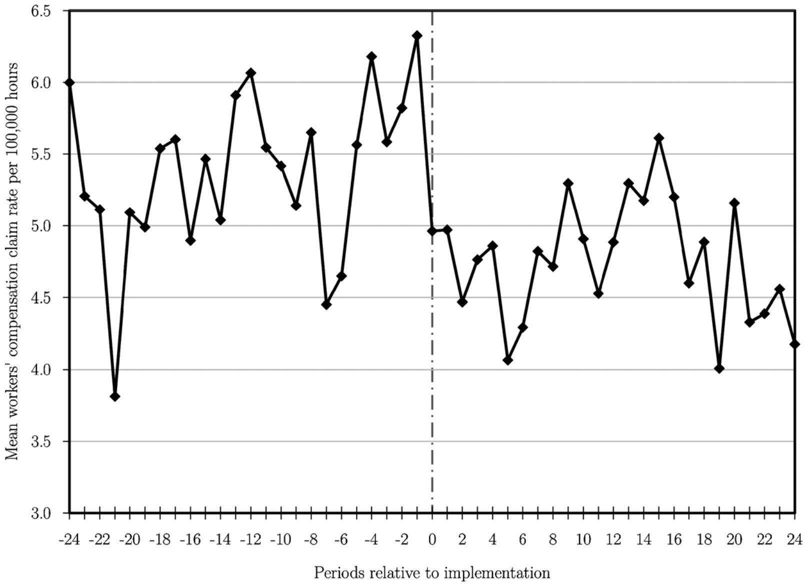

Figure 1. Mean rate of workers' compensation claims filed by period relative to post-accident drug testing program implementation for a large retail chain, 2002–2007. Notes: Each point on the above graph represents the mean rate of total workers' compensation claims filed (per 1000,000 hours worked) in all treatment districts at a given period relative to PADT implementation. (Means were first computed within each district × period cell; then for each period, means were computed across all districts.) The cut of the data reflected in the above graph: (1) excludes data from facilities that closed before any treatment districts implemented the PADT program; (2) includes data from nonimplementing districts in the treatment divisions; and (3) includes claims associated with cumulative injuries.

(which pertains only to treated districts) should be interpreted relative to all omitted event-time periods in a given district, i.e., all periods *prior to* twelve periods before implementation. For example, ϕ_2 would reflect the factor change in reported claims between two time increments in treatment districts: twelve periods and earlier before implementation; and the second period after the phase-in of PADT. In order to visually identify changes over time, Figures 2–5 graphically plot the estimated values of ϕ_τ (with their associated confidence intervals) for each of the periods examined. Plotting

period and reran the identical set of models. The particular number of periods used in defining the preimplementation period did not materially affect our results.

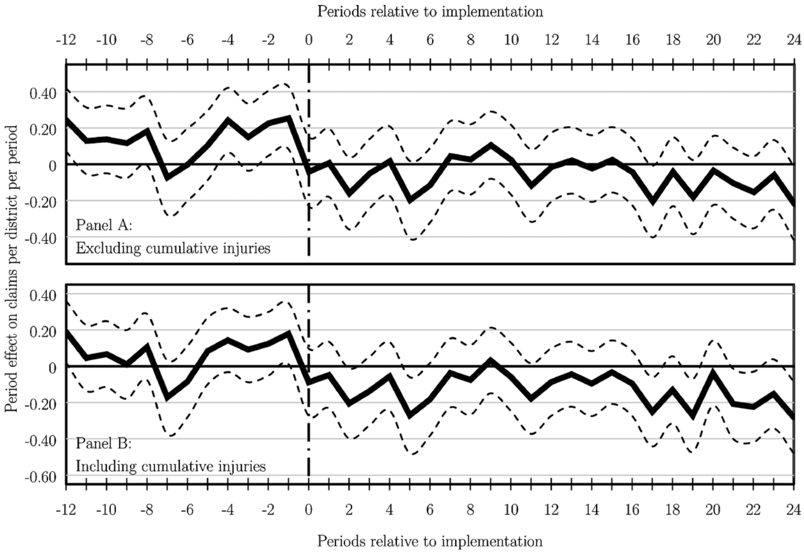

Figure 2. Effects of post-accident drug testing over time on TOTAL workers' compensation claims for a large retail chain, 2002–2007. The figures above are derived from negative binomial models in which the dependent variable is the number of total workers' compensation claims filed, and the unit of observation is the district $\times$ time period. Each model includes district-level fixed effects, the log of hours worked, "calendar" time periods (which correspond roughly to calendar months, as described in footnote 41), and "relative" time periods (calculated from the date of each district's implementation of the PADT program). The model estimated in Panel A excludes all cumulative injuries from the sample, while the model estimated in Panel B includes them. The thick center line in each figure represents the coefficients on the "relative" time-period dummies, whose values range from -12 to $+24$. The two thin lines "bracketing" each thick line represent the boundaries on the smoothed 95 percent confidence interval. In effect, then, each displayed coefficient represents the effect of "relative" time period on the frequency of claims in treatment districts, as compared to: (1) the frequency of claims among treatment districts earlier than twelve periods before PADT implementation; and (2) the frequency of claims among control districts. (Since the control districts never implemented the PADT program, the "relative" time-period dummies for such districts always take on a value of zero. Similarly, the "relative" time-period dummies from treatment districts earlier than twelve periods before PADT implementation are always coded as zeros). Both of the above models were estimated on a cut of the data, which: (1) excluded data from facilities that closed before any treatment districts implemented the PADT program; and (2) included data from nonimplementing districts in the treatment divisions.

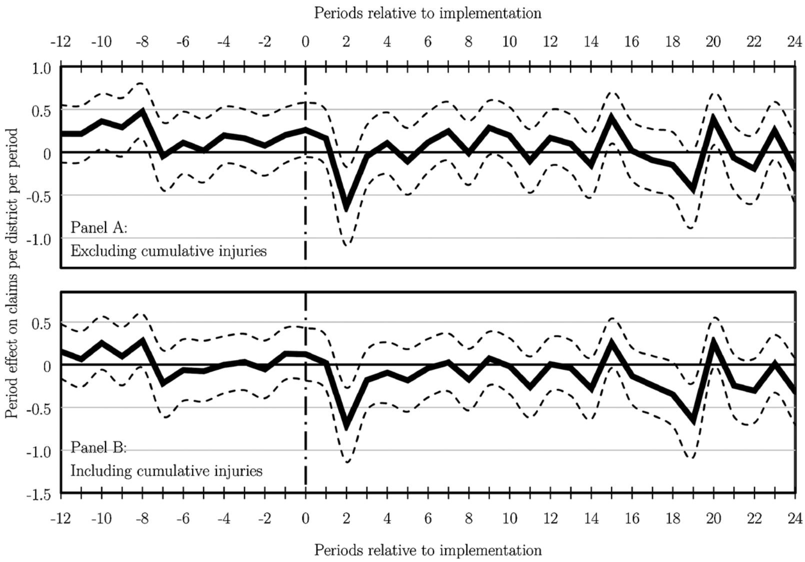

Figure 3. Effects of post-accident drug testing over time on INDEMNITY workers' compensation claims for a large retail chain, 2002–2007. The figures above are derived from negative binomial models in which the dependent variable is the number of indemnity workers' compensation claims filed, and the unit of observation is the district × time period. Each model includes district-level fixed effects, the log of hours worked, "calendar" time periods (which correspond roughly to calendar months, as described in footnote 41), and "relative" time periods (calculated from the date of each district's implementation of the PADT program). The model estimated in Panel A excludes all cumulative injuries from the sample, while the model estimated in Panel B includes them. The thick center line in each figure represents the coefficients on the "relative" time-period dummies, whose values range from -12 to $+24$. The two thin lines "bracketing" each thick line represent the boundaries on the smoothed 95 percent confidence interval. In effect, then, each displayed coefficient represents the effect of "relative" time period on the frequency of claims in treatment districts, as compared to: (1) the frequency of claims among treatment districts earlier than twelve periods before PADT implementation; and (2) the frequency of claims among control districts. (Since the control districts never implemented the PADT program, the "relative" time-period dummies for such districts always take on a value of zero. Similarly, the "relative" time-period dummies from treatment districts earlier than twelve periods before PADT implementation are always coded as zeros). Both of the above models were estimated on a cut of the data, which: (1) excluded data from facilities that closed before any treatment districts implemented the PADT program; and (2) included data from nonimplementing districts in the treatment divisions.

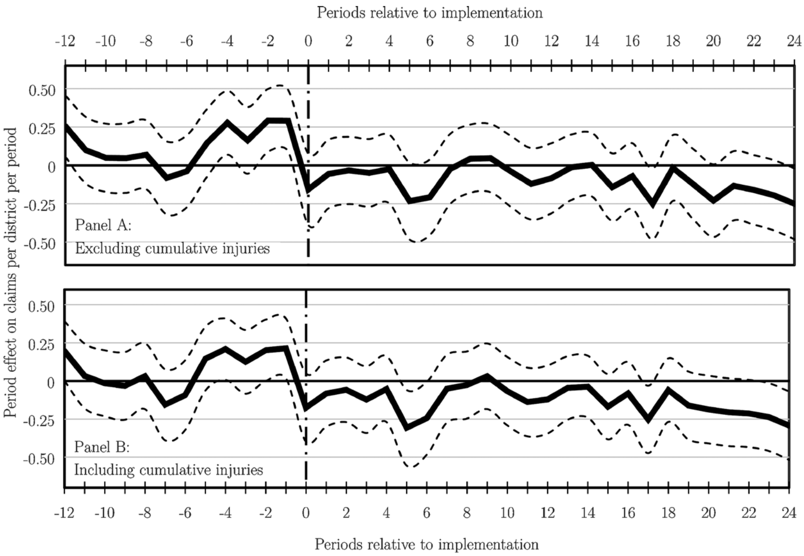

Figure 4. Effects of post-accident drug testing over time on MEDICAL-ONLY workers' compensation claims for a large retail chain, 2002–2007. The figures above are derived from negative binomial models in which the dependent variable is the number of medical-only workers' compensation claims filed, and the unit of observation is the district × time period. Each model includes district-level fixed effects, the log of hours worked, "calendar" time periods (which correspond roughly to calendar months, as described in footnote 41), and "relative" time periods (calculated from the date of each district's implementation of the PADT program). The model estimated in Panel A excludes all cumulative injuries from the sample, while the model estimated in Panel B includes them. The thick center line in each figure represents the coefficients on the "relative" time-period dummies, whose values range from −12 to +24. The two thin lines "bracketing" each thick line represent the boundaries on the smoothed 95 percent confidence interval. In effect, then, each displayed coefficient represents the effect of "relative" time period on the frequency of claims in treatment districts, as compared to: (1) the frequency of claims among treatment districts *earlier than* twelve periods before PADT implementation; and (2) the frequency of claims among control districts. (Since the control districts never implemented the PADT program, the "relative" time-period dummies for such districts always take on a value of zero. Similarly, the "relative" time-period dummies from treatment districts *earlier than* twelve periods before PADT implementation are always coded as zeros). Both of the above models were estimated on a cut of the data, which: (1) excluded data from facilities that closed before any treatment districts implemented the PADT program; and (2) included data from nonimplementing districts in the treatment divisions.

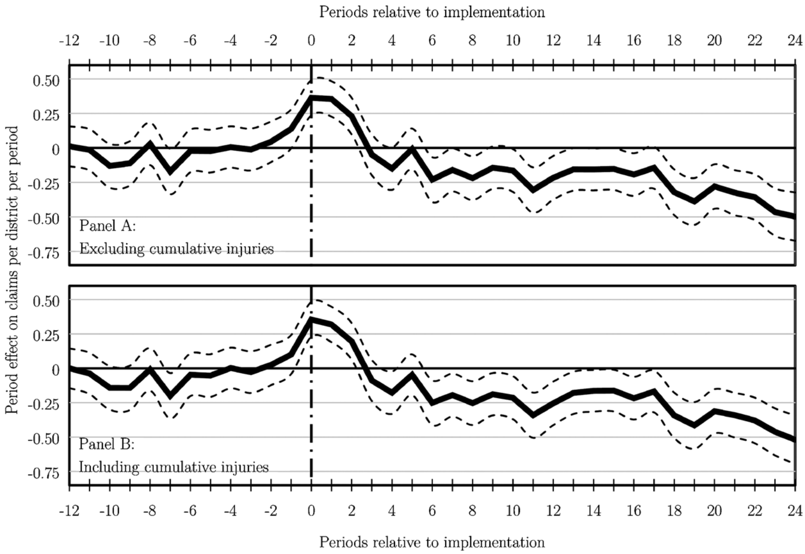

Figure 5. Effects of post-accident drug testing over time on FIRST AID reports for a large retail chain, 2002–2007. The figures above are derived from negative binomial models in which the dependent variable is the number of first-aid reports filed, and the unit of observation is the district × time period. Each model includes district-level fixed effects, the log of hours worked, "calendar" time periods (which correspond roughly to calendar months, as described in footnote 41), and "relative" time periods (calculated from the date of each district's implementation of the PADT program). The model estimated in Panel A excludes all cumulative injuries from the sample, while the model estimated in Panel B includes them. The thick center line in each figure represents the coefficients on the "relative" time-period dummies, whose values range from −12 to +24. The two thin lines "bracketing" each thick line represent the boundaries on the smoothed 95 percent confidence interval. In effect, then, each displayed coefficient represents the effect of "relative" time period on the frequency of claims in treatment districts, as compared to: (1) the frequency of claims among treatment districts *earlier than* twelve periods before PADT implementation; and (2) the frequency of claims among control districts. (Since the control districts never implemented the PADT program, the "relative" time-period dummies for such districts always take on a value of zero. Similarly, the "relative" time-period dummies from treatment districts *earlier than* twelve periods before PADT implementation are always coded as zeros). Both of the above models were estimated on a cut of the data, which: (1) excluded data from facilities that closed before any treatment districts implemented the PADT program; and (2) included data from nonimplementing districts in the treatment divisions.

these values for a wide range of periods allows us to detect not only whether there were any interesting dynamics after implementation, but also whether claims were trending prior to implementation. As before, we estimate Model (2) separately for different categories of claims.

Table 3. Percentiles of reference distributions of post-accident drug testing treatment effects for a large retail chain, 2002–2007

Non-parametric randomization tests of the parameters of a fixed-effects negative binomial model

Percentile	(1) All divisions	(2) Control divisions only
0.01	−0.1101	−0.1174
0.05	−0.0798	−0.0842
0.10	−0.0595	−0.0677
0.25	−0.0303	−0.0361
0.50	0.0032	0.0020
0.75	0.0346	0.0394
0.90	0.0654	0.0722
0.95	0.0874	0.0927
0.99	0.1406	0.1333
Observed treatment effect	−0.123	
(from Table 2, Model 1(a), row 1)	−0.123	

Notes: The table above summarizes the results of a nonparametric test of the robustness of the results presented in Model 1(a) in table 2, in which the dependent variable is the total number of workers' compensation claims, and the coefficient of interest is that on the post-implementation dummy. For each percentile of the reference distribution, the table displays cutoff values of the coefficient on "simulated" post-implementation dummies. Each iteration of the randomization procedure proceeded as follows: (1) a random subset of seventeen districts (the actual number of treatment districts) were chosen as "pseudo-treatment" districts; (2) each pseudo-treatment district was randomly assigned a pseudo-implementation date from a uniform distribution of dates spanning our entire sample period; and (2) we re-estimated Model 1(a) using these seventeen pseudo-treatment districts as treatment groups. (As in our original models, each pseudo-treatment district was coded as having implemented the PADT program on the randomly chosen implementation date as well as in all subsequent periods.) The reference distribution displayed above was generated by repeating the latter procedure one thousand times. A reference distribution for which many percentile cutoff values exceeded the magnitude of the coefficient on the true post-implementation dummy (presented in table 2) would cast doubt on the specificity of our model and on the validity of our findings.

Two variations of the above exercise were conducted. In the version presented in Column (1), the seventeen "pseudo-treatment" districts were chosen from among all forty-four districts in our sample, with all remaining districts in the sample serving as controls. In the version presented in Column (2), the seventeen "pseudo-treatment" districts were chosen only from among the twenty-seven control districts, with the remaining ten control districts serving as controls.

In all respects besides those described above, the models estimated to obtain the reference distribution were identical to those estimated in table 2. As robustness checks, we performed the exercise on three alternative cuts of data which (respectively): (1) included claims and hours data associated with stores that closed before any district implemented PADT; (2) excluded claims and hours data associated with the five nonimplementing districts within the treatment divisions; and (3) included claims associated with cumulative injuries. All three of these robustness checks yielded reference distributions similar to those presented above.

Although the analysis described thus far sheds light on whether the PADT program has triggered a significant decline in injuries over time, it does not distinguish among different types of employees. Claimants with different demographic characteristics are likely to respond differently to PADT. As noted earlier, both drug use and the frequency of workers' compensation claims have been shown to vary by characteristics such as age and gender.[45] Moreover, *ceteris paribus*, the opportunity cost of job loss may be particularly high for higher-tenure and full-time workers.[46] If those workers with the most to lose are indeed the most responsive to the policy, this would suggest that drug users are "rational cheaters" in the sense described in the Shapiro and Stiglitz (1984) efficiency wage model. Workers with the highest wages (relative to their expected outside earnings) have the greatest incentive to avoid misbehaviors that might cost them their jobs, and thus are likely to respond most to the increased probability of detection.

In table 4, therefore, we test whether the average responsiveness to PADT differs among employees with different characteristics. In so doing, we construct a disaggregated cell-level dataset, in which each observation represents the number of incidents reported within a detailed demographic subgrouping. These subgroupings include all possible permutations of the following six characteristics: district, period, at least/under 35 years of age, at least/under three years of tenure, male/female, and full/part-time status. By re-estimating Equation (1) and interacting the post-implementation dummy with one or more of the above characteristics, we test whether the behavioral response differs systematically among different groups of workers.[47]

The third dimension of our analysis exploits our ability to identify employees who test positive for illegal drugs. Since our data enable us to identify all those employees to whom post-incident drug tests were administered along with their test results, we can compare the characteristics of employees who tested positive to those who tested negative. We estimate a simple probit model in which the binary dependent variable takes on the value of 1 if the worker tested positive for drugs, and 0 otherwise. We present

45. See footnote 23.

46. If the opportunity cost of job loss and the likelihood of using drugs are positively correlated, it may be difficult to distinguish each factor's relative contribution to the observed programmatic response.

47. We chose the cutoff values of three years of tenure, and 35 years of age, because they were close to the median values for the entire sample. Although we considered including the employee's salary, occupation, and department—all of which were recorded in the dataset—in our models, each suffered from a major practical shortcoming. The salary field, and the occupation field, were both of very poor quality and often missing. Since there were literally dozens of departments in each facility, there were simply too few observations to make including department worthwhile. Therefore, we decided to omit these fields from the analysis.

Table 4. The effects of post-accident drug testing and employee characteristics on claim frequency for a large retail chain, 2002–2007

Pooled, fixed-effects negative binomial models of claim frequency disaggregated by gender, age, full-time status, and tenure

	(1) Total workers' compensation claims		**Indemnity workers' compensation claims: (2) Excluding cumulative injuries**		**Indemnity workers' compensation claims: (3) Including cumulative injuries**		**(4) Medical-only workers' compensation claims**		**(5) First-aid reports**	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
Post-implementation	0.185**	0.120*	0.173^	0.0791	0.0783	−0.0128	0.0984^	0.0512	0.434**	0.343**
	(0.049)	(0.047)	(0.098)	(0.096)	(0.095)	(0.093)	(0.054)	(0.052)	(0.038)	(0.036)
Post-implementation × male	−0.254**	−0.257**	−0.194*	−0.196*	−0.175*	−0.176*	−0.283**	−0.287**	−0.173**	−0.172**
	(0.043)	(0.043)	(0.078)	(0.077)	(0.075)	(0.075)	(0.050)	(0.049)	(0.035)	(0.034)
Post-implementation × older	0.0136	0.00379	0.146^	0.147^	0.142^	0.144^	0.0428	0.0337	−0.357**	−0.379**
	(0.042)	(0.041)	(0.082)	(0.082)	(0.079)	(0.079)	(0.047)	(0.047)	(0.034)	(0.033)
Post-implementation × full-time	−0.663**	−1.136**	−0.674**	−1.118**	−0.651**	−1.136**	−0.618**	−1.125**	−0.896**	−1.391**
	(0.050)	(0.12)	(0.084)	(0.22)	(0.080)	(0.21)	(0.059)	(0.14)	(0.042)	(0.094)
Post-implementation × higher tenure	−0.113**	0.0000659	−0.130^	0.0154	−0.156*	−0.0295	−0.0576	0.0348	−0.358**	−0.244**
	(0.042)	(0.047)	(0.075)	(0.086)	(0.073)	(0.083)	(0.047)	(0.054)	(0.034)	(0.038)
Post-implementation × full-time × higher tenure		0.575**		0.511*		0.573*		0.607**		0.752**
		(0.13)		(0.24)		(0.23)		(0.15)		(0.11)
Full-time × higher tenure		1.618**		1.609**		1.602**		1.603**		1.684**
		(0.039)		(0.065)		(0.062)		(0.047)		(0.035)
Male	−0.408**	−0.409**	−0.479**	−0.479**	−0.542**	−0.542**	−0.378**	−0.379**	−0.495**	−0.498**
	(0.017)	(0.016)	(0.027)	(0.027)	(0.026)	(0.026)	(0.020)	(0.020)	(0.015)	(0.015)

(continued overleaf)

Table 4. *(Continued)*

			Indemnity workers' compensation claims							
	(1) Total workers' compensation claims		(2) Excluding cumulative injuries		(3) Including cumulative injuries		(4) Medical-only workers' compensation claims		(5) First-aid reports	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
Older	0.271**	0.274**	0.813**	0.816**	0.843**	0.846**	0.0171	0.0188	0.0439**	0.0372**
	(0.017)	(0.016)	(0.029)	(0.029)	(0.027)	(0.027)	(0.020)	(0.019)	(0.015)	(0.014)
Full-time	−0.702**	−1.681**	−0.420**	−1.497**	−0.395**	−1.490**	−0.845**	−1.767**	−0.763**	−1.759**
	(0.018)	(0.032)	(0.027)	(0.055)	(0.026)	(0.053)	(0.021)	(0.038)	(0.016)	(0.029)
Higher tenure	0.168**	−0.331**	0.506**	−0.0438	0.580**	0.0309	−0.00609	−0.468**	0.0821**	−0.436**
	(0.017)	(0.020)	(0.027)	(0.034)	(0.026)	(0.032)	(0.020)	(0.024)	(0.015)	(0.018)
Log of hours	0.570**	0.846**	0.848**	1.007**	0.755**	0.904**	0.652**	0.870**	0.240**	0.401**
	(0.12)	(0.12)	(0.22)	(0.23)	(0.21)	(0.21)	(0.15)	(0.15)	(0.092)	(0.096)
Observations	44,720	44,720	43,680	43,680	43,680	43,680	44,720	44,720	45,760	45,760
Districts	43	43	42	42	42	42	43	43	44	44

Notes: Standard errors are presented in parentheses. Levels of significance are as follows: **1%, *5%, ^ 10%. The dependent variable in the above models is the number of injury claims/reports. The unit of observation is district × time period × sex × age × full-/part-time status × higher/lower tenure. "Older" refers to employees of age 35 or older, and "higher tenure" refers to employees with three or more years' tenure with the company. Although the array of control variables varies slightly between specifications (a) and (b), all models include district-level fixed-effects and time-period dummies. The number of observations and districts fluctuates slightly across the models because one very small district reported no indemnity claims, and another very small district reported no indemnity claims or medical-only claims, during our entire sample period, and therefore were automatically dropped from the applicable models. All models were estimated on a cut of the data, which excluded data from facilities that closed before any treatment districts implemented the PADT program; included data from nonimplementing districts in the treatment divisions; and (with the exception of Models 3(a) and 3(b)) excluded claims associated with cumulative injuries.

Besides the specifications presented above, we conducted the following additional robustness checks:

(1) We included claims and hours data associated with stores that closed before any of the treatment districts implemented the PADT program. This robustness check produced parameters exhibiting significance levels identical to or more significant than those displayed above, with identical signs on statistically significant parameters.

(2) We excluded claims and hours data associated with the five nonimplementing districts within the treatment divisions. Using this cut of the data induced the following changes: (1) the "post-implementation" dummy in Specification (1b) became significant at a 10% level, and the post-implementation dummy in Specification 4(a) became insignificant; (2) the "post-implementation × full-time × higher tenure" dummy in Specification (2b) became significant at 10% level; (3) the "older" dummy variable in Specification (5b) became significant at a 5% level; and (4) the "log of hours" variable in Specification 5(a) became significant at a 10% level.

(3) We included claims associated with cumulative injuries in all models (in addition to Models 3(a) & 3(b)). The estimated parameters in all such models exhibited significance levels identical to or more significant than those presented above, with identical signs on statistically significant parameters.

(4) We re-ran all models using facility × time period as the unit of observation. All estimated parameters exhibited significance levels identical to or more significant than those displayed above, with identical signs on statistically significant parameters.

four alternative specifications of the model, each with a slightly different array of explanatory variables.[48] Since we are analyzing only claims that were subject to PADT, the sample used for this analysis is all claims filed in treatment districts after the implementation of the program.

Importantly, our data only permit us to analyze the probability of a positive test result *conditional on a test being administered*. This is not the same as the unconditional likelihood of an employee using illegal drugs. The reason is simple: if certain demographic groups are more likely than others to underreport their injuries or to take care on the job, then the sample of employees to whom the PADT is administered will exhibit selection bias.[49] Observed differences in the percentage of positive drug tests across groups, therefore, may not correspond to rates of drug use in the underlying population. They may, however, reflect differences in how highly different groups value Company jobs and, consequently, the lengths to which they are willing to go to avoid drug testing.

Most of the explanatory variables included in the model—such as tenure, age, gender, claim type (indemnity versus medical only versus first aid), and region—are straightforward. Three, however, require a bit more explanation. First, as noted earlier, the Company's preemployment drug testing policy was implemented in the early 1990s. Employees with more than 15 years' tenure at the Company, therefore, may have never taken a drug test. Prior research suggests that exposure to drug testing may increase employee support for the policy,[50] raising the possibility that employees who have never been

48. To probe the robustness of our findings, we estimated all possible model specifications whose independent variables included tenure as well as some permutation of the other eight parameters. (The indemnity and medical dummies were included or excluded from the model as a pair, since together they indicate a three-way categorical variable.)

49. Imagine, for example, that the workforce is 50 percent female; all employees are equally injury-prone (regardless of drug use); a constant proportion of injuries are "hideable," and women are much more likely than men to use illegal drugs. In this situation, we would expect to observe two phenomena. First, we would expect fewer women to report injuries (for fear of losing their jobs). Second—even though there are more female than male drug users in the underlying population—we would expect the proportion of positive drug tests to be about the same among men and women.

50. See, for example, Bennett et al. (1994), who find a higher level of support for drug testing among employees whose firms conduct the practice and interpret this as evidence that exposure increases acceptance, although they are unable to rule out the possibility that employees who approve of the policy self-select into firms that adopt it; and Crant and Bateman (1989), who suggest that prior exposure increases employee acceptance.

tested experience higher "testing anxiety" than their peers.[51] In an effort to detect whether there is any link between prior testing experience and the conditional likelihood of a positive test, we include a dummy variable for "exposure to preemployment test" in one specification.

Second—as the following section explores in greater detail—an employee's discretion over whether to report an injury is likely to vary by the type of injury. Since lacerations involve a loss of blood, they may be less "hideable" than contusions (bruises) and abrasions (scrapes), the other major injury categories. If employees who use illegal drugs find it more difficult to hide lacerations than other types of injuries, then one might expect a higher proportion of lacerations to result in positive drug tests. To evaluate this hypothesis, we include a dummy variable for "laceration" in one specification. (Although cumulative injuries could be subject to similar reporting effects, we do not include them in this model since they are technically outside the scope of the program.[52])

Finally, the opportunity cost of losing one's job may vary with tenure. If the slope of the wage–tenure profile remains steep over a significant number of years, yet employees' productivity levels off rapidly, then higher-tenure employees may value their jobs more than their lower-tenure counterparts. Yet the magnitude of this disparity may not increase (with seniority) at a constant rate. For example, employees at the middle of the wage–tenure profile may value their jobs less than their higher-tenure peers who have

51. Even if "inexperienced" employees (i.e., those who have never been tested before) perceive the test as more invasive or burdensome, it is uncertain how this would affect the conditional likelihood of an employee testing positive. On one hand, if inexperienced employees who do *not* use drugs are extremely averse to taking the test, this could increase the fraction of drug users among those inexperienced workers that are tested. On the other hand, inexperienced drug users may also be more averse to testing than their experienced counterparts if, for example, they are less confident of their abilities to "beat the test." In other words, even if prior exposure affects employee's aversion to taking a drug test, the responses of drug users and non-drug-users within the inexperienced population could be mutually offsetting.

52. As noted earlier, company records reveal that an employee in a treatment division who received first aid treatment in 2005 for a cumulative hand injury tested positive and was terminated the following week. However, to our knowledge, this is the only case on record in which an employee has been fired for drug use after reporting a cumulative injury. Since we are uncertain of precisely how many cumulative injuries are in fact tested—and if only a small proportion are drug tested, including *all* cumulative injuries could bias our results—we decided that the safer course was to exclude cumulative injuries from the model.

vested into lucrative pension plans. To capture such possible nonlinearities in the effect of seniority on the opportunity cost of job loss, we include a quadratic polynomial as an explanatory variable in one specification of the third-stage models presented in table 5.

The fourth and final dimension of our analysis probes whether underreporting is likely to explain some of the observed drop in claims. In the hopes of finding "circumstantial evidence" that may shed light on the question, we analyze the data in three additional ways. First, *ceteris paribus*, one would expect minor injuries to be the most prone to underreporting, since they probably are the easiest to hide, rarely are disabling, and may not require immediate professional attention. With this assumption in mind, we compare trends among indemnity workers' compensation claims, the severest category, to trends among "medical only" workers' compensation claims and "first aid" claims.

Second, we compare lacerations and nonlacerations, on the logic that the former are generally more difficult to hide (since they involve a loss of blood), and thus may be the least prone to underreporting. The animating assumption of this second empirical inquiry is that a disproportionate drop in minor and/or nonlaceration injuries may be suggestive of some underreporting.

Our third and final test of underreporting, a comparison of cumulative and noncumulative injuries, requires careful explanation. Of all injury types, those classified as "cumulative" are probably, by far, the easiest to hide. Moreover, even if the symptoms themselves are readily apparent, the inherent difficulty of ascertaining whether such injuries are "work related" (in the requisite statutory sense) gives employees tremendous *de facto* discretion in deciding whether to report them.[53] Consequently, *ceteris paribus*, one might expect cumulative injuries to be the most prone to underreporting. As noted earlier, although cumulative injuries are formally excluded from the Company's PADT program, both anecdotal and documentary evidence gleaned from the Company casts doubt on whether this exemption is widely understood or enforced. In light of these countervailing factors, the results of our third test should be interpreted with caution. If cumulative injuries do

53. For an overview of the special problems involved in compensating cumulative injuries through the workers' compensation system, see, e.g., Shulman and Hofflander (1980).

Table 5. Probit models of the determinants of positive drug tests for a large retail chain, 2004–2007

	Specification (1)	**Specification (2)**	**Specification (3)**	**Specification (4)**
Full-time	0.000168	−0.0000622	−0.00223	−0.00197
	(0.0029)	(0.0029)	(0.0024)	(0.0026)
Tenure (expressed as share of five-year intervals)	−0.00298**	−0.00289**	−0.00439**	−0.00440*
	(0.00094)	(0.00092)	(0.0012)	(0.0020)
Tenure squared				0.000380
				(0.00040)
Laceration	0.000243			
	(0.0020)			
Male		0.00284	0.00307	0.00313
		(0.0019)	(0.0019)	(0.0019)
Age (expressed as share of five-year intervals)			0.000354	0.000470
			(0.00032)	(0.00033)
Indemnity			−0.00128	−0.00138
			(0.0030)	(0.0030)
Medical			0.000631	0.000578
			(0.0020)	(0.0021)
Division 1			−0.00122	0.00108
			(0.0019)	(0.0017)
Hired prior to implementation of preemployment drug testing			0.00799ˆ (0.0042)	
Observations	10,730	10,754	10,731	10,731

Notes: The sample period in the above models is 2004–2007 (as opposed to 2002–2007, the sample period for all other tables) because these models were estimated using data from positive drug tests recorded after the implementation of the post-accident drug testing program. The earliest-implementing company districts implemented the program in 2004.

The dependent variable in the above models is the likelihood of a positive drug test given that a test has been administered through the PADT program. Displayed estimates are marginal effects probabilities. The unit of observation is the drug test administered to each tested employee. Standard errors are presented in parentheses, and levels of significance are as follows: **1%, *5%, ˆ10%. All models were estimated on a cut of the data, which excluded data from facilities that closed before any treatment districts implemented the PADT program, and included data from nonimplementing districts in the treatment divisions. Claims and reports associated with cumulative injuries were excluded from the analysis. The number of observations fluctuates slightly across model specifications because we lack information on certain characteristics for some employees and some drug-tested accidents.

In order to limit the number of decimal places reported, age and tenure have been rescaled as a share of five-year time intervals. For example, a worker who has been employed for 1,095 days has a tenure value of 1,095/(365 × 5) = 0.6. The variable "hired prior to implementation of preemployment drug testing" is 1 for workers who were hired before preemployment drug-testing began in the applicable division.

To probe the robustness of the coefficient on "tenure," we estimated all possible model specifications whose independent variables included tenure as well as some permutation of the other eight parameters. (The indemnity and medical dummies were included or excluded from the model as a pair, since together they indicate a three-way categorical variable.) In all specifications that excluded "tenure squared," tenure was significant at the 1% level. In all specifications that included "tenure squared," tenure was significant at either a 5% or 1% level.

not exhibit larger declines than traumatic injuries, this could indicate simply that most employees and supervisors correctly understand such claims to be outside the scope of the program. If, on the other hand, cumulative injuries were to fall disproportionately, this would suggest that the net effect of the "hideability" factor probably outweighs the effect (if any) of their formal exclusion.

4. Results

Figure 1 provides an initial "raw" look at the frequency of workers' compensation claims before and after PADT implementation, plotting the average frequency of both cumulative and noncumulative claims (per one hundred thousand worker hours) across all treatment districts, for two years before and after the date of PADT implementation. Although the data are quite noisy, casual visual inspection suggests that the frequency of claims may have declined roughly contemporaneously with the phase-in date.[54]

Results from the first phase of the formal analysis are presented in tables 2 and Table 3. 2 presents two variations of each model, which are labeled Specification (a) and Specification (b). Specifications (a) contain estimates of the effect of PADT implementation, respectively, on total workers' compensation claims, indemnity workers' compensation claims, medical-only workers' compensation claims, and first-aid reports. Each specification of each model was estimated on two different "cuts" of the data: one that excluded cumulative injuries, and one that included them. Models 2a and 3a in table 2 present the results for each of these "cuts," respectively, for indemnity workers' compensation claims. Given space constraints, we present only the specification *excluding* cumulative injuries for all other injury types. (As described in the notes underneath table 2, the coefficient on the "post-implementation" dummy was always—if anything—*more* statistically significant when cumulative injuries were included.)

Table 2 reveals that total claims of all types fell significantly after the phase-in of PADT (although in the case of indemnity claims, the significance

54. The cut of data from which figure 1 was derived excludes data from facilities that closed before any treatment districts implement the PADT program; includes data from the small handful of nonimplementing districts in treatment divisions; and (as noted above) includes cumulative injuries. Although we do not present them here, similar figures derived from alternative cuts of the data exhibit similar overall trends.

of the effect disappears when cumulative injuries are excluded).[55] Focusing on Model 1a and 5a, for example, the coefficients on the "post-implementation" dummies suggest that PADT is associated with a decline in total workers' compensation claims of about 12 percent, and a decline in first-aid reports of about 18 percent.[56]

Specifications (b) of table 2 present the results of our first robustness check. As described earlier, we estimate composite models containing *both* the actual treatment and a fake "placebo" treatment that takes effect two, three, four, five, or six periods (respectively) beforehand. The models presented in table 2 include the "placebo" dummy taking effect five periods beforehand. As can be seen from the table, when both the real and placebo dummies are included, only the real treatment dummy is associated with significant declines.[57]

Table 3 presents the results of our second robustness check described earlier, involving two variations of a randomization inference technique.[58] By conducting each simulation exercise one thousand times, we estimate

---

55. We chose to include the log of hours as an independent variable rather than treating it as an exposure variable and constraining its coefficient (β) in Equation (1) to be 1. Table 2 reveals that β is approximately equal to 1 in all specifications, suggesting that a given percentage change in hours translates into a similar (or perhaps for workers' compensation claims, slightly higher) percentage change in claims.

56. In negative binomial models, elasticities can be calculated by exponentiating the estimated coefficient and subtracting one. Using Model 1a as an example, therefore, the percentage decline $= (\exp(-0.123) - 1) * 100 \approx 11.6$ percent. In Model 5a, the estimated percentage decline $= (\exp(-.198)-1) * 100) \approx 18.0$ percent.

57. Interestingly, some types of claims appear to be significantly *increasing* five months prior to implementation, an unanticipated result that might suggest some increased vigilance on the part of store managers to log injuries in the months leading up to PADT implementation. The models that included "placebo" starting dates of two, three, four, and six periods prior to actual implementation yielded identical results, with one exception. When the "-2" placebo dummy was included in the model, the *P*-value on the post-implementation dummy became significant at the 5 percent level in Model 1(b), and the post-implementation dummy became significant at the 10 percent level in Model 4(b). This decline in significance is not surprising (and in our view, does not undermine the robustness of our findings), since one would expect anticipation effects to be relatively pronounced two periods before the true implementation date.

58. In the version presented in column (1), the seventeen "pseudo-treatment" districts were chosen from among all forty-four districts, with the remaining districts in the sample serving as controls. In the version presented in column (2), the seventeen "pseudo-treatment" districts were chosen only from among the twenty-seven control districts and the remaining ten control districts served as controls.

the distribution of "treatment effects" obtained by randomly assigning fake implementation dates to randomly chosen groups of "pseudo-treatment" districts. Doing so enables us to compare the observed treatment effects ($\hat{\Phi}$) presented in table 2 to the randomization distribution of treatment coefficients. Each row of table 3 presents the "cutoff" coefficient value for the corresponding percentile from the randomization distribution, obtained from repeatedly estimating Model 1a presented in table 2 (in which the dependent variable is total workers' compensation claims). Comparing the coefficient on the post-implementation dummy (-0.123) with the percentiles of that coefficient's randomization distribution, it is evident that the observed decline is much larger than one would expect to see if the treatment districts and dates were randomly assigned. Regardless of which assignment mechanism we use in choosing pseudo-treatment districts, we observe a treatment effect that resembles the true value (i.e., is negative and at least as large in magnitude) in less than 1 percent of our simulations. In short, we have little reason to believe that our standard errors are biased or that the results presented in table 2 result from model misspecification.

An important limitation of the models presented in table 2 is that they do not distinguish between short-term and long-term effects. If, for example, claims frequency were to change immediately upon implementation, but then quickly revert back to its long-term preimplementation levels, then the significance of the post-implementation dummy might be of lesser real-world import. Figures 2–5 seek to explore this important time dimension. Each data point of each graph (derived from a negative binomial model of injury claims/reports) represents the factor change in reported injuries in all treatment districts at a given number of periods, ranging from -12 to $+24$, relative to PADT implementation. Those estimates above (below) zero on the y-axis indicate that the frequency of claims was higher (lower) than the baseline claims frequency. All models account for general time-patterns in claims and for the log of total labor hours. The thin lines that "bracket" each graph represent 95 percent confidence intervals on the estimated coefficient. Panel A (the upper panel) of each figure excludes cumulative injuries, while Panel B (the lower panel) includes them.

The figures contain several interesting findings. First, juxtaposing Panels A and B suggests that for any given claim type, including cumulative injuries tends to magnify the drop (if any) in claim frequency following PADT implementation. Second, the figures suggest that the persistence of the PADT effect differs among claim types. Figure 2 reveals a modest yet apparently persistent decline in total workers' compensation claims in the two years following PADT implementation, although the decline is statistically insignificant for many individual periods. Examining workers' compensation trends separately by level of severity, however, brings important disparities to light. Figure 3 indicates a rather noisy and ambiguous trend in the frequency of indemnity workers' compensation claims, with most post-implementation periods exhibiting only a slight (and statistically insignificant) drop. In contrast, the post-implementation decline in medical-only workers' compensation claims shown in figure 4 is not only larger in most periods, but also appears more stable and persistent. First-aid reports, depicted in figure 5, exhibit unique trends. Although the frequency of such claims spiked just before and after the implementation date, they fell precipitously over the next several periods and continued to decline in the following two years.

The first stage of the empirical analysis, then, suggests that PADT has significantly lowered injury claims/reports, with the less serious categories (medical-only workers' compensation claims and first-aid reports) tending to exhibit the largest and most persistent declines. In light of these findings, it is worth considering the possibility that rather than hiding their injuries, some employees may simply *delay* reporting them until they can "get clean" or take steps to ensure that they will pass a drug test. To test this hypothesis, we examined whether implementation of PADT changed the promptness with which employees filed claims after sustaining an occupational accident. Specifically, we modeled the number of elapsed days between an accident and the filing of a claim, from three periods before, through three periods after, PADT implementation.[59] Interestingly, the results suggest that the rapidity with which employees filed claims actually *increased* just prior to, and immediately following, the implementation of PADT. The most likely explanation for this phenomenon—which Company officials confirmed—was that although the mandatory post-accident reporting policy had been in place for many years beforehand, the roll-out of the PADT program made it more salient to both employees and supervisors, thereby increasing compliance. In short, there is little evidence that PADT

59. We used a logit model and a Cox proportional hazard model to test the effect of PADT on the average elapsed days. Although the results are not presented in this paper, they are available from the authors upon request.

implementation caused employees to *delay* filing claims. Rather, employees that declined to file claims after sustaining injuries probably chose not to report them at all.

The second stage of the analysis explores whether employees with different demographic characteristics react differently to PADT implementation. Specifically, our data enable us to disaggregate employee responsiveness by four different attributes: sex, age, tenure, and full-time/part-time status.[60] Since the Company could not provide us with total hours at the facility-level broken down by all of these characteristic variables, we could not compare the overall *rate* of claims (i.e., number per worker-hour) separately for each group.[61] Therefore, we modeled changes in the total *number* of claims as a function of key worker characteristics and the log of total hours worked during the same district and period. As long as each group's relative contribution to total worker hours remained stable before and after implementation—an assumption which the Company informed us was reasonable during this time period—the interaction term for each group should still reflect its relative responsiveness to PADT.

For each injury category, table 4 models the number of claims as a function of a post-implementation dummy, the four employee characteristics listed above, and interactions between them. (All models also include district-level fixed effects, time-period dummies, and the log of hours.) Specification (a) of each model interacts each employee attribute with the post-implementation dummy. Specification (b) of each model includes, further, interaction effects between full-time and higher tenure; and between both these characteristics and the post-implementation dummy. The goal of Specification (b) was to test the hypothesis that full-time workers with higher tenure—who earn the most from their jobs at the Company—would have the most to lose from job loss and therefore would be the *most* responsive

60. The Company's database does contain fields for average weekly wage, occupation, and department. The data on wage and occupation were problematic because they frequently contained missing values and were generally of poor quality. The data on department, although reasonably complete, was problematic since there were dozens of departments with very few observations in each, and many regressions did not converge given the very small sample sizes. Therefore, we chose not to include these three variables in the analysis.

61. Facility-level hours were available broken down by full-time/part-time status, but not by sex, age, or tenure.

to PADT implementation. In all models, the coefficients of interest are the interaction terms between the post-implementation dummy and employee demographics. (The significant coefficients on the basic worker attribute variables—"male," "older," etc.—simply reflect the relative proportions of these groups in the Company's total workforce.[62])

Table 4 reveals several interesting patterns. First and foremost, full-time workers (who constitute about a fifth of the Company's workforce) are much more responsive to PADT than their lower tenure counterparts. This finding seems to bear out our hypothesis that workers with the most to lose from failing a drug test will, *ceteris paribus*, be the most responsive.[63]

The other robust finding is that across all claim types, the frequency of claims drops much more sharply among male workers. Although the magnitude of this gender disparity differs somewhat across claim types, it is uniformly statistically significant at the 5 percent or 1 percent level. The most plausible explanation for this finding, in our view, is that male employees in our sample are more likely to use illegal drugs than their female counterparts, and therefore are more likely to alter their behavior following PADT implementation.[64]

Interestingly, the patterns of response among higher-tenure workers (those with three or more years' tenure) are rather complex and puzzling. Viewed in isolation, Specifications (a) suggest that higher-tenure workers are only slightly more responsive than their lower-tenure peers, at least for some claim types. This relatively attenuated effect (as compared to full-time status) could be at least partly explained by the positive correlation, demonstrated by prior scholars, between drug use and employee turnover (Kandel

62. We requested, but were unable to obtain, detailed breakdowns of the Company's entire workforce by age, gender, etc. The Company informed us, however, that these proportions did not change appreciably (or systematically) during the time period examined.

63. Formally speaking, of course, it is not a worker's wages as such that determine how much he or she has to lose from job loss, but the wage earned in the current job as compared to his or her next-best alternative. Our data do not permit us to quantify each worker's opportunity cost. However, the number of companies in this industrial sector that offer unionized jobs with comparable fringe benefits for relatively low-skilled jobs (such as those performed by the Company's employees) is declining. We believe that in practice, most higher-tenure and full-time workers in our sample would have trouble securing a comparable package of wages and benefits on the outside labor market.

64. For prior studies indicating that men are more likely to use illegal drugs, see footnote 23.

and Yamaguchi, 1987). In other words, if higher-tenure workers are less likely to use illegal drugs in the first place, it could counterbalance the fact that they (presumably) have more to lose from job loss.

At the same time, however, the positive and significant coefficients on the "post-implementation × full-time × higher-tenure" dummies in Specifications (b) suggest that among full-time workers, it is *lower tenure* workers that are the most responsive to PADT. Although our data do not permit us to fully explain this intriguing result, a number of possible hypotheses come to mind. Perhaps the prevalence of drug use is particularly low among workers who are *both* full-time and higher-tenure, for reasons unrelated to PADT testing. Another possibility is that full-time workers with *low* tenure fear that losing a job so soon after being hired would convey a negative stigma to future employers, and therefore are particularly strongly deterred from using drugs (or from reporting their injuries). This same group of workers—having undergone a job search more recently—may also face more severe credit constraints, and could be more cognizant of the increasing scarcity of full-time unionized jobs in this industrial sector.

The insignificance of the "post-PADT × older worker" term—except for first-aid claims, in which the coefficient is significantly negative—suggests that age as such generally has little effect, *ceteris paribus*, on workers' behavioral response to PADT. It is unclear why first-aid claims are the sole exception to this rule. Since US citizens over the age of 35 are generally less likely to use illegal drugs than their younger peers, we do not believe that this finding is plausibly explained by older workers' higher rates of drug use (Executive Office of the President, Office of National Drug Control Policy, 2008). If workers over the age of 35 are more likely to suffer from medical conditions that they wish to keep confidential, perhaps they are more reluctant to risk the potential disclosure of such information in the process of obtaining first-aid treatment for relatively minor injuries.

Viewed as a whole, then, the second stage of the empirical analysis suggests that different types of workers respond differently to PADT implementation. Although prior work has identified age as an important predictor of both drug use and occupational injuries, it explains little of the observed variation in our study. Rather, full-time status and sex emerge as the most important predictors of program responsiveness. Higher tenure also generally tends to enhance PADT responsiveness (although, rather surprisingly, the opposite relationship seems to hold for the minority of full-time workers). Overall, we interpret these results as supporting the hypothesis that it is workers who have the "most to lose" from job loss that are most likely to alter their behavior following PADT implementation. In this sense, the evidence seems broadly consistent with a "rational cheater" model of misbehavior first proposed by Shapiro and Stiglitz (1984).

The third stage of the empirical analysis pursues a new line of inquiry by exploring which factors are most predictive of a positive drug test result among those claimants who are drug tested as part of the PADT program. As described earlier, these models (four of which are presented in table 5) include a range of factors that conceivably could possibly affect the likelihood of a positive test. Interestingly, the *only* variable that is a significant and robust predictor of drug test outcomes is employee tenure: the coefficient is uniformly negative and statistically significant across all models specifications.[65] Viewed in isolation, this finding does *not* necessarily imply that the rate of drug use declines with tenure. Since we can only observe the probability of a positive test result *conditional on a test being administered*, it is possible that higher-tenure workers are simply more likely to underreport their injuries and/or take more caution on the job. In our view, it is likely that the *confluence* of several factors is driving this result: not only do higher-tenure workers have more to lose from job loss, but the positive correlation between drug use and employee turnover (documented by other scholars) also tends to reduce the "base rate" of drug use among higher-tenure employees. Since both of these effects work in the same direction, it is not surprising that tenure is the only robustly significant predictor of a negative test result.

In the fourth and final stage of the empirical analysis, we address the question of mechanisms: what, if anything, do the data suggest regarding the relative likelihoods that lower drug use and/or increased caution on one hand, or underreporting on the other, is driving the observed declines? In pursuing this line of inquiry, we started with two assumptions: (1) some employees—especially those who use illegal drugs—may decline to report their injuries for fear of losing their jobs and/or to avoid undergoing a drug test; and (2) the less serious an injury, and the more easily it can be

65. The coefficient on "tenure" is usually statistically significant at a 1 percent level. It occasionally becomes significant at a 5 percent level when "tenure squared" is added to the model, although "tenure squared" itself has no further predictive value.

either hidden or characterized as non-work-related, the greater a worker's incentives (and capacity) to avoid reporting it.

In light of these assumptions, we examine the extent to which programmatic responsiveness varies by three factors: (1) the relative severity of the claim (i.e., indemnity versus medical-only versus first aid); (2) whether the injury was a laceration, and thus particularly *difficult* to hide from management and co-workers; and (3) whether the injury was cumulative, and thus particularly *easy* to hide from management and co-workers. If the drop in claims following PADT implementation is unusually pronounced among minor, nonlaceration and/or cumulative injuries, one might plausibly view this as "circumstantial evidence" that some of the observed decline is due to underreporting.

The results presented in Figures 3–5 and in table 2, discussed earlier, seem broadly consistent with the underreporting hypothesis. Generally speaking, as injury severity increases, the magnitude, significance, and persistence of post-implementation effects tend to decline. Comparison between Panels A and B in 3–5 also suggests that cumulative injuries are especially responsive to PADT implementation.[66] In tables 6 and 7, we attempt to probe the available evidence in even greater depth, focusing in turn on lacerations and cumulative injuries.

Table 6 compares post-implementation effects for laceration versus nonlaceration injury types. As the theory would predict, nonlaceration claims generally tend to fall more markedly in the wake of PADT implementation than laceration claims. However, the disparity is only significant for first-aid reports, and the *opposite* relationship holds among medical-only claims. It could be that our assumption that lacerations are less "hideable" is incorrect, or that such disparities are trivial for most claim types. Overall, however, we interpret these results as providing, at best, only moderate support for the underreporting hypothesis.[67]

---

66. In theory, the disproportionate fall in minor and easily "hideable" accidents could be caused by some other unobservable factor unrelated to the likelihood of reporting. For example, if drug users are inherently more likely to sustain minor, and "hideable" injuries—and PADT encourages them to take more care on the job or kick their habit—this could cause a larger proportional decline in the specified injury types. Although we cannot rule out this possibility, we are aware of no data corroborating that drug use affects the likelihood of injury in this particular manner.

67. The insignificance of the laceration variable in table 5 also arguably undermines the existence of a reporting effect. If we are correct that lacerations are easier to hide—and

Table 6. The effects of post-accident drug testing and laceration versus nonlaceration status on claim frequency for a large retail chain, 2002–2007

Pooled, fixed-effects negative binomial models of claim frequency disaggregated by laceration versus nonlaceration status					
	Indemnity workers' compensation claims				
	(1) Total workers' compensation claims	**(2) Excluding cumulative injuries**	**(3) Including cumulative injuries**	**(4) Medical-only workers' compensation claims**	**(5) First-aid reports**
Post-implementation × laceration (I)	−0.0730	−0.0630	−0.122	−0.175**	0.176**
	(0.050)	(0.17)	(0.17)	(0.053)	(0.035)
Post-implementation × nonlaceration (II)	−0.133**	−0.0633	−0.174**	−0.147**	−0.312**
	(0.030)	(0.053)	(0.050)	(0.036)	(0.026)
Nonlaceration	1.596**	3.020**	3.150**	1.210**	1.348**
	(0.020)	(0.061)	(0.062)	(0.022)	(0.017)
Log of hours	1.125**	1.133**	1.074**	1.096**	0.862**
	(0.13)	(0.23)	(0.22)	(0.16)	(0.12)
Difference between laceration and nonlaceration claims (I–II)	0.0605 (0.050)	0.000379 (0.17)	0.0518 (0.17)	−0.0283 (0.053)	0.488** (0.035)
Observations	5,590	5,460	5,460	5,590	5720
Number of districts	43	42	42	43	44

Notes: This table presents the results of five negative binomial models of claim/report frequency, estimated on a dataset disaggregated by the laceration/nonlaceration status of the associated injury. Standard errors are presented in parentheses. Levels of significance are as follows: **1%, *5%, ^10%. The dependent variable in the above models is the number of injury claims/reports. The unit of observation is district × time period × laceration/nonlaceration status. All models include district-level fixed-effects and time-period dummies. The number of observations and districts fluctuates slightly across the models because one very small district reported no indemnity claims, and another very small district reported no indemnity claims or medical-only claims, during our entire sample period, and therefore claims from these districts were automatically dropped from the applicable models. All models were estimated on a cut of the data, which excluded data from facilities that closed before any treatment districts implemented the PADT program; included data from nonimplementing districts in the treatment divisions; and (with the exception of Model 3) excluded claims associated with cumulative injuries.

Besides the specifications presented above, we conducted the following additional robustness checks:

(1) We included claims and hours data associated with stores that closed before any of the treatment districts implemented the PADT program. This robustness check produced parameters exhibiting significance levels identical to those displayed above, with identical signs on statistically significant parameters.

(2) We excluded claims and hours data associated with the five nonimplementing districts within the treatment divisions. This robustness check produced parameters exhibiting significance levels identical to those displayed above, with identical signs on statistically significant parameters.

(3) We included claims associated with cumulative injuries in all models (besides Model 3). The resulting parameters exhibited significance levels identical to or more significant than those presented above, with identical signs on statistically significant parameters.

(4) We re-ran all models using facility × time period × laceration/nonlaceration status as the unit of observation. All estimated parameters exhibited significance levels identical to or more significant than those displayed above, with identical signs on statistically significant parameters.

Table 7. The effects of post-accident drug testing and cumulative injuries on claim frequency for a large retail chain, 2002–2007

Pooled, fixed-effects negative binomial models of claim frequency disaggregated by cumulative versus noncumulative status

	(1) Total workers' compensation claims	**(2) Indemnity workers' compensation claims**	**(3) Medical-only workers' compensation claims**	**(4) First-aid reports**
Post-implementation × cumulative (I)	−0.435**	−0.538**	−0.279**	−0.515**
	(0.079)	(0.12)	(0.099)	(0.10)
Post-implementation × noncumulative (II)	−0.141**	−0.123*	−0.159**	−0.205**
	(0.029)	(0.052)	(0.034)	(0.025)
Noncumulative	2.458**	2.016**	2.779**	3.557**
	(0.028)	(0.039)	(0.039)	(0.041)
Log of hours	1.057**	1.039**	1.025**	0.890**
	(0.13)	(0.22)	(0.15)	(0.12)
Difference between cumulative and noncumulative claims (I–II)	−0.294** (0.079)	−0.415** (0.12)	−0.119 (0.098)	−0.310** (0.10)
Observations	5,590	5,460	5,590	5,720
Number of districts	43	42	43	44

Notes: This table presents the results of four negative binomial models of claim/report frequency, estimated on a dataset disaggregated by the cumulative/noncumulative status of the associated injury. Standard errors are presented in parentheses. Levels of significance are as follows: **1%, *5%, ^10%. The dependent variable in the above models is the number of injury claims/reports. The unit of observation is district × time period × cumulative/noncumulative status. All models include district-level fixed-effects and time-period dummies. The number of observations and districts fluctuates slightly across the models because one very small district reported no indemnity claims, and another very small district reported no indemnity claims or medical-only claims, during our entire sample period, and therefore were automatically dropped from the applicable models. All models were estimated on a cut of the data, which excluded data from facilities that closed before any treatment districts implemented the PADT program; included data from nonimplementing districts in the treatment divisions; and included claims associated with cumulative injuries.

Besides the specifications presented above, we conducted the following additional robustness checks:

(1) We included claims and hours data associated with stores that closed before any of the treatment districts implemented the PADT program. This robustness check produced parameters exhibiting significance levels identical to those displayed above, with identical signs on statistically significant parameters.

(2) We excluded claims and hours data associated with the five nonimplementing districts within the treatment divisions. This robustness check produced parameters exhibiting significance levels identical to those displayed above, with identical signs on statistically significant parameters.

(3) We reran all models using facility × time period × cumulative/noncumulative status as the unit of observation. All estimated parameters exhibited significance levels identical to or more significant than those displayed above, with identical signs on statistically significant parameters.

Finally, table 7 explores the responsiveness of cumulative injury claims to PADT implementation. As discussed earlier, cumulative injuries are the most “hideable” of all injuries, although they are also technically exempt from PADT testing. The coefficients presented in table 7 should therefore be construed as the *net effect* of these countervailing factors on employee behavior. Somewhat ironically—given the technical exclusion of cumulative injuries from the PADT program—these results are *more* suggestive of underreporting than those focusing on lacerations. Across all claim categories, cumulative injuries fall more than noncumulative injuries, and the differential is statistically significant at a 1 percent level in all of the models (with the exception of medical-only claims).

Although our results are far from conclusive, this final stage of the analysis suggests that some of the observed effect of PADT may well be due to underreporting. Of course, our evidence is at best “circumstantial.” Since our data do not enable us to pin down causation with certainty, factors besides underreporting could explain the pattern of disparities that we observe. It is possible, in fact, that the declines we observe are driven by some combination of lower drug use, greater care on the job, underreporting, and even (conceivably) greater investment in commercial drug-testing aids. Taken as a whole, however, we construe our findings as lending some (albeit qualified) support to the view that PADT may encourage the underreporting of relatively minor and easy-to-hide injuries.

5. Conclusion

The scarcity of empirical scholarship on the real-world effects of workplace drug testing is surprising in light of its salience to state policymakers. The goal of this study is to contribute to the literature on this issue by examining the effect of PADT on accident claims in a large Fortune 100 company.

Our empirical methodology was designed to explore four interrelated issues: whether the frequency of claims has fallen significantly in the wake of PADT implementation; which worker characteristics are associated with responsiveness to the policy; which factors (if any) are predictive of a positive

---

if underreporting is widespread—then one would expect lacerations, *ceteris paribus*, to be associated with a higher rate of positive drug tests.

drug test result; and whether some portion of the observed drop in claims may be attributable to underreporting. Although we cannot answer these questions definitively, our empirical results do shed important light on each.

First, we find substantial evidence that the Company's employees did, in fact, alter their behavior in response to the PADT program. Total workers' compensation claims fell significantly in the wake of PADT implementation, although the effect was stronger and more persistent for less severe claims (those involving medical payments but no compensation for lost work time). The frequency of first-aid reports involving minor injuries fell even more dramatically and persistently in the wake of the program. We performed two robustness checks to test the likelihood that the apparent decline in workers' compensation claims and first-aid reports following PADT implementation is a statistical artifact resulting from model misspecification. Both of these exercises appeared to confirm the validity and robustness of our findings.

The second stage of the analysis explores the importance of demographic factors and job characteristics in explaining employees' patterns of behavioral response. We find that the frequency of claims falls most dramatically among full-time workers, apparently bearing out our hypothesis that workers who value their jobs most highly respond the most to the policy. We also find that male workers are more responsive to the policy than their female co-workers, a disparity that we believe is best explained by the higher prevalence of drug use among male employees.

In the third stage of our analysis, which examines the results of PADT drug tests, we find a negative correlation between employee tenure and the likelihood of testing positive for illegal drug use. No other employee attributes (or injury characteristics) significantly predict drug testing outcomes. We suggest that the unique significance of tenure is best explained by two reinforcing factors: the relatively high cost of job loss for higher-tenure workers, and the fact that higher turnover rates among drug users (found in other studies) will tend to lower the "base rate" of drug use among higher-tenure workers.

The fourth and final stage of the analysis, which probes the underlying question of mechanisms, provides some tentative "circumstantial evidence" of a reporting effect. If employees are induced by the policy to reduce their drug consumption and/or take greater care—and if these behavioral changes, in turn, make them less accident-prone in performing all occupational tasks—one might expect to see across-the-board declines in injuries.

The fact that by and large, the declines that we document are most pronounced among the least severe and most "hideable" injury categories, seems at least consistent with the hypothesis that some portion of the observed decline is due to underreporting.[68]

The fact that claims and reports fell significantly after PADT implementation—notwithstanding the existence of two other drug-testing policies at the Company—has important practical implications. At least in the short term, the drop in claims may reduce the Company's direct workers' compensation costs.

In our data, for example, the observed reductions in workers' compensation claims translated roughly into a direct per-period cost savings, across all treatment districts, of $16,938 for medical-only claims, and $76,552 for indemnity claims.[69] In comparison, the direct cost of running the program in all districts (given sample collection and lab analysis costs of about $20

---

68. To the extent that our results do reflect an underreporting effect, the coefficient estimates are likely to represent a lower bound on its magnitude. Of all forms of drug testing, the "cheek swab" method used by the Company is likely to be perceived as the least objectionable—far less intrusive, for example, than the collection of a urine or blood sample. Therefore, it would be surprising if a large number of employees—other than those who feared the consequences of a positive test result—were deterred from reporting accidents merely because of the psychic costs of undergoing the test.

69. To calculate these cost savings, we used baseline per-district/per-period frequencies of 2.62 and 7.32 for indemnity and medical-only claims, the mean frequencies (respectively) of claims during the four periods prior to implementation in the treatment districts. (Calculating the means over three, five, and six periods prior to implementation, respectively, yielded similar frequencies.) Since there were seventeen treatment districts, this resulted in an average number of total claims of 44.5 indemnity claims and 124.5 medical-only claims during each period prior to implementation. For "medical only" claims, we estimate a 15.38 percent reduction in the number of claims following implementation (using the coefficient on the "post-implementation" dummy in a model of medical-only workers' compensation claims similar to that presented in column 4(a) of table 2 but including cumulative injuries). Given these assumptions, "medical only" claims fell, on average, by 19.15 claims per period across all treatment districts after implementation. Since the mean cost of a "medical only" claim is $884.50 (calculated over the entire post-implementation period in the treatment districts), the estimated per-period cost reduction from reduced "medical only" claims for treatment districts was calculated to be about $16,938. Meanwhile, we assumed a 15.72 percent reduction in indemnity claims after PADT implementation (calculated from column 3(a) in table 2, which includes cumulative injuries), translating into about 7.0 fewer claims per period. Using the mean cost figure of $10,936 per indemnity claim (encompassing both medical and indemnity payments), the estimated per-period cost reduction in indemnity claims across all treatment districts was calculated to be about $76,552.

per claim) is only about $7000 per period. Since the policy had only been implemented in a small fraction of the company's facilities nationwide at the time of this study, the potential reduction in workers' compensation costs could be considerably larger.

If a substantial portion of the observed drop in claims is driven by underreporting, however, PADT's net effect on the Company and its employees is less clear. Not only may the administration of the PADT program itself be costly to the Company, but unreported workplace hazards could fester and, over the long term, impose even higher costs. Meanwhile, PADT may make accident reporting so costly for some workers that they opt to pay for medical care out-of-pocket or simply endure injuries that would otherwise be treatable through workers' compensation. If many workers are covered by health insurance plans—particularly if they are covered on a family member's plan—the costs of treatment could be shifted from the Company onto other benefits providers.

At the very least, however, our study suggests that PADT programs raise an important set of policy concerns that other forms of occupational drug testing do not. The marked decline in work-related injuries and workers' compensation claims since the early 1990s has puzzled policymakers for well over a decade (Conway and Svenson, 1998). This apparently secular decline seems surprising in light of the previously assumed pro-cyclicality of injury rates.[70] Although some scholars have proposed an increase in underreporting as a possible explanation, few studies have attempted to rigorously assess the likelihood of such effects.[71] In the workers' compensation arena, not only has the frequency of claims continued to fall since the year 2000, but researchers have observed that the drop has been the most pronounced for smaller claims (Robinson *et al.*, 2005). It is an open question whether the fall in claims might stem, at least in part, from underreporting. Our study suggests that examining a possible causal relationship between the decline

70. Conway and Svenson (1998) (citing Peter Dorman, *Markets and Mortality*).

71. OSHA conducted a study in 1987, and again in 1996, designed to determine whether underreporting had increased during the intervening decade. The study concluded that although underreporting on OSHA "200 logs"—which the BLS used to calculate its injury statistics—was prevalent, it had not increased during the preceding decade (Conway and Svenson, 1998). However, to the best of our knowledge, no similar follow-up study of BLS data has been conducted since 1996. Moreover, we know of no study that has focused on the likelihood of underreporting of workers' compensation claims.

in workers' compensation claims and the roughly contemporaneous proliferation of PADT, as well as other programs that may increase the costs of claiming, could be a profitable area for future research.

References

Ackerman, D. L. 1991. "A History of Drug Testing" in Robert H. Coombs and Louis J. West, eds., *Drug Testing: Issues and Options*. New York: Oxford University Press.

American Civil Liberties Union. 1999. "Drug Testing: A Bad Investment," ACLU mimeo 8.

Barkume, Anthony J., and John W. Ruser. 2001. "Deregulating Property-Casualty Insurance Pricing: The Case of Workers' Compensation," 44 *Journal of Law & Economics* 37–63.

Barnum, Darold T., and John M. Gleason. 1994. "The Credibility of Drug Tests: A Multi-Stage Bayesian Analysis," 4 *Industrial and Labor Relations Review* 610–21.

Becker, Gary S., and Kevin M. Murphy. 1988. "A Theory of Rational Addition," 96 *Journal of Political Economy* 675–800.

Bennett, Nathan, Terry C. Blum, and Paul M. Roman. 1994. "Employee Attitudes Toward Drug Testing: Effects of Individual Characteristics and Employment Setting," 7 *Employee Responsibilities and Rights Journal* 117–28.

Bensinger, Peter B. 1988. "Drug Testing in the Workplace," 498 *Annals of the American Academy of Political and Social Science* 43–50.

Betts, Kenneth C. 1990. "Fourth Amendment—Suspicionless Urinalysis Testing: A Constitutionally 'Reasonable' Weapon in the Nation's War on Drugs?" 80 *Journal of Criminal Law & Criminology* 1018–51.

Blank, David L., and John W. Fenton. 1989. "Early Employment Testing for Marijuana: Demographic and Employee Retention Patterns" in Steven W. Gust and J. Michael Walsh, eds., *Drugs in the Workplace: Research and Evaluation Data*. Rockville, MD: National Institute on Drug Abuse.

Boden, Leslie I., and John W. Ruser. 2003. "Workers' Compensation 'Reforms,' Choice of Medical Care Provider, and Reported Workplace Injuries." 85(4) *Review of Economics and Statistics* 923–929.

Bureau of Labor Statistics. 1990–2008 (annual publication). "State Labor Legislation Enacted in ___." *Monthly Labor Review*. Washington, DC: US Department of Labor, Bureau of Labor Statistics.

Butler, R. J., and J. D. Worrall. 1983. "Workers' Compensation: Benefit and Injury Claims Rates in the Seventies," 65 *Review of Economics and Statistics* 580–99.

Butler, R. J., and J. D. Worrall. 1985. "Work Injury Compensation and the Duration of Nonwork Spells," 95 *Economic Journal* 714–24.

Card, David, and Brian P. McCall. 1996. “Is Workers’ Compensation Covering Uninsured Medical Costs? Evidence from the ‘Monday Effect’,” 49 *Industrial & Labor Relations Review* 690–706.

Chait, L. D., M. W. Fischman, and C. R. Schuster. 1985. “‘Hangover’ Effects the Morning after Marijuana Smoking,” *Drug and Alcohol Dependence* 229–38.

Comer, Debra R. 1994. “A Case against Drug Testing,” 5 *Organizational Science* 259–67.

Conway, Hugh, and Jens Svenson. 1998. “Occupational Injury and Illness Rates, 1992–1996: Why They Fell,” 121 *Monthly Labor Review* 36–58.

Coombs, Robert H., and Louis Jolyon West, eds. 1991. *Drug Testing: Issues and Options*. New York: Oxford University Press.

Crant, J. Michael, and Thomas S. Bateman. 1989. “A Model of Employee Responses to Drug-Testing Programs,” 2 *Employee Responsibilities and Rights Journal* 173–90.

Crouch, Dennis J., Douglas O. Webb, Lynn V. Peterson, Paul F. Buller, and Douglas E. Rollins. 1989. “A Critical Evaluation of the Utah Power and Light Company’s Substance Abuse Management Program: Absenteeism, Accidents and Costs” in Steven W. Gust and J. Michael Walsh, eds., *Drugs in the Workplace: Research and Evaluation Data*. Rockville, MD: National Institute on Drug Abuse.

De Bernardo, Mark A., and Gina M. Pedro. 2006. *2004–2005 Guide to State and Federal Drug-Testing Laws*. Washington, DC: Institute for a Drug-Free Workplace.

Donohue, John J., and D. E. Ho. 2007. “The Impact of Damage Caps on Malpractice Claims: Randomization Inference with Differences-in-Differences,” 4 *Journal of Empirical Legal Studies* 69–102.

Ehrenberg, Ronald G. 1989. “Workers’ Compensation, Wages and the Risk of Injury” in John F. Burton, Jr., ed., *New Perspectives on Workers’ Compensation*. Ithaca, NY: Cornell University Press.

Elmuti, Dean. 1993. “Effects of Drug-Testing Programme on Employee Attitude, Productivity and Attendance Behavior,” 14 *International Journal of Manpower* 58–69.

Elster, John, and Ole-Jorgen Skog, eds. 1999. *Getting Hooked: Rationality and Addiction.* Cambridge, UK: Cambridge University Press.

Executive Office of the President, Office of National Drug Control Policies. 2006. *Drug Use Trends: October 2002*. http://www.whitehousedrugpolicy.gov/publications/factsht/druguse/ (accessed March 16, 2008).

Feinauer, Dale, and Stephen Havlovic. 1993. “Drug Testing as a Strategy to Reduce Occupational Accidents: A Longitudinal Analysis,” 24 *Journal of Safety Research* 1–7.

French, Michael T., M. Christopher Roebuck, and Pierre Kebreau Alexandre. 2001. “Illicit Drug Use, Employment, and Labor Force Participation,” 68 *Southern Economic Journal* 349–68.

Fricker, Mary. 1997. "Widespread Fraud: Bogus Claim," *Press Democrat Online*. http://www.pressdemo.com/workerscomp/day1/fraud2.html (accessed August 16, 2006).

Frone, Michael R. 2006. "Prevalence and Distribution of Illicit Drug Use in the Workforce and in the Workplace: Findings and Implications from a US National Survey," 91 *Journal of Applied Psychology* 856–69.

Gerber, Jonathan, and George S. Yacoubian. Jr. 2001. "Evaluation of Drug Testing in the Workplace: Study of the Construction Industry," 127 *Journal of Construction Engineering and Management* 438–44.

Gust, Steven W., and J. Michael Walsh, eds. 1989. *Drugs in the Workplace: Research and Evaluation Data*. Rockville, MD: National Institute on Drug Abuse.

Hartwell, T., P. Steel, M. French, and N. Rodman. 1996. "Prevalence of Drug Testing in the Workplace," 111(11) *Monthly Labor Review* 35–42.

Herning, Ronald I., Barbara J. Glover, Bonnie S. Koeppl, and Jerome H. Jaffe. 1989. "Cocaine and Workplace Performance: Inferences from Clinical Studies," Paper presented at *NIDA Conference on Drugs in the Workplace*, Bethesda, MD.

Heyman, Gene. 1996. "Resolving the Contradiction of Addiction," 4 *Behavior and Brain Sciences* 561–610.

Hoffman, John, and Cindy Larison. 1999. "Drug Use, Workplace Accidents and Employee Turnover," 29 *Journal of Drug Issues* 341–64.

Holcom, Melvin L., Wayne E. K. Lehman, and D. Dwayne Simpson. 1993. "Employee Accidents: Influences of Personal Characteristics: Job Characteristics, and Substance Use in Jobs Differing in Accident Potential," 24 *Journal Safety Research* 205–21.

Jacobson, Mirielle. 2003. "Drug Testing in the Trucking Industry: The Effect on Highway Safety," 46 *Journal of Law Economics* 131–56.

Jobs, Sarah M., Fred E. Fiedler, and Chad T. Lewis. 1990. "Impact of Moderate Alcohol Consumption on Business Decision Making" in Steven W. Gust and J. Michael Walsh, eds., *Drugs in the Workplace: Research and Evaluation Data,* vol. II. Rockville, MD: National Institute on Drug Abuse.

Joseph, Paul R. 1987. "Fourth Amendment Implications of Public Sector Work Place," 11 *Nova Law Review* 605–45.

Kaestner, Robert. 1991. "The Effect of Illicit Drug Use on the Wages of Young Adults," 9 *Journal of Labor Economics* 381–412.

Kaestner, Robert. 1994. "New Estimates of the Effect of Marijuana and Cocaine Use on Wages," 47 *Industrial and Labor Relations Review* 454–70.

Kaestner, Robert, and Michael Grossman. 1998. "The Effects of Drug Use on Workplace Accidents," 5 *Labour Economics* 267–94.

Kandel, Denise, and Yazuo Yamaguchi. 1987. "Job Mobility and Drug Use: An Event History Analysis," 92 *American Journal of Sociology* 836–78.

Kelly, Thomas H., Richard W. Foltin, and Marian W. Fischman. 1990. "Effects of Alcohol on Human Behavior: Implications for Workplace Performance" in Steven W. Gust and J. Michael Walsh, eds., *Drugs in the Workplace: Research and Evaluation Data*, vol. II. Rockville, MD: National Institute on Drug Abuse.

Kesselring, Robert G., and Jeffrey R. Pittman. 2002. "Drug Testing Laws and Employment Injuries," 23 *Journal of Labor Research* 294–301.

Kopstein, Andrea, and Joseph Gfroere. 1990. "Drug Use Patterns and Demographics of Employed Drug Users: Data from the 1988 National Household Survey on Drug Abuse," in Steven W. Gust and J. Michael Walsh, eds., *Drugs in the Workplace: Research and Evaluation Data*, vol. II. Rockville, MD: National Institute on Drug Abuse.

Krueger, Alan B. 1990. "Workers' Compensation Insurance and the Duration of Workplace Injuries," NBER Working Paper No. 3253.

Krueger, Alan B., and John F. BurtonJr. 1990. "The Employers' Costs of Workers' Compensation Insurance: Magnitudes, Determinants, and Public Policy," 72 *Review of Economics and Statistics* 228–241.

Leigh, J. Paul, Steven B. Markowitz, Marianne Fahs, and Philip J. Landrigan. 2000. *Cost of Occupation Injuries and Illnesses*. Ann Arbor, MI: University of Michigan Press.

McDaniel, Michael A. 1989. "Does Pre-Employment Drug Use Predict on-the-Job Suitability?" in Steven W. Gust and J. Michael Walsh, eds., *Drugs in the Workplace: Research and Evaluation Data*. Rockville, MD: National Institute on Drug Abuse.

MacDonald, Scott, and Samantha Wells. 1994. "The Impact and Effectiveness of Drug-Testing Programs in the Workplace" in Scott MacDonald and Paul Roman, eds., *Research Advances in Alcohol and Drug Problems Volume II: Drug Testing in the Workplace*. New York: Plenum Press.

McManis, Sam. 1999. "Drug Testing Takes a Hit." *San Francisco Chronicle*.

Mehay, Stephen L., and Rosalie Liccardo Pacula. 1999. The Effectiveness of Workplace Drug Prevention Policies: Does 'Zero Tolerance' Work? NBER Working Paper 7383.

Mell, Christina Louise. 1987. "Employee Drug Testing: Guilty until Proven Innocent?" 52 *Monthly Law Review* 625–45.

Meyer, Bruce D., W. Kip Viscusi, and David L. Durbin. 1995. "Workers' Compensation and Injury Duration: Evidence from a Natural Experiment," 85 *American Economic Review* 322–40.

Morrow, Tod T. 1989. "Drug Testing in the Workplace: Issues for the Arbitrator." 4 *Ohio State Journal on Dispute Resolution* 273–94.

Murray, J. D. "Marijuana's Effects on Human Cognitive Functions, Psychomotor Functions, and Personality," 113 *Journal of General Psychology* 23–55.

National Highway Traffic Safety Administration. 2003. "Physician's Guide to Assessing and Counseling Older Drivers." http://www.nhtsa.dot.gov/People/injury/olddrive/OlderDriversBook/pages/Ch9-Section5.html (accessed April 6, 2006).

National Institute of Occupational Safety and Health. 2004. *Worker Health Chartbook*. NIOSH Publican No. 2004–146. http://www.cdc.gov/niosh/docs/chartbook (accessed August 7, 2008).

Normand, Jacques, S. D. Salyards, and J. J. Mahoney. 1990. "An Evaluation of Pre-Employment Drug Testing," 75 *Journal of Applied Psychology* 629–39.

Normand, Jacques, Richard O. Lempert, and Charles P. O'Brien, eds. 1994. *Under the Influence? Drugs and the American Work Force*. Washington, DC: National Academy Press.

Robinson, John, Tom Sheppard, and Sandra Starnes. 2005. "Workers' Compensation Claim Frequency Continues to Decline, Particularly for Smaller Claims," *National Council on Compensation Insurance Research Brief.* http://www.ncci.com/media/pdf/claims-researchbrief2.pdf (accessed April 6, 2006).

Ruser, John W. 1991. "Workers' Compensation and Occupational Injuries and Illnesses," 9 *Journal of Labor Economics* 325–50.

Shapiro, Carl, and Joseph P. Stiglitz. 1984. "Equilibrium Unemployment as a Worker Discipline Device," 74 *American Economic Review* 433–44.

Sheridan, John, and Howard Winkler. 1989. "An Evaluation of Drug Testing in the Workplace" in Steven W. Gust and J. Michael Walsh, eds., *Drugs in the Workplace: Research and Evaluation Data*. Rockville, MD: National Institute on Drug Abuse.

Shulman, David, and Alfred E. Hofflander. 1980. "Cumulative Trauma in the Context of Workers' Compensation," 1 *Journal of Occupational Behavior* 119–27.

State ex rel. Ohio AFL-CIO v. Ohio Bur. of Workers' Comp., 97 Ohio St.3d 504, 2002-Ohio-6717.

Taggart, Rober W. 1989. "Results of the Drug Testing Program at Southern Pacific Railroad" in Steven W. Gust and J. Michael Walsh, eds., *Drugs in the Workplace: Research and Evaluation Data*. Rockville, MD: National Institute on Drug Abuse.

Texas Department of Insurance. 1998. "Fraud in the Texas Workers' Compensation System." http://www.tdi.state.tx.us/wc/regulation/roc/fraud.html (accessed August 16, 2006).

Thompson, Joel A. 1981. "Outputs and Outcomes of State Workmen's Compensation Laws," 43 *Journal of Politics* 1129–52.

Thomason, Terry, Timothy P. Schmidle, and John F. Burton. 2001. *Workers' Compensation: Benefits, Costs, and Safety under Alternative Insurance Arrangements*. Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Todd, Steven O'Neal. 1987. "Employee Drug Testing—Issues Facing Private Sector Employers," 65 *North Carolina Law Review* 832–47.

Viscusi, W. K., and Michael J. Moore. 1987. "Workers' Compensation: Wage Effects, Benefit Inadequacies, and the Value of Health Losses." 69 *Review of Economics and Statistics* 249–61.

Walstatter, Matthew S. 2001. "Pissed Off: Drug Tests Conducted Under the Workers' Compensation and Unemployment Systems Are Government Searches Subject to Constitutional Scrutiny," 5 *Journal of Small and Emerging Business Law* 161–82.

Westphal, Edward E. 1987. "Public-Sector Employer Drug-Testing Programs: Has Big Brother Finally Arrived?" 20 *John Marshall Law Review* 769–93.

White, Tony. 2003. "Drug Testing at Work: Issues and Perspectives," 38 *Substance Use and Misuse* 1891–1902.

Yesavage, J. A., V. O. Lehrer, M. Denari, and L. E. Hollister. 1985. "Carry-Over Effects of Marijuana Intoxication on Aircraft Pilot Performance: A Preliminary Report," 142 *American Journal of Psychiatry* 1325–9.

Zwerling, C., J. Ryan, and E. J. Orav. 1990. "The Efficacy of Preemployment Drug Screening for Marijuana and Cocaine in Predicting Employment Outcomes," 264 *Journal of the American Medical Association* 2639–43.