WORKERS COMPENSATION AND ECONOMIC CYCLES: A LONGITUDINAL APPROACH

ROBERT P. HARTWIG

RONALD C. RETTERATH

TANYA E. RESTREPO

AND

WILLIAM J. KAHLEY

Abstract

This paper uses cross-sectional time series techniques in an econometric framework to model workers compensation frequency, severity, and loss ratios over the course of the business cycle. Empirical evidence from 37 states over the 1979–1993 period strongly suggests that frequency is strongly pro-cyclical, tending to increase during periods of economic expansion and fall during periods of economic decline or sluggishness. Similarly, the analyses reported in this study indicate that the economic determinants of indemnity and medical severity and loss ratios can be characterized by large pro-cyclical and small counter-cyclical components. The latter finding is contrary to conventionally held beliefs concerning this topic.

At the time this study was performed, the authors were employed by the National Council on Compensation Insurance: Robert Hartwig as Senior Economist; Ronald Retterath as Senior Vice President and Chief Actuary; Tanya Restrepo as Economist; and William Kahley as Assistant Vice President for Economic Research.

1. INTRODUCTION

The explosion in workers compensation costs during the 1980s and early 1990s continues to be the subject of much debate among actuaries, economists, and regulators even though the crisis began to show signs of abating in 1993 and has continued to improve since then. This ongoing dialogue has been constructive inasmuch as it has identified many of the cost drivers responsible for the tumult in the industry over the past fifteen years.

Much has been written about some important cost drivers—such as medical inflation, medical cost shifting, attorney involvement, and fraud—from many angles, including premium avoidance, employee malingering, and medical/legal workers compensation mills. The popular media have even sensationalized some fraud-related activities in print and television.

In contrast, important relationships between workers compensation costs and changes in the economic environment generally have been ignored or discussed only anecdotally in the literature. The objective of this longitudinal study is to demonstrate empirically that workers compensation costs across states are fundamentally dependent on the economic environment. Specifically, empirical evidence suggests that frequency is strongly pro-cyclical, tending to increase during periods of economic expansion and fall during periods of economic decline or sluggishness. Similarly, the analyses reported in this study indicate that the economic determinants of indemnity and medical severity and loss ratios can be characterized by large pro-cyclical and small counter-cyclical components. The latter finding is contrary to conventionally held beliefs concerning this topic.

The remainder of this paper is structured as follows. Section 1 concludes with some background information and a brief literature review on the economic factors affecting workers compensation results. Section 2 contains definitions of the economic variables used in the study and their hypothesized impact on the workers compensation market. Section 3 describes the modeling methodology and defines the dependent variables. In Section 4, empirical results from analyses of the impact of business cycle effects on workers compensation frequency, severity, and loss ratios are presented. Section 5 concludes the paper with a summary of results presented in this study and directions for future research.

The term "business cycle," as it is typically used by business analysts, describes the periodic, but irregular, ups and downs of the real economy over time.[1] The cumulative processes at work during business cycles ensure that at some point every industry and virtually every firm will be either directly or indirectly affected. The workers compensation insurance industry is no exception. In almost all states, workers compensation insurance (or its equivalent) is compulsory, and policies are purchased by businesses in all industries. The fact that workers compensation premiums change with companies' payrolls ensures that fluctuations in economic activity will have direct impacts on the workers compensation industry. Losses, of course, will be affected by shifting levels of exposure during the different phases of the cycle, and also by the changing claim filing incentives facing workers in a dynamic economic environment.

Much of the economic literature in workers compensation insurance has focused on the factors that motivate individual workers to use the workers compensation system, and contrasted employee incentives with the incentives employers have to maintain a safer workplace to hold down workers compensation costs. The preponderance of evidence indicates that higher benefits, particularly indemnity benefits, precipitate increased claim filing by workers. Other research suggests that during an expansion, em-

[1] It is important to distinguish between the meaning of the term "cycle," as it is frequently applied in insurance, versus the economic meaning intended here. The insurance usage of the term "cycle" generally refers to the so-called "underwriting cycle," where "hard markets," characterized by high prices and profits and limited availability of coverage, are followed by "soft markets," when prices and profits are low and the availability of coverage increases.

ployers' incentives to increase revenue may outweigh incentives to contain increases in workers compensation costs through investments in safety. The influence of economic incentives on workers and employers is explored in Moore and Viscusi [20], Butler and Appel [6], and Worrall and Butler [30, 31].[2]

Research into the costs and determinants of workers compensation medical expenditures has burgeoned in recent years in response to deteriorating market conditions through much of the 1980s and early 1990s. The Clinton administration's effort in 1993 and 1994 to enact national health care reform also provided a forum and impetus for such research. Some research findings suggested that price discrimination by medical care providers was at the root of workers compensation medical cost inflation [2]. Others have argued that higher costs for workers compensation cases (relative to non-occupational injuries) are due to the different mix and intensity of treatments necessary to hasten return to work [11]. The range and effectiveness of various medical cost containment strategies such as fee schedules, provider choice, bill/utilization review, and anti-fraud initiatives have also been analyzed.

The potential influence of the macroeconomy on workers compensation frequency and severity was recognized in a 1991 study performed by the Insurance Services Office (ISO). Using policy year data for thirty states, the ISO study found that higher interest rates and real growth in gross national product (GNP) are associated with higher claim frequency and that medical claim severity increased faster than an ISO-modified consumer price index [17].[3] Similarly, a 1996 National Council on Compensation Insurance (NCCI) study found a strong, positive association

[2]A more detailed literature survey of the relationship between workers compensation frequency, severity, and benefit structures is found in Butler [5]. One early effort to model workers compensation losses and premiums econometrically is given in Lommele and Sturgis [19].

[3]The results of the ISO study were derived using data for just seven policy years (1980–1986). In contrast, the sample period used in this study runs from 1979 to 1993. The ISO study therefore fails to include observations over the full course of a business cycle.

between workers compensation claim frequency, total claim costs per worker, and real growth in gross domestic product (GDP) over the 1980 to 1995 period [14].

Apart from the ISO and NCCI analyses, the implications of studies dealing with economic factors generally are confined to the microeconomic influences on claim frequency and severity. In other words, they focus on the incentives facing individual workers, employers, medical service providers, and insurers. Current research, reported here, complements these efforts through empirical analysis of the important *macroeconomic* factors that can affect frequency, severity, and loss ratios in the workers compensation line. For this reason, the economic explanatory variables used in the present analysis are confined to overall state- and national-level measures of economic performance.

Workers Compensation and the Economic Environment

The complex relationship between the workers compensation industry (measured in this study by frequency, severity, and loss ratios) and the economic environment can be crudely decomposed into the several broad categories listed below. Each of the first three categories is based on our prior expectations of rising frequency, severity, and loss ratios during the expansionary phase of the business cycle. Other forces, as suggested in the final category, may work in the opposite direction. Empirical evidence, however, suggests that these countervailing effects are relatively small. The various effects are discussed briefly here, and then again in more detail in subsequent sections of this study as they specifically apply to the empirical results.

- Workers compensation is influenced by both the level and rate of growth of production. Higher levels and/or growth rates of production tend to be associated with an increase in losses that is in excess of wage increases.
- Periods of rising capacity utilization are associated with adverse impacts on workers compensation. This is in part due

to pressure to increase the speed and volume of production, which may lead to a decreased emphasis on workplace safety. Moreover, worker fatigue due to increased overtime also contributes to a general worsening in the line, as does the reuse of older, less efficient, machinery.

- Changes in the composition of employment also play an important role in workers compensation insurance. During the initial phases of economic expansions, for example, employment in many hazardous industries, such as construction, manufacturing, and trucking, expands rapidly.
- The incentives facing workers and employers shift over the course of the business cycle. Workers, for example, are less likely to file workers compensation claims during an economic expansion when the opportunity cost of being out of work is relatively high.[4]

It is important to recognize that the four broad relationships discussed in this section are themselves influenced by innumerable other socio-economic, demographic, and regulatory factors. The result is that there can be a great deal of variability between the strength and speed with which economic influences are transmitted to the workers compensation insurance line over time and across states. Nevertheless, certain important fundamental relationships that are the subject of this study are useful when discussing the impact of economic conditions on workers compensation frequency, severity, and loss ratios.

2. BUSINESS CYCLE EFFECTS ON WORKERS COMPENSATION

The econometric models used to determine business cycle effects on the workers compensation market typically include a measure of overall economic activity (employment), a measure

[4]The concept of *opportunity cost* is used by economists to denote the foregone value of the next best alternative which is *not* chosen. Hence, it is more costly for a worker to file a workers compensation claim during an economic expansion when overtime work is more available and wage increases are greater due to tightening labor market conditions.

of the health of the labor market (the unemployment rate), and possibly a measure of crisis labor market conditions (business failures). The waiting period and a measure of cost containment are also included in some cases to control for non-business cycle effects specifically related to workers compensation. The medical severity and medical loss ratio models include hospital cost per stay as a measure of medical costs. Sources and definitions of the explanatory variables are listed in Table 1.

The hypothesized effects of the explanatory variables on the workers compensation market are discussed below. In some cases, such as with the unemployment rate, there are countervailing effects, and the dominant impact is determined empirically. The expected sign of some explanatory variables may change depending on the model's dependent variable: either frequency, severity, or the loss ratio.

The employment variable is included in the models as a broad-based measure of economic activity within the state. Gross state product (GSP) is the broadest measure of economic activity at the state level. However, federal statistics on GSP from the Bureau of Economic Analysis are available only with a significant lag (through 1991 at the time of this analysis). Thus, employment is used as a proxy for GSP or the measure of production.[5] The expected sign on the estimated coefficient of the employment variable is positive. In other words, economic activity (as proxied by employment growth) and frequency, severity, and the loss ratio are expected to move in the same direction. Such an expectation is consistent with both previous research and economic theory. For example, during an expansion when employment rises, overtime increases, less-experienced workers are hired, and employment in hazardous industries increases. All these factors could be expected to lead to increased frequency, severity, and loss ratios.

[5]Historically, the correlation between GSP and employment has been extremely high—roughly 98 percent.

TABLE 1

EXPLANATORY VARIABLE DEFINITIONS AND SOURCES

Variable Name	Definition	Source
NAGEMP	Annual average nonagricultural employment measured by persons on establishment payrolls. It excludes proprietors, the self-employed, unpaid volunteer or family workers, farm workers, and domestic workers. Persons who worked in more than one establishment during the reporting period are counted each time their names appear on payrolls [24].	U.S. Department of Labor, Bureau of Labor Statistics, Current Employment Statistics, Survey of Establishments
UNRATE	Annual average unemployment rate measured by the number of unemployed persons as a percent of the labor force. Unemployed persons are those who had no employment during the reference week of the household survey, were available for work, and had made specific efforts to find employment some time during the 4-week period ending with the reference week. The labor force includes all unemployed persons and employed persons. In the household survey, a person is considered employed if they did any work at all (at least 1 hour) as a paid employee, worked in their own business, profession, or on their own farm, or worked 15 hours or more as unpaid workers in an enterprise operated by a member of the family. Each person employed is counted only once, even if he or she holds more than one job [24].	U.S. Department of Labor, Bureau of Labor Statistics, Current Population Survey of households

BUSFAIL	Total business failures measured as businesses that ceased operations and were involved in court proceedings or voluntary actions involving losses to creditors. This does not include business discontinuances which are defined as businesses that cease operations for reasons such as loss of capital, inadequate profits, ill health, retirement, etc., if creditors are paid in full. Although business failures represent only a percentage of total closings, they have the most severe impact upon the economy [10].	The Dun & Bradstreet Corporation, *Business Failure Record*
WAITPER	State-mandated waiting period defined as the time that must elapse during which income benefits are not payable [23].	U.S. Chamber of Commerce, *Workers Compensation Laws*
COSTCON	Dummy variable proxy for cost containment initiatives, equal to unity for years 1991 through 1993 in all states and zero otherwise. See Appendix A for a discussion of dummy variables.	NCCI
PERSTAY	The product of the average daily hospital charge and the average duration of a hospital stay. The average daily charge is expenses incurred for inpatient care divided by inpatient days. Average length of stay is the average stay of inpatients derived by dividing the number of inpatient days by the number of admissions [1].	American Hospital Association, *Hospital Statistics*

The unemployment rate is included in the models as a gauge of the health or tightness of the state's labor market. While there is some small degree of correlation between employment and the unemployment rate, the latter is a better indicator of general labor market conditions.[6] Employment, on the other hand, is an indicator or proxy of the health of the overall economy. For example, during the early stages of an economic recovery, the unemployment rate often increases despite strong gains in employment. This is because the number of new entrants to the labor force exceeds the number of jobs created.

It is difficult to determine the impact of the unemployment rate on the workers compensation market a priori. Various effects are present, and the models will determine the dominant impact. For instance, an increase in the unemployment rate could have an inverse impact on the workers compensation market because inexperienced workers, who often have higher accident rates, are laid off first and workers may defer filing claims during a recession for fear of losing their jobs. Conversely, a decrease in the unemployment rate could increase frequency or severity since those inexperienced workers are rehired and workers who deferred filing a claim may file at the first sign of economic recovery.

On the other hand, the unemployment rate could have a direct impact on the workers compensation market. There are several hypotheses suggesting that frequency, severity, and the loss ratio should rise as the unemployment rate increases during a recession. These include increased duration and frequency due to diminished job opportunities, increased claim-filing incentives due to layoffs (i.e., workers substitute relatively generous workers compensation benefits for unemployment insurance benefits), jumps in the number of claims filed following business layoffs and failures, and increased incentives for workers to take time off to heal nagging injuries. Conversely, as the unemployment

[6]The correlation coefficient between employment and the unemployment rate over the entire 37-jurisdiction, 15-year sample is 0.035.

rate decreases in an economic expansion, the opportunity cost of filing a claim rises, reducing frequency, severity, and the loss ratio. The dominant impact of the unemployment rate on workers compensation loss ratios will be determined empirically.

The business failure variable is sometimes included in the models to measure crisis labor market conditions. The a priori expectation is that the sign on this variable will be positive. The preponderance of anecdotal evidence from industry executives, business owners, risk managers, and others suggests that business failures and plant closings provide a direct and special motivation for employees to file workers compensation claims because the benefits generally are larger and paid over a longer period than unemployment benefits. Moreover, workers compensation benefits are non-taxable. By one estimate, approximately forty to fifty percent of laid-off workers will file workers compensation claims against their employers within six months of termination [3, 4]. An increase in business failures is also expected to lead to an increase in severity because the employee's objective is to obtain a total workers compensation benefit that exceeds the expected unemployment benefit. Severity may also be higher for these claims as a result of the type of injury. Some chronic injuries may be concealed for extended periods of time, only to be revealed upon layoff.

Two variables that are independent of the business cycle, but are included in the models, are waiting period and cost containment. The underlying rationale for including the waiting period variable in the frequency model is straightforward. Longer waiting periods are a barrier and disincentive for workers to file claims, especially for minor injuries. Hence, an inverse relationship between frequency and the waiting period is expected. However, the expected sign on the waiting period coefficient in the indemnity severity regression is positive. Because longer waiting periods are a disincentive to file minor claims, the average severity of the remaining claims will be higher in states with longer waiting periods (holding all other factors constant), particularly if after a retroactive period the employee collects back to the date of injury.

An increase in the waiting period should have no statistically significant impact on the loss ratio. This is because the effects on losses and premiums associated with the waiting period increase should cancel out after on-leveling. However, on-level adjustments do not take into consideration the fact that law changes alter the incentives for workers to use the system. The sign on this variable in the loss ratio equation will depend on whether workers respond to the altered incentives and is discussed in more detail in Section 4.

The cost containment dummy variable is included as a proxy for the vigorous efforts adopted by many states and insurers in the early 1990s to attack rapidly rising workers compensation costs. The logical expectation is that such efforts reduce system costs. Thus, a negative sign on the cost containment coefficient is anticipated. Commonly employed reform initiatives, such as medical fee schedules, anti-fraud campaigns, managed care, and utilization review, have met with widely varying degrees of success across the states. Moreover, because the savings generated through legislative reform or insurer cost containment initiatives are not always readily observable, the dummy variable approach is a practical alternative way to quantifying these initiatives in dollar terms.[7]

Hospital cost per stay is included in the medical severity and medical loss ratio models as a measure of medical costs. The relationship between this variable and medical severity and the medical loss ratio is expected to be direct or positive. Ideally, the medical models would include a physician cost variable. Unfortunately, these data are not available at the state level.

Indemnity severity is included in the medical severity equation and the indemnity loss ratio is in the medical loss ratio equation

[7]A detailed discussion of dummy variables is contained in Appendix A.

to account for some of the variables that indirectly affect medical losses. NCCI believes that medical costs are driven, in part, by workers' demands to file indemnity claims. Beyond the obvious observation that an increase in overall claim severity drives up both indemnity and medical severity, explanations for this expectation include the possible longer observed durations of claims for higher wage earners, greater expectations from medical care by the high earners, and the tendency of these high earners to be in more urban areas where access to medical specialists and state-of-the-art technologies are more prevalent [22].

3. MODELING METHODOLOGY AND DESCRIPTION OF THE DATA

Modeling Methodology

The specifications of the frequency, severity, and loss ratio econometric models used in this study were determined after evaluating and testing families of models containing alternative specifications. These families of models used alternative variables and lag structures to measure the production and labor market effects discussed in the previous section. During this testing phase of the analysis, candidate explanatory variables were selected and evaluated.[8] Variables were rejected when they either failed to achieve statistical significance, behaved erratically over time, or were inconsistent with economic theory. The models themselves were rejected when they did not meet specific goodness-of-fit criteria based on the adjusted-R^2 statistic.

In the cross-sectional time series study reported here, parameters of the model were estimated using ordinary least squares (OLS) multiple regression analysis. All variables were in natural logarithmic form. Where appropriate, estimates were corrected

---

[8] The testing of many models may overstate at times the statistical significance of the chosen models since, for example, out of every 100 invalid models tested, one would likely pass significance tests at the one percent level. However, in the case of this study, the various families of models were similar and measured the same effects, but with different variables.

for autocorrelation and heteroscedasticity.[9] State dummy variables were included to control for omitted state-specific effects such as fraud, administrative variations, or other differences such as the propensity to litigate claims.[10]

NCCI maintains an extensive data base containing insurance industry, economic, demographic, and regulatory data in order to support its modeling efforts. Initial choices of explanatory variables were determined by researching the historic and current economic conditions nationally and in individual states. The data base information on economic conditions was obtained from four sources: federal and state government agencies, state universities, and private forecast and database concerns, such as Dun & Bradstreet. This information has enabled NCCI to account for certain state-specific cost drivers as well as more regional and national developments that have affected workers compensation frequency, severity, and loss ratios.

There is no compelling practical or theoretical reason to expect that the final specification of any of the models will be identical. For example, the model that best estimates indemnity severity generally will not result in the best estimates of indemnity loss ratios, even though both dependent variables are statistically correlated.

Structural stability of the estimated models is often an issue in econometric analyses. In this study, Chow tests were performed to measure the stability of the models across states and over time [12]. When performing Chow tests, the regression is rerun

[9]Durbin–Watson statistical tests were performed to check for the presence of autocorrelation. Autocorrelation arises when regression error terms are correlated through time. The consequences of failing to account for the presence of autocorrelation include inefficient ordinary least squares estimates of the regression model parameters (i.e., standard errors are inflated) and misleading hypothesis tests.

Heteroscedasticity is the situation where the model error terms do not have the same variance. When heteroscedasticity is present and the problem is not corrected, hypothesis tests will also be misleading. See Appendix B for a discussion of the method used to correct for autocorrelation and heteroscedasticity.

[10]See Appendix A for a detailed discussion of the role and interpretation of dummy variables.

on subsamples of the overall sample, and the estimated models are compared to determine if they are significantly different. In this analysis the sample was split by state randomly and based on population. The sample was also divided in half by year. The results from these various tests were mixed. In order to control for the differences among states and years, state dummy variables and the cost containment dummy variable discussed above were added to the initial models.

Dependent Variable Definitions

The dependent variables used in the analyses were computed using statewide (voluntary and residual market) premium, loss, and claim count data collected from workers compensation carriers through financial data calls. The loss and claim count data used are on an accident year basis, while accident year premium data are derived from a weighted average of two policy years. In the models, the accident year insurance data are paired with the economic data for the corresponding calendar year. Data are for 37 jurisdictions over the 15-year period 1979–1993.[11]

All data used in this study were thoroughly validated. Severities and frequencies by state were compared to Unit Statistical Plan data, and the analysis excluded all companies with suspect claim counts. This is the same data used in NCCI's rate level analyses.

The claim count data and indemnity and medical losses were developed to ultimate. Development factors varied by state based on a by-state analysis to determine which method would produce the most accurate estimate of ultimate losses. In general, the de-

[11]The jurisdictions included in the study are: Alabama, Alaska, Arizona, Arkansas, Colorado, Connecticut, Florida, Georgia, Hawaii, Idaho, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Michigan, Mississippi, Missouri, Montana, Nebraska, New Hampshire, New Mexico, North Carolina, Oklahoma, Oregon, Rhode Island, South Carolina, South Dakota, Tennessee, Utah, Vermont, Virginia, Wisconsin, and the District of Columbia.

velopment methodology is comparable to that selected in each state's rate filing. The data were also brought on-level. Indemnity and medical losses were brought to current benefit levels, and the measure of premium used in this study, designated statistical reporting (DSR) level standard premium, was brought to current bureau loss cost level.[12] All loss data used in the calculation of the dependent variables exclude all loss adjustment expenses.

The following describes how the frequency variable used in this analysis was calculated from the data elements of the financial call:

$$\text{Frequency} = \frac{\text{indemnity claims developed to ultimate}}{\text{workers (in hundred thousands)}}.$$

The number of workers (in hundred thousands) is estimated by year by state from premium as follows:

$$\left(\frac{\text{on-level DSR premium}}{\text{DSR average rate}} \times 100\right) \div (\text{average weekly wage} \times 52 \times 100{,}000).$$

The DSR average rate is the weighted average of the DSR rate by class multiplied by the Unit Statistical Plan (USP) payroll by class for the state. The average weekly wage is from the Current Population Survey performed by the Bureau of Labor Statistics, adjusted to exclude businesses not generally covered by workers compensation.

The severity dependent variables are defined as follows:

Indemnity severity

$$= \frac{\text{real indemnity losses on-level and developed to ultimate}}{\text{indemnity claims developed to ultimate}},$$

[12]DSR standard premium essentially represents the premium that would have been charged before adjustments such as company deviations, loss cost multipliers, premium discounts, retrospective rating, or schedule rating.

and

$$\text{Medical severity} = \frac{\substack{\text{real medical losses (excluding medical-only)} \\ \text{on-level and developed to ultimate}}}{\text{indemnity claims developed to ultimate}}.$$

In this study, the nominal or current (unadjusted) indemnity severity was converted to real terms using an average weekly wage index constructed for each state. Medical severity was deflated using the medical component of the Consumer Price Index (CPI) produced by the United States Bureau of Labor Statistics. For the medical severity dependent variable, a factor derived from USP data was applied to total medical losses to remove all medical-only dollars. Since claim counts in the financial calls exclude medical-only claims, this was necessary to achieve consistency between the numerator and denominator of medical severity.

The loss ratio dependent variables are defined as follows:

$$\text{Indemnity loss ratio} = \frac{\text{indemnity losses on-level and developed to ultimate}}{\text{loss cost portion of on-level DSR premium}},$$

and

$$\text{Medical loss ratio} = \frac{\text{medical losses on-level and developed to ultimate}}{\text{loss cost portion of on-level DSR premium}}.$$

In some states DSR premium is loss costs, while in others it is rates. In this analysis, however, it was adjusted to loss costs for all states. While the loss ratio typically uses full premium in the denominator, in this analysis the term "loss ratio" is used to describe the ratios defined above.

To summarize, the data base constructed for this study permits the analysis of a significant number of potential cost drivers affecting frequency, medical and indemnity severity, and medical and indemnity loss ratios. The ultimate goal has been to ensure that the final model is statistically sound, logically consistent in terms of economic theory, and that it is the best among alternative specifications.

4. RESULTS AND INTERPRETATION

Claim Frequency and the Business Cycle

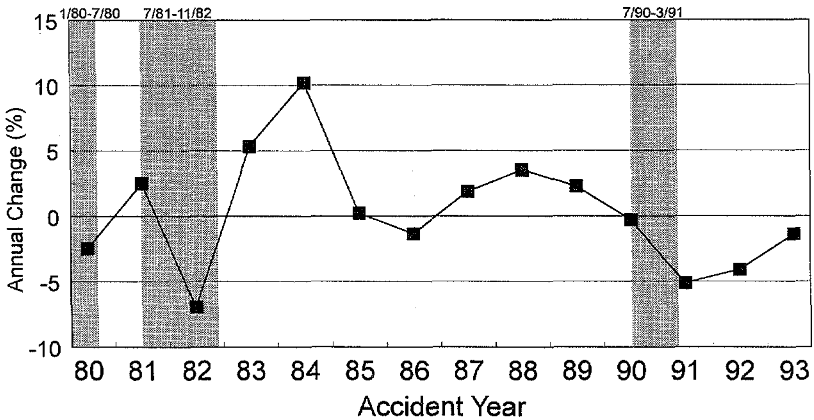
Numerous hypotheses suggest a relationship between workers compensation claim frequency and the economic environment, as discussed in Section 2. The a priori expectation in this study is that claim frequency will tend to rise during periods of economic expansion and fall during contractions.

Figure 1 shows the average annual percentage change in work-related claim frequency for the 37 jurisdictions in our sample.[13] The frequency decreases shown for 1980, 1982, 1990, and 1991 are coincident with national economic recessions (shaded regions) during those years. Frequency increased sharply during the economic recovery that began in 1983 before reaching a plateau in 1985. Modest further increases in frequency were recorded as the economic recovery matured during the late 1980s. The period beginning in 1983 was the longest sustained recovery in the post-World War II era, but eventually gave way to recession during the second half of 1990 and into 1991. The modest decrease in claim frequency in 1992 and 1993 is consistent with the most recent recovery's unusually slow rate of job creation during the first two years of expansion. Job creation and

[13] A simple average of the claim frequency data for the 37 jurisdictions was calculated for each year, 1979 to 1993. The annual percentage changes included in Figure 1 were calculated based on these simple averages. A simple average was used instead of a weighted average to more closely match the data in the study since it is cross-sectional time series rather than total national.

FIGURE 1

ANNUAL CHANGES IN FREQUENCY (%)



an expected upswing in frequency were limited because 18 of the 37 jurisdictions in the sample remained mired in recession even after the national recession ended in March 1991. Frequency increases were also limited because of the workers compensation system reforms that were initiated in these years in several states.

Table 2 summarizes the results from two econometric estimates of claim frequency. The absolute *t*-statistic is shown in italics below the estimated coefficient.[14]

The dependent variable in both regressions, *l*FREQPW, is the natural logarithm of frequency. The "*l*" preceding NAGEMP, UNRATE, and WAITPER indicates that these variables have also been converted to natural logarithms.

The empirical results for both models shown in Table 2 are consistent with prior expectations. The coefficient on the employment variable is positive, while the coefficients on the waiting

[14]The *t*-statistics are used to test the null hypothesis that the respective coefficient is equal to zero.

TABLE 2

WORKERS COMPENSATION FREQUENCY MODELS
COEFFICIENTS AND *t*-STATISTICS
Dependent Variable = *l*FREQPW

Independent Variable/Constant	Models (1)	Models (2)
Constant	3.64 *5.59*	3.78 *5.81*
*l*NAGEMP	0.58 *7.16**	0.59 *7.31**
*l*UNRATE	−0.09 *3.18**	−0.09 *3.19**
COSTCON	−0.06 *3.48**	−0.05 *3.25**
*l*WAITPER	—	−0.16 *2.54**
State Dummies	Yes	Yes
Adjusted-R^2:	0.907	0.909
N:	555	555

*Significant at the 1% level.

period and cost containment variable are negative. As discussed in Section 2, it is difficult to determine a priori the expected sign on the unemployment rate variable. The sign on the unemployment rate coefficient in the frequency regression is negative. That is, a *decrease* in the unemployment rate is associated with an increase in frequency. The interpretation is that the marginal hires are more likely to be injured [28]. This effect dominates any possible countervailing effects as discussed in Section 2.

It is worth noting that the nonagricultural employment and unemployment rate variables tend to move in opposite directions. For this reason, the opposite signs on the two coefficients reinforce each other, resulting in an amplification effect. Hence, claim frequency is expected to increase during economic expansions and decline during contractions or periods of sluggish growth. Changes in the cost containment and waiting period variables are independent of the business cycle.

Because the models were estimated in logarithms rather than in levels, the variable coefficients can be interpreted as elasticities or sensitivities. For example, the coefficient on the *l*NAGEMP variable in model (1) indicates that a 10 percent increase in nonagricultural employment leads (approximately) to a 5.8 percent increase in claim frequency. Likewise, a 10 percent decrease in the unemployment rate is associated with a 0.9 percent increase in frequency.

This example can be made more realistic by using actual 1995 national forecast values from Regional Financial Associates (RFA), an econometric forecasting organization.[15] Using Model 1, RFA's forecast for 2.5 percent employment growth nationally is expected to lead to a 1.45 percent increase in frequency. Similarly, the projected 8.2 percent decline in the unemployment rate (from 6.1 percent to 5.6 percent) is associated with a frequency increase of 0.74 percent.

Based on the results from the above models, Table 3 presents specific examples of how cyclical economic factors can influence the workers compensation line. Because the arguments presented here generally are symmetric with respect to the phase of the business cycle, discussion is limited to the case of economic expansion. While the cited factors should not be construed as an exhaustive list, they are representative of some of the more important developments that affect frequency when employment and the unemployment rate change.

An increase in employment and decrease in the unemployment rate during an economic expansion is expected to increase

---

[15] *Précis*, Vol. 3, No. 6, Regional Financial Associates, West Chester, PA, June 1995.

TABLE 3

ECONOMIC RATIONALE FOR CHANGES IN FREQUENCY DURING ECONOMIC EXPANSIONS

Employment Increases (increases frequency)	Unemployment Rate Decreases (increases frequency)
• Employment increases in hazardous industries • Overtime increases, leading to increased worker fatigue • Less machine maintenance diminishes job safety • Older, less safe and less efficient equipment may be reused • Inexperienced workers, more prone to injuries, are hired or rehired • Workers who deferred claims in recession for fear of losing job may file now	• Overtime increases, leading to increased worker fatique • Inexperienced workers, more prone to injuries, are hired or rehired • Workers who deferred claims in a recession for fear of losing job may file now

claim frequency for several reasons, some of which are listed in Table 3 and discussed here.

First, employers must hire new workers to meet increased demand during an economic expansion. In general, these workers tend to be younger and less experienced, resulting in more frequent injuries. In this case, it is the level of economic activity that compels employers to expand employment.

Increased demand for goods and services also leads to higher workers compensation costs indirectly. For example, during an economic expansion, producer shipments will rise and business-related vehicle traffic increases. Increasing vehicle travel will lead to an increase in claim frequency. Moreover, motor vehicle accidents are the leading cause of on-the-job fatalities [4].

Another reason why an economic expansion may lead to higher claim frequency is linked to employment growth in hazardous or risky industries, such as in the highly cyclical construction, heavy manufacturing, and mining industries. In this case, it is the shifting mix of employment that is the key transmittal mechanism from the economy to the workers compensation system.

Increased utilization of the system by workers also contributes to an increase in frequency during an economic expansion. Employees who feared they would lose their jobs if they filed a claim during the recession may file once the economy recovers. A 1993 NCCI survey of twenty of the largest workers compensation carriers found that workers' concerns over job security can override the incentive to file a workers compensation claim. The same observation has been made by some in the risk management community [26].

Finally, as employment rises and the pace of economic activity quickens, machine usage increases, and less maintenance and overall safety may accompany the higher capacity [25]. Also, workers' overtime hours increase, leading to fatigue and an increase in accidents. During a 1994 strike, workers at General Motors Corporation cited excessive overtime as the primary reason for their walkout. GM ended the dispute by agreeing to hire hundreds of new workers [27].

The association between higher claim frequency in general and economic expansion has been documented most recently in California. The California Workers Compensation Institute (CWCI) recently reported a 21 percent increase in indemnity claims frequency and an 11 percent increase in medical-only claims frequency; it attributes these increases to the state's recovering economy [8].

The converse of the above arguments is also generally true. In other words, the economic factors that contribute to increasing claim frequency during an economic expansion work to reduce frequency during a contraction.

Claim Severity and the Business Cycle

Both indemnity and medical severities increased rapidly during the early 1980s. Over the 1980–1982 period, the average annual increase in indemnity severity was 10.2 percent and 15.0 percent in medical severity. It is important, however, to recognize that the early 1980s was an inflationary period and that the behavioral and economic decisions made by workers, employers, insurers, and others are responses to real (inflation adjusted) changes in benefits and costs. Moreover, the true impacts of the explanatory variables on severity cannot be observed without first controlling for inflation. For these reasons, it is more informative to use inflation adjusted or real severity data in this analysis.[16]

Figure 2A and Figure 2B compare annual changes in nominal and real indemnity and medical severity, respectively. Indemnity severity increases recorded in 1981, 1983, 1984, 1992, and 1993 were actually declines when measured in real terms. The interpretation is that the average cost of an indemnity claim fell relative to the mean increase in average weekly wages during those years. These years approximately correspond to the recession induced frequency declines discussed above.[17] The decreases in severity are consistent with the decline in hazardous industry employment and the previously discussed incentive for workers to defer filing claims during recessionary periods out of fear of losing their jobs.

[16]As discussed in Section 3, nominal or current (unadjusted) indemnity severity was converted to real terms using an average weekly wage index constructed for each state, while medical severity was deflated using the medical component of the Consumer Price Index (CPI).

[17]Recession induced severity declines, as shown in Figure 2A and Figure 2B, are not precisely coincident with national economic recessions because the experience of our 37-jurisdiction sample differs from the nation's experience. For example, 27 states in the sample were in recession for some part of 1983, even though the national recession ended in November 1982.

FIGURE 2A

ANNUAL CHANGES IN INDEMNITY SEVERITY (%)

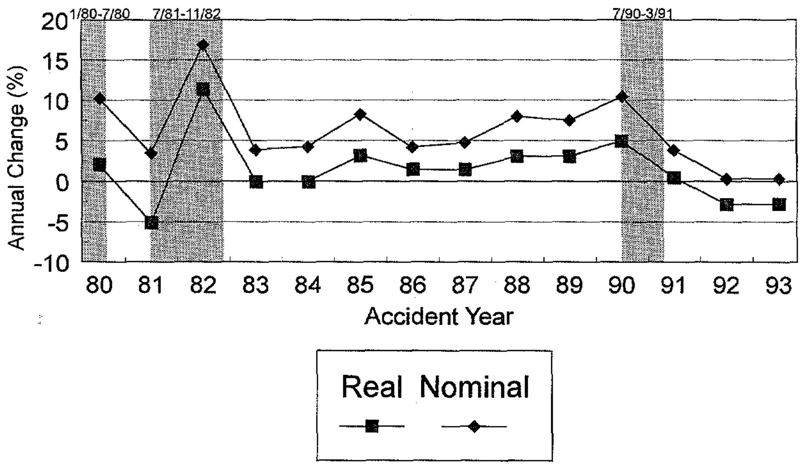
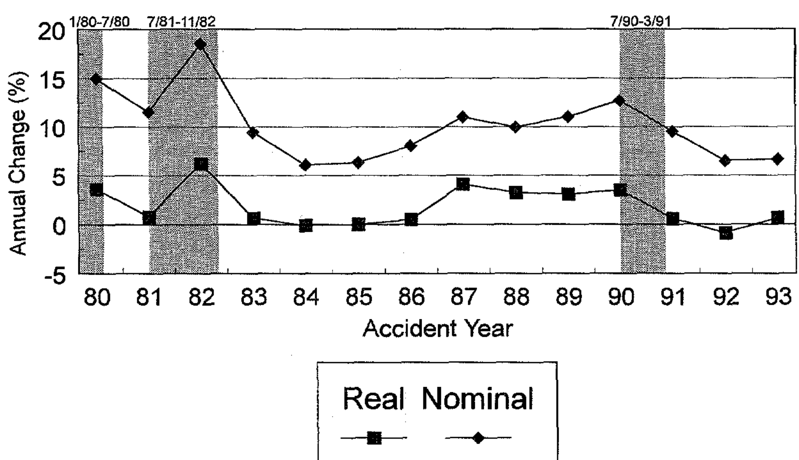


FIGURE 2B

ANNUAL CHANGES IN MEDICAL SEVERITY (%)



Changes in real medical severity, on the other hand, have remained positive for all years over the sample period except 1984 and 1992. Increases have occurred despite adjusting for inflation using the medical component of the CPI, which is significantly higher than the overall CPI. Thus, workers compensation medical severity not only grew faster than the general rate of inflation (as measured by the CPI), but also faster than medical inflation nationwide.

The hypothesis that real indemnity and medical severity vary over the business cycle can be tested econometrically. Table 4 summarizes the results from the econometric estimate of indemnity and medical claim severity. The absolute *t*-statistic is shown in italics below the estimated coefficient.

All variables have been converted to natural logarithms. As in the frequency regressions, estimation of the severity models in logarithmic form implies that the variable coefficients can be interpreted as elasticities. For example, the coefficient on the *l*NAGEMP variable in the indemnity model implies that a 10 percent increase in nonagricultural employment leads to a 6.8 percent increase in claim severity. Similarly, a 10 percent increase in hospital cost per stay is associated with a 2.5 percent increase in medical claim severity.

Unlike the frequency regressions, the unemployment rate variable in the severity regressions has a positive sign, indicating that a direct effect is the dominant impact. The signs on employment, business failures, waiting period, and hospital cost per stay variables are consistent with prior expectations discussed in Section 2.

As discussed previously, employment and the unemployment rate generally move in opposite directions. The positive sign on the coefficients of both variables means that a decrease in the unemployment rate during an economic expansion, for example, tends to dampen the severity increasing influence of expanding employment. Empirical evidence indicates that the employment

TABLE 4

WORKERS COMPENSATION REAL SEVERITY MODELS
COEFFICIENTS AND *t*-STATISTICS
Dependent Variable = *l*INDSEV, *l*MEDSEV

Independent Variable/Constant	Models: Indemnity	Models: Medical
Constant	2.11 *1.54*	2.07 *4.77*
*l*NAGEMP	0.68 *3.86**	—
*l*UNRATE	0.13 *2.25***	—
*l*BUSFAIL	0.04 *2.59**	—
*l*WAITPER	0.35 *4.16**	—
*l*PERSTAY	—	0.25 *5.74**
*l*INDSEV	—	0.44 *8.51**
State Dummies	Yes	Yes
Adjusted-R^2:	0.908	0.906
N:†	370	555

*Significant at the 1% level.
**Significant at the 5% level.
†Because business failures data were available beginning only in 1984, the number of observations in the indemnity model is less than for the medical model, which spans the entire 1979–1993 sample period.

effect is dominant. To illustrate this point, RFA's 1995 forecast for 2.5 percent employment growth would be expected to lead to a 1.7 percent increase in indemnity severity, while the projected 8.2 percent decrease in unemployment would be associated with a 1.1 percent *decrease* in severity. In combination, there is an expected 0.6 percent increase in severity, holding all else constant.

TABLE 5

ECONOMIC RATIONALE FOR CHANGES IN INDEMNITY CLAIM SEVERITY DURING ECONOMIC EXPANSIONS

Employment Increases (increases severity)	Unemployment Rate Decreases (decreases severity)
• Employment increases in hazardous industries	• Opportunity cost of claim rises, reducing duration
• Truck shipments increase	
• Overtime increases, leading to increased worker fatigue	
• Less maintenance; older machinery reused	
• Unfamiliarity of new employees with machinery	

Some of the important economic factors that contribute to increasing indemnity claim severity during an economic expansion are shown in Table 5. The converse is generally true during economic contractions.

Indemnity and Medical Loss Ratios and the Business Cycle

The statistical evidence presented thus far in this paper has documented that econometric techniques are an important tool that can be used to quantify the relationship between the economic environment and workers compensation frequency and severity. In this section, the econometric methodologies developed in the previous two sections are extended and modified to model indemnity and medical loss ratios over the course of the business cycle.

Loss ratios provide a more complete picture of the economy's impacts on the workers compensation industry because they incorporate information on both losses and premiums. Unlike other lines of insurance, workers compensation premiums are assessed as a proportion of total payroll. Payroll at the firm level often is subject to considerable variability over the course of the business cycle, thereby affecting premium collections. Variability in losses is due to changes in both the corresponding exposure base and the complex interactions between economic variables as they impact frequency and severity.

In general, the factors that best explain changes in medical and indemnity loss ratios will be different from those that best explain frequency and medical and indemnity severity. This is because modeling loss ratios incorporates the interactions between frequency and severity, while modeling frequency and severity separately does not incorporate these effects.

Table 6 summarizes the results of the estimated indemnity and medical loss ratio models.

The positive sign on the employment variable, *l*NAGEMP, in the indemnity model is consistent with a priori expectations and with the frequency and severity findings presented above. As employment rises and the pace of economic activity quickens, the indemnity loss ratio tends to deteriorate. This is partly because the accelerated rate and higher level of production cause machine usage to increase and older, less safe machinery to be brought back on line as capacity constraints are approached. Increased worker fatigue due to a faster pace of production and abundant overtime opportunities is likely to contribute to higher injury severities as well as increased claim frequency.

As in the severity model, the sign on the unemployment rate variable in the loss ratio regression is positive. A decline in the unemployment rate during an expansion is associated with a probable slight decline in the loss ratio. The decline is principally the result of the increasing opportunity cost to the employee (and employer) of a workers compensation claim. Rising wages and overtime opportunities diminish the incentive of the worker to file or stay out on a claim. Concomitantly, the cost to the

TABLE 6

WORKERS COMPENSATION LOSS RATIO MODELS COEFFICIENTS AND t-STATISTICS Dependent Variable = *l*INDLRAT, *l*MEDLRAT

Independent Variable/Constant	Models: Indemnity	Models: Medical
Constant	−12.32 *6.42*	−8.02 *29.90*
*l*NAGEMP	1.44 *5.98***	—
*l*UNRATE	0.08 *1.34*****	—
*l*WAITPER	0.22 *2.26**	—
COSTCON	−0.06 *1.56****	—
*l*PERSTAY	—	0.91 *29.41**
*l*INDLRAT	—	0.53 *16.05**
State Dummies	Yes	Yes
Adjusted-R^2:	0.757	0.898
N:	555	555

*Significant at the 1% level.
**Significant at the 5% level.
***Significant at the 10% level.
****Significant at the 20% level.

employer of losing a worker increases during periods of strong demand. Hence, the increasing opportunity cost of labor income and the rising value of the worker to the employer contribute to a decline in the loss ratio during an economic expansion.

As discussed in Section 2, an increase in the waiting period should have no statistically significant impact on the loss ratio since the effects on losses and premiums associated with the waiting period increase should cancel out after on-leveling. However, the coefficient on this variable in the above equation is positive and significant.

One reason for this result is that on-level adjustments do not take into consideration the fact that law changes alter the incentives for workers to use the system. Specifically, some proportion of workers will be sufficiently motivated to stretch the duration of their claim to meet the new waiting period, thereby increasing average claim severity. Importantly, the new (longer) waiting period will be several days closer to the retroactive date (assuming the retroactive period is not also increased), at which point benefits are paid retroactively to the injury date. Because the time interval between the waiting period and retroactive period has been reduced, the cost to the worker of extending the claim to the retroactive period is also reduced. Ironically, the net result of an increase in the waiting period may be to increase average severity and total system costs.[18]

Finally, the negative sign on the cost containment coefficient indicates that cost containment initiatives are successful in reducing system costs.

Overall, however, the employment effect is dominant, leading to the conclusion that the indemnity loss ratio will increase during an economic expansion and decrease during an economic contraction.[19]

The numerical interpretation of the results is analogous to the frequency and severity models discussed previously. Estima-

[18] Some industry observers believe that the on-level (law amendment) factors were generally overestimated. Systematic overestimation of these factors would lead to a perceived increase in severity. We thank an anonymous referee for bringing this possibility to our attention.

[19] Previous research examining the incentive effects contributing to workers compensation loss ratios includes Butler and Worrall [7]. Butler [5] also contains a comprehensive review of incentive effect studies, including several using NCCI data.

tion of the models in logarithmic form permits interpretation of the estimated coefficients as elasticities. The 1995 employment growth forecast of 2.5 percent implies a 3.6 percent increase in the indemnity loss ratio. Similarly, the projected 8.2 percent decrease in the unemployment rate would be expected to lead to a 0.7 percent decline in the indemnity loss ratio. In combination, they would lead to an expected 2.9 percent increase in the indemnity loss ratio, holding all else constant.

The medical loss ratio model is analogous to the medical severity model. The 2.9 percent increase in the indemnity loss ratio discussed above implies a 1.5 percent increase in the medical loss ratio.

The coefficients on the economic variables displayed in Table 6 indicate how loss ratios are expected to change over the course of the business cycle—rising during expansions and falling during contractions. Failure to consider these factors may result in systematic and cyclical bias when actuarial trend indications are used.

5. SUMMARY AND CONCLUSIONS

The empirical results presented in this study provide significant statistical evidence that econometric modeling is a powerful explanatory and diagnostic tool for explaining variability in the workers compensation line over the course of the business cycle. Specifically, this study demonstrates that cross-sectional time series techniques can be used to estimate the relationship between the economic environment and workers compensation claim frequency, indemnity and medical severity, and indemnity and medical loss ratios over the 15-year, 37-state, sample period. The estimated frequency, severity, and medical loss ratio models explain ninety percent or more of the variability in the respective dependent variable. More than seventy-five percent of the variability in the indemnity loss ratio is explained.

This study's major empirical finding is that there is a strong association between economic growth and rising frequency, severity, and loss ratios.[20] The converse is also true. This finding displaces the conventional wisdom on this subject, which has held that the workers compensation line generally improves during economic expansions and worsens in recessions. A possible consequence of not accounting for the impacts of economic factors on loss ratios is rate inadequacy. During a prolonged economic expansion, the inadequacy problem is compounded. Conversely, during recessionary periods, rates may become redundant.

Extensions and Directions for Future Research

One logical application of this study's findings is to employ econometric time series techniques at the individual state level to generate econometric estimates of trend. Actuarial methodologies traditionally have relied on curve fitting techniques that depend on historical values of the loss ratios and time to estimate trend. Econometric trend models, in contrast, draw upon a rich base of economic, demographic, and insurance variables.[21] Along with their intuitive appeal, econometric models can incorporate predicted future changes in the economy that may affect the direction and growth of losses and premiums. Moreover, econometric forecasts use all available data points to capture the impacts of the economy on loss ratios over the full course of the business cycle.

Econometric trend estimates have been filed by NCCI in several states, using models similar to those explored in this study. The models typically include a measure of overall economic activity, a measure of the health of the labor market, and possibly

---

[20]The conclusion is demonstrated for the business cycles during the data period used in this analysis, but may not follow over different periods of time.

[21]For a more complete discussion of econometric trending methodologies, see Hartwig, Kahley, and Restrepo [15].

variables relating to crisis labor market conditions or the share of employment in risky industries in the state. As in this study, logarithmic regression models are estimated separately for indemnity and medical loss ratios. Forecast values of the independent variables are then used to forecast expected loss ratios (the dependent variable). Applying these techniques to workers compensation claim frequency and severity at the state level is another logical extension that NCCI plans to explore.

NCCI plans to continue its study of the relationship between the business cycle and the workers compensation industry. To date, the results of these analyses are promising. Research and methodological refinements are ongoing, and the knowledge gained from modeling a larger number of states across a broad spectrum of economic and workers compensation experience should prove to be invaluable.

REFERENCES

[1] American Hospital Association, *Hospital Statistics*, annual issues.

[2] Baker, Laurence C. and A. B. Krueger, "Twenty-Four-Hour Coverage and Workers Compensation Insurance," *Health Affairs*, Supplement 1993, pp. 271–281.

[3] *Business Insurance*, "Avoiding Workers Comp Claims Crush Following Layoffs and Plant Closings," April 13, 1992.

[4] *Business Insurance*, "Planning for Downsizing: Reducing Potential for Workers Comp Claims Requires Teamwork," May 8, 1995.

[5] Butler, Richard J., "Economic Determinants of Workers Compensation Trends," *Journal of Risk and Insurance* 61, 1994, pp. 383–401.

[6] Butler, Richard J. and David Appel, "Benefit Increases in Workers Compensation," *Southern Economic Journal* 56, 1990, pp. 594–606.

[7] Butler, Richard J. and John D. Worrall, "Premium and Loss Cycles in Workers Compensation," in Philip S. Borba and David Appel, eds., *Benefits, Costs, and Cycles in Workers Compensation*, Boston, MA: Kluwer Academic Press, 1990.

[8] California Workers Compensation Institute, "Workers' Comp Claim Frequency Jumps as California Economy Improves," report dated September 9, 1994.

[9] Doan, Thomas A., *RATS User's Manual*, Version 4, Evanston, IL: Estima, 1995.

[10] Dun & Bradstreet Corporation, *Business Failure Record*, annual issues.

[11] Durbin, David L., Daniel Corro, and Nurhan Helvacian, "Workers Compensation Medical Expenditures: Price vs. Quantity," *Journal of Risk and Insurance* 63, 1996, pp. 13–33.

[12] Gujarati, Damodar N., "Comparing Two Regressions: The Chow Test," *Basic Econometrics*, New York, NY: McGraw-Hill Publishing Company, 1988, pp. 443–448.

[13] Hansen, Lars P., "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica* 50, 1982, pp. 1029–1054.

[14] Hartwig, Robert P., "Business Cycle-Induced Volatility in Workers Compensation Frequency, Severity, and Claim Cost per Worker," *Digest*, National Council on Compensation Insurance, 11, 2, November 1996.

[15] Hartwig, Robert P., William J. Kahley, and Tanya E. Restrepo, "Workers Compensation Loss Ratios and the Business Cycle," *Digest*, National Council on Compensation Insurance, 9, 2, December 1994, pp. 1–14.

[16] Hsiao, Cheng, *Analysis of Panel Data*, Cambridge University Press, 1993.

[17] Insurance Services Office, Inc., "Changes in the Economic Environment: Insurance Applications," *ISO Insurance Issue Series*, October 1991.

[18] Kennedy, Peter, *A Guide to Econometrics*, Cambridge, MA: MIT Press, 1992.

[19] Lommele, Jan A. and Robert W. Sturgis, "An Econometric Model of Workmen's Compensation," *PCAS* LXI, 1974, p. 170.

[20] Moore, Michael J. and W. Kip Viscusi, "Have Increases in Workers Compensation Costs Paid for Themselves?," in Philip S. Borba and David Appel, eds., *Benefits, Costs, and Cycles in Workers Compensation*, Boston, MA: Kluwer Academic Press, 1990.

[21] Newey, W. and K. West, "A Simple Positive-Definite Heteroscedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* 55, 1987, pp. 703–708.

[22] Retterath, Ronald, "Total Payroll/High Wage Payers," *Digest*, National Council on Compensation Insurance, September 1991, pp. 13–20.

[23] United States Chamber of Commerce, *Workers Compensation Laws*, annual issues.

[24] United States Department of Labor, Bureau of Labor Statistics, *Employment and Earnings*.

[25] *Wall Street Journal*, "An Auto Worker Earns More Than $100,000, But at a Personal Cost," August 1, 1996, p. A1.

[26] *Wall Street Journal*, "Employers Grab the Initiative on Workers Compensation," November 15, 1994.

[27] *Wall Street Journal*, "GM, in a Switch, Agrees to Hire New Workers," October 3, 1994, p. A3.

[28] *Wall Street Journal*, "Truckers' Pay Rises Amid Labor Crunch," September 23, 1996, p. A2.

[29] White, Halbert, "A Heteroscedasticity-Consistent Covariance Matrix Estimator and Direct Test for Heteroscedasticity," *Econometrica* 48, 1980, pp. 817–838.

[30] Worrall, John D. and Richard J. Butler, "Heterogeneity Bias in the Estimation of the Determinants of Comp Loss Distributions," in Philip S. Borba and David Appel, eds., *Benefits, Costs, and Cycles in Workers Compensation*, Boston, MA: Kluwer Academic Press, 1990.

[31] Worrall, John D. and Richard J. Butler, "Workers Compensation: Benefits and the Duration of Claims," in John D. Worrall and David Appel, eds., *Workers Compensation Benefits: Adequacy, Equity, and Efficiency*, Ithaca, NY: ILR Press, 1985.

APPENDIX A

DUMMY VARIABLES IN ECONOMETRIC ANALYSIS

Variables used in regression analysis are usually, but not always, continuous over some range. There are situations, however, where explanatory or independent variables are qualitative in nature, taking on two or more distinct values. For example, individuals are male or female, and gender may explain certain behaviors or outcomes. Similarly, investment and consumption decisions may be influenced by whether a country is at war or peace. In such instances, a proxy variable, referred to as a "dummy variable" in econometrics, must be constructed. The dummy variable takes on the value unity whenever the qualitative phenomenon it represents occurs, and zero otherwise. For estimation purposes, the dummy variable is treated no differently than any other explanatory variable, and no modifications to the chosen estimation technique are necessary.[22]

Fixed Effects Models

Qualitative variables are frequently not dichotomous and may assume more than two distinct values (female/male). In the context of the present study, the possibility that the intercept varies across the $N = 37$ cross-sectional units (states) is recognized and incorporated in the frequency, severity, and loss ratio models by constructing dummy variables for $(N - 1)$ states in the sample.[23] One state is omitted to avoid perfect multicollinearity. With perfect multicollinearity, the regression matrix is singular and, therefore, coefficients for the explanatory variables cannot be estimated. The individual contribution of the omitted state is given by the value of the constant term in the regression equation. Such observation specific dummy variables are equal to unity

[22]A more detailed discussion on dummy variables is given in Kennedy [18] and most other econometric textbooks.

[23]State dummy variables in this study are generally statistically significant at the one percent level.

for a specific observation (e.g., Alabama) and zero for all other observations (i.e., every state except Alabama).

The fact that the 37-state sample is observed over a 15-year period suggests another possible application of dummy variables. Dummy variables could be constructed for $(T - 1)$ years, under the hypothesis that the intercept varies over time, as well as across states (for a total of $(N - 1) + (T - 1)$ dummies). Time dummies would assume the value of one for a specific year and zero for all other years. This hypothesis was tested and rejected using the data in this study.

The use of cross-sectional and/or time dummy variables in a longitudinal (or panel) data context is commonplace in econometrics. This method of analyzing longitudinal data is known as *fixed effects* modeling.[24] Essentially, the dummy variable coefficients reflect ignorance. In other words, dummy variables are introduced for the purpose of measuring shifts in the regression line arising from unknown variables. Of course, other explanatory variables, with hypothesized functional relationships to the dependent variable, generally are included in such models as well.

Non-Observable and Difficult-to-Quantify Variables

Many variables believed to influence the level and growth rate of workers compensation frequency, severity, and loss ratios are difficult to quantify and/or cannot be observed directly or in a practical fashion. For example, governments, insurance companies, and employers embarked upon a myriad of legislative reforms, cost containment initiatives, and loss control programs across the country during the early 1990s. From a practical standpoint, it is impossible to survey and quantify all of these changes in state workers compensation systems. Indeed, insurers are applying their acquired knowledge and experience to states

[24] A detailed and rigorous discussion of fixed effects modeling is found in Hsiao [16].

where no major legislative reforms have been approved. Moreover, from an econometric perspective, modeling such a large variety of reforms and initiatives would require far more degrees of freedom than are available in this study sample of 37 states over 15 years.

A practical solution to modeling the many and diverse cost containment initiatives across states individually is to model them collectively using a dummy variable as a proxy. In this study, a single dummy variable is constructed, *COSTCON*; it is equal to one for the years 1991–1993 in every state and zero otherwise.[25] The coefficient on the dummy variable can be interpreted as an average estimate of the net impact of legislative reform and other initiatives on workers compensation frequency and the indemnity loss ratio (the variable was not statistically significant in the severity regressions).

[25] Anecdotal evidence suggested that the "turning point" for workers compensation came in the early 1990s. The choice of 1991 as the beginning point is based on the statistical strength of *COSTCON* in that year relative to other years. The ending point is 1993 because that is the last year of data in the sample used in this study.

APPENDIX B

AUTOCORRELATION AND HETEROSCEDASTICITY CORRECTION

When correcting for autocorrelation and heteroscedasticity, Regression Analysis of Time Series (RATS), the programming software used in this study, computes the regression using least squares, but then computes a consistent estimate of the covariance matrix allowing for heteroscedasticity and serial correlation up to a first-order moving average.

Ordinary least squares provides a consistent estimate for β in the regression model $\mathbf{Y} = \mathbf{X}\beta + \mathbf{u}$ in a large number of settings where the standard assumption that the residuals satisfy the equation $\mathbf{V} = \mathbf{E}(\mathbf{uu}') = \sigma^2\mathbf{I}$ is violated. Although least squares may provide consistent estimates of the coefficients, $s^2\mathbf{X}'\mathbf{X}^{-1}$ is not a consistent estimate of the variance of the coefficient estimates. Therefore, tests based on the regression output will be incorrect.

To correct the problem, RATS computes consistent estimators for the covariance matrix of estimators using a procedure that imposes little structure on the matrix $\mathbf{V}$. The estimators for least squares are $(\mathbf{X}'\mathbf{X})^{-1}\text{mcov}(\mathbf{X},\mathbf{u})(\mathbf{X}'\mathbf{X})^{-1}$ where mcov($\mathbf{X},\mathbf{u}$) refers to the following matrix

$$\sum_{k=-L}^{L} \sum_{t} u_t X_t' X_{t-k} u_{t-k},$$

and u_t is the residual at time t. Serial correlation is handled by making L non-zero. This corrects the covariance matrix for serial correlation in the form of a moving average of order L [9].[26]

[26]Additional detail regarding the correction can be found in Hansen [13], Newey and West [21], and White [29].