

PROPERTY-CASUALTY INSURANCE CYCLES, CAPACITY CONSTRAINTS,
AND EMPIRICAL RESULTS

by

ANNE GRON

B.A. Economics and Computer Science, Williams College
(1984)

SUBMITTED TO THE DEPARTMENT OF ECONOMICS
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
September, 1990

© Anne Gron, 1990. All rights reserved.

The author hereby grants to MIT permission to reproduce and to
distribute copies of this thesis document at whole or in part.

Signature of Author _____

Department of Economics
July, 1990

Certified by _____

Professor Paul Joskow
Thesis Supervisor

Accepted by _____

Professor Peter Temin
Chairman, Department Graduate Committee

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

OCT 19 1990

LIBRARIES

ARCHIVES

Property-Casualty Insurance Cycles, Capacity Constraints,
and Empirical Results

by
Anne Gron

Submitted to the Department of Economics
in partial fulfillment of the requirements
for the Degree of Doctor of Philosophy

ABSTRACT

Since World War II, the property-casualty insurance industry has been subject to periodic market disruptions, usually called insurance crises. These crises are marked by rapidly increasing prices and shrinking markets. The most severely affected insureds may be unable to renew expiring policies or may even have their current policies canceled. Others are faced with high premium increases, decreased limits, or both. These disruptions are part of a repeating series of market conditions affecting the property-casualty industry known as the insurance cycle.

This dissertation investigates the capacity-constraint hypothesis of property-casualty insurance cycles. Chapter one presents a model of price determination in competitively structured insurance markets with regulatory restrictions specifying a minimum level of capital to back policies. It is assumed that there are adjustment costs associated with large increases to insurer capital, and that firms are monitored by insurance regulators whose goal is to achieve a low probability of bankruptcy and maintain the ability of insurers to meet claims made by policyholders. Under these assumptions, large random shocks to industry net worth produce market conditions similar to those observed over the cycle.

Chapter two analyzes industry time series data for the period 1947-87 in the context of the capacity constraint hypothesis detailed in chapter one. The focus is on two implications of this hypothesis: the relation between low capacity and high premium growth and the nature of net worth adjustment costs. Statistical analysis shows that the probability of unusually high premium growth is significantly larger when observed capacity is unusually low in the preceding year. The chapter also considers the relative magnitudes of adjustment costs associated with increasing or decreasing net worth. Analysis of the distribution of annual growth in capacity and premium volume shows that the industry adjusts faster to positive than to negative capacity shocks. Both results are consistent with the capacity-constraint hypothesis.

In contrast to chapter two which examines the empirical significance of two components of the capacity-constraint theory, chapter three tests the capacity-constraint hypothesis against the alternative that accounting profit cycles are predominantly the outcome of regulatory and institutional factors and not due to changes in insurer capacity. Two

kinds of evidence in support of the capacity constraint theory are found. First, the relation of insurer accounting profitability and variations in insurance capacity is estimated using industry time series data over the 1952-1986 period. The estimation controls for factors such as expected inflation, expected interest rates and unexpected cost increases which affect accounting profit measures through institutional and regulatory procedures. The results are consistent with the capacity constraint hypothesis.

A second form of evidence is found by examining the change in the stock market response to two recent catastrophes, Hurricane Hugo and the San Francisco earthquake. Under the alternative that profitability cycles are predominantly the product of regulatory and institutional features, large, unexpected loss events should reduce the stock market value of insurers liable for the losses. The capacity-constraint hypothesis indicates that there will be two offsetting effects, a decline in market value from the increase in liabilities, and an increase based on the expected increase in future returns to the remaining net worth in the industry. Daily and cumulative excess of market returns show that average firm values remained constant or increased after these two events. These results support the hypothesis that the stock market response has two offsetting effects over the alternative that there is only a negative effect.

Thesis Supervisors: Dr. Paul Joskow, Dr. Nancy Rose

ACKNOWLEDGEMENTS

Many people have contributed to the production of this thesis in some way, and no list would be complete. While I take this opportunity to acknowledge those with the greatest tangible influence, I am, of course, deeply grateful to the others.

I would first like to thank Paul Joskow who initially suggested studying the property-casualty insurance industry. His critical commentary and advice have improved not only this thesis, but more importantly, my understanding of economics.

I have also learned much from Nancy Rose as her student, research assistant, and advisee. Our numerous discussions have taught me much about economic research. Most importantly, she provided support and encouragement when I became frustrated. A special thanks goes to Jim Poterba for generously taking me on as yet another advisee in my last year. In addition, I am grateful to help from the members of the Department of Economics faculty, Peter Diamond, Andrea Shepard, Jeff Wooldridge and Peter Temin in particular.

It is often said that the students are one, if not the, most important assets of the department. My fellow classmates and contemporaries at M.I.T. have been invaluable as friends, instructors and colleagues. Of the many who helped, Mitch Petersen, Doug Staiger, Debbie Swenson and Chris Velturo were particularly instrumental in the development and completion of this thesis.

Several people and organizations provided financial and technical support. I am grateful to the Department of Economics for generous financial support in my first two years. Support from the Bradley Foundation and the Olin Foundation is also gratefully acknowledged. A special debt is owed to Jean Lucey and the staff of the Insurance Library Association of Boston. Without them this research would not have been possible.

Finally, I would like to thank my friends and family who have put up with me all this time.

TABLE OF CONTENTS

Abstract	2
Acknowledgements	4
Introduction and Overview	6
Chapter One: Price Determination in Insurance Markets with Regulatory Constraints	11
Chapter Two: Capacity and Property-Casualty Insurance Cycles: Time Series Evidence	47
Chapter Three: The Role of Capacity in Accounting Profitability Cycles	97

INTRODUCTION AND OVERVIEW

Since World War II, the property-casualty insurance industry has been subject to periodic market disruptions, usually called insurance crises. These crises are marked by rapidly increasing prices and shrinking markets. The most severely affected insureds may be unable to renew expiring policies or may even have their current policies canceled. Others are faced with high premium increases, decreased limits, or both. These disruptions are part of a repeating series of market conditions affecting the property-casualty industry known as the insurance cycle.

The typical property-casualty insurance cycle has four stages, defined by changes in accounting profits, prices and quantity. The first stage, known as the crisis, marks the transition from low to high profitability. There are substantial premium increases, reduced availability and improving accounting profitability, often occurring quite suddenly. The hard market, a period of relatively low availability, high premium levels and high accounting profitability follows the crisis stage. The third stage is marked by a gradual erosion of high premium and high accounting profitability; premium levels decline, profitability declines and availability restrictions ease. Low premiums, low accounting profits and relatively abundant quantity typifies the fourth and final stage of a typical cycle, referred to as the soft market. Both price and quantity are unobserved

over the cycle, only anecdotal evidence is available. Therefore cycles are usually described and analyzed in terms of changes in accounting profitability.

This dissertation investigates the capacity-constraint hypothesis of property-casualty insurance cycles. Chapter one presents a model of price determination in competitively structured insurance markets with regulatory restrictions specifying a minimum level of capital to back policies. It is assumed that the industry is competitively structured and monitored by insurance regulators whose goal is to achieve a low probability of bankruptcy and maintain the ability of insurers to meet claims made by policyholders. To this end, regulators require insurers to hold a minimum financial cushion in excess of expected liabilities. The sources of capital to back policies are either net worth or the difference between price and expected loss. The funding available to back policies is related to the industry's capacity to accept risk in that it defines the maximum quantity of insurance that can be sold.

Large shocks to net worth, such as higher than anticipated cost inflation or catastrophic events, reduce net worth and decrease industry capacity. In the presence of net worth adjustment costs, the industry will not immediately return to long run equilibrium. Insurers currently in the industry will earn a higher return on their remaining capital, reflecting the temporary increase in the opportunity cost of capital. Over time, net worth expands due to the accumulation of retained earnings and capital from outside the firm, and premiums decline toward

long run equilibrium until another shock occurs.

Chapter two analyzes industry time series data for the period 1947-87 in the context of the capacity constraint hypothesis detailed in chapter one. Three issues are addressed. First, a definition of property-casualty insurance cycles based on capacity data is compared to the traditional criterion based on accounting profitability. Next, the behavior of other time series, in particular the relationship between low capacity and high premium growth, is examined. Finally, capacity and premium growth data are analyzed to determine whether they are consistent with the types of adjustment costs assumed in the capacity constraint hypothesis.

The timing of cyclical episodes identified using a capacity-based criterion is found to coincide closely with those identified using the more traditional accounting profit criterion in data from 1950-87. Such a finding is of interest since the prior work on the role of capacity in insurance cycles has relied on profit based criteria to define the cycles. Since these theories interpret changes in profitability as caused by changes in capacity, a criterion based on capacity is preferable because it is more direct. This work shows that cyclical periods identified using this alternative criterion do not substantially alter the cyclical episodes found.

Examination of the behavior of other aggregate time series finds support for the capacity constraint hypothesis. Periods of unusually low capacity growth coincide with or precede episodes of unusually high

premium growth. The low frequency of the annual data makes exact timing of the two events impossible, but statistical analysis shows that the probability of unusually high premium growth is significantly larger when observed capacity is unusually low in the preceding year. This relation describes the data particularly well when the 1970s period is omitted.

Finally, chapter two also considers the relative magnitudes of adjustment costs associated with increasing or decreasing net worth. The path of capacity adjustment is found to be asymmetric. Analysis of the distribution of annual growth in capacity and premium volume shows that the industry adjusts faster to positive than to negative capacity shocks. When equity from external sources is more costly than equity generated from within firm activities, insurers will not issue sufficient equity to return to long run equilibrium after negative capacity shocks. Instead, equity will be issued until the expected return within the firm is equal to the cost of externally generated funds. The rest of the adjustment will come from lower cost internal sources such as retained earnings.

In contrast to chapter two which examines the empirical significance of two components of the capacity-constraint theory, chapter three tests the capacity-constraint hypothesis against the alternative that accounting profit cycles are predominantly the outcome of regulatory and institutional factors and not due to changes in insurer capacity. Two kinds of evidence in support of the capacity

constraint theory are found. First, the relation of insurer accounting profitability and variations in insurance capacity is estimated using industry time series data over the 1952-1986 period. Four lines of insurance are considered: automobile physical damage, automobile liability, homeowners' multiple peril and other liability. The estimation controls for factors such as expected inflation, expected interest rates and unexpected cost increases which affect accounting profit measures through institutional and regulatory procedures. The results are consistent with the capacity constraint hypothesis.

A second form of evidence is found by examining the change in the stock market response to two recent catastrophes, Hurricane Hugo and the San Francisco earthquake. Under the alternative that profitability cycles are predominantly the product of regulatory and institutional features, large, unexpected loss events should reduce the stock market value of insurers liable for the losses. The capacity-constraint hypothesis indicates that there will be two offsetting effects, a decline in market value from the increase in liabilities, and an increase based on the expected increase in future returns to the remaining net worth in the industry. Daily and cumulative excess of market returns show that average firm values remained constant or increased after these two events. These results support the hypothesis that the stock market response has two offsetting effects over the alternative that there is only a negative effect.

CHAPTER ONE:

Price Determination in Insurance Markets with Regulatory Constraints

I. INTRODUCTION

The property-casualty insurance industry has experienced a series of cycles in industry profitability since the 1950s.¹ A typical insurance cycle is characterized by several years of low profitability and low premiums, followed by a transition marked by increasing prices and declining availability of coverage, to a hard market where premiums are high but no longer rising and the market has stopped shrinking. Then prices begin to decrease and availability problems disappear as the industry moves towards a soft market once more. The conventional term for these repeating market conditions is the property-casualty insurance cycle.

Traditional economic models of insurance pricing in competitive markets do not explain the substantial variation in profitability and prices. In models where equilibrium premiums are determined by the capital asset pricing model or discounted cash flow approach in competitive markets, price changes occur only when there are changes in expected losses or in the systematic risk of losses. Such models might explain changes in the levels of premiums, but not the associated reduction in quantity nor the gradual easing of prices and quantity

¹Insurance can be divided into three branches- health, life and property-casualty (also called property-liability). The discussion here applies only to property-casualty insurance.

restrictions which are observed during the cycle.²

Recently, attention has been given to the hypothesis that property-casualty insurance cycles are caused by collusive behavior among insurers.³ Proponents of this theory contend that the current form of regulation, which partially exempts insurers from antitrust statutes, facilitates collusion among insurers. Because the collusion is not perfectly enforced, the industry is subject to periodic price wars. Alternating periods of collusion and price wars creates the observed cycle of prices and profitability.

This chapter proposes an alternative explanation of insurance cycles. It assumes that insurers operate in a competitive market but are constrained by regulators who impose conditions on firm supply to reduce the likelihood of bankruptcy. Since loss outcomes are uncertain, regulators require that insurers have a minimum financial cushion in excess of expected liabilities. The sources of this cushion of capital are net worth and the difference between price and expected loss. This cushion is related to the industry's capacity in that it defines the maximum quantity of insurance that can be sold.

Due to informational asymmetries between insurers and equity holders or other factors, large additions to net worth are costly. When

²For models using capital asset pricing theory, see Hill [1979] and Fairley [1979]. A discounted cash flow model (an extension of the Myers-Cohn model) is used by Harrington [1988].

³See J. Angoff [1988]. G.K. Bernstein [1988] analyzes both the legal complaints as well as the conspiracy arguments associated with recent antitrust suits against some property-casualty insurers.

events occur that reduce net worth such as unexpectedly high claims or drops in asset values, industry net worth will not immediately adjust. Since net worth is reduced, a higher premium is required to support the same magnitude of expected liabilities. Over time, net worth expands due to the accumulation of retained earnings and capital from outside the firm, and premiums decline toward long run equilibrium until another shock occurs. These events produce the observed cycles.⁴

The remainder of the chapter is organized as follows. In the next section, I review the structure of the property-casualty insurance industry and current regulation. A model of competitive insurance markets with capacity constraints is developed in section III. Section IV summarizes the model's results and suggests directions for further research.

II. INDUSTRY STRUCTURE AND REGULATION

This section provides background on the property-casualty insurance industry and its regulatory structure, which can be used to evaluate the plausibility of alternative insurance cycle explanations. It begins with an overview of property-casualty insurance, identifying buyers and sellers and the important features of property-casualty insurance. Then I discuss current regulation and actions aimed at reducing the

⁴An alternative model with similar implications is presented by Ralph A. Winter [1989]. The model presented has three features which prevent the market from adjusting to long run equilibrium each period. These are: limited liability of insurers, adjustment costs (due to double taxation of dividends) and, like this chapter, the assumption that capital raised from within the firm is less costly to the firm than external funds.

probability of insurer insolvency.

A. THE INDUSTRY

A typical insurance contract specifies a premium to be paid by the buyer in return for promise of payment should certain events occur during the term of the agreement. While premiums are paid at the beginning of the period, claims for the insured events may be settled during or after the policy period has expired.⁵ Types of insurance, such as property damage or liability, vary in the median time between event occurrence and claim settlement. Insurance also may be distinguished by the type of buyer and the type of seller. This section describes these distinctions in more detail.

Property-casualty insurance, as the name implies, can be classified by the type of loss event covered. If the loss is to the insured party, the insurance is known as first-person coverage. Most first-person insurance covers property loss. If a third party is protected from loss, it is called third-person or liability insurance. The average payout period, or tail, of a first-person claim is significantly shorter than that of liability coverages (where detection lags and litigation to determine liability often extend the time between the event that caused the loss and payment by the insurer).⁶

⁵Premiums are not always paid in full at the beginning of the policy period but this assumption simplifies the exposition without loss of generality. This form of contract is known as occurrence-based insurance because the policy pays according to when the event happened, not when the claim is filed.

⁶The following statistics provide an indication of the relative size, in

Property-casualty insurance also can be classified by who buys the coverage. Insurance contracts purchased by households and individuals are known as personal lines. Insurance sold to all others-- firms; professionals; state, local and municipal governments; and nonprofit entities-- is known as commercial insurance. Industry premium volume is almost evenly divided between commercial and personal lines.

Premium volume for the property-casualty industry was approximately \$200 billion in 1988. Industry assets totaled \$477 billion with net worth of \$118 billion.⁷ Although there were about 3600 property-casualty insurers operating in the US in 1986, most of the premium volume was written by about 900 of these firms.⁸ As table 1 shows, concentration as measured by the Herfindahl-Hirschman Index (HHI) is low for the industry and fairly low for most lines of insurance.⁹ As a reference point to interpret these data, note that the Department of Justice merger guidelines consider an industry with a post-merger HHI

terms of premium volume, of first person and third person business in the past two decades. Of the largest lines of liability business, auto liability, other liability, and medical malpractice comprise roughly one-third of total premiums, while automobile physical damage (pure property coverage) accounts for about 16% of premiums. Homeowners multiple peril and commercial multiple peril, coverages that combine property and liability risks, account for about 15% of the industry premium.

⁷ Best's Aggregates and Averages, 1989 edition.

⁸ Insurance Facts [1986-7 edition].

⁹ The HHI by line or by state will understate the true degree of competition because of the large number of potential entrants. Entry by an existing property-casualty insurer into specific lines or states is not very restrictive: if an insurer meets regulatory standards in one state for some lines, it is usually not very difficult to meet minimum capital requirements for other states and/or lines.

less than 1000 unconcentrated to the extent that the merger and its effects on competition are not significant. Entry barriers, discussed in the regulatory section, are low. Property-casualty insurance is structurally competitive: there are many sellers, low concentration and low barriers to entry.¹⁰

TABLE 1

HHI FOR SELECTED LINES, 1980-85

year:	1980	1981	1982	1983	1984	1985
ALL P-C	213	213	215	225	226	229
MED. MALPRACTICE	513	523	500	567	663	622
OTHER LIABILITY	224	219	220	220	236	278

source: Clarke, Warren-Boulton, Smith and Simon (1988).

Most of the larger insurers operate in all or most states and sell several lines of insurance. They differ in the types of insureds they concentrate on, the form of corporate structure and the marketing system.¹¹ The types of buyers targeted-- personal lines versus commercial lines-- are described above. The two major types of firm organization are covered below.

Stock companies are the largest form of firm organization in terms of assets, premium volume and net worth. Like other publicly held corporations, stock companies are owned by their shareholders who receive returns for their investment either through dividends or through

¹⁰See Harrington [1988] and references therein.

¹¹The three ways to market insurance is through a broker, an independent agent, or direct mailings and firm personnel. Commercial insurance is usually sold through a broker or by firm personnel.

stock appreciation. Equity issues to the public can be used to finance company expansion. Stock insurers have traditionally been the most prevalent type of corporate structure; since the 1960s stock insurers have accounted for at least 70% of industry assets, premiums and net worth.¹²

The other common form of corporate structure in the property-casualty industry is the mutual company. Mutual insurers are owned by their policyholders who, in theory, control the corporation by voting for management similar to shareholders in a stock company.¹³ Policyholders may receive their share of profits through dividend issues which effectively lower their premium level. During the period considered here, mutuals' share of industry assets, premium volume and net worth ranged from 20 -25%.

B. CURRENT REGULATION

The insurance industry is regulated at the state level. A primary goal of regulators is limiting the probability of insurer default. To accomplish this most states require companies writing policies in a state to be licensed by the insurance department. The solvency of licensed insurers is regulated and monitored in three ways: by capital and net worth requirements, review of annual statements, and by periodic

¹²Data on premium volume, assets and net worth are from Best's Aggregates and Averages.

¹³It has been suggested that mutual policyholders have more control over mutual insurers than than stockholders have over stock insurers. There exists anecdotal evidence, however, that in some mutuals policyholders have (or at least exercise) very little control, see Tobias [1984].

in-depth examinations. In some states, insurance departments also regulate policy forms and rates, although the primary motivation for this is consumer protection rather than insurer solvency.¹⁴

One way insurance departments promote financial stability of insurers is by setting minimum capital and net worth levels.¹⁵ To be licensed or admitted to write insurance within a given state, an insurer must meet the state's minimum capital requirements for that line of business. Non-admitted insurers may write business within states subject to numerous restrictions for both sellers and buyers.

Depending upon the form of corporate organization and the lines of insurance an insurer wants to sell, the state insurance department sets minimum initial capital and net worth requirements and a minimum standard for continuing operation. While these requirements are not, in general, related to premium volume, they do differ across types of insurance.

Table 2 presents the minimum capital and net worth standards for

¹⁴In general, policy forms are regulated to protect buyers. Rate control has been used in some states, primarily in personal lines; Most states moved toward less restrictive rate regulation in the late 1960s and early 70s. Many states do not have extremely restrictive rate regulation. This chapter is concerned with describing an industry-wide phenomenon, observed at the national level, which changes in state rate regulation do not appear to have substantially altered. The form and effect of rate regulation will not be considered further here.

¹⁵Net worth is usually called surplus or policyholders' surplus in the insurance industry. It is defined as the difference between admitted assets and liabilities. Admitted assets and liabilities are those recognized by regulators under Statutory Accounting Principles. In general this means that nonliquid assets such as office furniture and buildings are not included, stocks are valued at market value and most bonds are valued at amortized book value.

selected lines for companies licensed in Massachusetts. An insurer licensed to sell fire insurance must have at least \$200,000 capital and \$400,000 net worth while a company selling other liability has minimum capital and net worth requirements of \$400,000 and \$800,000. In general, companies selling insurance lines with longer tails and more variable results, such as reinsurance or liability must have greater capital and surplus relative to those selling less variable lines such as fire.

TABLE 2
MINIMUM CAPITAL AND NET WORTH REQUIREMENTS FOR MASSACHUSETTS

INSURANCE SOLD	CAPITAL	NET WORTH	TOTAL
FIRE	200,000	400,000	600,000
OTHER LIABILITY	400,000	800,000	1,200,000
REINSURANCE	500,000	1,000,000	1,500,000
FIRE AND REINSURANCE	700,000	1,400,000	2,100,000
OTHER LIABILITY AND REINSURANCE	900,000	1,800,000	2,700,000

source: Company Licensing Requirements: Massachusetts.
NAII publication.

Insurance departments also monitor insurers' financial health by reviewing information submitted in annual reports. The National Association of Insurance Commissioners (NAIC) has developed a two-part diagnostic procedure to help insurance departments identify insurers in greatest danger of insolvency using information from annual statements. The first stage of the Insurance Regulatory Information System, as the tests are called, consists of calculating eleven ratios that measure

overall solvency, profitability, liquidity and reserve adequacy. In the second stage, insurers who fall outside of established ranges on four or more tests are examined in more detail. Insurers determined to be financially troubled may undergo closer monitoring, submit more frequent financial reports, be placed in regulators' control for rehabilitation, or be declared insolvent and forced to liquidate the company.

The NAIC emphasizes that "valid interpretation of ratio data depends to a considerable extent on the judgment of financial examiners."¹⁶ The analysis does not account for the product mix or change in product mix which may cause the results to fall outside the usual range standard. The NAIC urges further consideration of a company's circumstances to aid in evaluating the firm's results. For example, the most widely cited criterion of solvency regulation, the premium to surplus ratio, has a stated usual range of up to 3, but the NAIC warns that other circumstances may have significant effects on the interpretation of this statistic. Profitability of the business being written, the distribution between property and liability lines and the exposure of a firm to large losses are noted as particularly important.¹⁷

¹⁶From NAIC [1985] p.4.

¹⁷The NAIC report notes that "[i]n general, companies with stable profits are able to sustain a higher ratio of premium to surplus without undue risk than companies with losses or unstable profits" and that "[c]ompanies with a larger portion of premium from Schedule P lines [composed of liability and multiple peril lines] should generally maintain a lower ratio of premium to surplus, due to the greater variability of losses and the difficulty of accurately estimating potential losses for these lines of business." NAIC [1985], p.7.

III. MODEL OF INSURANCE PRICING RESULTING IN CYCLES

This section develops a model of price determination in competitively structured insurance markets with solvency regulation similar to that described above. The model of insurance pricing is presented in three stages. In the first stage I present the simplest case, in which there is only one line of insurance and the mean loss for the risks is known. Stage two expands this framework to insurers that supply several types of insurance. The third stage incorporates the effects of uncertainty about the mean loss in the analysis.

A. Assumptions and set-up

I assume there are initially M firms selling insurance, monitored by a regulator who oversees solvency. The firms operate within a competitive market. The regulator achieves a desired low probability of bankruptcy by specifying a maximum relationship between quantity sold and the firm's net worth, S .¹⁸ Firms are initially endowed with net worth equal to \bar{S} .

Net worth can be adjusted through dividends, retained earnings or equity issues. In the short run, however, large additions to net worth

¹⁸Quantity sold is a confusing measure. It could be the number of policies sold or the number of risks covered. For this discussion, quantity refers to the number of exposure units sold-- for example, automobile insurance is sold at a rate times the number of cars (the exposure unit) covered. For daycare general liability the exposure unit is the number of pupils, for apartment houses liability is sold per 1000 sq. feet, for clothing manufacturers and gas stations the unit is per \$100 payroll (Illinois Department of Insurance, [1986]). The premium quoted by insurers to their policyholders will be the rate times the number of exposures plus administrative costs such as expenses involved in underwriting, taxes and fees. This model simplifies the exposition by ignoring these costs.

are costly. Large additions to net worth are usually desired after periods of relatively low industry profitability and net worth has declined relative to demand. Although firm managers may know that the next period's prices will result in relatively good profitability, they cannot credibly convince equity holders of that fact. Equity holders therefore will require a premium on funds invested in insurance. This information asymmetry causes large additions to net worth to be very costly to insurers. It is assumed that the cost is so high that net worth is fixed in the short run.

Insurers collect premiums, P , at the beginning of the policy term. Although a policy covers only one period, claims arising from any policy may take several periods to be filed and settled. Payments on claims arising from a given policy year may occur from that period forward in time. The pattern of claims payments is known as the loss development, or loss emergence pattern. The average time between the policy period and claim payment is known as the "tail" of the line.

The relevant loss cost at the outset of the policy period is the present value of the mean loss. For now I assume insurers know the present value of the population mean loss for period t , \bar{L}_t , with certainty. This will be relaxed in section IIID. \bar{L}_t is related to \bar{L}_{t-1} by

$$\bar{L}_t = \gamma \bar{L}_{t-1} \quad (1)$$

where γ is a trend factor for expected inflation.

The actual mean loss realized by an individual insurer may differ from the population mean. Each exposure or risk covered for insurance

line k has a mean, $\bar{L}_{t,k}$ and a variance σ_k^2 . There are R_k policyholders distributed uniformly over the M firms. Some firms may have mean losses greater and others less than $\bar{L}_{t,k}$ depending upon the actual losses for their sample of $\frac{R_k}{M}$ ($= r_k$) policyholders. For a given firm, m , actual mean loss $\bar{L}_{t,k,m}$ is described by

$$\bar{L}_{t,k,m} = \bar{L}_{t,k} + \nu_{k,m} \quad (2)$$

where $E[\nu_{k,m}] = 0$ and $\text{var}[\nu_k] = \frac{\sigma_k^2}{r_k}$. For now I will consider only one line of insurance and so drop the k subscript until section IIIC.

Regulators recognize that an individual firm's mean loss may vary from the population mean. To reduce the probability of firm insolvency, regulators require that each insurer has a minimum net worth and a minimum capital above the present value of the expected loss available to back the policies.¹⁹ The "safety margin" available will be the total assets (net worth + premium income) less total liabilities (the present value of expected losses). The model abstracts from expenses, fees, and premium taxes.

In this simple model, only the realization of ν will cause the mean loss for some firms to be differ from the population mean. Since the source of uncertainty is related to characteristics of the line of insurance, the probability that losses will exceed premiums plus surplus also will be related to the specific type of insurance. Lines whose

¹⁹ Insolvency occurs when the size of the losses exceeds the premium income plus net worth (surplus).

loss distributions have large variance (σ^2) relative to the number of policyholders (R) will have more uncertainty of results across firms (var ν).

Let regulators require a minimum safety margin of U dollars per policy above the expected loss to back each policy.²⁰ U is increasing function of the variance of ν , $U(\text{var } \nu)$. Later, in section IIID, the U function will be expanded to include other factors which affect the uncertainty of expected losses. The regulatory supply constraint for the representative firm when one line of insurance is provided is:

$$q_{t,m} U(\text{var } \nu) \leq (S_{t,m} + q_{t,m} P_t) - q_{t,m} \bar{L}_t \quad (3)$$

or
$$P_t - \bar{L}_t \geq U(\text{var } \nu) - S_{t,m} / q_{t,m}$$

where q is quantity, P is premium and S is the level of net worth. Under this set-up, the first-best solution would be a regulated monopoly.²¹ It is assumed, however, that the regulator operates in a second-best world where he is confronted with M firms and cannot impose firm specific constraints.

B. Pricing with one line of insurance

The short run problem for a representative firm in a world where only one line of insurance is produced is to choose quantity to maximize the present discounted value of the expected profits, subject to the regulatory constraint (3).²² The losses of each insured are identically

²⁰ It will be assumed that each policy is equivalent to one exposure.

²¹ The first-best solution is described in Appendix A.

²² Maximizing the expected present value of all future streams of profit

and independently distributed with equation (1) specifying the mean loss for the risks.²³ In the short run, when S is fixed at \hat{S} , the problem for the firm is

$$\begin{aligned} \max_q \quad & qP - PV(E(\sum_{i=1}^q L_i)) - qP - q\bar{L} \\ \text{s.t.} \quad & q U(\text{var } \nu) \leq (\hat{S} + qP) - q\bar{L} \end{aligned} \quad (4).$$

When $P \geq \bar{L}$ the firm will supply policies. For $P = \bar{L}$ quantity may range from 0 to $\frac{\hat{S}}{U(\cdot)}$.²⁴ When $P > \bar{L}$, the firm will want to supply as many policies as possible, so $q = \frac{\hat{S}}{\bar{L} + U(\cdot) - P}$. The short run supply curve described above is depicted in figure 5.²⁵

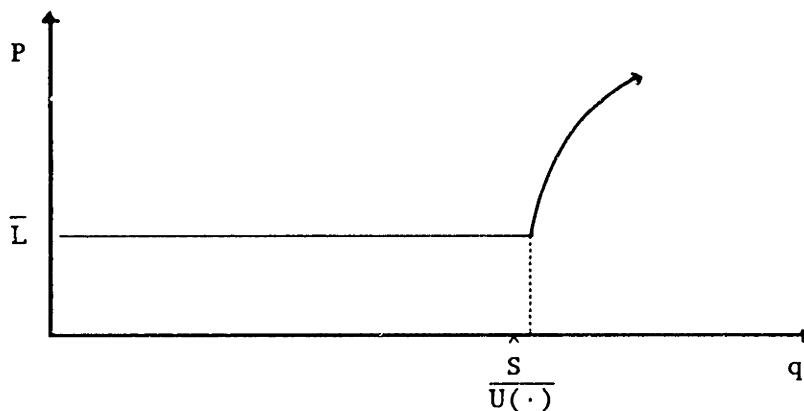
 reduces to maximizing the firm's next period expected profit. Nothing the firm can choose today will affect its future expected profits except the level of net worth it takes into future periods. Maximizing next period's expected profit will maximize the expected net worth available in future periods.

²³This setup assumes that the representative firm cannot distinguish among different risks in a heterogeneous risk group, so it is as if all insureds' have identically and independently distributed loss distributions.

²⁴where $U(\cdot) = U(\text{var } \nu)$.

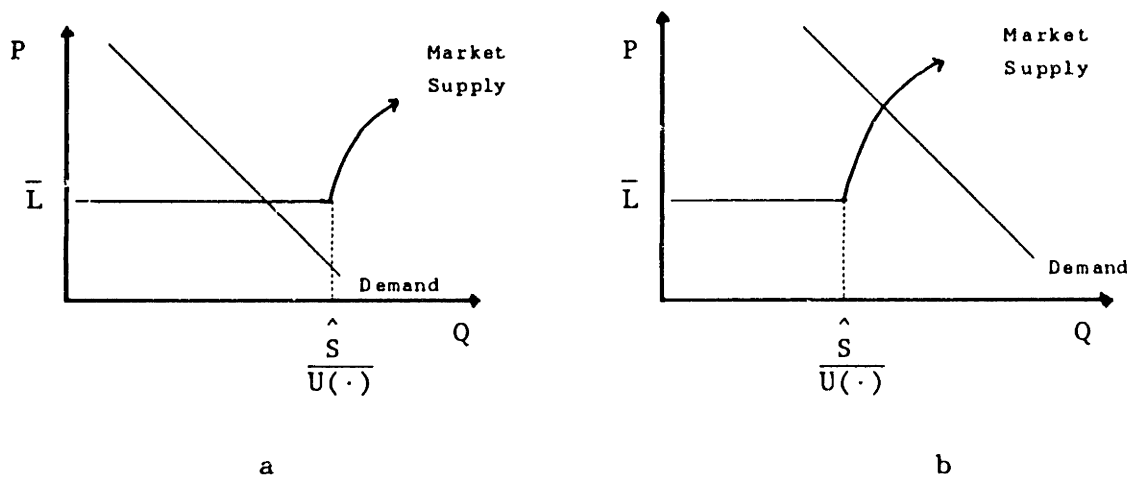
²⁵The model in Winter (1989, described in footnote 4) generates the same shape for the industry supply curve. In Winter's model, however, the perfectly elastic section of the supply curve occurs at a price greater than the mean payout, the quantity at which the supply becomes upward sloping is determined by the limited liability and no bankruptcy assumption, and, unlike the model here, there is a maximum price at which the supply becomes perfectly elastic once more.

FIGURE 5



The premium level is determined by the market clearing condition. The market is in short run equilibrium when market supply, the horizontal summation of firm supply, is equal to market demand. Figure 6 shows the equilibrium prices when there is excess capacity (a) and when the capacity constraint is binding (b). If demand is low relative to capacity, then $P = \bar{L}$ as in figure 6a; if demand is large relative to industry capacity, price will be greater than \bar{L} , reflecting the short run opportunity cost of scarce capacity, as in Figure 6b.

FIGURE 6



An insurer's long run profit is revenue less total cost. Total long run cost includes both losses and the opportunity cost of holding surplus, S , to back the policies. It is assumed that the rate of return that the insurer earns by investing S , r^* , is less than the opportunity cost of the funds, r . Restrictions by regulators on what assets net worth can be invested in means that the investment, S , will earn a lower rate of return than in its best alternative use. Let ϕ denote the difference $r - r^*$.²⁶ The firm will maximize

$$PV (E[\Pi]) = Pq - \bar{L} q - \phi S \quad (5).$$

In the long run insurers will choose S to maximize the present value of profit each period. Since long run profit cannot be maximized unless short run profit also is maximized, firms will operate on their short run supply curves. As firms are attempting to earn the greatest total return from their investment of S , the capacity constraint will be binding.²⁷ Therefore the representative firm's long run problem is:

$$\begin{aligned} \max_s \quad & Pq - \bar{L} q - \phi S \\ \text{s.t.} \quad & P - \bar{L} = U(\text{var } \nu) - \frac{S}{q} \end{aligned} \quad (6).$$

²⁶This assumption is also made by Joskow [1973]. This condition guarantees that the constraint will bind in some periods. Without this assumption, there is no cost to holding excess capital and firms will hold excess funds to avoid having too little capital in the future.

²⁷The capacity constraint must bind in the long run for the firm to earn the largest possible return from a given level of S . For example, when the constraint does not bind the firm could have supplied the same quantity while maintaining a lower surplus. At a lower surplus the cost of maintaining that net worth, ϕS , is lower and so profits will be greater.

In addition to this condition, long run competitive equilibrium for the industry under a free entry and exit assumption implies that expected profit is zero. The long run equilibrium is defined by the price and capacity combinations for which

$$\partial(PV(E[\Pi]))/\partial S = 0 \quad (7a)$$

and

$$PV(E[\Pi]) = 0. \quad (7b).$$

Using the fact that q is a function of S , substituting the constraint into the equation and maximizing over S yields the first order condition for long run maximization:

$$P - \bar{L} = \phi (\bar{L} + U(\cdot) - P). \quad (8)$$

The optimal size of the firm, then, is indeterminate, a usual feature of models with constant returns to scale cost functions.

In a competitive industry with constant costs, the optimal size of the firm is indeterminate, the optimal quantity supplied by each firm will be defined by the level of net worth if $P > \bar{L}$, and will again be indeterminate when $P = \bar{L}$. Premium level will be determined by equating market supply and demand. Since the aim here is to discuss pricing of insurance, no generality is lost if we consider the industry as one large price-taking firm operating under conditions of a competitive market where S will now denote total industry net worth.

Condition (7b), that the market is in long run equilibrium, is satisfied when $(P - \bar{L})q - \phi S = 0$, which holds when equation (5) is met. From equation (8), an expression for the long run equilibrium

price is found to be

$$P^* = \bar{L} + \frac{\phi}{1+\phi} U(\cdot) \quad (9)$$

where P^* = long run equilibrium price.

Since $\frac{\phi}{1+\phi} U(\cdot) > 0$, the long run premium will be greater than the present value of the population mean loss. The second term on the right hand side of (9) arises because insurers incur a cost for holding net worth to back policies. Insurers must recoup this cost in the long run if they are to continue writing insurance. The long run premium, although it is above the mean losses alone, is not above the long run cost of supplying insurance incurred by the firm which include the cost of capital and is consistent with zero expected profit.

Changes in the mean loss are not the only source of change in the long and short run premiums. If any of the variables that change the present value of the mean loss are altered, the premium will reflect that change. If the interest rate rises but all other factors remain constant, the short and long run equilibrium premiums will fall. The effect of an increase in the tail, the number of years after the policy period where nonzero losses are paid, will also decrease \bar{L} .

This model generates an equilibrium long run price for the market based on mean loss and the opportunity cost of committing net worth to insurance, ϕ .²⁸ The firm's short run supply curve is a function of these variables as well as the short run level of surplus. In the long run

²⁸The long run equilibrium price is independent of demand due to the constant return to scale specification.

insurers can choose S optimally, but the optimal size of the firm is indeterminate.

C. Firms selling multiple lines of insurance

Most insurers do not just supply one type of coverage, but several. Consider an insurer who supplies two types of insurance, lines 1 and 2, to illustrate how providing several types of insurance alters the previous results. Both lines, as in the previous model, are composed of identical policyholders. Each line has mean loss \bar{L}_k , quantity q_k , and premiums P_k (where $k = 1, 2$).

The representative firm's maximization problem in the short and long run is:

$$\begin{aligned} \max \quad & P_1 q_1 - \bar{L}_1 q_1 + P_2 q_2 - \bar{L}_2 q_2 - \phi S \\ \text{s.t.} \quad & S_1 + S_2 = S \end{aligned} \tag{10}^{29}$$

$$S \geq (\bar{L}_1 + U_1(\cdot) - P_1) q_1 + (\bar{L}_2 + U_2(\cdot) - P_2) q_2$$

The long run equilibrium prices will be the same as equation (9) because the production technology is constant cost and there are no scope economies in this model,

$$P_1^* = \bar{L}_1 + \frac{\phi}{1+\phi} U_1(\cdot) \qquad P_2^* = \bar{L}_2 + \frac{\phi}{1+\phi} U_2(\cdot). \tag{11}.$$

²⁹ In the short run the firm will maximize over q_1 and q_2 with S fixed at \hat{S} . In the long run, the firm will maximize over S , as well. S_1 and S_2 the amount of capital allocated to each line, are not explicitly calculated by the insurer. Explicit S_1 and S_2 terms help highlight the relationship between capital and quantity supplied for each line.

In the short run, when the firm is not capacity constrained, equilibrium premium equals \bar{L}_k . The interesting case is the short run equilibrium when the regulatory constraint binds, where the firm must allocate net worth between the two lines. It is straightforward to show the maximization of (10) with S fixed implies that when the industry is capacity constrained the equilibrium return per unit net worth is equated across lines.

$$(P_1 - \bar{L}_1) \partial q_1 / \partial S_1 = (P_2 - \bar{L}_2) \partial q_2 / \partial S_2 \quad (12)$$

Since the change in quantity sold from a change in the net worth dedicated to that line ($\partial q / \partial S$) is independent of S, price must adjust to equate these expressions.

The relation between P_1 and P_2 that must exist for short run equilibrium can be found by manipulating the first order conditions. This yields:

$$\frac{(P_1 - \bar{L}_1)}{(P_2 - \bar{L}_2)} = \frac{\bar{L}_1 + U(\text{var } v_1) - P_1}{\bar{L}_2 + U(\text{var } v_2) - P_2} \quad (13).$$

The right hand side of equation (13) is the ratio of $\partial S_1 / \partial q_1$ to $\partial S_2 / \partial q_2$. The relative markup will be larger for the line of insurance whose incremental exposure requires more capital. The first order conditions also yield expressions for short run premiums and the shadow price of capacity, λ :

$$P_1' = \bar{L}_1 + \frac{\lambda}{1 + \lambda} U_1(\cdot), \quad P_2' = \bar{L}_2 + \frac{\lambda}{1 + \lambda} U_2(\cdot),$$

and
$$\lambda = \frac{q_1(P_1 - \bar{P}_1) + q_2(P_2 - \bar{P}_2)}{S} \quad (14).$$

When $\lambda = \phi$ these prices are simply long run equilibrium prices. When $\lambda > \phi$, short run prices are above the long run equilibrium.³⁰ The difference between long and short run price, $P'_k - P_k^*$, is a function of U_k , λ , and ϕ . The change in price is $(\frac{\lambda}{1+\lambda} - \frac{\phi}{1+\phi})U_k$. Since U is a function of the sample variance only, the line with the greater sample variance will have larger price increases.

The above discussion looks at the short run equilibrium when there are two lines of insurance assuming that both lines are supplied. The model gives insight into pricing under these assumptions but cannot determine whether $S_1 > S_2$ without more information. Using the equations for short run prices and the conditions for capacity constraints to be binding, it is found that

$$\frac{q_1}{q_2} = \frac{S_1/U_1(\cdot)}{S_2/U_2(\cdot)} \quad (15).$$

The ratio of the quantity supplied is equal to the ratios of dedicated net worth to the required capital cushion, U , for each line.³¹ If $S_1/U_1(\cdot) = S_2/U_2(\cdot)$, then the quantities are equal. Otherwise, so long as $S_1/U_1(\cdot)$ and $S_2/U_2(\cdot)$ are constant, quantities for both lines will

³⁰To be accurate, if demand for insurance is growing S does not need to decrease in order to go from a long run equilibrium to a situation where the capacity constraint binds, all that is needed is that S increases by less than is required to maintain long run equilibrium.

³¹This condition will also be true in the long run equilibrium.

fall proportionately. Therefore the line with the largest long run equilibrium quantity will realize the greatest absolute reduction in quantity when the capacity constraint binds. If the right hand side of equation 15 is not constant, then more information is needed to determine anything about q_1 , q_2 and the relative changes in quantity for the two lines.

There is also the possibility that P_1 and P_2 which solve (14) are not equilibrium prices. It might be that demand for either or both lines is zero at that price. In such cases, gains from trade have disappeared and no insurance will be sold.

D. Uncertain Population Means

Until now the discussion has assumed that the mean loss for various risks is known by the insurers at the time insurance is sold. In reality, however, the policy is sold before the actual costs are observed and insurers must base decisions on the expected mean loss.³² Insurers estimate the expected mean loss from information on past claims, trends in costs and claims, and any new information that is relevant.³³

In this section it will be assumed that the expected population mean loss per exposure in line k for policy year t , $\bar{L}_{t,k}$, is related to

³²Expectation is used to refer to the conditional expectation, mean will refer to the unconditional expectation.

³³Relevant information may include changes in the risk environment due to tort reform legislation or a court ruling that the existing tort reform legislation is unconstitutional.

$\bar{L}_{t-1,k}$ in the following manner:

$$\begin{aligned}
 E[\bar{L}_{t,k}|\Omega_t] &= \gamma E[\bar{L}_{t-1,k}|\Omega_t] & (16) \\
 &= \gamma (E[\bar{L}_{t-1,k}|\Omega_{t-1}] + \epsilon_{t-1,k}) \\
 &= \gamma E[\bar{L}_{t-1,k}|\Omega_{t-1}] + \gamma \epsilon_{t-1,k}
 \end{aligned}$$

where ϵ is a random variable shock to the loss distribution with $E[\epsilon_{t,k}|\Omega_t] = 0$ and Ω_t is the set of all information known at the beginning of period t .³⁴ γ is still a trend factor. The mean expected loss for year t will be the mean expected loss for year $t-1$ (as of the beginning of that period), plus the effect on the loss distribution of new information revealed between $t-1$ and t , adjusted for trends in costs and claims. The actual population mean loss may differ from its expectation depending upon the realization of ϵ in current and future periods. Any unpredictable changes in the underlying loss distributions (which affect the mean) will be captured by the realizations of ϵ .

Nonzero realizations of ϵ are assumed to affect all unsettled claims. Let β_1 be the proportion of expected losses not settled i years after the policy period. For policies written in year t the change in

³⁴The realization of $\epsilon_{t,k}$ becomes known at the end of period t , so

$E[\bar{L}_{t-1,k}|\Omega_t] = E[\bar{L}_{t-1,k}|\Omega_{t-1}] + \epsilon_{t-1,k}$.
 The two random variables, ν_t and ϵ_t are orthogonal. The realization of ϵ_t affects the mean loss for the population, hence industry, and affects the mean loss for the firm. The realization of ν_t affects only the firm mean, not the population mean.

expected losses will be $\beta_0 q_t (E[\bar{L}_t | \Omega_{t+1}] - E[\bar{L}_t | \Omega_t]) = q_t \epsilon_t$, where $\beta_0 = 1$. For policies written in year $t-1$, $\beta_1, \beta_1 \leq \beta_0$, of the expected losses are unsettled at the end of period t . The change in expected losses will be:

$$q_{t-1} \beta_1 (E[\bar{L}_{t-1} | \Omega_{t+1}] - E[\bar{L}_{t-1} | \Omega_t]) = q_{t-1} \beta_1 \epsilon_t \quad (17).$$

Summing over all prior years, the total change in expected claims at time t will be $\sum_{i=0}^{\infty} q_{t-i} \beta_i \epsilon_t$ where $\lim_{i \rightarrow N} \beta_i \rightarrow 0$ (it is assumed that policies from N years back are no longer active). Long tailed lines will have large N relative to short tailed lines. Thus the cumulative addition to expected claims from $\epsilon_t > 0$ will be larger for long tailed lines, *ceteris paribus*.³⁵

There are now three sources of uncertainty that may cause a firm's actual mean losses to differ from expected population mean losses at the time the policy is written. Two of these are incorporated in ϵ , unanticipated changes in the population mean, from unexpected changes in inflation or other trend factors, or in the loss distribution itself.³⁶ The third source of deviations from the expected value is the realization of ν , the difference between the population mean and the realized sample mean. All these sources of uncertainty are related to characteristics of the line of insurance, so the probability that losses

³⁵ β_{ik} and N_k are assumed to be fixed.

³⁶ The fact that trends in claims and costs are uncertain could be illustrated by making γ an expectation rather than a constant. In this chapter, however, any unexpected effects of inflation or other trends are captured in ϵ .

will exceed premiums plus net worth also will be related to the specific types of insurance provided.

The regulator will alter the required margin, U , to account for these sources of uncertainty. In addition to being a function of ν , U also will be an increasing function of the policy limit and the loss tail, N . An insurer can reduce exposure to adverse realizations of both ϵ and ν by setting lower policy limits. The expanded regulatory constraint will be written as $U(\text{LIMIT}, N, \text{var } \nu)$.³⁷ All the previous results concerning short and long run equilibrium premiums are unchanged. The differences are that \bar{L} is replaced with $E(\bar{L}_t | \Omega_t) - \gamma E[\bar{L}_{t-1} | \Omega_{t-1}] + \gamma \epsilon_{t-1}$ in the relevant equations and that the U function includes two additional arguments.

V. SUMMARY AND APPLICATIONS

The model presented in section III demonstrates that the market conditions characterizing property-casualty insurance cycles can be generated in a competitively structured market. The key assumptions for this result are regulators who require a minimum level of capital backing each unit of insurance and costs to increasing net worth from sources external to the firm.

The model also implies that, when the industry becomes capacity constrained, premium increases will vary across lines. Lines requiring more financial backing per unit sold will have larger premium increases

³⁷Where $0 < U() < 1$, $\frac{\partial U}{\partial n} > 0$; $\frac{\partial U}{\partial(\text{var } \nu)} > 0$; and $\frac{\partial U}{\partial \text{LIMIT}} > 0$.

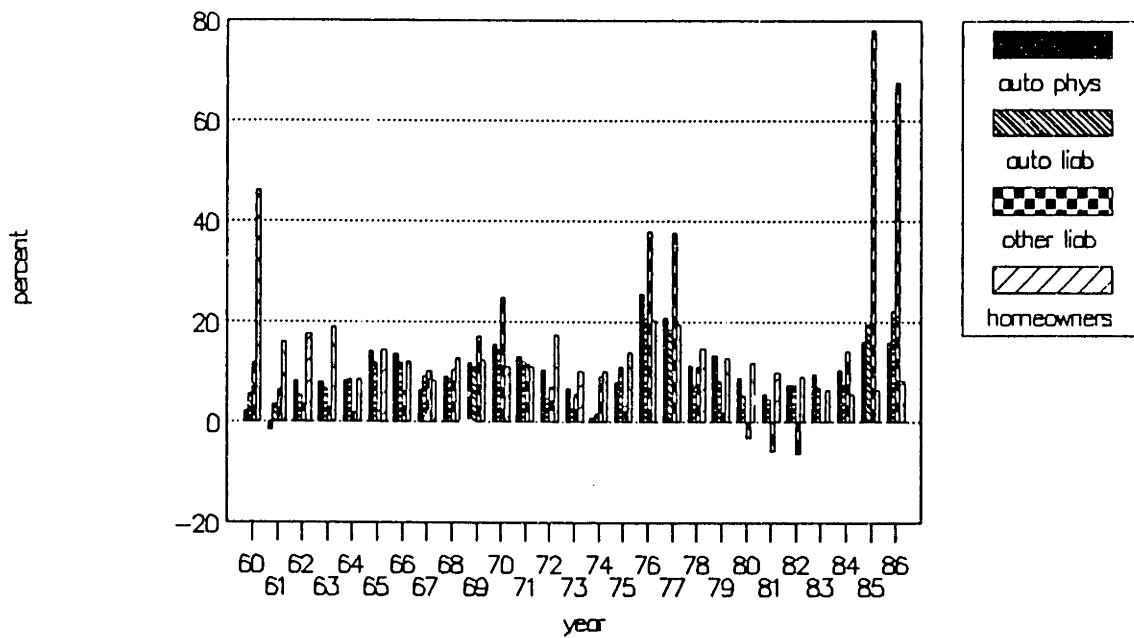
as will lines with increases in expected losses. This is consistent with data on premium increases from the two most recent periods of rapidly increasing insurance prices, 1976-77 and 1985-86. As figure 7 demonstrates, lines with longer loss tails and greater heterogeneity among risks such as other liability has larger premium growth. During the mid 1980s liability lines, and commercial liability lines in particular, were perceived as having much greater variance and longer loss tails which is consistent with the relatively higher premium growth observed in liability lines during the 1985-6 episode relative to the 1976-77 experience.³⁸

The results also are consistent with the observed price increases and availability problems observed within commercial lines during 1982-1986. In a survey of 232 firms, Weber [1987] found that the limits on firms' primary product and professional liability insurance coverage were relatively constant over the 1982-86 period but the premiums paid for that coverage increased.³⁹ In contrast, the observed level of expanded coverage, where firms pay an additional premium for limits

³⁸Clarke, Warren-Boulton, Smith and Simon [1988] provide some evidence suggesting accounting loss and profit data from liability lines is more variable.

³⁹The sample of 232 companies was composed of "major U.S. manufacturing, trade and service" companies with annual sales in excess of \$100 million, excluding financial firms. The survey reported the percent of firms with coverage limits in several intervals. The proportion of firms with primary coverage limits in the various intervals was relatively stable. The survey also reported ranges of premium payments. From 1982-1984 15-16% of firms paid less than \$25,000 for primary coverage, while in 1986 only 7% of firms fell into that category. In 1986 32% of firms paid between \$100-500,000 and 36% paid more than \$500,000 for primary coverage compared to an average 28% and 18% respectively for the 1982-84 period.

FIGURE 7
% CHANGE IN PREMIUM VOLUME
1960-1986, selected lines



above primary coverage, declined over the period.⁴⁰ As shown in table 3, more than 65% of firms had excess coverage limits under \$50 million in 1986 compared to an average of less than 39% in the preceding four years. At the same time, the proportion of firms paying more than \$500,000 for expanded coverage rose to 48% from the 8-9% range in 1982-84. These figures suggest that premium increases and availability

TABLE 3

PRODUCT (OR PROFESSIONAL) LIABILITY INSURANCE
EXPANDED COVERAGE SURVEY RESULTS

year	1982	1983	1984	1985	1986
percent of respondents with limits:					
< \$25 million	25	22	19	20	41
\$25-49.9 million	17	17	16	18	26
\$50-99.9 million	26	24	26	27	17
>\$100 million	26	31	33	29	11
NO COVERAGE	5	4	4	4	6
percent of respondents with payments:					
< \$25,000	20	20	17	10	3
\$25-49,999	18	18	17	9	6
\$50-99,999	18	21	22	15	5
\$100-499,999	22	21	23	34	23
> \$500,000	8	9	9	20	48
NO RESPONSE	14	12	11	11	15

source: Weber [1987]

⁴⁰Coverage above the usual primary policy limits is usually called excess, or surplus, coverage.

problems were more severe in the higher limit segment of the market.⁴¹ Higher limit policies generally have longer tails and more exposure to unexpected inflation, either from changing liability standards or from economic costs, than do lower limit policies. When the market still exists for coverage, the analysis here predicts that insurance for greater limits will have larger absolute increases in price, as observed.

Prior insurance market disruptions, where the problem lines were not commercial liability coverages, also are consistent with this framework. In the mid 1960s, for example, the lines most affected by large price increases and quantity restrictions were inner city property insurance and auto insurance for high risk drivers. During this period commercial liability was small in terms of premium volume, and had relatively stable losses while auto insurance losses were rising quickly and urban riots made inner city property coverage very uncertain.⁴²

Finally, the model can also be employed to evaluate the effects of tax reform on property-casualty insurance. The major changes that affect property-casualty insurers increase taxable income by decreasing allowable losses, and by adding part of the income from premiums paid in advance to revenues, reducing the exclusion for stock dividends, and taxing 15% of some previously tax-exempt income. For example, prior to

⁴¹There is not enough information to determine whether the observed reduction in excess limits is due to very high prices for higher limits or higher limits not being offered.

⁴²Insurance Facts, 1966 edition.

the 1986 tax reform insurers offset revenues with undiscounted losses; many insurers paid essentially no taxes. Under the new tax structure losses must be discounted to present value, increasing taxable income. Insurers who still have very low taxable income will be subject to a tighter alternative minimum tax. The model can be expanded to include taxes and the effect of tax reform on short and long run premium.

Although most economic work on pricing in competitive insurance markets does not explain the empirical regularity of price and profit cycles, the structure of the industry makes collusion theories implausible. This chapter presents a theory of pricing in a competitive insurance industry where cycles are the result of shocks to capacity which cannot immediately adjust. The theory explains why some lines experience greater premium increases and limit reductions than others.

Appendix A

The first-best solution to the model with independent losses is a regulated monopoly. When losses are correlated, either within lines or between lines, the first-best optimum depends upon the relative size of the parameters. This appendix describes the optimal industry structure in the case where losses, across and within lines, are independent and describes the relevant trade-offs when losses are correlated.

The regulator's goal is to maintain a high degree of solvency. To attain this objective, insurers are required to hold assets in excess of expected losses. In the text, it is assumed that determining the distribution of the mean loss and setting a minimum excess-of-loss capital level for each firm is too costly because such computations will vary with the quantity supplied of each type of insurance as well as the lines of insurance sold. Therefore, regulators require a minimum amount of funds excess of expected losses per unit sold in each line, denoted by U .

In this appendix regulators monitor solvency by defining the cushion, U , on the characteristics of aggregate firm losses. Each loss is assumed to be distributed $N(\mu, \sigma^2)$. Since there is a flow cost, ϕ , to holding excess capital to back policies, the first-best solution will minimize the capital which backs policies for a given level of solvency. This is analogous to minimizing the variance of the mean aggregate loss. First I consider the case of independent losses, then the case of correlated losses.

Consider the case where there is only one line of insurance and losses are independently distributed. Let L_i , distributed $N(\mu, \sigma^2)$, describe the loss of policy i . R is the total number of insureds and M is the total number of firms. If insureds are evenly distributed across firms, each firm's mean loss is distributed $N(\mu, \frac{\sigma^2}{n})$ where $\frac{1}{n} = \frac{R}{M}$. U , the cushion needed to maintain a given degree of solvency, is an increasing function of the variance, $\frac{\sigma^2}{n}$. U is minimized when $\frac{\sigma^2}{n}$ is minimized, which is achieved at $M = 1$.

The same result holds in the case of two or more uncorrelated lines of insurance under most conditions. Consider the case of two lines, line 1 and 2. Losses from line 1, L_{11} , are distributed $N(\mu_1, \sigma_1^2)$ while losses from line 2 are distributed $N(\mu_2, \sigma_2^2)$. The firm has n_1 line 1 policies and n_2 line 2 policies.⁴³ The firm's aggregate mean loss will be distributed $N\left((\mu_1 + \mu_2), \frac{1}{(n_1 + n_2)}(\alpha \sigma_1^2 + (1-\alpha) \sigma_2^2) \right)$, where $\alpha = \frac{n_1}{n_1 + n_2}$. The variance of the firm's aggregate mean loss is decreasing in quantity when σ_2^2 is relatively close in size to σ_1^2 or the quantities sold are not far apart. As total quantity increases, the decrease in variance from an additional policy sold becomes very small.⁴⁴

⁴³ R_k is the total number of insureds buying line k policies, then $\frac{1}{n_k} = \frac{R_k}{M}$.

⁴⁴The condition that $\partial \text{variance} / \partial (\text{quantity line } i) < 0$ becomes $n_i \sigma_i^2 + 2n_j \sigma_j^2 > n_j \sigma_j^2$. This can be rewritten as $\frac{n_i}{n_j} + 2\frac{\sigma_j^2}{\sigma_i^2} > 1$ which will always be true if $\sigma_j^2 \geq \frac{1}{2}\sigma_i^2$ or when n_i is close to n_j .

The optimal industry structure depends upon σ , n , and the correlation between lines, ρ , when losses are positively correlated.⁴⁵

In this situation, with one line of insurance, the distribution of the firm's mean loss is $N(\mu, \frac{\sigma^2}{n} + \frac{(n-1)}{n}\rho)$. Whereas an increase in quantity unambiguously reduced the variance of the firm mean when losses were independent, there are now two offsetting effects from increasing quantity, as shown in equation A1.

$$\frac{\partial \text{variance}}{\partial n} = \frac{\rho}{n^2} - \frac{\sigma^2}{n^2} \quad (\text{A1})$$

If $\rho > \sigma^2$ then increasing quantity increases the variance. It is precisely this type of risk that is not "insurable" in the sense that insurance works because aggregating risks reduces the variance of the sample mean. In the "insurable" case where the optimal $\rho < \sigma^2$, increasing quantity decreases the variance.

When there are several lines of insurance the outcome will depend upon the relative sizes of line's variance and its correlation with other lines. The first-best market structure will depend upon the actual parameters of the problem.

⁴⁵The case of negative correlation within or across lines is straightforward. In this case the variance reducing effect of increasing quantity is reinforced by the negative correlation.

REFERENCES

- A.M. Best Company Best's Aggregates and Averages Property/Casualty edition.
- Angoff, J., "Insurance Against Competition: How the McCarran-Furgeson Act Raises Prices and Profits in the Property-Casualty Insurance Industry," Yale Journal on Regulation, summer 1988.
- Bernstein, G.K., "Antitrust Agnostics," Best's Review (P/C edition) October 1988.
- Clarke, Richard N., Frederick Warren-Boulton, David D. Smith and Marilyn J. Simon, "Sources of the Crisis in Liability Insurance: An Economic Analysis," Yale Journal on Regulation, Summer 1988.
- Fairley, William B. "Investment Income and Profit Margins in Property-Liability Insurance: Theory and Empirical Results," Bell Journal of Economics, spring 1979 p. 192-210.
- Gottheimer, George M. Jr. "Crisis in Confidence," Crisis Avoidance: Insurance Responsibilities CPCU 1988.
- Harrington, Scott E. "Prices and Profits in the Liability Insurance Market," in Robert Litan and Clifford Winston, eds. Liability: Perspectives and Policy. Brookings. 1988.
- Harrington, Scott E. and Robert Litan "Causes of the Liability Insurance Crisis" Science February 1988 p. 737-741.
- Hensley, Roy F. Competition, Regulation and the Public Interest in Nonlife Insurance, 1962.
- Hill, Raymond D. "Profit Regulation in Property Liability insurance," Bell Journal of Economics, spring 1979 p. 172-191.
- Illinois Department of Insurance, Report on the Insurance Crisis, John E. Washburn, director, April 1986.
- Insurance Information Institute, Insurance Facts (insurance fact book) various years.
- Joskow, Paul L., "Cartels, Competition and Regulation in the Property-Liability Insurance Industry," Bell Journal of Economics autumn 1973.
- Michelbacher, G. F. and Nestor Roos, Multiple-Line Insurers, 1970.
- Myers, Stuart C. and R. A. Cohn, "Insurance Rate of Return Regulation

and the Capital Asset Pricing Model,"

Myers, S. and N. Majluf, "Corporate Financing and Investment Decisions when Firms Have Information That Investors Do Not Have," Journal of Financial Economics, vol 52, 1984.

National Association of Insurance Commissioners Using the Insurance Regulatory Information System Property and Liability Edition, 1985.

Priest, George L. "The Current Insurance Crisis and Modern Tort Law," Yale Law Journal, vol 96 June 1987 p. 1521-90.

Risk and Insurance Management Society, Inc., 1989 Insurance Availability Survey, management report, 1989.

Stewart, Barbara D. "Profit Cycles in Property-Liability Insurance," John D. Long and Everett D. Randall (eds) Issues in Insurance vol 2 1984 American Institute for Property and Liability Underwriters, Inc. USA.

Strain, Robert, W. (ed.) Property-Liability Accounting, 1986.

Tobias, Andrew, The Invisible Bankers, 1984.

Tort Policy Working Group, Report on the Causes, Extent and Policy Implications of the Current Crisis in Insurance Availability and Affordability, Department of Justice, Inter-Agency report, Feb. 1986.

-----, An Update on the Liability Crisis, Department of Justice, March 1987.

Troxel, Terrie and Cormick Breslin, Property-Liability Insurance Accounting and Finance, 1983.

Weber, Nathan Product Liability: The Corporate Response. The Conference Board, report #893, 1987.

Winter, Ralph A., "The Liability Crisis and the Dynamics of Competitive Insurance Markets," Yale Journal on Regulation vol 5, 1988, p. 455-500.

-----, "The Dynamics of Competitive Insurance Markets," mimeo, University of Toronto, July 1989.

CHAPTER TWO:

Capacity and Property-Casualty Insurance Cycles: Time Series Evidence

I. INTRODUCTION

In chapter one, I developed a model of price determination in competitively structured insurance markets with capacity constraints. In this chapter, I analyze industry time series data for the period 1947-1987 in the context of the capacity constraint hypothesis. Three issues are addressed. First, a definition of property-casualty insurance cycles based on capacity data is compared to the traditional criterion based on accounting profitability. Next, the behavior of other time series, in particular the relationship between low capacity and high premium growth, is examined. Finally, I analyze capacity and premium growth data to determine whether they are consistent with the types of adjustment costs assumed in the capacity constraint hypothesis.

The timing of cyclical episodes identified using a capacity-based criterion is found to coincide closely with those identified using the more traditional accounting profit criterion in data from 1950-87. Such a finding is of interest since the prior work on the role of capacity in insurance cycles has relied on profit based criteria to define the cycles. Since these theories interpret changes in profitability as caused by changes in capacity, a criterion based on capacity is preferable because it is more direct. This work shows that cyclical periods identified using this alternative criterion do not substantially alter the cyclical episodes found.

Examination of the behavior of other aggregate time series finds

support for the capacity constraint hypothesis. Periods of unusually low capacity growth coincide with or precede episodes of usually high premium growth. The low frequency of the annual data makes exact timing of the two events impossible, but statistical analysis shows that the probability of unusually high premium growth is significantly larger when observed capacity is unusually low in the preceding year. This relation describes the data particularly well when the 1970s period is omitted.

Finally, the relative magnitudes of adjustment costs associated with increasing or decreasing net worth were considered. The path of capacity adjustment is found to be asymmetric. Analysis of the distribution of annual growth in capacity and premium volume shows that the industry adjusts faster to positive than to negative capacity shocks. When equity from external sources is more costly than equity generated from within firm activities, insurers will not issue sufficient equity to return quickly to long run equilibrium after negative capacity shocks. Instead, equity will be issued until the expected return of within the firm is equal to the cost of externally generated funds. The rest of the adjustment will come from the accumulation of retained earnings over time.

The cost differential between internal and external equity can be explained by an adaptation of the Myers and Majluf [1984] model to the insurance industry. In this application, managers are better informed about the true magnitude of net worth, while investors base their decisions on the expected value of insurer net worth. Since insurers

with larger unacknowledged losses are more likely to issue equity, there is an equilibrium where insurers with lower unexpected losses may forego profitable expansion of their insurance activities. The implied cost from this information asymmetry may be greater during periods when large unanticipated losses reduce capacity because of the greater probability that an insurer has additional losses that have not been incorporated into the current financial statement. This model also explains the uncharacteristic delay between low capacity and high premium growth observed in the mid 1970s. A substantial portion of the reduction in capacity that occurred in 1973 and 1974 was due to an unprecedented drop in stock market values. This type of a shock will be fully incorporated into reported net worth and will not cause any information asymmetry between managers and investors. As a result high growth in premium volume is not observed until 1976, after an episode of unusually high growth in financial liabilities in 1975.

The chapter is organized in six sections. Section II reviews the standard definition of property-casualty insurance cycles and the cycles it defines during the period 1950-1987. The next section describes capacity measures and compares cycles identified by this criterion with the cycles determined by accounting profitability. Behavior of other aggregate time series data are presented in section IV. Section V describes the explanations of the cost differential between internal and external capital and tests whether adjustment is symmetric with respect to the sign of capacity shocks. An example illustrating how the Myers and Majluf asymmetric information model applied to insurers can explain

the observed asymmetry is included at the end of section V. Summary and conclusion are in section VI.

II. BACKGROUND: ACCOUNTING PROFITABILITY CYCLES AND THEORIES

A typical property-casualty insurance cycle, measured from trough to trough in accounting profit, is usually characterized as having four phases, as described in chapter one and reviewed here. The crisis or hard market marks the transition from low profitability. There are substantial premium increases, reduced availability and improving accounting profitability, often occurring quite rapidly. A period of relatively low availability, high premium levels and high accounting profitability follows. The third stage is marked by a gradual erosion of high premium and high accounting profitability; premium levels decline, profitability declines and availability restrictions ease. Low premiums, low accounting profits and relatively abundant quantity characterizes the soft market which is the fourth and final stage of a typical cycle.

Both price and quantity are unobserved over the cycle and are somewhat imprecise terms in the context of insurance. What are referred to as quantity restrictions, or availability problems, have several dimensions. Some insureds may be offered coverage with lower limits and higher deductibles; while other insureds may have their policies canceled (or not renewed) and find coverage only in the nonadmitted or

excess and surplus market.¹ Although these features of the cycle are often discussed as quantity reductions they reflect price effects as well. An insurance policy with lower limits and higher deductibles for the same premium is coverage at a higher price. Easing quantity restrictions often mean that previously excluded insureds can find coverage again in the admitted market. This is equivalent to a substantial decrease in price.²

The usual method of identifying insurance cycles over a specified time period is from trough to trough in accounting profit. In property-casualty insurance accounting profits are often expressed in ratio form. Two commonly reported ratios are the loss ratio and the combined ratio. Both will be used in this analysis. The loss ratio is defined as incurred losses divided by earned premiums. It is the undiscounted estimate of expected losses from premiums earned in this period as well as revisions to past expectations normalized by premiums earned in the accounting period. The combined ratio is the loss ratio

.....
¹Nonadmitted insurers are those not licensed to write insurance within a given state. There are restrictions on buyers of nonadmitted insurance, usually they must show that no admitted insurer could provide the desired coverage. If commercial buyers meet certain size criteria, they are exempt from these restrictions. Excess and surplus insurers are insurers who sell primarily umbrella and specialty coverages.

²Actual cycles vary in length and in price and quantity effects. The last two cycles, beginning in 1975 and 1985, had particularly high impact on liability lines. In 1975 the price and quantity changes for professional malpractice lines were particularly large, while in 1985 commercial liability lines were most severely affected. Coverage for some perils, such as pollution liability, were withdrawn from the market and have not returned to any appreciable extent. It is the opinion of this author that complete withdrawal of coverage is caused by additional factors, beyond those which underlie the typical repeating cyclic phenomenon. These issues are beyond the scope of this chapter.

plus the expense ratio, which is calculated as total underwriting expenses divided by written premiums. Both ratios measure costs relative to revenues and are commonly reported as percent. To make the profit ratios increase and decrease with profitability, the loss ratio and the change in the both ratios are multiplied by negative one before they are graphed. The combined ratio is graphed as (1 - combined ratio) in percent; this is just a normalization of the statistic around 100, the value where accounting inflows equal accounting outflows.³

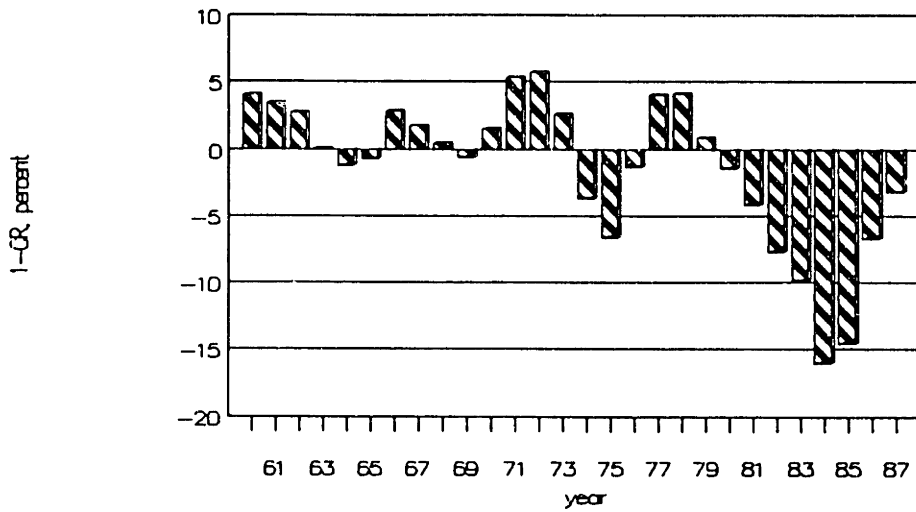
Figure 1a shows industry aggregate profit over the period 1960-1987, measured by (1 - combined ratio), in percent. Figure 1b shows the inverse loss ratio for the same period.⁴ The same statistics are shown in figure 2 for stock insurers only, where data for a longer time period, 1945-1987 are available. In both figures the loss ratio has a more pronounced time trend than the combined ratio. The inverse combined ratio does not reflect the downward trend found in the inverse loss ratio because that trend is offset by an upward trend in the expense ratio, demonstrated in figure 3. The trends in both series are

.....
³Refer to the appendix on accounting profits for further discussion of insurance profit data.

⁴Note that Stewart [1984] indicates that cyclic episodes in data with investment income included are quite similar to those found in data without investment income.

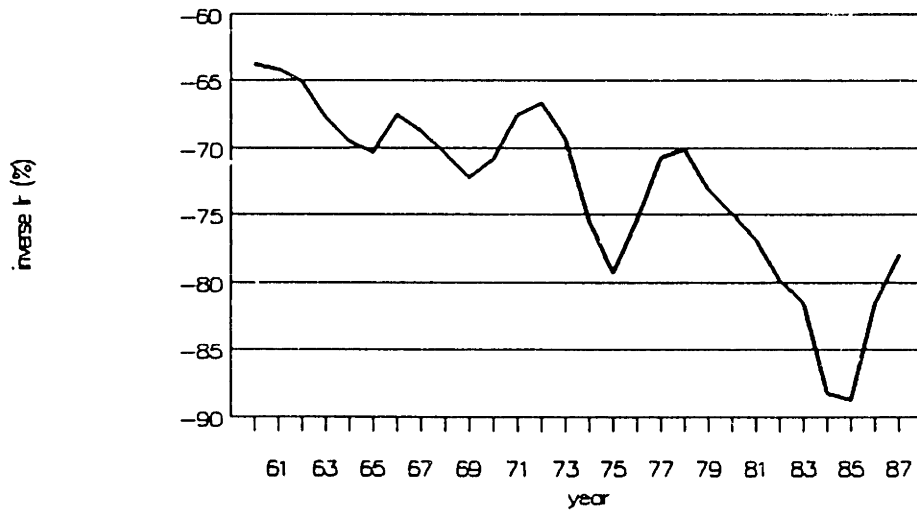
FIGURE 1

1 - combined ratio (%), industry
years 1960 to 1987



(a)

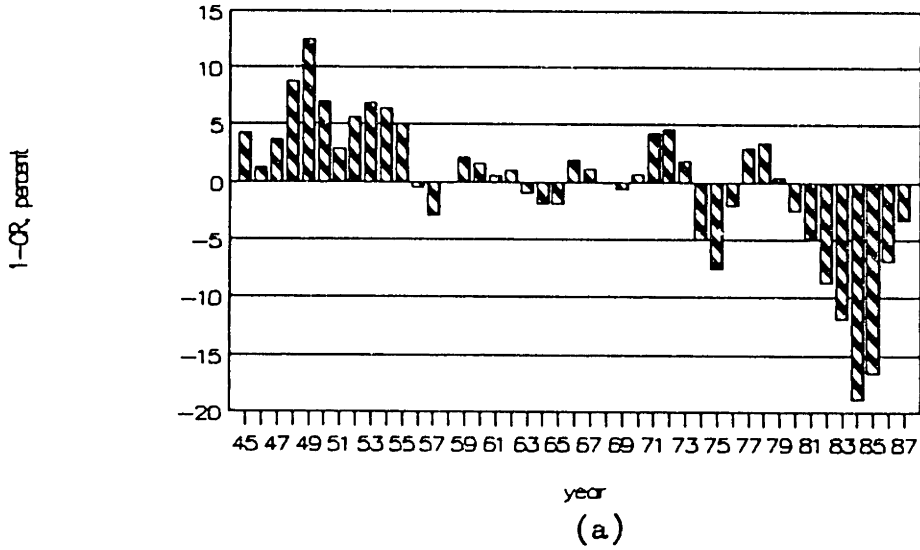
inverse loss ratio, industry
years 1960 to 1987



(b)

FIGURE 2

1 - combined ratio (%), stock
years 1945 to 1987



inverse loss ratio, stock
years 1945 to 1987

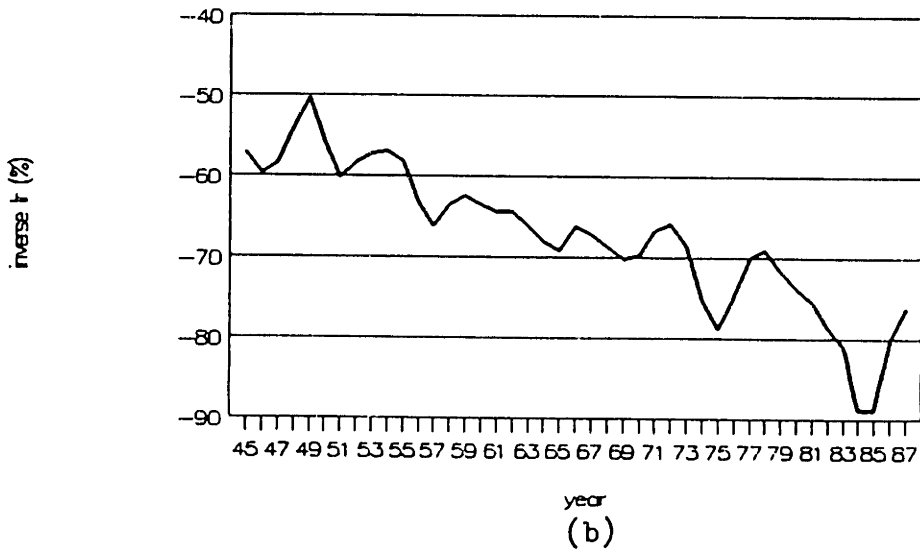
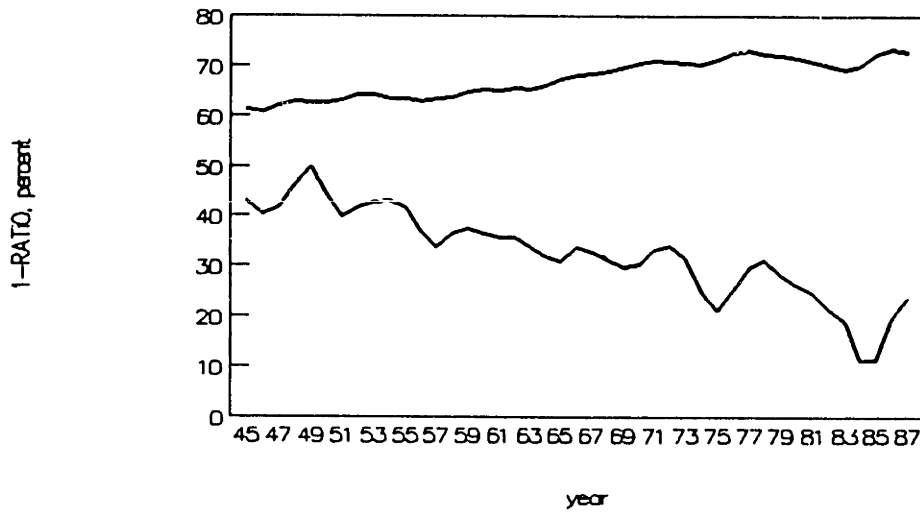


FIGURE 3
1 - RATIO (%), stock
years 1945 to 1987



more pronounced after the late 1950s.⁵

Annual change in profitability removes the time trend and displays the cycles more clearly. Since accounting profitability declines when the change in profitability is negative, the last year of negative change is the year accounting profitability is at a relative minimum. Figures 4 and 5 show the annual change in aggregate accounting profitability from 1961-1987 for the industry and from 1946-1987 for stock insurers. Measuring a cycle as the period between troughs in profitability, there are five complete cyclical episodes identified in the data, with the timing varying slightly depending upon which profit measure is used. The loss ratio identifies the cycles as 1951-57, 1957-65, 1965-69, 1969-75, 1975-84/5; the combined ratio identifies the cycles as 1951-57, 1957-64/5, 1964/5-69, 1969-75, 1975-84.^{6,7} Turning points in the loss ratio data are sometimes a year later than those in the combined ratio. If the change in accounting profitability comes predominantly from higher premiums (relative to costs) then

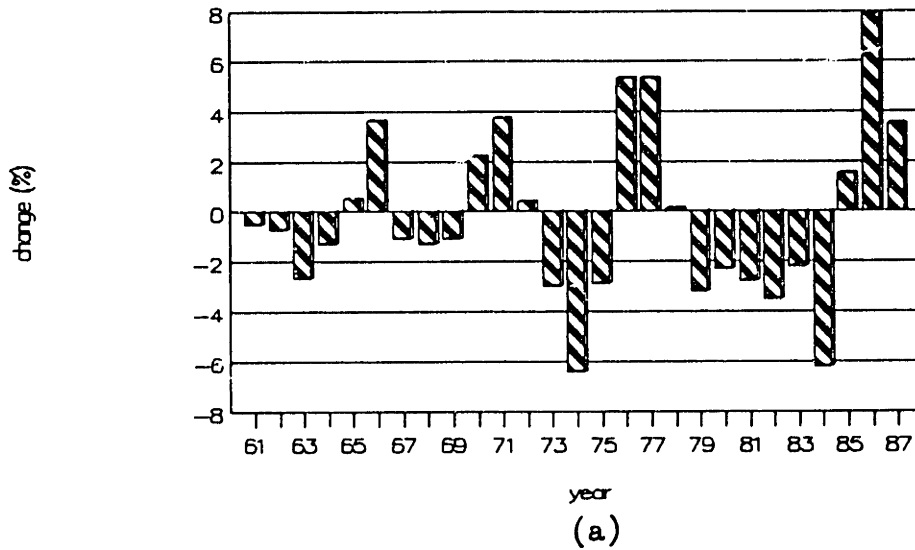
.....
⁵The downward trend in the negative aggregate loss ratio (which translates to an upward trend in the loss ratio) reflects increasing interest rates, loss tails, inflation, and long-tailed premium volume over the period (see the appendix on the accounting measures). The trend in the expense ratio is likely due to increasing market share of direct writers, and increasing automation. The contribution of each of these to the trends in the loss ratio and expense ratio are beyond the scope of the current chapter.

⁶The combined ratio has a small blip in 1961 not identified in the loss ratio data, suggesting perhaps that there is are two "lesser cycles", from 1957-61 and 1961-64/5, within the 1957-64/5 cycle.

⁷Stewart [1984] uses the combined ratio to define the cycles. In data that cover 1945-1983, she identifies five cyclical episodes: 1946-52, 1952-57, 1957-64, 1964-9, and 1969-75. (Stewart p. 286)

FIGURE 4

neg change in combined ratio
industry, 1961-1987



neg change in loss ratio
industry, 1961-1987

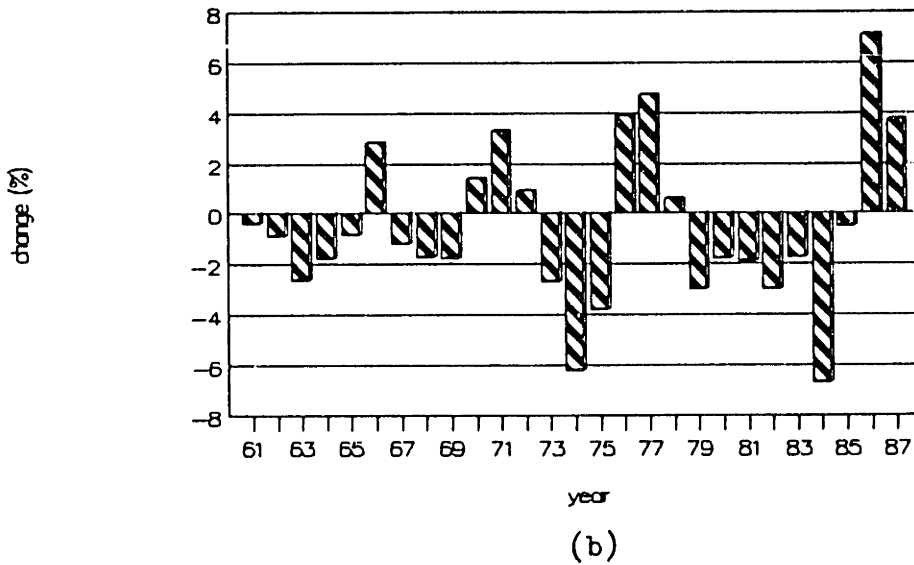
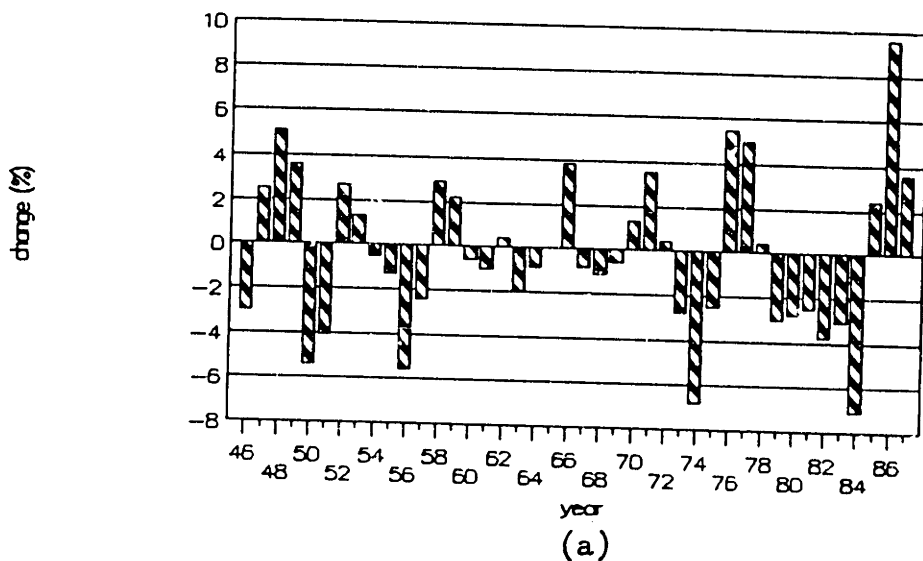
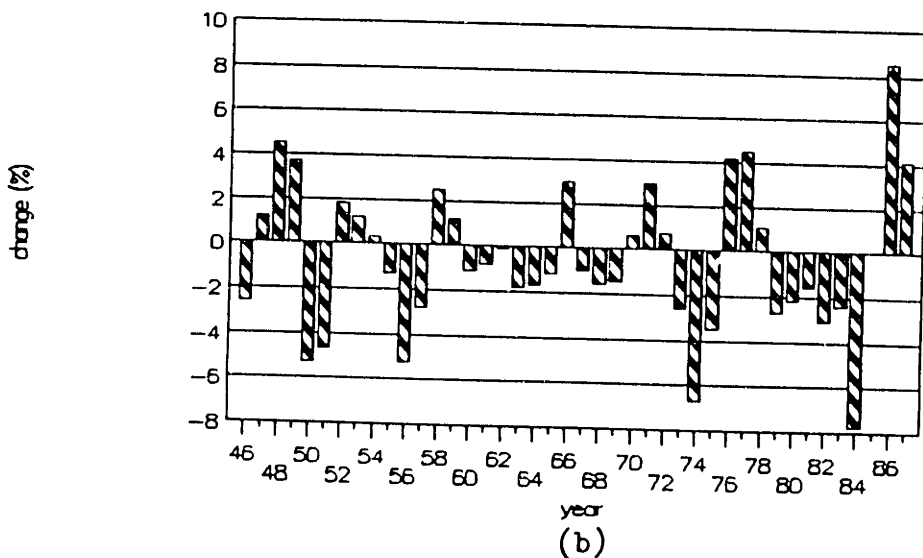


FIGURE 5

neg change in combined ratio
stock insurers, 1946-1987



neg change in loss ratio
stock insurers, 1946-1987



profitability measured by the combined ratio is likely to improve before profit measured by the loss ratio. Increases in premiums have a large, immediate impact on written premiums and thus the expense ratio, while premium increases affect earned premiums with a lag since the premium is only earned as the policy period elapses.

III. CAPACITY DEFINED CYCLES

The cause of property-casualty insurance cycles is the subject of some debate. One theory, the capacity-constraint hypothesis, is presented in chapter one of this thesis. The capacity constraint hypothesis suggests an alternative capacity-based criterion for delineating cycles in time series data. In this section I develop a capacity-based criterion for marking cycles. This definition is applied to four capacity series. The cyclical episodes identified by the capacity-based definition are then compared to those identified with the traditional accounting profit method described in the previous section. Cycles identified by the two definitions in data from 1950-87 are quite similar.

The term capacity refers to the maximum amount of insurance insurers are willing to write. Insurance premiums are based on expected costs, loss frequency and severity as well as expected returns on funds invested. Actual outcomes may differ from the expected levels due to uncertainty in predicting many of these variables. When realized costs exceed their expectations, the difference is paid from insurer net worth or policyholders' surplus. Therefore the amount of risk an insurer can

support and still maintain a low probability of bankruptcy is related to its level of net worth. Insurance regulation is aimed at maintaining a low probability of insurer bankruptcy by monitoring insurers' risk exposure relative to net worth.⁸ To accomplish this regulators require insurers to submit annual financial data and evaluate several statistics which relate levels of liabilities and revenues to assets.⁹

It is assumed that insurers hold net worth in order to keep the probability of bankruptcy low and to maintain their ability to meet claims made by policyholders. Large adverse shocks to claims, such as losses from unanticipated inflation or large catastrophes, or shocks to asset values, such as unexpected increases in interest rates or large declines on the stock market, can substantially reduce industry capacity.¹⁰ Capital does not immediately readjust to long run equilibrium levels due to a cost differential between various sources of capital. Capital from internal sources, such as retained earnings, is assumed to be less costly than external capital. Therefore, when

.....
⁸Net worth, called policyholders surplus or surplus in the insurance industry, is the difference between assets and liabilities reported in financial statements using Statutory Accounting Principles. Some items included in assets measured by Generally Accepted Accounting Principles are not permitted under Statutory Accounting Principles. Such items are known as nonadmitted assets and include some equipment and furniture, leasehold improvements, balances due from agents over 90 days due and amounts recoverable from salvage and subrogation. [Strain, p. 321]

⁹There is also some external monitoring by capital markets, reinsurers, and information companies such as A.M. Best and Co.

¹⁰Unexpected increases in demand may also result in low industry capacity, but the capacity constraint hypothesis usually assumes that demand for insurance is relatively stable and most changes in capacity are driven by shocks to supply.

capacity is unexpectedly reduced, some external capital will enter the industry, but the industry will not fully adjust until firms can use their retained earnings to augment net worth as well.¹¹

During the period when capacity is below long run equilibrium, the opportunity cost of internal capital will be relatively high. Short run premiums will rise, reflecting the increase in costs (accounting profit will also increase). As net worth returns to long run equilibrium, the opportunity cost of capital will decline causing short run equilibrium premiums and accounting profits to decline.

This theory seeks to explain the timing of insurance cycles as being caused by large, unanticipated reductions in industry capacity. Thus the high premiums and high accounting profitability are the symptoms of a short run equilibrium with industry capacity below the long run optimum. High premiums, high profits persist a substantial period of time, often a year or more. In order for the capacity constraint theory to be consistent with this observation some mechanism must prevent net worth from immediately readjusting.

Before turning to the data, there is an element of property-casualty insurance which may change the relation between capacity and net worth which is not considered in this analysis due to the lack of the necessary data. The supply of reinsurance will also alter the equilibrium level of net worth. Reinsurance expands insurance capacity by providing a means for insurers to sell off a portion of

.....
¹¹ There is additional discussion of adjustment costs and their causes in section v.

their expected loss (often the higher end), portions of riskier policies, or simply by improving geographic diversification. Consistent and comprehensive data on the level and use of reinsurance are not available, although it has been suggested that the supply of reinsurance expanded substantially during the late 70s and early 80s and then contracted suddenly in the mid 1980s.¹²

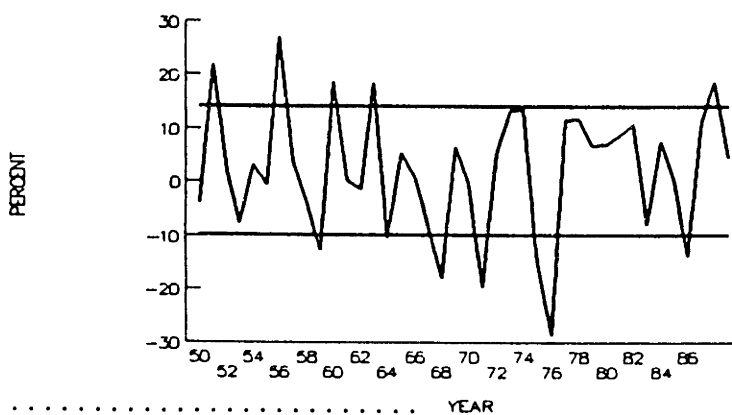
In the context of the capacity constraint hypothesis, an alternative method of marking insurance cycles is to identify periods where capacity is unexpectedly low. To do this requires some assumptions about equilibrium capacity which can be used as a reference point to determine low capacity. I assume that there is an optimal level of net worth relative to demand, and that to maintain equilibrium there is an ideal (and constant) growth rate for net worth relative to demand.¹³ Insurers try to achieve this growth rate, but uncertainty over realizations of losses, asset values and returns on investments, may cause actual relative growth rate to diverge from the optimal rate. Under these assumptions periods of low capacity are defined by years in which the relative growth of industry net worth is unusually low compared to the mean for the period.

.....
¹²The types of insurance written may also affect the equilibrium capacity. If, as is hypothesized, capital is held to keep the probability of bankruptcy low, riskier lines of insurance will require more capital per dollar of expected loss sold than less risky lines. A significant shift in the risks of coverages written --either from a shift in the risk associated with losses or a shift in the types of coverages written will result in a different equilibrium level of industry net worth.

¹³Industry net worth is measured as unconsolidated aggregate policyholders surplus reported in Best's Aggregates and Averages.

Four hypotheses concerning the constant relative growth between net worth and demand are considered. The first two assume demand growth is constant, the second two assume demand grows at a rate proportional to GNP. The first hypothesis is that nominal net worth grows at a constant rate relative to demand, therefore the mean growth in nominal net worth over the sample is proportional to the growth of demand. The second is that real net worth (calculated using the GNP deflator) grows at a constant rate relative to demand. Demand growth is assumed to be proportional to nominal GNP growth under the third hypothesis, so that the difference between the ideal growth in nominal net worth and nominal GNP is constant. The fourth hypothesis is similar to the third except that demand is assumed to be proportional to real GNP.¹⁴ Capacity measured as the difference in real growth rates is graphed in Figure 6. The other three measures of capacity produce similar graphs.

FIGURE 6
DIFFERENCE REAL GROWTH, NET WORTH &
GNP, WITH 1 STD DEV. MARKED

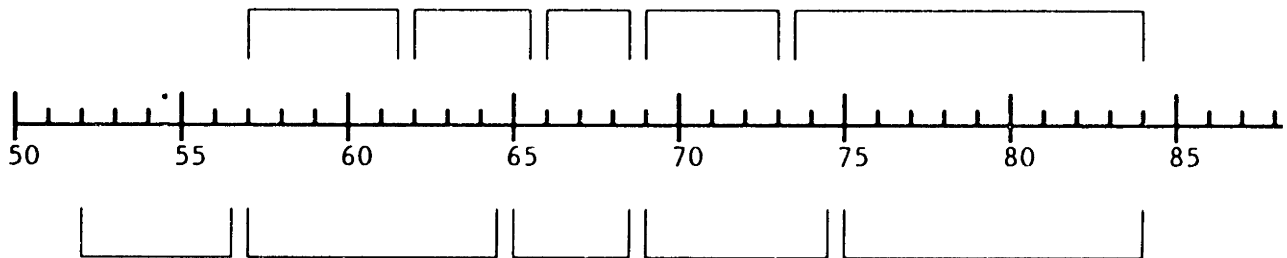


¹⁴The change in capacity measures are not adjusted for time trends since a regression of each series on a constant and a time trend found no statistically significant time trend. (The highest t-statistic was .6)

The time period under consideration is 38 years, 1950-87.¹⁵ Using the criterion that observations below one sample deviation less than the sample mean define periods of low capacity, six episodes (seven years) of low capacity are found in the data: 1957, 1962, 1966, 1969, 1973-4, 1984.¹⁶ The results are generally robust to the different hypotheses about demand growth. Only one year, 1981, is identified by only one measure and was therefore left off the list.

FIGURE 7

CAPACITY MEASURED CYCLES



ACCOUNTING PROFIT MEASURED CYCLES

Figure 7 compares cycles based on capacity measures with those based on profit measures. The capacity data identify five cycles as

.....
¹⁵Some industry statistics, profit ratios in particular, are not available for the full time period. The industry aggregate loss ratio and combined ratio data is available from 1958, Stewart constructs longer series by combining stock and mutual data back until 1951 and casualty company data only prior to that.

¹⁶The year 1951 is almost one standard deviation below the mean in the difference in growth rate measures. Both observations are 82% of a standard deviation below the mean. There is some question about the timing of the capacity variable since net worth is reported at end of year levels. For this reason the four variables that measure changes in capacity were also calculated using net worth at t-1. This method is equivalent to the one used here in that identifies the same episodes lagged one year.

measured from low capacity to low capacity: 1957-62, 1962-66, 1966-69, 1969-73/4, 1973/4-84. The profitability data also identify five cyclical episodes: 1952-57, 1957-65, 1965-69, 1969-75, 1975-84. The cycles from the two methods coincide closely after 1965. The pre-1965 results are also quite similar if the profitability trough in 1961 (found only in the combined ratio data) is included with the accounting profit cycles and 1951, a year where capacity is .82 standard deviations below the mean, is included in the low capacity cycles.

IV. OTHER TIME SERIES DATA

The capacity constraint hypothesis states that price increases should follow episodes of low capacity. Under the assumption that the effect of changing prices outweighs the effect of changing quantity on insurer revenues, large increases in premium volume should follow low capacity periods. This section examines the behavior of premium volume over the period. Other variables which may be related to low capacity, such as growth in financial liabilities, growth in net financial assets and cash flows, are also examined. Observations more than one standard deviation above (below) the mean determined unusually high (low) levels of the variables. The exception is cash flows where values less than zero were classified as unusually low.

Rapidly increasing premium volume is usually interpreted as the effect of price increases, where price increases refer to the price of a policy relative to a measure of quantity of insurance i.e. the expected

payout.¹⁷ Premium volume may also increase for various non-price reasons: inflation in the value of objects to be insured, other factors that increase the present value of nominal costs, or an increase in the number of objects insured (number of policies sold). An attempt was made to adjust for quantity and inflation effects employing a set of hypotheses similar to those used in analyzing net worth. Four hypotheses were again considered. On the assumption that quantity and inflation effects alone would cause net written premiums to grow at a constant rate, deviations from nominal and real mean annual growth rates were calculated. The mean difference in nominal and real annual growth rates of net written premiums and GNP were also computed under the alternative assumptions that, in the absence of price effects, net written premiums grow at a rate proportional to the growth of nominal or real GNP.¹⁸

The all lines industry data indicate four episodes of high premium volume growth: 1952, 1969, 1976-7, and 1985-6. The 1969 episode of high growth is not identified in the real and nominal annual growth series, only in the series measuring the difference between premium growth and GNP growth.

.....
¹⁷This is particularly true of the periods when prices are reported to be rapidly increasing and quantity available is reported to be decreasing. In these cases increases in premium revenues will understate the underlying price increases because they are not adjusted for the quantity reduction.

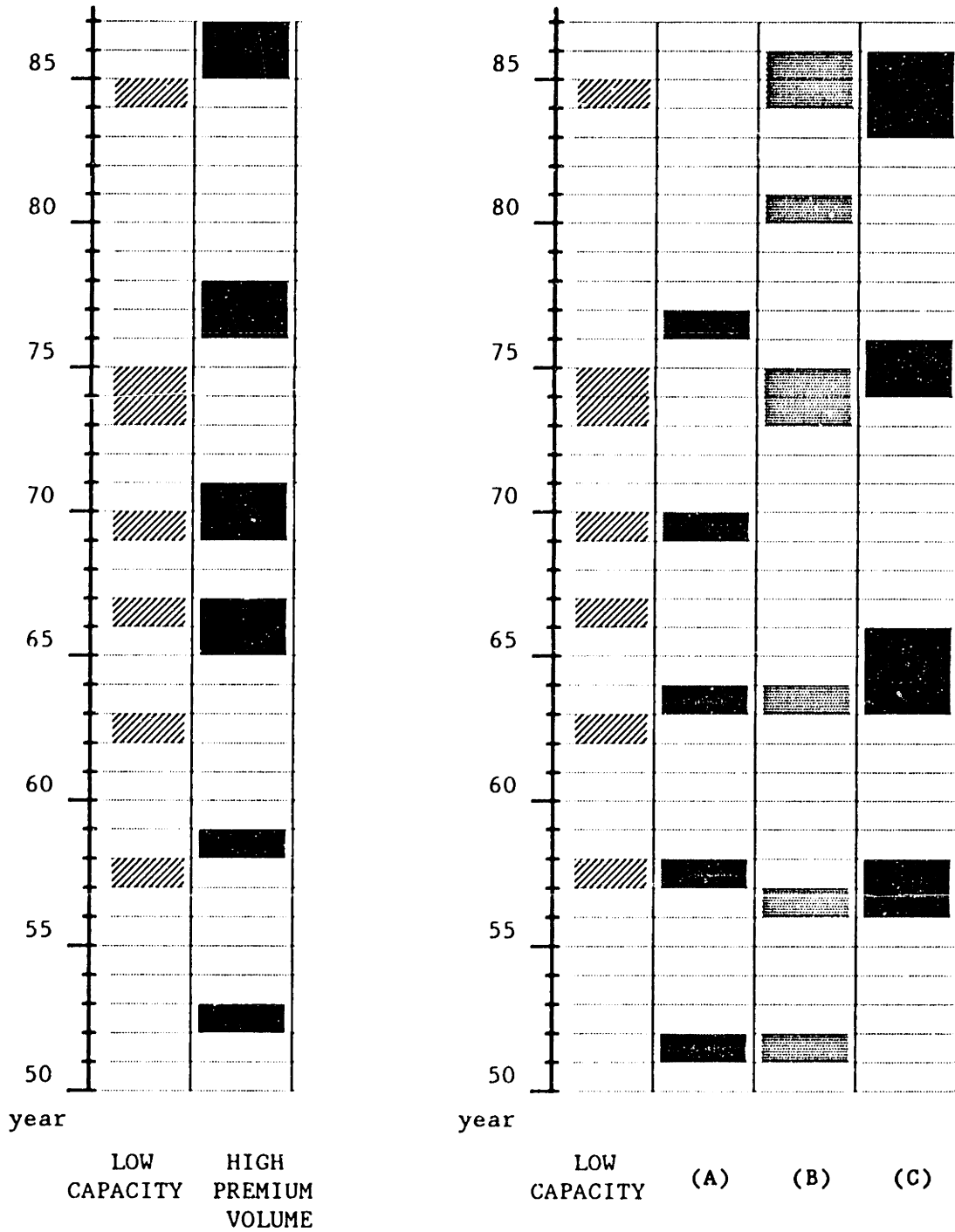
¹⁸Data were not adjusted for a time trend. Regression of each variable on a constant and a time trend showed that one series, annual growth in nominal premium volume, does have a significant, positive time trend. The other three series fail to reject the hypothesis that the coefficient on the time trend is zero.

Premium volume of the two automobile lines (auto liability and auto physical damage) was also examined since they comprise approximately 42-44% of property-casualty insurance revenues during the period.¹⁹ In these data results for stock insurers as well as industry aggregates were considered. The auto liability data have one additional period of high premium growth in 1958 and the timing of the second (1969) high growth event shifts to 1970. Automobile physical damage premiums showed substantial growth in 1965-6 as well as the four periods found in the all-lines data.

As depicted in Figure 8, large premium volume increases coincide with or immediately follow most of the low capacity years. Although the capacity constraint hypothesis indicates that premium increases will follow drops in capacity, the data show the relative timing of low capacity and premium increases varies. In the data, premium increases coincide with, follow and in one case precede capacity declines. Some of the variance in timing may be due to the relatively long period between observations and insurers' ability to delay full realization of large losses. For example, the drop in capacity may occur in the beginning of the year and premium increases may take place in the second half of the year or the premium growth may occur at the beginning of one

.....
¹⁹Data are from Best's Aggregates and Averages, 1987 edition. The data, covering 1957-86, show that auto liability accounts for 25-30% of net written premiums and auto physical damage 14-18%.

FIGURE 8



KEY: (A) HIGH GROWTH IN FINANCIAL LIABILITIES
 (B) LOW GROWTH IN NET FINANCIAL ASSETS
 (C) NEGATIVE CASH FLOWS

year but insurers may be able to delay realization of losses until later.²⁰ The longer delay between drop in capacity in 1973-4 and the increase in premium volume beginning in 1976 is less easily explained and is left for further research.

Years with large increases in financial liabilities, low asset growth and low cash flows are also marked in figure 8.²¹ The years 1951, 1957, 1963, 1969, 1976, 1984-5 had large increases in reported financial liabilities, while 1951, 1956, 1963, 1973-4, and 1980 were, by most measures, years of low growth in net acquisition of financial assets. The accounting statistics should be interpreted with caution since much of insurers' financial liabilities are estimated reserves and there is some opportunity for insurers to manipulate reported liabilities. These data do provide some additional information about cycles: reported liabilities tend to have larger increases around periods of low capacity and asset growth is particularly low in years of low capacity.²² The unusually large increase in liabilities may be the cause of the low

.....
²⁰Observed premium growth may precede the observed drop in capacity if insurers have some opportunity to delay adjusting reported net worth. For example, Stewart [1984] cites 1965's Hurricane Betsy as a major source of significant decline in insurer net worth which helped cause the change in market conditions (p. 304). It may be that insurers could postpone full accounting of the effect of these events on net worth until 1966 while the premium increases occurred after the actual events and were included in that accounting period.

²¹These data are from the Flow of Funds and cover "Other Insurance" described as "fire, casualty, and other companies covered in Best's Aggregates and Averages," life insurance companies are covered in a separate category [Board of Governors of the Federal Reserve System, June 1980, p. 36].

²²There are two exceptions, liabilities in 1976 and asset growth in 1984.

capacity and low asset growth. If liabilities are unusually high, funds that might have been allocated to net worth will be required to for the additional liabilities.

The Flow of Funds data can also provide a crude measure of cash flow: net sector surplus plus profit taxes payable. An important problem with this measure is that it may reflect non-cash flow events. For example, if financial liabilities change significantly due to revisions on outstanding liabilities from past polices, net sector surplus will reflect that change regardless of whether cash flows changed. Also, financial assets are the majority of property-casualty assets, and while bonds are primarily recorded at book value, stocks are recorded at market value. Movements in net financial acquisitions due to changes in the value of owned stocks will also change the cash flow measure without having any role in actual cash flow. With these caveats in mind, it is still interesting that the four periods of negative cash flows: 1956-7, 1963-5, 1974-5 and 1983-4 coincide with years of low capacity and that 1969, a year where cash flow is essentially zero, is also a year of low capacity.²³

The following empirical regularities are found in the time series data. Periods of low capacity generally coincide with unusually low growth in financial assets and coincide with or are closely followed by

.....
²³There are also data on net equity issues in the flow of funds data, but since they are increasing each year (even in real terms) there does not seem to be as strong a relation between net equity issues measured in these data and the property-casualty insurance cycle. Years with growth in real net equity issues greater than a standard sample deviation above the sample mean are 1950, 1953, 1958, 1961, 1963, 1968-9, 1982, and 1984.

large increases in reported financial liabilities. Unusually high growth in premium volume follows the low capacity years. Negative cash flows are often associated with the periods of low capacity.

Statistical analysis of the series in figure 8 indicates that low capacity has a positive effect on the probability of high premium growth. A series of dummy variables were generated to mark the years of significant change in the variables shown in figure 8. The probability of high premium growth was estimated using a probit specification, the results are in table 1. The first two columns of the table are results from estimating the probability of high premium growth in $t+1$ conditional on low capacity at t . Columns 1 and 2 differ in that column 2 includes the marginal years 1981 and 1951 in addition to the 7 years of low capacity marked in figure 8. Low capacity increases the probability of high premium growth in the following year, but the results are significant only at the 15% level.

TABLE 1

PROBIT RESULTS FROM ESTIMATING THE PROBABILITY OF
LARGE PREMIUM VOLUME GROWTH

$$\Pr(Y_t = 1) = \alpha + \beta_1 D_{t-1} + \epsilon_t$$

where $Y_t = 1$ if premium volume growth is above 1 sample deviation

	(1)	(2)	(3)	(4)	(5)
β_1	.64 (.54)	.75 (.48)	1.24 (.63)	1.27 (.56)	2.25 (.65)
N	41	41	36	36	41
years	1946-87	1946-87	(omit 1973-77)		1946-87
Log Likelihood	-22.1	-21.6	-17.6	-16.4	-15.2
$\Pr(Y=1 D_{t-1}=1)$.43	.44	.60	.57	.86

standard errors in parentheses

columns:

- (1), (3) $D_t = 1$ if capacity is low in year t-1 (as shown in figure 8).
 $D_t = 1$ in when t = 1957, 1962, 1966, 1969, 1973, 1974, 1984.
- (2), (4) $D_t = 1$ if capacity is low (as in columns 1 and 3) with the addition of two marginal years, 1951 and 1981.
- (5) $D_t = 1$ if growth in financial liabilities is unusually high in year t, as shown in figure 6.

Columns 3 and 4 also predict the probability of high premium growth as do the first 2 columns, but omit the 5 years from 1973-77. The probability of high premium growth conditional on low capacity is larger than in columns 1 and 2 and the coefficients are estimated with greater precision. Column 5 estimates the probability of high premium growth as

a function of large increases in net financial liabilities. High increases in liabilities are related to low capacity; unusually large expected losses or upward revision of past losses causes the change in liabilities to be large and is likely to reduce growth in net worth as well.²⁴

Table 1 results are consistent with the hypothesis that the cause of the drop in capacity in year t has an effect on the probability of high premium growth in year $t+1$. The low capacity-high premium growth relation appears to break down or change during the 1970s, however, the high growth in financial liabilities-high premium growth relation does not. The 1970s drop in net worth is unique in that much of the drop in net worth occurring in 1973 and 1974 is due to a substantial decline in the stock market. The large increase in financial liabilities is not observed in the data until 1976 the year premium volume increases substantially. It may be that drops in net worth caused by

.....
²⁴The regressions in columns 1 through 4 were also estimated including D_t and D_{t-2} . In both cases the coefficients on the additional variable was positive, but statistically insignificant at the 30% level. The coefficient on D_{t-1} changed little with the inclusion of D_{t-2} and was sometimes slightly larger with the inclusion of D_t .

The regression in column 5 was also estimated including D_t and D_{t-2} . The inclusion of D_{t-2} left the results in column 5 essentially unchanged. When D_t was included, the estimated coefficient, 1.06 (standard error of .63) was significant at the 10% level, and the coefficient on D_{t-1} , 2.46, was larger in magnitude.

Reestimating column 5 omitting 1973-77 leaves the result in column 5 essentially unchanged. When D_t is included the estimated coefficient is reduced to .63 and statistically insignificant at the 30% level.

unanticipated losses have different effects than reductions in net worth caused by low asset values.²⁵ Since only one episode of low capacity is associated with a large drop in the stock market, this hypothesis cannot be explored further using these data. Clearly, more research on the topic of the sources of shocks and outcomes in terms of market conditions is needed.

V. ADJUSTMENT COSTS

The central implication of the capacity constraint hypothesis is that when industry capacity is unexpectedly low, short run capacity constraints will result in premiums which are temporarily above the long-run equilibrium levels. The industry does not immediately readjust to long-run equilibrium because there are costs which prevent firms from issuing sufficient equity or taking on enough debt. In this section dynamic and static properties of observed capacity are examined. The results indicate that the costs to increasing unusually low net worth are greater than the costs to decreasing abnormally high net worth. The section ends with an example of how the Myers and Majluf model applied to insurance may explain the observed asymmetry. First, however, there is a brief discussion of the regulatory environment that strongly favors equity as a source of capital.

.....
²⁵The 1976 increase in liabilities may not show up as a decline in net worth because net worth was unusually low due to the stock market decline. It could be that the large increase in liabilities slowed expected growth in net worth, resulting in low capacity due to increased liabilities, but without the usual visible drop in capacity since the value was already low.

Regulatory restrictions on firm debt make equity the dominant source of additional net worth. Stock and mutual companies may borrow funds, but the debt must usually "meet stringent state requirements [relating] to interest rates and repayments and must be subordinated to all other company obligations."²⁶ There are often regulatory limits on the size of debt relative to net worth. Mutual insurers may issue a debt-like security known as guaranty funds or guaranty notes to increase net worth, but the conditions associated with this source of capital are also closely regulated. Given the restrictions placed upon debt as a source of capital, the preferred source of firm capital is equity financing.²⁷

In the models in chapter one and in Winter [1988, 1989] it is hypothesized that information asymmetries between equity holders and insurance managers make external capital more expensive than internal equity such as retained earnings. While insurers will issue equity when capacity is low, the cost differential prevents them from issuing enough equity to return to long run equilibrium. Accumulated retained earnings, or some other source of internal capital are needed for the industry to fully adjust.

.....
²⁶Strain [1986], p. 130.

²⁷Mutual insurers cannot issue stock (unless it is through a stock subsidiary). A mutual insurers sources of capital are predominantly retained earnings and guaranty notes. It is quite likely that capital from outside the firm is more costly for mutual insurers than for stock insurers because they cannot issue equity. The discussion on the sources of adjustment costs therefore relates primarily to stock insurers, since stock insurers are most likely the least constrained group and account for about 70-75% of the industry during the period considered.

Winter [1988, 1989] also includes costs to decreasing net worth. Regulations which limit the maximum annual dividend and the double taxation of capital (due to dividend tax) are two sources of downward adjustment costs.²⁸ These costs cause a dollar paid out by the firm to be worth less than a dollar to the recipient. Therefore, then capital is unexpectedly high, the firm will not fully adjust net worth to long run equilibrium, but will stop where the expected return from capital within the firm is equated with its value outside the firm. I will summarize these effects as the cost of trapped equity.²⁹

While the exact sources and magnitude of net worth adjustment costs cannot be determined from the accounting data, examining data on capacity and premium volume over several cycles provides some indication of the nature of these costs. The asymmetric information argument generates a cost to increasing net worth from external sources. The trapped equity hypothesis suggests there is a cost to decreasing net worth.

This implies that comparing responses to positive and negative capacity shocks can yield information about the relative size of net

.....
²⁸ New York insurance law prohibits insurers from issuing an annual dividend in excess of the lesser of 10% net worth from the last financial statement, or 100% of the adjusted investment income without the insurance superintendent's approval [New York insurance laws, section 1405, part a].

²⁹ Previous empirical work, in particular that by Poterba and Summers [1985] suggests that any costs from "trapped equity" are likely to be small. Their work tests three hypotheses about firm investment and dividend behavior, one of which they call the tax capitalization view which generates the trapped equity phenomenon. Poterba and Summers find no support for the tax capitalization view in their data.

worth adjustment costs. Under the maintained hypothesis that random shocks to net worth are drawn from a symmetric distribution, the path of the capacity variable should be symmetric if both types of costs have effects of similar magnitude. Evidence of symmetry indicates that the costs associated with adjusting net worth in either direction are of similar magnitudes. It cannot determine whether there are any adjustment costs at all. Asymmetry, however, provides support for the existence of adjustment costs at least in one direction.

A test of whether the path of a variable is symmetric is found in the literature on business cycles. The methodology tests whether the correlation properties of a variable differ depending upon the sign of the shock. In this context I test whether the path of the capacity variable is symmetric with respect to positive and negative shocks.

The test is as follows. Let (X_t) be the path of the variable of interest. Define (I_t) as:

$$I_t \begin{cases} 1 & \text{if } \Delta X_t > 0 \\ 0 & \text{if } \Delta X_t \leq 0 \end{cases} \quad (1).^{30}$$

I_t will be one when capacity is increasing and zero when capacity is declining. Assuming that (I_t) follows a first order Markov process, a symmetric path is described as one where the transition probabilities

.....
³⁰ I_t is assumed to be stationary. Since I_t is based on the change in X_t , it is invariant to monotonic transformations of X_t . Neftçi [1984] points out that this type of procedure will work even if (X_t) is not stationary. All that is needed is that I_t can "plausibly be stationary even when X_t is not" (p. 314).

given last period's change are the same.³¹ The test of whether adjustment is symmetric is then a test of whether

$$\Pr(I_t = 1 | I_{t-1} = 1) = \Pr(I_t = 0 | I_{t-1} = 0) \quad (2).$$

All four definitions of change in capacity were used to construct an $\{I_t\}$ from which to test this type of dynamic symmetry. I_t was constructed as described by equation 1. The transition probabilities were estimated by the following equation:

$$\Pr(I_t = 1) = \beta_1 D_{1t} + \beta_2 D_{2t} + \epsilon_t \quad (3)$$

where $D_{1t} = 1$ if $I_{t-1} = 1$ and 0 otherwise, and $D_{2t} = 1$ if $I_{t-1} = 0$ and 0 otherwise. In this formulation, $\Pr(I_t = 1 | I_{t-1} = 1) = F(\hat{\beta}_1)$ and $\Pr(I_t = 0 | I_{t-1} = 0) = 1 - F(\hat{\beta}_2)$ where $F(\)$ is the cumulative normal distribution. The test for symmetry becomes a test of whether $\hat{\beta}_1 = \hat{\beta}_2$.

The results of the test of whether $\hat{\beta}_1 = \hat{\beta}_2$ are presented in table 2. The null hypothesis of symmetric transition probabilities is rejected at or above the 7 % confidence level for all capacity measures. The results indicate that the adjustment process is asymmetric.

.....
³¹ A first order Markov process was chosen since it is likely that the current observation is correlated with that of the last period. This is particularly likely if insurers have some discretion in the timing of revisions to expected losses and use their discretion to smooth accounting profitability.

TABLE 2
ESTIMATION OF EQUATION 3³²

capacity measure:	annual growth in nominal net worth	annual growth in real net worth	difference in annual growth rates, net worth - GNP	
			nominal	real
β_1	.89 (.21)	.70 (.25)	.21 (.26)	.21 (.26)
β_2	1.07 (.59)	.52 (.42)	.62 (.35)	.62 (.35)

$H_0: \beta_1 = -\beta_2 \Rightarrow F(1,37)$	9.36	6.30	3.71	3.71
N	39	39	39	39
Log Likelihood	-18.3	-22.1	-25.0	-25.0

DeLong and Summers [1986] criticize the above procedure for "eliminating the quantitative information in the data." They analyze the distribution of output growth rates to test whether business cycles are asymmetric. A similar methodology can be applied here. If costs to increasing net worth are greater than those associated with decreasing net worth, the industry should show more rapid adjustment to positive shocks. Given the low frequency of observations, it may be that the industry will substantially readjust to positive shocks within the observation period of a year. An asymmetry in the magnitude of adjustment costs to positive and negative costs may then be reflected in an asymmetry in the distribution of changes in capacity.

.....
³²The results in columns three and four are the same although the two series are slightly different.

Figure 9 is representative of the distribution of annual growth in capacity using the four measures described in section III. All four distributions are skewed to the left. Table 3 reports the mean, median, mean-median and skewness for each distribution. These statistics confirm the indications of figure 9, the distribution of annual growth in capacity is not symmetric over the period 1948-87.

The capacity constraint theory indicates that insurance premiums will increase (decrease) if capacity does not readjust rapidly after negative (positive) shocks. The distribution of changes in capacity implies that capacity adjusts slowly after negative shocks but relatively rapidly after positive shocks. If large shocks to capacity is one of the primary causes of large changes in premium volume (through the effect on premium levels) then the distribution of changes in premium volume should also be asymmetric. Large positive changes in premium volume will be observed, but large negative changes will not since positive shocks to capacity appear to be less persistent.

A representative distribution of the annual growth in premium volume is shown in figure 10 and the mean, median, and skewness statistics for all the premium volume measures are reported table 3. The distribution of annual growth in premium volume is skewed to the right, showing that unusually large premium growth in some periods is not offset by low growth in others.

FIGURE 2

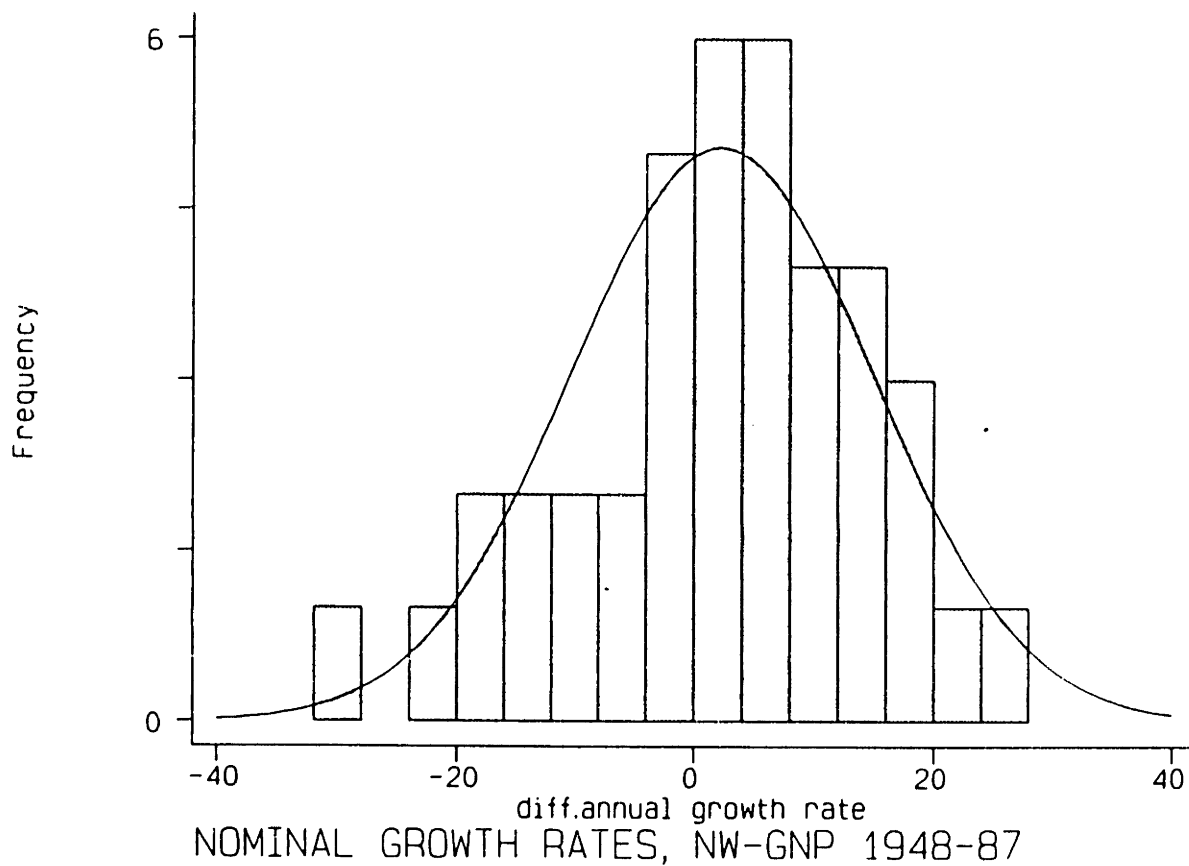


FIGURE 10

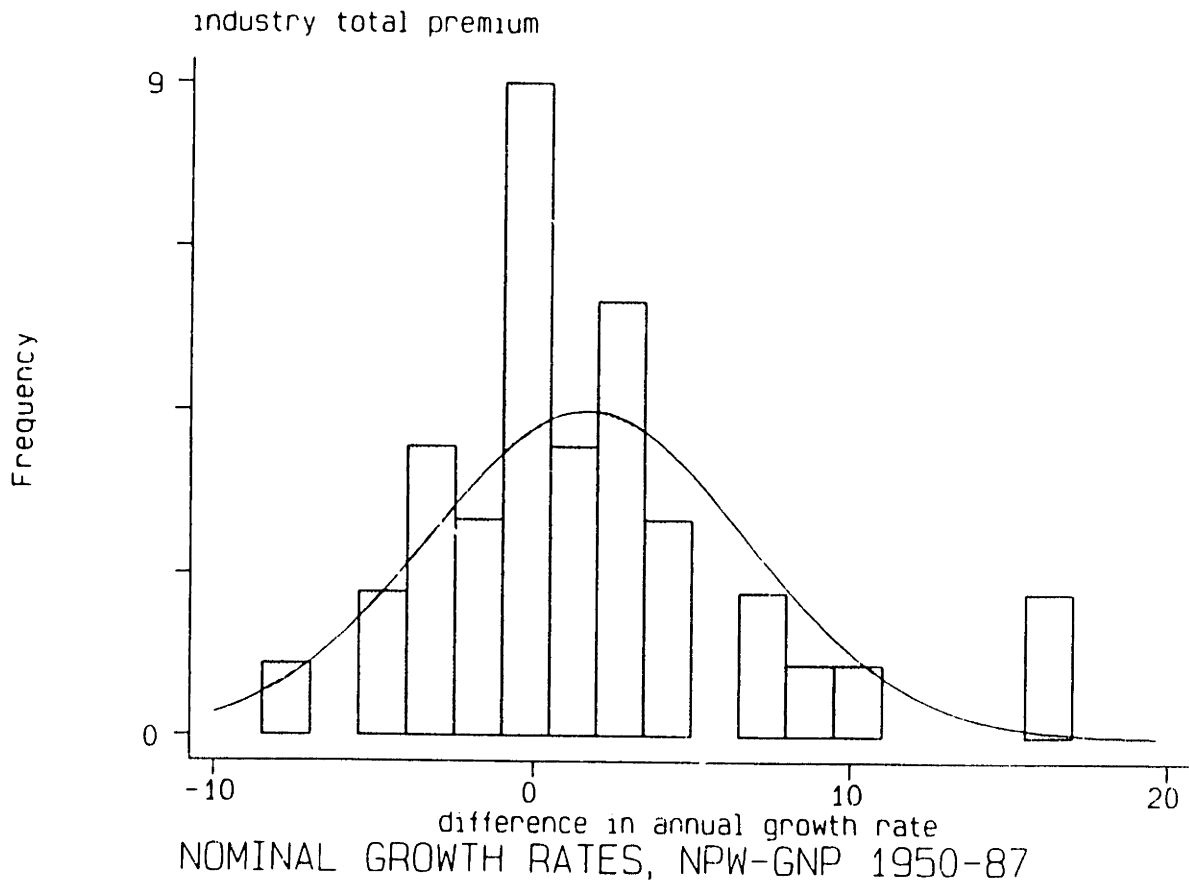


TABLE 3

variable	mean	median	mean - median	skewness
annual growth in nominal net worth	9.95	10.75	-.80	-.69
annual growth in real net worth	5.01	6.75	-1.74	-.74
difference in annual growth rates, nominal net worth - nominal GNP	2.25	3.57	-1.32	-.42
difference in annual growth rates, real net worth - real GNP	2.23	3.55	-1.32	-.34
annual growth in nominal net premium written	9.52	8.70	.82	1.21
annual growth in real net premium written	4.99	5.00	-.01	.79
difference in annual growth rates, nominal net premium - nominal GNP	1.69	.35	1.32	1.19
difference in annual growth rates, real net premium - real GNP	1.64	.36	1.28	1.25

The analysis of the dynamics of the path of capacity reveal an asymmetry over the cycle. The distribution of annual capacity growth is also asymmetric and skewed to the left. Assuming that shocks to net worth are drawn from a symmetric distribution, this implies that the costs of adjusting net worth upward after large negative shocks are

substantially larger than the costs associated with decreasing net worth after large positive shocks. The distribution of changes in premium volume is also skewed, but to the right rather than the left. This supports the capacity constraint hypothesis where unusually large shocks to net worth will have effects on industry prices. Large negative shocks persist due to relatively high adjustment costs and result in relatively large increases in policy prices, which cause large increases in premium volume. The asymmetry indicates that the costs of decreasing net worth, from restrictions on dividend payments and the double taxation of dividends, are relatively small. Such a finding is consistent with the empirical work on dividend decisions which indicates double taxation of dividends is not distortionary.³³

One explanation for higher costs to increasing net worth when capacity is unusually low can be found by applying the Myers and Majluf [1984] model to property-casualty insurance. In this model, asymmetric information concerning the value of firm's assets and investment opportunities between managers and investors generates a cost differential between equity from internal sources and equity from external sources. With some modifications, the model can be used to show that asymmetric information about the true extent of insurer's liabilities between insurance managers and outside equity holders will give the same result. An example from the Myers and Majluf paper is reworked in the context of the insurance industry in Appendix B to

.....
³³See, for example, Poterba and Summers [1985].

illustrate this point. The results are summarized here.

Managers and investors have the same information and beliefs about the current profitability of insurance being sold. They differ in their knowledge of insurers capital. New information about outstanding claims, such as increased inflation and claims frequency, is learned first by managers. The full impact of the new information on a particular insurer's net worth is a function of the characteristics of past policyholders. For these reasons, there is likely to be a lag between the time managers and outsiders realize the full impact of unexpected changes in claims on particular insurer's net worth.

Since insurers with low capital relative to premiums are more likely to attract closer regulatory scrutiny, insurers with larger unexpected losses will be more likely to issue equity in order to increase net capital before the unexpected losses are fully realized. Investors recognize this fact and require a higher expected return on their investment. There may be a greater likelihood that firm's reported net worth does not fully reflect all revisions to estimated losses when many insurers have suffered unexpected losses. Under these conditions the penalty to "good" insurers for issuing stock will be greater, and insurers will not issue enough equity to return the industry to equilibrium.

VI. CONCLUSION

This paper has examined three topics related to the capacity-constraint hypothesis. First, a capacity-based criterion for

cycles in time series data was compared to the traditional accounting profit criterion. The two methods define essentially the same cycles in aggregate industry data for the 1947-87 period.

Second, the behavior of other aggregate time series, in particular that of premium volume, was also analyzed. Exceptionally high premium volume growth is more likely when capacity is unusually low in the preceding year. This relation is stronger when the mid-1970's period is omitted or when unusually high growth in financial liabilities is used to capture changes in losses which might reduce capacity.

Finally aggregate time series data were examined to determine the relative magnitude of net worth adjustment costs. The path of capacity adjustment and the distribution of annual capacity growth were found to be asymmetric. The distribution of annual capacity growth was skewed to the left, consistent with relatively greater costs to increasing net worth. The distribution of premium volume growth was skewed in the opposite direction. This finding is also supportive of a theory where large negative capacity shocks cause temporarily higher prices but large positive capacity shocks, because they do not persist, do not result in offsetting low premium growth.

The relatively slow adjustment to negative capacity shocks can be explained by a cost differential between internal and external equity. Such a differential may arise from asymmetric information between managers and investors, as illustrated by an application of the Myers and Majluf model to property-casualty insurance. This view also explains the breakdown in the low capacity-high premium growth observed

in the 1970s while the large increase in financial liabilities-high premium growth relation does not have a similar breakdown. An unusually large drop in the stock market caused much of the 1973-74 decline in industry net worth. In property-casualty accounting, net worth held in stocks are recorded at market values, so these declines were immediately incorporated into reported net worth and did not exacerbate any information asymmetry between managers and investors. The large increase in financial liabilities observed in 1975, however, may have signaled upward revisions in outstanding liabilities which would increase the information asymmetry.

The conclusions of this paper are preliminary, they suggest further research in a number of areas. Two of the most important for understanding the nature of insurance cycles concern the causes of the slow adjustment to negative capacity shocks. The data in this paper indicate that the cause of the shock to net worth may have an effect on industry response. Unfortunately, there is only one period where a reduction in asset values can be readily identified as a major determinant of a decline in aggregate net worth. Additional work on this topic would hopefully clarify this relation. Investigation into the causes of the discrepancy between the cost of internal and external capital would also improve understanding of insurance cycles. For example, variation in the relative size of unexpected losses over different lines might be used to identify insurers where asymmetric information problems are more severe. These differences could then be exploited to learn more about costs of increasing net worth.

APPENDIX A

ACCOUNTING PROFITABILITY

Accounting profitability is often expressed as a ratio of revenues to costs. The two profitability measures examined in this paper are the loss ratio and the combined ratio. The loss ratio is the ratio of expected losses to premium revenues, while the combined ratio includes expenses relative to premiums as well. Both statistics deal exclusively with insurance activities and ignore income from investment activities.³⁴

The loss ratio is calculated as incurred losses divided by earned premiums. Incurred losses are defined as losses paid during the accounting period plus the change in the loss reserve. Since the loss reserve is an undiscounted estimate of all outstanding liabilities regardless of when the corresponding revenues were earned, revisions to expected losses from expired policies as well as the estimated liabilities for the policies earned during the current period are included in incurred losses.³⁵

The denominator of the loss ratio, earned premiums, is the revenue from all policies active in the accounting period which corresponds to

.....
³⁴ Stewart [1984] compares accounting profitability statistics with and without investment income and concludes that operating results exhibit the same cycles as underwriting results. (there appears to be a different trend in the operating data, but the up and down movement in profitability is quite similar between the two series).

³⁵ There are three sources of changes in incurred losses: changes in current payments made for expired policies, changes in the estimated future costs associated with expired policies, and changes in estimated and current payments on policies earned over the period.

the fraction of the policy period which coincides with the accounting period. For example, one half of the premium from a one year policy is included in earned premiums if six months of that policy term is within the accounting period. If only three months of the policy coincided with the accounting period, one-quarter of the premium would be included in earned premiums.

The loss ratio is sometimes described as measuring the dollar of expected payouts for each dollar of premium income earned during the period. Incurred losses, however, may reflect information from expired policies as well as those earned in the current period, making the loss ratio at best a noisy measure of expected payouts per dollar premiums.

The combined ratio is the sum of the loss ratio and the expense ratio. Total underwriting expenses, which include acquisition costs such as commissions, premium taxes, and fees, divided by written premiums compose the expense ratio. Although the loss ratio and the expense ratio have different denominators, the combined ratio is usually interpreted as insurance costs relative to revenues.

APPENDIX B

An Application of the Myers and Majluf Model to Insurance

This appendix provides an example showing how the Myers and Majluf [1984] framework can be adapted to explain the differential between internal and external capital. Some extensions which explain why it is particularly costly to issue equity during certain periods are also discussed. Myers and Majluf develop a model where asymmetric information between managers and investors results in a cost differential between internally and externally generated equity. In equilibrium firms may forego investment opportunities with positive net present value because of the cost differential.³⁶

Myers and Majluf make the following assumptions: (a) there is a valuable real investment opportunity, (b) the firm must issue common shares to raise all or part of the funds needed, (c) managers know more about the value of assets and opportunities than do outside investors, (d) management acts in the interest of old stockholders, (e) existing stockholders do not rebalance their portfolios in response to firm actions, (f) the investment project is indivisible, (g) there are no transactions costs to issuing equity, and (h) the value of shares is the expected future value of the firm conditional on all available

.....
³⁶In Myers and Majluf managers have asymmetric information about the returns from investment as well as the asset in place. In this formulation, the return on the investment is not in doubt, it is the value of the asset in place (net worth) which is in doubt.

information. These assumptions are made here with the following modifications. Investors are investing in the general business of insurance. Both managers and investors have the same beliefs about the return to capital which supports insurance policies.

The asymmetric information in this example of the model is due to the nature of the insurance business. Insurance is priced before costs are known. Insurers hold net worth in order to make the promise to pay credible. If realized losses exceed the ex-ante expected losses, the difference is made up by insurer net worth.³⁷ New information about losses, such as increased inflation, claims frequency and so forth, is learned first by managers who observed the losses coming in to the firm. Later, outsiders learn about the changes in losses as the statistics (such as inflation) become publicly available, the insurer readjusts loss estimates, other insurers revise their loss estimates or the sources of the increased claims become apparent through other means. The full impact of the new information on a specific firm's expected claims will be a function of the firm's policyholders. For this reason, there is likely to be a lag between the time managers and outsiders realize the extent of the unexpected losses and their effects on the insurer's net worth.

The following example illustrates how insurance is analogous to the situation examined in Myers and Majluf. The example is similar to the one presented in the text, the difference being the specific application
.....

³⁷This assumes that there is no bankruptcy.

to insurance.

Consider an insurer whose net worth is currently reported to be \$100 million. The return to funds backing insurance policies is 10% and regulators require insurers to have a minimum net worth of \$100 million in order to operate. Either of two cases applies: the state value of insurer net worth is equal to the true value of net worth or the insurer will be revealed to have additional losses that reduce net worth to \$60 million.

First consider a situation where the insurer issues \$100 million no matter what the true value of net worth. Table 4 shows the value of firm net worth and return on invested equity at the end of the period under the two cases.

	case 1 value of Net Worth is accurately stated	case 2 value of Net Worth is really only \$60
true value of net worth	110	66
return to additional equity invested	10	10

Table 4

The value of the firm held by old shareholders in both cases when the firm does and does not issue equity is shown in Table 5.³⁸ Always

.....

³⁸Let P' denote the market value of a share at the beginning of the period if stock is issued. V is the total value of the firm, V^o is the value of the firm held by old shareholders, V^e is the value of the firm held by the new equity holders. $V = V^o + V^e$, also $V^e = \frac{P'}{P'+E} * V$ and $V^o = \frac{E}{P'+E} * V$.

issuing equity is not an equilibrium strategy since old shareholders are better off if the firm does not issue equity in the case where net worth is correctly stated. Never issuing is also not an equilibrium strategy since old shareholders do better when the insurer issues in the case where true net worth is lower than reported.

		firms issue	firms do not issue
value of firm to old shareholders	case 1	108.9	110
	case 2	87.1	66

Table 5

The decision to issue equity is only optimal in the case where the true value of net worth is less than the reported value. Therefore the decision to issue equity will signal that the insurer's net worth is lower than reported. In the situation where the decision whether or not to issue perfectly reveals the insurer's type, only the boxes with \$110 and \$87.11 in table are relevant. An insurer with correctly reported net worth will thus forego expanding insurance operations by issuing equity even though the activity has a net positive value.

The above example demonstrates how the Myers and Majluf framework can be applied to insurance companies. The analysis can be extended to allow the probability of case 1 or case 2 to vary with industry conditions. When many insurers have suffered unexpected losses and liabilities have been revised upwards, there may be a greater likelihood that reported net worth does not fully reflect revisions to estimated

losses. Under these conditions situation the cost to "good" insurers of issuing equity will be larger, or the penalty for issuing stock will be greater, so the firm will not issue enough equity to return the industry to equilibrium. Since large unexpected losses are correlated with large reductions in industry capacity, the cost of increasing net worth through equity issues is likely to be greatest when it is needed the most.

REFERENCES

- A.M. Best Company Best's Aggregates and Averages Property/Casualty edition.
- Angoff, J., "Insurance Against Competition: How the McCarran-Furgeson Act Raises Prices and Profits in the Property-Casualty Insurance Industry," Yale Journal on Regulation, summer 1988.
- Bernstein, G.K., "Antitrust Agnostics," Best's Review (P/C edition) October 1988.
- Board of Governors of the Federal Reserve System Flow of Funds Accounts 1946-75 December 1976 Annual total flows and year-end assets and liabilities
- Board of Governors of the Federal Reserve System Flow of Funds Accounts second quarter 1988 September 30, 1988 Seasonally adjusted and unadjusted, Annual Revisions
- Board of Governors of the Federal Reserve System Flow of Funds Accounts second quarter 1982 September 1982 Seasonally adjusted and unadjusted, Annual Revisions
- Board of Governors of the Federal Reserve System Introduction to Flow of Funds, June 1980
- Clarke, Richard N., Frederick Warren-Boulton, David D. Smith and Marilyn J. Simon, "Sources of the Crisis in Liability Insurance: An Economic Analysis," Yale Journal on Regulation, Summer 1988.
- Cummins, J. David and J. François Outreville, "An International Analysis of Underwriting Cycles in Property-Liability Insurance," Journal of Risk and Insurance, 1986, pp. 246-262.
- DeLong, J. Bradford and Lawrence H. Summers "Are Business Cycles Symmetrical?" in The American Business Cycle, Robert J. Gordon, ed., 1986.
- Harrington, Scott E. "Prices and Profits in the Liability Insurance Market," in Robert Litan and Clifford Winston, eds. Liability: Perspectives and Policy. Brookings. 1988.
- Harrington, Scott E. and Robert Litan "Causes of the Liability Insurance Crisis" Science February 1988 p. 737-741.
- Myers, S. and N. Majluf, "Corporate Financing and Investment Decisions when Firms Have Information That Investors Do Not Have," Journal of

Financial Economics, vol 52, 1984.

Neftçi, Salih N. "Are Economic Time Series Asymmetric Over the Business Cycle?" Journal of Political Economy, vol. 92, 1984 pp. 307-328.

Poterba, James and Lawrence H. Summers "Economic Effects of Dividend Taxation," in E. Altman and M. Subrahmanyam, eds. Recent Advances in Corporation Finance, 1985.

Stewart, Barbara D. "Profit Cycles in Property-Liability Insurance, " John D. Long and Everett D. Randall (eds) Issues in Insurance vol 2 1984 American Institute for Property and Liability Underwriters, Inc. USA.

Strain, Robert, W. (ed.) Property-Liability Accounting, 1986.

Winter, Ralph A., "The Liability Crisis and the Dynamics of Competitive Insurance Markets," Yale Journal on Regulation vol 5, 1988, p. 455-500.

-----, "The Dynamics of Competitive Insurance Markets," mimeo, University of Toronto, July 1989.

CHAPTER THREE

The Role of Capacity in Accounting Profitability Cycles

I. INTRODUCTION

This chapter analyses the phenomenon of cycles in property-casualty accounting profits to determine whether variations in industry capacity affect accounting profitability. Insurance "capacity" is the industry's ability to credibly accept risk without incurring high bankruptcy probabilities. Insurance capacity is measured as net worth relative to total demand. Insurers hold net worth in order to comply with regulatory requirements established to keep the probability of bankruptcy low and to maintain the ability of insurers to meet claims made by policyholders.

Large shocks to net worth, such as higher than anticipated cost inflation or catastrophic events, reduce net worth and decrease industry capacity. In the presence of net worth adjustment costs, the industry will not immediately return to long run equilibrium. Insurers currently in the industry will earn a higher return on their remaining capital, reflecting the temporary increase in the opportunity cost of capital. The relation of insurer accounting profitability and variations in insurance capacity is estimated using aggregate time series data for stock insurance companies over the 1952-1986 period. Four lines of insurance are considered: automobile physical damage, automobile liability, homeowners' multiple peril and other liability. The results support the capacity constraint hypothesis: unanticipated decreases in insurer capacity are found to cause increases in future profitability.

The findings are applicable to several policy issues, particularly price regulation and operation of state guaranty funds. Insurance regulators are interested in controlling the large premium swings and reduction in coverage that occur during the property-casualty cycle. Under the capacity constraint hypothesis, short run variation in the expected return to net worth is part of the industry adjustment process. The recent trend toward increased price regulation, as exemplified by California's Proposition 103, will inhibit large price movements and may have substantial impact on the dynamics of the property-casualty industry. Insurance regulators also operate state guaranty funds which protect policyholders from insurer insolvency. The analysis here suggests that the current form of guaranty funds may exacerbate the market disruptions described above.

The chapter is organized as follows: section II reviews features of the industry which are relevant for the analysis. It also gives some background on property-casualty cycles, and reviews explanations for the cycles and previous empirical work. A specification for estimation and the data that will be used are covered in section III. Section IV presents the regression results and section V includes a preliminary analysis of the stock market reaction to two large loss events: Hurricane Hugo and the San Francisco earthquake. Finally section VI summarizes the results and discusses their relevance for current issues in insurance regulation.

II. THE PROPERTY CASUALTY INSURANCE INDUSTRY AND CYCLES

Part A of this section provides some background of the basic

features of property-casualty insurance that is relevant for this analysis. The structure of the industry and the types of insurers are covered in earlier chapters and only briefly reviewed here. Some facts and theories about property-casualty profitability cycles are covered in part B.

A. OVERVIEW OF PROPERTY-CASUALTY INSURANCE

Property-casualty premium volume was approximately \$200 billion in 1988. Industry assets totaled \$477 billion with net worth of \$118 billion.¹ Although there were about 3600 property-casualty insurers operating in the US in 1986, most of the premium volume was written by about 900 of these firms.² Concentration is low for the industry and fairly low for most lines of insurance. Entry barriers are low. Property-casualty insurance is structurally competitive: there are many sellers, low concentration and low barriers to entry.³

A typical insurance contract specifies a premium to be paid by the buyer in return for promise of payment should certain events occur during the term of the agreement. While premiums are paid at the beginning of the period, claims for the insured events may be settled during or after the policy period has expired.⁴ Property-casualty

¹ Statistics are from Best's Aggregates and Averages, 1989 edition.

² Insurance Facts.

³ See Harrington [1988] p. 43 and references therein.

⁴ Premiums are not always paid in full at the beginning of the policy period but this assumption simplifies the exposition without loss of generality. This form of contract is known as occurrence-based insurance because the policy pays according to when the event happened, not when the claim is filed.

insurance accounting procedures, known as Statutory Accounting Principles (SAP), require insurers to set aside reserves covering the estimated value of all outstanding claims. If the actual payout exceeds the estimated value, the difference comes out of insurer's net worth. Net worth is the residual when SAP liabilities are subtracted from SAP assets.

Property-casualty insurance can be described by the type of loss event covered. If the loss is to the insured party, the insurance is known as first-person coverage, while if a third party is protected from loss, it is called third-person or liability insurance. The average payout period, or tail, of a first-person claim is significantly shorter than that of liability coverages, where detection lags and litigation often extend the time between the event and payment by the insurer. Of the four lines of insurance examined in this chapter, auto physical damage and homeowners' multiple peril are short-tailed lines (mostly property losses). Auto liability and other liability have significantly longer tails. During the period 1977-86 over 95 % of claims for homeowners' multiple peril were paid within the first three years after the policy period, within the first 5-6 years for auto liability and within 9-10 years for other liability.⁵

⁵These numbers are from the author's calculations from a sample of 25 insurer groups' schedule P data. Loss emergence data on auto physical damage is not included in Schedule P because auto physical damage is not considered a long-tailed line. The assumption made here is that the loss tail for auto physical damage is at most as long as that for homeowners' multiple peril. (data are from Best's Convention Statements)

Most of the larger insurers operate in most or all states and sell several lines of insurance. One significant difference among insurers is their form of corporate structure. There are two primary forms of organizational structure in the insurance industry: the stock and mutual company. Together they account for about 95% of property-casualty insurance assets, net worth, and sales from 1965 to 1986.⁶

Stock companies are the largest segment of the industry. Like other publicly held corporations, stock companies are owned by their shareholders who receive a return on their investment either through dividends or through stock appreciation. Equity issues to the public can be used to finance company expansion. Stock insurers have historically been the most prevalent type of corporate structure; since the 1960s stock insurers have accounted for at least 70% of industry assets, premiums, and net worth.⁷

The data used in this chapter are for stock insurers only. As noted in section IV of chapter two, it is likely that the different organizational forms lead to different costs of generating equity. Since this analysis examines whether capacity, which is based on industry net worth, has an effect on profitability, the sample is restricted to stock insurers to control for cost of equity. This

⁶These data are from Best's Aggregates and Averages. The other two corporate structures are the reciprocal exchange and the lloyds organization.

⁷The other common form of corporate structure is the mutual company. Mutual insurers are owned by their policyholders who, in theory, control the corporation by voting for management similar to shareholders in a stock company. During the period considered here, mutuals' share of industry assets, premium volume and net worth ranged from 20 -25%.

restriction is unlikely to bias the results since the omitted group, mutual insurers, is likely to have a higher cost of external capital due to greater regulatory restrictions on increasing capital and has a substantially smaller market share than stock insurers.

Insurance regulation is covered in chapter one. A primary goal of regulators is limiting the probability of insurer default. Regulators monitor the solvency of licensed insurers three ways: by capital and net worth requirements, review of annual statements, and by periodic in-depth examinations. Insurers who fall below the minimum solvency standards may be liquidated or placed in rehabilitation where the insurance department dictates required changes in business practices. Other insurers who are below standards in several categories are put on regulatory "watch" lists and are subject to more frequent scrutiny. In addition to the oversight activities, insurance regulators also restrict firms' ability to distribute capital.⁸

B. PROPERTY-CASUALTY INSURANCE CYCLES

1. The Property-Casualty Insurance Cycle

A typical insurance cycle, as described in chapters one and two, has four stages. The crisis stage is characterized by rapidly

⁸Some insurance departments limit the size of dividend payments. New York's insurance laws, section 4105, limits the annual total dividend payment to be no greater than the lesser of 10% of net worth reported in the last financial statement filed with the department or 100% of adjusted net investment income. (New York Insurance Laws, 1986). In some states, insurance departments also regulate policy forms and rates, although the primary motivation for this is consumer protection rather than insurer solvency.

increasing premiums and accounting profits with decreasing availability. There follows a period of high premiums, profitability and low availability. Over the next period prices begin to decline and availability problems disappear as the industry enters the fourth stage of low accounting profits and low premiums.

The ideal data to describe and study the cycle would include price and quantity data. These data, however, are not available.⁹ Usually the cycle is described using accounting profitability data. Accounting profitability generally rises when premiums increase and falls as premium levels decline. Anecdotal evidence indicates that large increases in premium volume (total sales revenues) coincide with rapidly increasing prices and lower availability.¹⁰

The two common measures of insurer profitability are the loss ratio and the combined ratio. Both statistics deal exclusively with cash flows from insurance activities and ignore insurer income from investment activities. The loss ratio is calculated as incurred losses

⁹ An ideal price measure would be a time series on the premium for a policy with fixed real limits. Such data are not collected. No time series on premiums per policy are available, even when the deductibles and limits are not controlled for. The quantity of insurance is also a difficult concept to measure. Quantity can be measured as the number of policies issued, the number of objects insured or the value of the expected losses covered. As with price, none of these measures are collected over time.

¹⁰ One measure of availability might be number of policies written or number of existing policies that are not renewed or canceled. These statistics are, unfortunately, not collected either. The anecdotal evidence referred to is from surveys collected around and following the period of rapidly increasing premiums and reduction in quantity. Two sources of anecdotal evidence from the most recent restricted market are the Tort Policy Working Group [1986, 1987] and Risk and Insurance Management Survey.

divided by earned premiums. The combined ratio is arrived at by adding the expense ratio (underwriting expenses divided by written premiums) to the loss ratio. Both are commonly reported in percent, a loss ratio greater than 100 indicates that accounting losses are larger than earned premiums. The two measures are highly correlated and, since the goal of the empirical analysis is to explain the movements in accounting profitability through capacity variables and factors which affect losses, I use loss ratio as a measure of property-casualty insurance profitability.¹¹

The numerator of the loss ratio, incurred losses, is defined as the losses paid out during the accounting period plus the net change in the loss reserve.¹² The denominator of the loss ratio is earned premiums.

¹¹Some variation in the loss ratio will also be due to changes in expenses over time. Expenses as a proportion of written premiums have fallen over the period for some lines due to the increase in direct writers and other changes. I assume this does not have a large effect on the loss ratio.

Clarke, et.al. [1988] examine the variability in the components of underwriting gains or losses and find that there is more variation in the loss ratio than the expense ratio over time. They look at four lines of insurance: homeowners' multiple peril, private passenger auto liability, other liability and medical malpractice from 1976-1986. They conclude that "the components of property-casualty insurers' financial statements most responsible for changes in its position are loss payments and [loss] adjustment expenses" (i.e. the components of the loss ratio numerator).

¹²"Losses" refers to both claims costs and loss adjustment expenses throughout this chapter. The loss reserve is an undiscounted estimate of all outstanding losses and loss adjustment costs. As the policy term elapses the insurer adds to the loss reserve an amount equal to the expected undiscounted losses and loss expenses associated with events that have occurred during the period already passed. These funds include an estimate of the final settlement for known events as well as an estimate for events which are likely to have occurred but have not yet been reported to the insurer. When the claims are paid the reserves for those events are deducted from the loss reserve.

The earned premium is the proportion of revenues from all policies in effect over any part of the accounting period that can be attributed to that accounting period. For example, one half of the premium for a twelve month policy effective on July 1, year t-1 will be counted as earned premiums for calendar year t. Likewise, one quarter of a policy effective October 1, t-1 will be included in the year t-1 earned premium. The loss ratio can be interpreted as a measure of the expected payouts per dollar of premium revenue, but it is a noisy measure since incurred losses include adjustments to the loss reserve for policies which expired prior to the accounting period and unexpected claims payments which occurred during the accounting period.

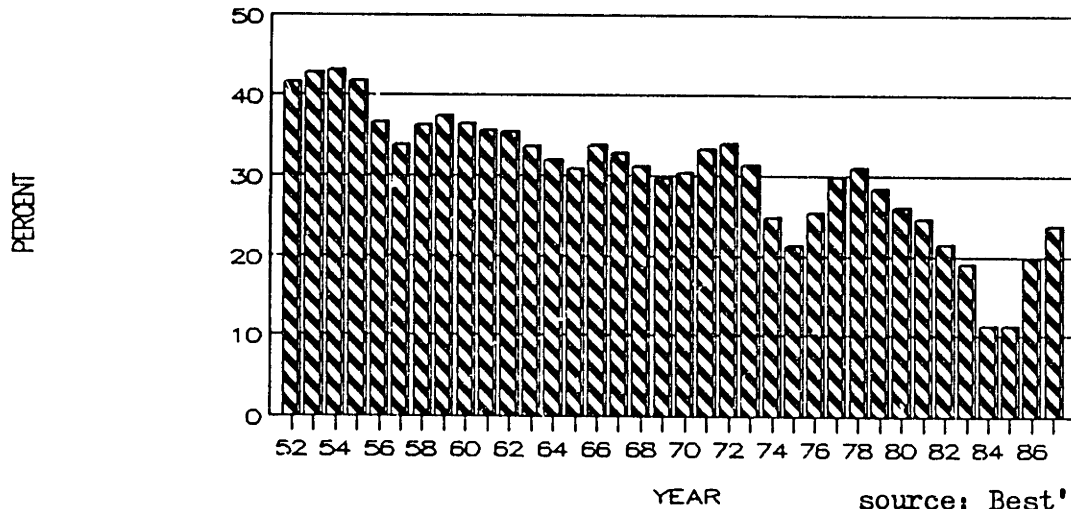
Figure 1 shows accounting profitability as measured by 1-loss ratio (in percent) for the property-casualty industry from 1952 to 1987. During this interval there are 6 episodes of peak profitability and 5 episodes of low profitability.¹³ Comparing this figure to figure 2 indicates a correlation between the number of insurer insolvencies and the observed cycle in accounting profitability. During the years covered, the low points in accounting profits, 1969-70, 1975, 1984-85,

Using the combined ratio instead of the loss ratio also introduces the need to control for changes in underwriting expenses over time. Market share of direct writers of insurance would be a useful variable to control for the level of underwriting expenses, but this measure is not available by line for the time period examined.

¹³The overall decline in accounting profitability is in large part due to increases in the level of expected inflation and expected nominal interest rates as well as an increase in the market share of long tailed lines where these variables have greater impact on accounting profitability. The effects of expected inflation and expected nominal interest rates are demonstrated in section III.

FIGURE 1

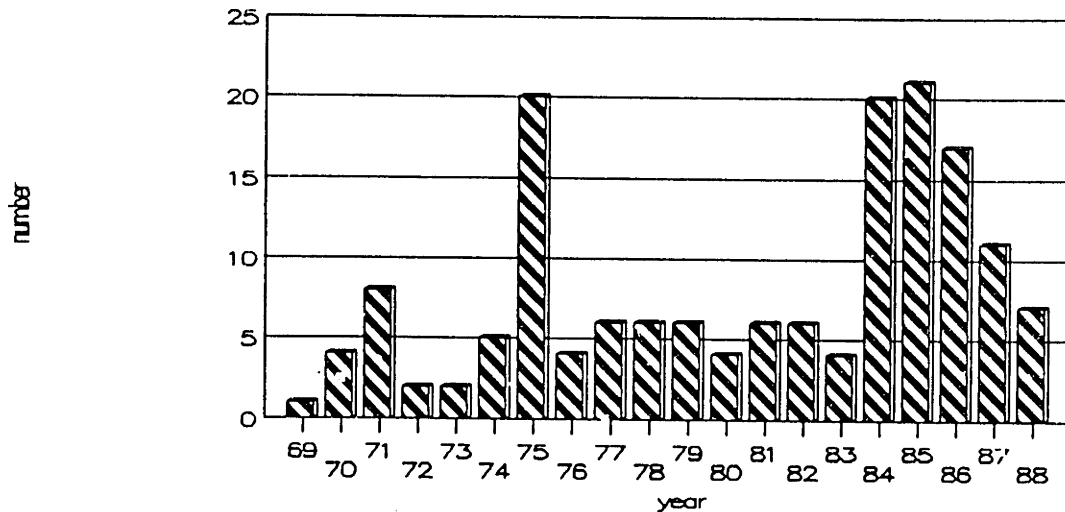
P-C ACCOUNTING PROFITABILITY 1952-1987



source: Best's Aggregates and Averages

FIGURE 2

Insurer Insolvencies 1969-1988



source: Operating Results ISO

roughly coincide with significant increases in the number of insurer insolvencies.

2. Explanations for the Observed Cycle and Previous Empirical Work

The literature on property-casualty insurance pricing falls into three categories: theoretical models of competitive premiums; descriptions and more formal models of property-casualty insurance profitability cycles; and analyses of the most recent insurance crisis.¹⁴ The strand of the literature devoted to property-casualty insurance cycles is covered in this chapter.

The work on property-casualty cycles can be further divided into three categories: models where profit cycles are caused by external factors; models where institutional and behavioral characteristics are the major determinants of insurance cycles; and models which combine elements of both. The "external features" theories generally describe profit cycles as the result of shifting supply which alters the short run equilibrium price and expected economic profitability. The "institutions" approach to modeling property-casualty insurance cycles has concentrated on institutional and behavioral aspects such as regulation, loss forecasting and accounting rules which will produce cyclical accounting profitability measures without necessarily affecting expected economic profit.

¹⁴For an overview of competitive premium models, see Fairley [1979] and Hill [1979], as well as Cummins and Harrington [1987]. Examples of papers that fall into the third category are: Harrington [1988], Priest [1987], and the various works directed at the crisis itself such as Tort Policy Working Group [1986, 1987], and Clarke, et.al. [1988].

Property-casualty trade journals and industry officials often cite the "lack of capacity" in the industry as the underlying cause of restricted availability and increasing premium levels. External features models are aimed at formalizing this idea. In these models the property-casualty industry is usually described as competitive and demand is assumed to be relatively inelastic. The key feature of these models is the nature of insurance supply and, in particular, the effect of external shocks on insurance supply.

The supply of insurance is determined by insurers' capacity or ability to accept risk. The usual measure of industry capacity is total net worth relative to total demand. Large adverse shocks to claims, such as losses from unanticipated inflation or unusually large catastrophes, or shocks to asset values, such as unexpected increases in interest rates or large declines on the stock market, can substantially reduce industry capacity and lead to higher premiums and future profitability. Unexpected increases in demand may also result in low industry capacity, but these models usually assume demand for insurance is relatively stable and most changes in capacity are driven by shocks to supply.

The unexpected decrease in capacity results in short run equilibrium premiums which are above the long run levels. The key assumption is that net worth adjustment costs cause the industry to be temporarily out of long run equilibrium. Industry profitability increases. As new capital and retained earnings increase industry net worth, industry capacity increases and both premiums and profitability tend toward their long run equilibrium levels.

Two informal models of capacity constraints in property-casualty insurance are presented in Stewart [1984] and Bloom [1987]. One unsatisfactory feature of both models is the absence of a mechanism which prevents net worth from quickly adjusting to long run equilibrium. Winter [1988, 1989] and chapter one present more formal models of price determination in competitive markets which address some of these problems. It is assumed that insurers hold net worth in order to keep the probability of bankruptcy low and to maintain the ability of insurers to meet claims made by policyholders.¹⁵ In these theories financial capital does not adjust immediately due to a cost differential between the various sources of capital. Capital from internal sources, such as retained earnings, is assumed to be less costly than external capital.¹⁶ Therefore when capacity is unexpectedly reduced, some external capital will enter the industry, but firms prefer to let the accumulation of retained earnings increase capacity.

The other approach to modeling property-casualty cycles focuses on the time-series structure of the profitability data. The common feature of these models is that institutional aspects of property-casualty insurance are shown to produce a certain time-series structure of

¹⁵Chapter one assumes this is done to comply with a regulatory requirement. Winter assumes that firms cannot become bankrupt.

¹⁶There are several explanations for the cost differential. If managers have better information on the firm's value and investment opportunities than equity holders, they will be more likely to issue equity when the value of the firm is lower than the market value. This will cause equity holders to reduce the market value of the firm when the firm issues equity [Mayers and Majluf, 1984]. In addition, equity holders may require a premium when firms issue equity due to restrictions on the firm's ability to pay out capital.

underwriting results. Any change in the underlying economic conditions is of secondary importance in such models.

Venezian [1985] presents a model where a simple loss forecasting technique in ratemaking "contributes to the fluctuations of underwriting profit margins." Insurers are assumed to use a relatively simple regression methodology to predict future losses. In a theoretical model of the insurance industry, this type of ratemaking produces a specific autocorrelation structure in underwriting profit margins. Unlike the previous models, insurers are assumed to unilaterally set prices and, due to inefficient loss forecasting, profitability varies over time.

A different model is developed by Cummins and Outreville [1987]. The authors state that the property-casualty cycle is "apparent in the sense that it has nothing to do with the underlying economic and statistical characteristics of insurance markets but rather [is] attributable to institutional and regulatory rigidities." Like Venezian, Cummins and Outreville develop a model based on institutional features, such as accounting practices, and regulation, which then predicts that underwriting profit data should follow an autocorrelated process. Cummins and Outreville are unique in their careful modeling of accounting profitability. They recognize that accounting profits measure premiums on an accrued basis which will generate some serial correlation independent of the underlying economic processes.

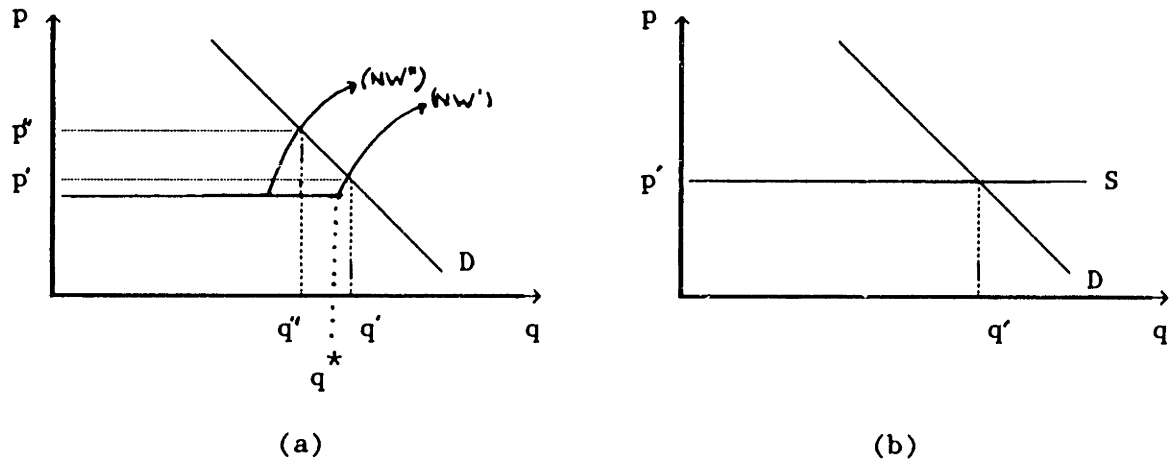
Both of these theories fail to explain important aspects of insurance cycles outside of the second order autocorrelation found in accounting data. Rapidly increasing prices are generally accompanied by a reduction in supply. A substantial number of insureds cannot find

coverage from licensed insurers, and others pay the same or greater premium for lower policy limits and higher deductibles. These models fail to address these features of insurance cycles.

The primary difference between the two approaches lies in the description of short run equilibrium expected return to net worth. The capacity constraint theory as developed in chapter one has an upward sloping short run supply curve, as pictured in figure 3a.¹⁷ If the industry, in long run equilibrium at net worth - NW' , quantity - q' and price - p' , experiences a significant reduction to net worth the short run supply curve will shift to the left ($NW' > NW''$). The new short run equilibrium is defined by NW'' , q'' and p'' . Over time retained earnings and new equity issues will return the industry to the long run equilibrium, defined by the intersection of the long run supply curve, S , and demand, D . Under the institutional features approach, there is no difference between short and long run equilibrium, figure 3b. The expected rate of return on net worth is always at the long run equilibrium, so the short and long run supply curves are horizontal at the long run equilibrium price. As demonstrated in the graphs, the two approaches are not mutually exclusive. Although it is not usually presented as such, the capacity constraint story is the institutional features model with an upward sloping short run supply curve.

¹⁷The assumption that net worth is fixed in the short run (due to adjustment costs) causes a positive relation between total risk (quantity) and price. For any given level of net worth, in order to support a larger quantity, the equilibrium expected return to net worth must be increase in order to maintain a specified low probability of bankruptcy.

FIGURE 3



The regression analysis in section IV of this chapter bridges these two approaches. It examines whether variations in capacity have a role in explaining the movement of accounting profitability for four lines of insurance in a specification which controls for many of the institutional features. The estimation controls for the effect of regulatory accounting conventions which report losses as undiscounted estimates of future liabilities and which report premium income only as it is accrued. Given these institutional features, I then test whether, *ceteris paribus*, measures of variations in capacity have an effect on insurer accounting profitability.

III. SPECIFICATION AND DATA

I have profitability data on four lines of insurance, three lines cover the 1952-1986 period and the fourth covers 1957-1986. The data are for stock insurers only and the four lines are automobile liability, automobile physical damage, homeowners' multiple peril and other liability. I expect some variation in profitability across lines

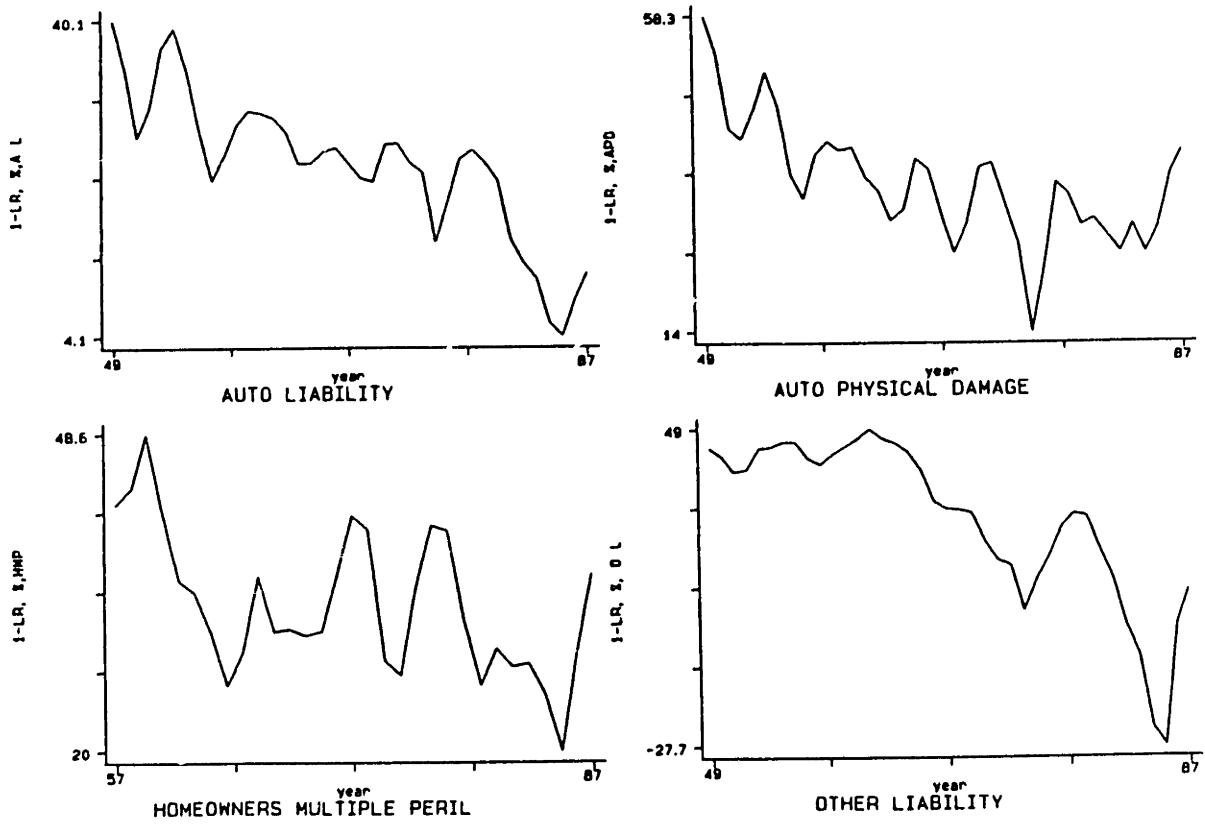
because 1-loss ratio is a measure of actual profitability and unexpected changes in costs will vary across lines. Profitability will also vary across lines if firms cannot costlessly shift capacity between lines.¹⁸ Figure 4 shows accounting profitability over time for each of the four lines considered here. In the two liability lines, there is a noticeable decline in profitability over time. Much of this is explained by the increase in expected inflation and expected nominal interest rates over the period which have larger effects on longer tailed lines.¹⁹

Premium share by line is shown in table 1. Automobile insurance comprises around 40% of the total property-casualty written premiums over the period with auto liability coverage accounting for about 20-25% and auto physical damage for 14-18%. Homeowners' multiple peril was introduced in the early 1950s but did not have substantial market share until the early 1960s. From 1960-1987 homeowners' multiple peril has been about 8-12% of total written premiums. Other liability, which

¹⁸There may be a cost associated with reducing net worth allocated to a specific line of insurance. If insurance buyers have search costs, insureds will tend to stay with their current supplier unless their policies are canceled or not renewed or the premium increases substantially. Under these conditions it may be costly to reduce the quantity supplied this period because it will effect the quantity a firm supplies in future periods when the rate of return to net worth allocated to that line of insurance may be higher.

¹⁹The connection between nominal interest rates, inflation and accounting profitability is laid out in section III.

Figure 4



PROFITABILITY BY LINES

includes medical malpractice insurance, accounts for 5-10% of premiums.²⁰

TABLE 1

PERCENT OF TOTAL STOCK INSURER NET WRITTEN PREMIUM BY LINE

	AUTO LIABILITY	AUTO PHYSICAL DAMAGE	HOMEOWNERS MULTIPLE PERIL	OTHER LIABILITY
1950	19.60	18.42		4.94
1955	23.37	17.52	.77	6.39
1960	25.24	13.68	5.86	7.19
1965	26.44	14.66	8.54	6.49
1970	26.24	14.60	8.74	7.84
1975	23.10	13.14	10.25	7.30
1980	20.36	13.32	11.21	8.22
1985	20.65	14.18	10.80	10.09

source: Best's Aggregates and Averages, 1987.

Other liability includes medical malpractice. Homeowners' multiple peril is 2.26% in 1957, the first year profitability data is available.

A. SPECIFICATION

Accounting profit is measured as 1-loss ratio, in percent. This section describes the components of the loss ratio in detail in order to develop a specification which controls for various factors that

²⁰Other liability insurance includes all commercial liability coverage other than that recorded under commercial multi peril, commercial automobile, and workers' compensation. It includes lines such as manufacturer's and contractor's liability; owners, landlords and tenants liability; and, as reported here, medical malpractice. Medical malpractice was not reported as a separate line until after 1975. In order to make the data consistent, other liability includes medical malpractice for the whole period. Medical malpractice accounted for 1-2% of total premium written by stock insurers from 1975-85. During the mid and late 1970s there was significant entry and exit in the medical malpractice line. This may have some effect on part of the other liability figure reported here, since some of the exiting companies may have been stock firms and some of the important new entrants are hospital mutuals.

influence accounting profitability. Earned premiums, the denominator of the loss ratio, are a weighted average of premiums sold in prior accounting periods and premiums sold in the current accounting period. Since premiums are generally priced when they are sold, earned premiums reflect market conditions over several periods. The numerator of the loss ratio includes revisions to loss reserves for premiums earned in earlier accounting periods and unexpected claims payments made during the current accounting period as well as an estimate of claims costs for premiums earned during the current accounting period. First I describe the factors which determine premium level. Next I examine the numerator of the loss ratio, incurred losses. Then I combine the two and generate an equation to estimate.

The equilibrium premium will be based on expected costs. Under the institutional features models the major expected cost of insurers is the present value of the expected payout. The expected mean nominal loss for year t , \bar{L}_t , is assumed to be the expected mean for year $t-1$ updated by any new information learned about the loss distribution over the last year, adjusted for trends in costs and claims. Equation 1 summarizes this relation:

$$\begin{aligned} \bar{L}_t = E [L_t | \Omega_t] &= (1+\gamma_t) E [L_{t-1} | \Omega_t] \\ &- (1+\gamma_t) E [L_{t-1} | \Omega_{t-1}] + (1+\gamma_t) \epsilon_{t-1} \end{aligned} \quad (1).$$

\bar{L}_t is the expected mean loss for premiums written in year t at time t unless otherwise explicitly noted. E is the expectations operator, L_t is the mean loss for policies written in year t , γ_t is a trend factor which

represents expected inflation as well as expected trends in accident frequency and severity. Ω_t is the information set at the beginning of period t , ϵ_t incorporates any changes in the expected nominal loss from new information in year t . $E[\epsilon_t | \Omega_t]$ is zero since ϵ_t is not known until the end of period t .

The present value of \bar{L}_t is the relevant cost for determining the equilibrium premium. Let c_i be the fraction of the expected loss that is paid i years after the policy is written and let $r_{i,t}$ be the average expected nominal interest rate for period t to $t+i$.²¹ The present value of the expected loss, $PV(\bar{L}_t)$, is then described by equation 2:

$$PV(\bar{L}_t) = \sum_{i=0}^{\infty} \frac{c_i (1+\gamma_t)^i (E[L_{t-1} | \Omega_{t-1}] + \epsilon_{t-1})}{(1+r_{i,t})^i} \quad (2).$$

According to models of insurance premiums with capacity constraints, another determinant of the equilibrium premium is the opportunity cost of industry capacity. Variations in capacity, ceteris paribus, will have an effect on short run premiums because adjustment costs prevent the industry from immediately returning to long run equilibrium. The short run premium reflects the opportunity cost of capacity as well as the present value of expected losses. When capacity

²¹Harrington and Litan [1988] note that the appropriate discount factor depends upon whether insurers view loss payments as riskless or not. If insurers consider loss payments riskless and the insurer bears all investment risk, the risk free rate is the appropriate discount rate. The discount rate should include a risk adjustment if insurers view loss payments as risky due to the probability that actual claims deviate from their expected levels.

is low, the opportunity cost of net worth is high. The short run premium increases due to the increased opportunity cost of net worth. An expression of the short run equilibrium premium is:

$$P_t = PV(\bar{L}_t) + h(CAP_t) \quad (3)^{22}$$

where $\frac{\partial h}{\partial CAP} < 0$.

1-loss ratio is calculated in such a way that the denominator is actually a weighted average of premiums written in year t and prior years. This occurs because revenues are based on *earned* premiums and not *written* premiums. In general, a premium is priced when it is written, or sold, but it is not recorded as income until it is earned. Therefore, if policies are written for one year periods, the denominator will include premiums written in year t-1 as well as year t.

Incurred losses, the numerator of the loss ratio, is calculated as the change in the loss reserve plus losses paid. The loss reserve in year t is an undiscounted estimate of all outstanding liabilities. The change in the loss reserve from the end of year t-1 to the end of year t has three components: an undiscounted estimate of liabilities for premiums in effect during year t net of payments made in year t, revisions to future liabilities for premiums earned in prior years and revisions to reserves to account for claims paid during year t on premiums earned prior to t. This is summarized by equation 4, where

22 The h() function varies across lines of insurance: lines of insurance which require more net worth per policy (those with more uncertainty in their expected losses such as long tailed lines have) incur a proportionately larger share of the increased opportunity cost of net worth and have a larger h() function.

ΔLSR is the change in the loss reserve between the end of year $t-1$ and t , $E[L_p | \Omega_t]$ is the undiscounted expectation as of period t of the loss from policies earned in year p and c_{t-p+1} is the fraction of total losses paid in year $t-p+1$ from premiums earned in year p .²³ The revision to expected payments in year $t-p+1$ on premiums earned in year p is $(c_{t-p+1} \epsilon_t)$.

$$\begin{aligned} \Delta LSR = & (E [L_t | \Omega_{t+1}] - c_0 E [L_t | \Omega_{t+1}]) + \sum_{p=1}^t \sum_{i=1}^{\infty} c_{t-p+1} (\epsilon_t) \\ & - \sum_{p=0}^t c_{t-p} E [L_p | \Omega_t] \end{aligned} \quad (4)$$

Adding losses paid during year t to equation 4 yields incurred losses, IL . Equation 5 describes incurred losses after simplifying terms.

$$IL_t = \sum_{p=0}^t c_{t-p} \left\{ \epsilon_t + \nu_t \right\} + \bar{L}_t + \sum_{p=1}^t \sum_{i=1}^{\infty} c_{t-p+1} (\epsilon_t) \quad (5)^{24}$$

This equation highlights the three components of incurred losses: unexpected payouts during year t on premiums earned prior to t , estimated liabilities for premiums earned in period t , and revisions to outstanding liabilities from premiums earned prior to year t . There are

²³Appendix A derives the underlying expressions for the loss reserve and paid losses in more detail. As explained in Appendix A, Ω_t is the information set at the beginning of period t . Since the loss reserve is reported at the end of period t , the appropriate information set is Ω_{t+1} , the information set at the end of period t .

²⁴Where \bar{L}_t is substituted for $E [L_t | \Omega_t]$.

two sources of unexpected payments in year t . One is new information about the loss distribution, ϵ_t . Period t payments may also differ from their expectations due to random error, ν_t , such as sampling error or large idiosyncratic losses such as a large hurricane or windstorm.

Using equations 3 and 5, 1-loss ratio can be written as:

$$1 \cdot \frac{EP_t}{IL_t} = 1 \cdot \left\{ \frac{\sum_{p=0}^t c_{t-p} \left\{ \epsilon_t + \nu_t \right\} + \hat{L}_t + \sum_{p=1}^t \sum_{i=1}^{\infty} c_{t-p+i} (\epsilon_t)}{(1-\alpha) P_t + \alpha P_{t-1}} \right\} \quad (6)$$

$$\text{where } P_t = \sum_{i=0}^{\infty} \frac{c_i (1+\gamma_t)^i (E[L_{t-1} | \Omega_{t-1}] + \epsilon_{t-1})}{(1+r_{i,t})^i} + h (CAP_t)$$

by substituting equation 2 into equation 3. α , the weight for premiums written in year $t-1$, is equal to $\frac{EP_{t,t-1}}{EP_{t,t} + EP_{t,t-1}}$ where $EP_{j,i}$ is the earned premium written in year i but recorded as revenue in year j . α and $(1-\alpha)$ are assumed to be constant over the period estimated.

Under these assumptions, 1-loss ratio is a function of variables that affect premium levels in years t and $t-1$ and variables that affect incurred losses in year t . The following describes the observed variables which affect 1-loss ratio.

Variations in capacity affect accounting profitability through the effect on short run premiums. Low capacity results in higher short run premiums, *ceteris paribus*. Higher premium levels reduce the loss ratio which increases accounting profitability. Since earned premiums are accounted on an accrued basis, year t earned premiums reflect period t

and $t-1$ market conditions if policies are in effect for 12 months or less. Therefore both period t and $t-1$ capacity variables are included in the specification.

The capacity constraint hypothesis indicates that period t capacity may be endogenous. The cost advantage of internally generated funds implies that retained earnings are an important source of capital when capacity is low. Low capacity in period $t-1$ induces a higher rate of return on net worth in period $t-1$. Retained earnings from $t-1$ will be used to increase capacity in period t . Since accounting profitability reflects profitability in period $t-1$ as well as period t , the coefficient on period t capacity may be biased upwards.

If period t capacity is endogenous, ordinary least squares estimation will produce biased results. Instrumental variables estimation, however, will be consistent. Variables which affect the value of insurers' assets are good instruments for period t capacity because they affect the level of net worth but are not influenced by the industry return to net worth in period $t-1$. Three possible instruments for period t capacity are: unexpected changes in nominal interest rates, the unexpected return on stock investments, and unanticipated losses in prior periods.

Accounting profitability varies with changes in the present value of expected losses. This is measured by the difference between expected inflation and expected nominal interest rates. When this term is positive, the present value of expected losses is higher, causing higher short run premiums, *ceteris paribus*. An increase in the premium level increases accounting profitability.

Changes in any of the three components of incurred losses also affect accounting profitability. Higher expected inflation means, all else equal, that \bar{L}_t^{\wedge} will be larger. Higher than expected inflation will also affect the present value of expected losses but the relevant measure there is the effect of changes in expected inflation net of changes in expected nominal interest rates. The level of expected inflation is expected to have a negative effect on accounting profitability.

Unexpected loss payments in period t and revisions of estimate future liabilities also impact l-loss ratio. When actual cost inflation in year t is greater than expected, payments in year t will be higher than anticipated. These unexpected payments will reduce accounting profitability. Similarly, when losses associated with catastrophic events are higher than expected, claims payments in year t and their associated costs will be higher than expected.²⁵ Revisions to expected cost inflation also affect incurred losses. Upward revisions to losses induce upward revisions in the firm's expected future liabilities. This will have a negative effect on accounting profitability.

.....

²⁵The costs of catastrophic events such as hurricanes, windstorms, other severe weather damage and urban riots are reported in Insurance Facts. The data are compiled only for property losses only, so they apply only to homeowners' insurance.

The basic specification to estimate is then:

$$(1-LR)_t = \alpha_0 + \beta_1 CAP_t + \beta_2 CAP_{t-1} + \beta_3 \Delta PV_t + \beta_4 \Delta PV_{t-1} + \beta_5 \Delta E[INFL]_t + \beta_6 ERRINF_t + \beta_7 E[INFL]_t + \epsilon_t \quad (7)^{26}$$

where the variables are defined in table 2. The estimated coefficients on the capacity measures will include the weights, α and $(1-\alpha)$. The net effect of a change in capacity on accounting profitability is the sum of the estimated coefficients on capacity period t and $t-1$, $\beta_1 + \beta_2$. The net effect of a change in ΔPV is also the sum of the coefficients on period t and $t-1$ variables, $\beta_3 + \beta_4$.

An important simplification has been made concerning the relative timing of premium and loss reserve determination. In this exposition premiums are assumed to be set at the beginning of the period and loss reserves calculated at the end of the accounting period. If the interval of observation were relatively brief, such as monthly or quarterly, this might not be problematic. The observation interval for available data, however, is annual. It is likely that information concerning the difference between actual and expected inflation revealed during the course of the year has additional effects not considered here: the equilibrium premium may increase since expectations of future

26 Estimation for homeowners' multiple peril includes $CATAS_t$, a measure of catastrophic losses. Some homeowners' policies are written for periods longer than a year so longer lags of variables which affect premium levels, variations in capacity and change in present value, may be significant. CAP and ΔPV from period $t-2$ were found to be significant. All the results for homeowners' multiple peril reported here include two lags of both these variables.

TABLE 2

DEFINITION OF VARIABLES

VARIABLES		EXPECTED SIGN
CAP_t	residual from regression of CAP1 on time trend and	
CAP_{t-1}	constant	<0
$\Delta E[INFL]_t$	$E[\text{cost inflation}]_{t+1} - E[\text{cost inflation}]_t$	<0
$ERRINF_t$	actual cost inflation _t - $E[\text{cost inflation}]_t$	<0
$E[INFL]_t$	$E[\text{cost inflation}]_t$ - predicted value for year t from a regression of annual inflation up to year t-1	<0
ΔPV_t	$E[\text{cost inflation}]_t - E[\text{nominal interest rate}]_t$	>0
ΔPV_{t-1}	$E[\text{cost inflation}]_{t-1} - E[\text{nominal interest rate}]_{t-1}$	>0

for homeowners:

$CATAS_t$ The residual from a regression of annual catastrophic losses in thous. of 1983 dollars (those disasters which result in greater than about \$4 million 1983 dollars of damage to insured property) on a constant and a time trend.

$$\begin{array}{rcc}
 CATAS = & \alpha & + \quad \gamma \text{ time} \\
 & 1.06 \cdot 10^5 & 3.50 \cdot 10^4 \\
 (t \text{ stats}) & (.61) & (4.63)
 \end{array}$$

instruments used for capacity

CAP_t : stock market returns last period, $ERRINF_{t-1}$, errors in expected nominal interest rates (actual - expected), $\Delta E[INFL]_{t-1}$

CAP_{t-1} : stock market returns last period, $ERRINF_{t-1}$, errors in expected nominal interest rates (actual - expected), $\Delta E[INFL]_{t-1}$

inflation rates will be revised upwards, nominal expected losses associated with premiums earned in period t will be higher (through greater expected inflation), and equilibrium premiums may increase further if the unanticipated payouts and revisions to loss reserves associated with higher than expected inflation reduce capacity. To the extent that these feedback effects between the variables are significant, the specification will only capture the net annual effect of these changes on accounting profitability. An implicit assumption made by this specification is that the predicted direct effect on accounting profitability will not be fully offset by any following feedback effects.

B. DATA

A measure of capacity by line is not available: insurers do not report net worth allocated to various lines of insurance. Several measures of industry capacity were used in estimation and the instruments available for CAP_t do vary by line. Variations in capacity are measured as the residual from a regression of relative net worth on variables which determine the equilibrium level of industry capacity. Relative net worth is calculated as the aggregate stock insurers' policyholders' surplus as of december 31, year $t-1$ divided by GNP of year t .²⁷ It is interpreted as the level of net worth the industry has

²⁷Relative net worth is the CAP1 variable referred to in table 2. It might be argued that another measure of capacity, one that includes mutual insurers' net worth as well or one that includes total industry net worth, may be preferred to the measure used here. The correlations between these other two measure and the one used here are .97 and .95

upon entering the new period relative to the quantity of goods and services which can be insured during that period.²⁸

Two capacity variables are reported here. The residual from regressing relative net worth on a constant and a time trend is the capacity measure used for the results reported in the text. The second measure controls for changes in economic activity that affect the level of demand as well using a measure of the GNP gap. Those results are reported in appendix C.

The ratio of net written premiums to net worth is often used as a measure of industry capacity. It measures "insurance exposure," or a "company's exposure to pricing errors in its current book of business."²⁹ If the ratio is two, for example, and premiums are revealed to be 10% below the actual costs, then 20% of net worth will be needed to settle claims. This usage assumes that expected costs are a constant proportion of equilibrium premiums. If the difference between premiums

respectively. The correlation between the residuals from regressing the capacity variables on a constant and time trend are above .99. Using aggregate industry policyholders' surplus divided by GNP instead of the measure reported here does not significantly alter the results.

²⁸ An alternative approach to this estimation would be to specify the full structural model, including an equation for capacity as well as an equation for accounting profitability for each line. I have chosen the present methodology because the primary relation of interest is the effect of variations in a measure of capacity on accounting profitability. A structural specification of industry capacity would require, among other things, specifying the adjustment mechanism, adjustment costs, and firm decisions concerning dividend distribution and equity issues in aggregate. The current specification avoids imposing significant structure on relations that are not well understood.

²⁹ The first quote is from Troxel and Breslin [1983], the second quote is from Best's Insurance Reports, P-C Edition [1985].

and expected claims costs varies across time as with the capacity constraint theory, then the interpretation of this measure as capacity and comparisons across time are unclear.

Net written premiums divided by net worth, also called the Kenney ratio, will not be used to measure industry capacity. In addition to the question about its meaning in the presence of short term capacity shortages, the statistic is endogenous to accounting profitability. Net written premiums are total revenues (net of reinsurance ceded and assumed). Following the conventional assumption that most of the variation in written premiums comes from changes in price and not quantity, increases in net written premiums reflect higher prices.³⁰ Higher prices, ceteris paribus, will result in higher profitability.³¹

The expected signs of the variables are summarized in Table 2. The inflation measures used vary by line. The index of medical care costs from the consumer price index was used to construct the inflation measures used for auto liability and other liability. The automobile maintenance and repair index was used for auto physical damage while for homeowners' multiple peril, the consumer price index was used. The interest measures are based on the nominal interest rate on medium term government bonds (three to five years prior to 1952, five year bonds after that). To calculate expected inflation in year t, equation 8 is

³⁰ See Harrington and Litan [1988].

³¹ It is unclear how restrictive the often-quoted net premiums written to net worth regulatory criterion is. There is some evidence that when premiums are generally considered to be at higher levels, higher Kenney ratios are tolerated by regulators (see Using the NAIC Insurance Regulatory Information System, 1985).

estimated using data up to year t-1.

$$I_s = \alpha_0 + \gamma_1 I_{s-1} + \gamma_2 I_{s-2} + \gamma_3 I_{s-3} + \nu_s \quad (8)$$

where I_s is the annual inflation rate for year s. Expected inflation for year t is then $(\hat{\gamma}_1 I_{t-1} + \hat{\gamma}_2 I_{t-2} + \hat{\gamma}_3 I_{t-3})$ where $\hat{\gamma}$ are the estimated coefficients from equation 8. Expected nominal interest rates are determined in the same manner. Errors in expected inflation and nominal interest rates are computed as the difference between actual and expected inflation.³²

The relevant horizon for forming expectations and making revisions to the loss reserve is from the time the new information becomes available to the time at which all currently outstanding liabilities are expected to be settled, i.e. less than or equal to the loss tail of the line. Expectations about effects for periods further in the future will have less impact on accounting profitability than expectations concerning periods closer to the present.³³ It is assumed that the net

³²It is assumed that premiums are priced at the beginning of the period, so they are based on information known at the beginning of year t (up through t-1). Since the loss ratio and loss reserves are from the annual statement as of the end of year t, it is assumed that the insurers calculate these measures using information known at the end of year t (that is, the information available at the beginning of year t+1). Therefore the change in expected inflation which affects incurred losses is determined as the inflation expected next period (t+1) less the expected inflation this period (t) while the difference between expected inflation and expected nominal interest in year t is the relevant measure which effects the present value of losses in year t so far as the premiums are concerned.

³³There are two reasons for this: in incurred losses, expectations concerning periods further in the future effect fewer losses (because the proportion of currently outstanding claims that will still be unsettled at date t+n is declining in n) and in earned premiums the

effect of all expectations and changes in expectations can be approximated by a model where the horizon is one period. The one period expectations are described in table 2. This one period horizon simplification is likely to be more restrictive for longer tailed lines (auto liability and other liability) than for shorter tailed lines such as auto physical damage and homeowners' multiple peril.³⁴

IV. ESTIMATION RESULTS

This section presents results from estimating equation 7 on data for four lines of insurance: auto physical damage, homeowners' multiple peril, auto liability and other liability. The estimated coefficients using ordinary least squares and instrumental variables techniques are consistent with the capacity constraint hypothesis for all lines except other liability. A specification test indicates that the endogeneity of period t capacity t is statistically significant for auto liability but not for auto physical damage or homeowners'. An alternative specification is estimated which tests whether insurers smooth accounting profitability by manipulating reported loss reserves. The data indicate that any evidence of smoothing is outweighed by capacity effects, particularly in the longer tailed auto liability line where the effect is most significant.

Results from ordinary least squares (OLS) and instrumental variables (IV) estimation of equation 7 are reported in table 3.

effects of expectations further in the future are discounted.

³⁴One extension of this work is to estimate alternative specifications which provide for longer horizons.

TABLE 3

REGRESSION RESULTS

(using Robust standard errors correction)

coeff. on	AUTO. PHYS. DAM.		HOMEOWNERS'		AUTO LIABILITY	
	OLS	IV	OLS	IV	OLS	IV
CAP	-.11 (.05)	-.14 (.05)	-.15 (.05)	-.14 (.04)	-.04 (.04)	-.08 (.04)
$\Delta E[\text{INFL}]_t$	-1.07 (.88)	-1.29 (.93)	.84 (1.28)	.58 (1.29)	1.67 (.94)	2.74 (.62)
ERRINF _t	-.41 (.54)	-.48 (.55)	-.63 (.86)	-.47 (.86)	-1.42 (.83)	-2.45 (.54)
E[INFL] _t	-2.60 (.49)	-2.76 (.57)	-1.98 (.36)	-2.01 (.36)	-1.37 (.45)	-1.41 (.55)
ΔPV	.95 (.23)	1.00 (.24)	1.65 (.32)	1.65 (.29)	2.36 (.51)	2.40 (.55)
CATAS _t			$-3.2 \cdot 10^{-6}$ ($1.5 \cdot 10^{-6}$)	$-3.1 \cdot 10^{-6}$ (-1.60)		
N	35	35	30	30	35	35
\bar{R}^2	.52	.50	.65	.65	.72	.68
SEE	5.02	5.09	3.94	3.95	4.23	4.52

standard errors are in parentheses

SEE: standard error of estimate

Variations in capacity are measured as the residual from a regression of relative net worth on a constant and time trend. Since the data are aggregate time series, serial correlation may bias the OLS standard errors.³⁵ The standard errors reported in table 3 are robust to the presence of serially correlated errors. The estimation procedure uses the Newey-West procedure with one lag to correct for the presence of a general form of serial correlation.

The results for the two lines with relatively short loss tails are consistent with the capacity constraint hypothesis. OLS results for both lines indicate that lagged capacity variables have a statistically significant negative effect on accounting profitability. Period t capacity for automobile physical damage has a fairly large positive effect on accounting profitability which is significant at the 30% level (in a two-tail test) although the sum is negative and significant above the 2.5% level (in a one-tailed t test). If the capacity hypothesis is correct, period t capacity (and period $t-1$ for homeowners') is endogenous and the coefficient will be biased upwards. IV estimation reduces both the magnitude and the significance of the period t capacity variable in the auto physical damage equation.

OLS results for auto liability accounting profits are less significant, but also provide results consistent with the capacity constraint hypothesis. The results improve substantially when IV estimation is performed on the equation. The estimated net effect of

³⁵The durbin-watson statistics for the different regressions ranged from 1.10 to 1.60.

the capacity variable increases from -.04 to -.08 and is statistically significant at the 5% level (in a one-tailed t test) under IV estimation.

The results for estimation of equation 7 on other liability accounting profitability, reported in Appendix B, are inconclusive. The industry capacity variable used here does not have a significant impact on accounting profitability. Although the IV estimation alters the coefficient on CAP_t in a direction consistent with the capacity constraint theory, the coefficients on both capacity variables are positive and the one for period t-1 is quite large.

There are several reasons why capacity for other liability may be more difficult to approximate with the industry capacity variable than for other lines. Unexpected payouts and revisions in loss reserves associated with events such as the mass toxic torts and catastrophes such as the 1981 collapse of the walkway in the Hyatt Hotel are likely to generate substantial movements in incurred losses which are not controlled for.³⁶ Reinsurance capacity, which is not reflected in the industry capacity variable, is likely to be a more important determinant of other liability capacity than for the other lines. For these reasons a different approach, such as a firm-level analysis, is required to investigate the effects of capacity on other liability profitability. The results for other liability will be omitted from the following

³⁶One hundred fourteen people were killed and two hundred injured when a walkway collapsed at the Hyatt Regency Hotel in Kansas City. Over 300 lawsuits and other liability claims were filed. (Insurance Facts)

tables and largely ignored in the discussion.³⁷

A test for the endogeneity of CAP_t was also performed. A regression form of a lagrange multiplier test which is robust to serial correlation and heteroskedasticity is outlined in Wooldridge [1989]. The null hypothesis, that CAP_t is exogenous, could not be rejected for auto physical damage and homeowners' multiple peril at the 10% level. For auto liability the null hypothesis was rejected at the 2.5% level.³⁸ This indicates that the feedback effects between profitability and industry capacity are empirically less significant for auto physical damage and homeowners' multiple peril than for the longer tailed line, auto liability. The test was repeated using a non-robust form of the test statistic since the robust form may be less likely to reject the null. The outcome was the same, CAP_t is endogenous only in the auto liability profitability estimation.

The estimated coefficients of the other variables are largely consistent with the expected effects. The coefficient on the change in expected inflation was expected to have a negative coefficient since higher expected inflation causes positive revisions to the loss reserve, ceteris paribus. Policies are written throughout the year, however, and increases in expected inflation will also increase equilibrium premiums

³⁷ An F test for change in regime between 1952 -1968 and 1969-1986 rejected the null hypothesis, that all the coefficients were the same for the two periods, at the 5% level for other liability accounting profitability. The 68/69 cutoff was chosen because this time marks the beginning of many of the significant changes in the tort system which are often cited as having a significant effect on liability insurance.

³⁸ A test on other liability rejected the null at just over 5%.

through its effect on the present value of expected losses. The coefficient on $\Delta E[\text{INFL}]_t$ ranges from negative to positive and statistically significant above the 1% level when the three lines are ordered from shortest to longest loss emergence pattern, i.e. auto physical damage, homeowners' multiple peril, auto liability. It appears that for longer tailed lines the premium effect dominates the increase in incurred losses. The coefficient on $\Delta E[\text{INFL}]_t$ may also be capturing capacity effects caused by the feedback from the errors to capacity. Large increases in expected inflation will cause adjustments to the loss reserves, reducing net worth. Longer tailed lines will have a larger proportion of claims outstanding so the revisions may have a larger impact on line capacity.

The other variables all have the expected signs and many are significant at the 10% level. The results across lines are consistent with the predicted effects. The ΔPV variable has a larger and more significant effect on lines with longer tails. At least some of the greater volatility observed in accounting profitability of longer tailed lines over the past 15 years in figure 4 can be attributed to greater swings in interest and inflation rates which have a larger effect on the accounting profitability of these lines. For example, an increase of 1 point in ΔPV increases the auto physical damage loss ratio by 1 point and the loss ratio of auto liability by 2.4 points.^{39,40}

³⁹The results of estimation using another measure of capacity are presented in Appendix C. Most of the above observations apply.

⁴⁰Including errors from equation 8 to measure unexpected cost inflation (ERRINF) may cause the estimated coefficient on ERRINF to be biased if

The estimation procedure thus far has assumed that insurers fully adjust loss reserves to new information as soon as it is known and that loss reserves are not discounted (either explicitly or implicitly).⁴¹ Insurers may find it optimal not to fully adjust loss reserves in the event of news about losses. Firms may understate incurred losses when income is low and overstate incurred losses when income is high in order to smooth underwriting profits over time. Weiss [1985] finds evidence consistent with income smoothing activity for automobile liability lines from 1955-1975.

Smoothing is most likely to occur in the longer tailed lines. Reserves for these lines cover longer periods, there is more uncertainty as to the final payout and thus more room for manipulation. If insurers influence underwriting results by only partially adjusting loss reserves as new information is revealed, lagged values of inflation errors and changes in expected inflation should have a negative effect on accounting profitability. As claims are settled and paid in subsequent periods insurers will adjust loss reserves to the "old" news. Errors in

there are omitted variables that are correlated with ERRINF, such as unexpected inflation in other costs. To the extent that ERRINF is included to control for all unexpected cost inflation, this may not present a significant problem since the coefficient itself is not being used to predict the effects of changes in a specific cost on accounting profitability. If ERRINF is orthogonal to the other variables in equation 7, possible correlation with the error term will not bias other coefficients. This indicates that care should be exercised in interpreting the coefficients in table 3.

⁴¹In general, discounting of loss reserves was not allowed by Statutory Accounting Principles prior to the Tax Reform Act of 1986, except for workman's compensation, which is not considered in this study (Strain, 1986).

inflation in period $t-1$ will have a negative effect on accounting profitability in period t if insurers are smoothing underwriting results. Note, however, that under the capacity hypothesis the sign on this variable will be positive: large errors in inflation expectations in period $t-1$ lead to adjustments in reserves which reduce net worth. Lower net worth reduces capacity which causes next period's profitability to increase.

Table 4 reports the results from equation 7 when lagged errors in inflation are included to test whether insurers smooth underwriting results. For auto liability, in the last two columns of table 4, the positive coefficient on lagged errors in expected inflation supports the capacity hypothesis, it suggests that any smoothing effects are dominated by the effect on capacity of unanticipated losses. The inclusion of lagged errors decreases the magnitude and significance of the capacity variable indicating that it may be measuring some of the same effects. The net effect of errors in inflation expectations on 1-loss ratio is negative, lagged errors have a positive effect, but the negative effect of current unexpected losses is substantially larger. When lagged errors are included in the auto physical damage and homeowners' multiple peril equations the estimated effects are smaller and not statistically significant. In the auto physical damage equation the coefficient on lagged errors in expected inflation is negative and its inclusion reduces the coefficient on change in expected inflation. The coefficient on lagged error in expected inflation is positive in the homeowners' multiple peril equation and its presence increases the estimated effect of both period t error in expected inflation and change

TABLE 4

ESTIMATION INCLUDING LAGGED ERRORS

using robust standard errors correction

coeff. on:	AUTO. PHYS. DAM.		HOMEOWNERS' *		AUTO LIABILITY	
	OLS	IV	OLS	IV	OLS	IV
CAP	-.11 (.05)	-.14 (.05)	-.14 (.05)	-.13 (.05)	-.02 (.04)	-.05 (.04)
$\Delta E[\text{INFL}]_t$	-.45 (.88)	-.49 (.84)	1.51 (1.41)	1.34 (1.32)	3.14 (.68)	3.77 (.87)
ERRINF _t	-.38 (.47)	-.47 (.46)	-1.12 (1.03)	-1.03 (.88)	-2.76 (.52)	-3.39 (.71)
$E[\text{INFL}]_t$	-2.39 (.60)	-2.52 (.68)	-1.99 (.37)	-2.01 (.37)	-1.71 (.39)	-1.69 (.43)
ΔPV	.98 (.23)	1.04 (.26)	1.53 (.27)	1.50 (.23)	2.14 (.38)	2.21 (.43)
ERRINF _{t-1}	-.66 (.51)	-.95 (.57)	.49 (.65)	.56 (.76)	1.92 (.44)	1.66 (.53)
N	35	35	30	30	35	35
\bar{R}^2	.52	.49	.63	.63	.80	.78
SEE	4.98	5.15	4.02	4.03	3.54	3.78

standard errors are in parentheses

SEE: standard error of estimate

* The coefficient on CATAS is similar in magnitude and precision to that in table 3.

in expected inflation. The sum of the coefficients on the error in expected inflation is approximately the same as the estimated effect from table 3 where only the period t variable is included.

The conclusion from these results is that variations in capacity do affect insurer profitability in a manner consistent with the capacity constraint hypothesis when other factors are controlled for. For short-tailed lines, which are predominantly property coverages, the industry measure of variations in capacity used here appears to approximate by line capacity reasonably well. A robust test for endogeneity of contemporaneous capacity failed to reject the null of exogeneity for auto physical damage and homeowners' multiple peril but rejects the null for auto liability and other liability. For automobile liability, the industry capacity variable had less effect on profitability, but the effects of other variables on accounting profitability are consistent with the capacity constraint story. In particular, the effects of lagged inflation errors and the effect of $\Delta E[\text{INFL}]$ provide support for the role of insurer capacity in automobile liability profitability.

V. STOCK MARKET EVIDENCE

Stock market reaction to two major catastrophes, Hurricane Hugo and the San Francisco earthquake provides, additional evidence that variations in industry capacity affect expected profitability. The institutional features approach indicates that the market value of insurers will decline following a major unexpected loss event, reflecting the decline in capital associated with the larger than

anticipated losses.⁴² In the capacity constraint hypothesis, however, there are two offsetting effects from a large loss event. The capital loss associated with higher than anticipated claims will decrease market value, while the change in the expected rate of return on net worth remaining in the industry will increase market value.

The magnitude of the change in expected rate of return will depend upon several factors, particularly where industry capacity is relative to its long run equilibrium, and the size of the unexpected loss. When capacity is relatively abundant there is less probability that an unexpected loss event will result in a capacity shortage and increase the expected return to net worth. If capacity is already scarce and profitability is above long run equilibrium, the loss event may not significantly increase the expected rate of return on net worth in the industry because the high profitability is rapidly generating retained earnings and attracting external capital to add to current capacity.

Firms associated with the property-casualty insurance industry can be separated into three groups based on the degree to which unexpected losses affect market value. Firms which are predominantly property-casualty insurers-- those whose primary activity is property-casualty insurance-- will be significantly impacted by both effects. Multiline firms, which have substantial resources allocated towards life and/or health insurance as well, will have impact from both effects as well. Since multiline firms have significant other

⁴²It is very possible that the reduction in stock market value is less than the total unexpected payout since the probability of a catastrophic loss may already be incorporated in the market value.

activities, they may have additional sources of internal capital. If this is the case, the increase in value from the higher expected return to capital in the industry will be larger for multiline firms than for property-casualty insurers. The third category, brokers, bear none of the losses but participate in the increased income associated with higher premiums. Commissions, usually a percentage of premiums, are the major source of revenues for brokers. Thus when premiums increase, so will brokers' revenues. The 18 firms used in this analysis are listed in Appendix D.

The response of the market is measured using as the excess return, ER_i , defined as return on stock i , R_i , minus the market return, RM :

$$ER_i = R_i - RM \quad (9).$$

Daily returns for each stock are calculated and the simple average for each of the three market segments is determined.⁴³ Then the market return, as measured by the return on the New York Stock Exchange composite index, is subtracted from the average return to arrive at the excess return. The t-statistics reported in tables 5 and 6 use a standard deviation of the excess return for the three groups calculated over five days prior to both events.⁴⁴

On September 18 and 19, 1989, Hurricane Hugo passed over the Caribbean Islands causing an estimated \$1-2 billion of damage. Hugo

⁴³The returns are calculated using the difference in logs method.

⁴⁴The five days used to calculate the standard deviations are September 12-18, 1989.

then hit Charleston, South Carolina late on September 21 and proceeded northward and inland. Total damage to insured property is estimated at \$4-5 billion.⁴⁵ As of the morning of the 21st it was still unclear where Hugo would hit the U.S. coastline and how costly it would be.⁴⁶ The losses from Hugo were larger than might have been expected just prior to landfall for two reasons: it struck land at Charleston where property values are high and it then moved inland causing additional damage by heavy rains and numerous tornadoes.⁴⁷

Table 5 shows net of market returns for the three industry groups. Although insurers sustained heavy losses from the storm, the stock returns reflect little decrease in market value. While the companies with the largest loss exposure, in particular State Farm, Allstate, and Nationwide, are not included in the sample and insurers' losses are limited by reinsurance contracts, firms in the sample did sustain significant losses.⁴⁸ It may be that much of the expected loss from the storm was anticipated and thus already incorporated into the market values by the time Hugo hit the continental United States. The cumulative excess of market returns, beginning 7 trading days prior to

⁴⁵ New York Times 10/22/89.

⁴⁶ Wall Street Journal 10/21/89

⁴⁷ Wall Street Journal 9/25/89

⁴⁸ The Wall Street Journal reported that Continental Corp. would take a \$50 million pretax charge against third quarter earnings for losses associated with Hugo, while CIGNA and USF&G might also take one-time charges against earnings [9/27/89 p. A13]. State Farm estimated its losses at about \$400 million, Nationwide at \$100 million, Allstate's exposure is estimated to be \$175 million. AIG and "other big writers of commercial lines" each face losses of \$30 million or more [Businessweek, 10/9/89, p. 46]. All estimates are net of reinsurance.

TABLE 5
HURRICANE HUGO

NET OF MARKET RETURNS

day	P/C STOCKS		MULTILINE STOCKS		BROKER STOCKS	
	daily	cumulative	daily	cumulative	daily	cumulative
-7	-.21		-.26		-.33	
-6	.83		-.07		.91	
-5	.10		.37		.31	
-4	-.29		.97		-1.22	
-3	.01		-.75		.21	
-2	-.01		.80		-.41	
-1	.11		.20		.43	
0	-.19		-.19		.28	
1	-.50 (-1.25)	-.50	.74 (1.28)	.74	.21 (.29)	.21
2	-.31 (-.78)	-.80	-.43 (-.74)	.31	.36 (.50)	.57
3	.25 (.63)	-.55	.33 (.56)	.64	.30 (.42)	.87
4	-.22 (-.56)	-.77	-.18 (-.31)	.46	-.72 (-1.01)	.15
5	-.06 (-.14)	-.83	.40 (.69)	.86	1.38 (1.93)	1.53

day 0 is September 21. Hugo hits Charleston the evening of the 21st.
t statistics are in parentheses

landfall (September 13) for property-casualty and multiline insurer stocks are shown in figure 5. Property-casualty stocks were up just under one percentage point over the market for most of the 7 days prior to landfall. The cumulative excess returns for multiline insurers varied more in the days before Hugo hit the South Carolina coast, but the range is quite narrow: from -.5 to just over 1 percentage point. There is no evidence of a significant and sustained drop in market value prior to the event date. The value weighted average three day cumulative excess returns are .43, .66 and 1.00 percent for property-casualty, multiline and broker stocks respectively. The corresponding average changes in market value are \$15 million for property-casualty firms, \$28 million for multiline insurers and \$22 million for broker firms. The data show that any decline in value from the capital loss is being offset, to varying degrees, by the expected increase in profitability from the reduction in capacity.

The market response to the San Francisco earthquake is much more striking, as shown in table 6. The quake occurred just after 5 p.m. on October 7, two weeks after Hurricane Hugo. The following day multiline insurance stocks had an excess return of 1% and the average excess return for broker stocks was 5.3%. Property-casualty stocks showed no effect of the earthquake on market value.

The positive effect on multiline and broker stocks persist in the three and five day cumulative returns, shown in figure 7. Property-casualty insurance stocks are essentially unchanged over the period, multiline stocks have a simple average cumulative increase of 2.5% the first three days and 3% for the five days after the earthquake.

FIGURE 5
CUMULATIVE EXCESS RETURN
HURRICANE HUGO

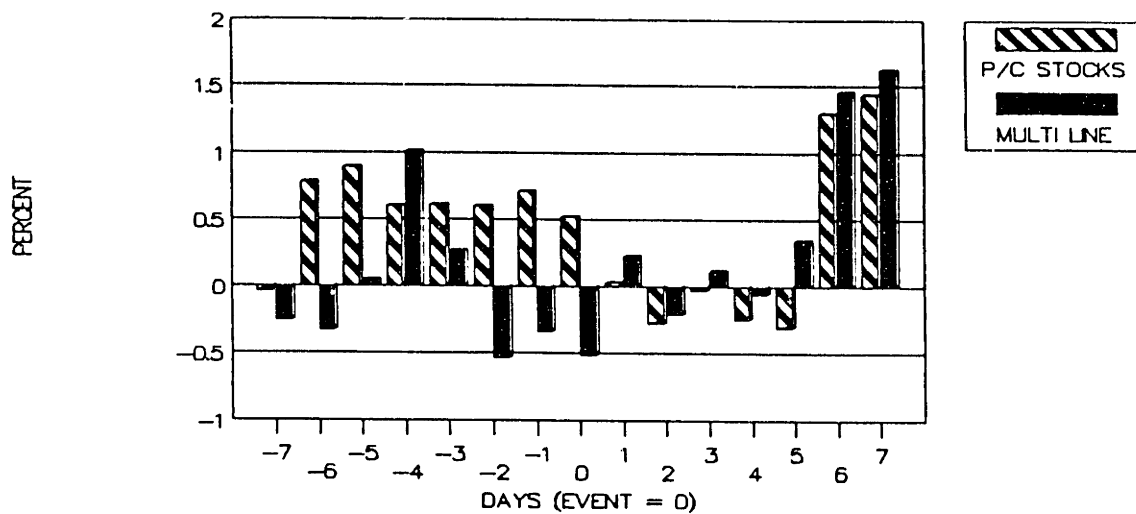


FIGURE 6
THREE AND FIVE DAY CUM. RETURN
HURRICANE HUGO

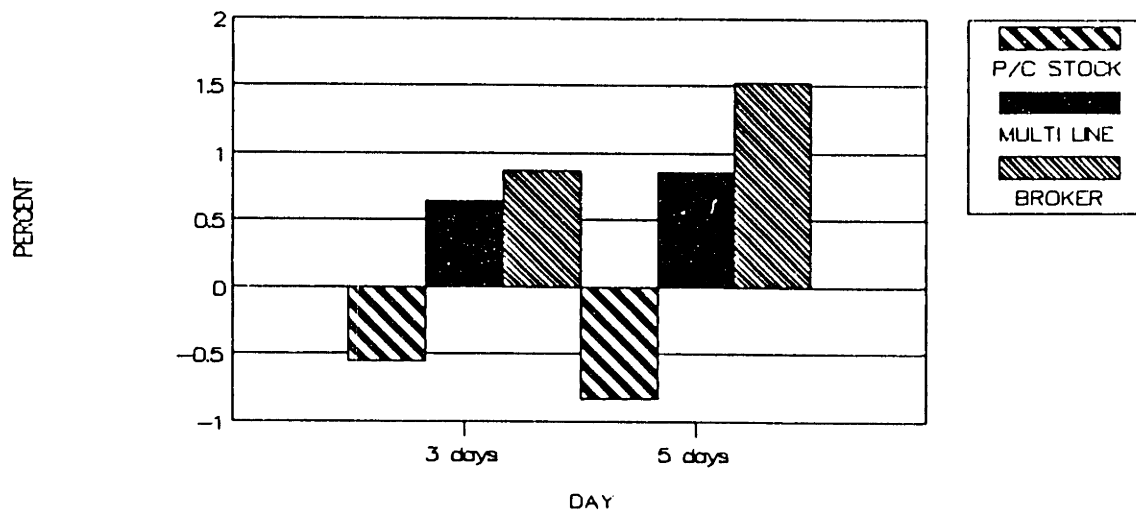


TABLE 6

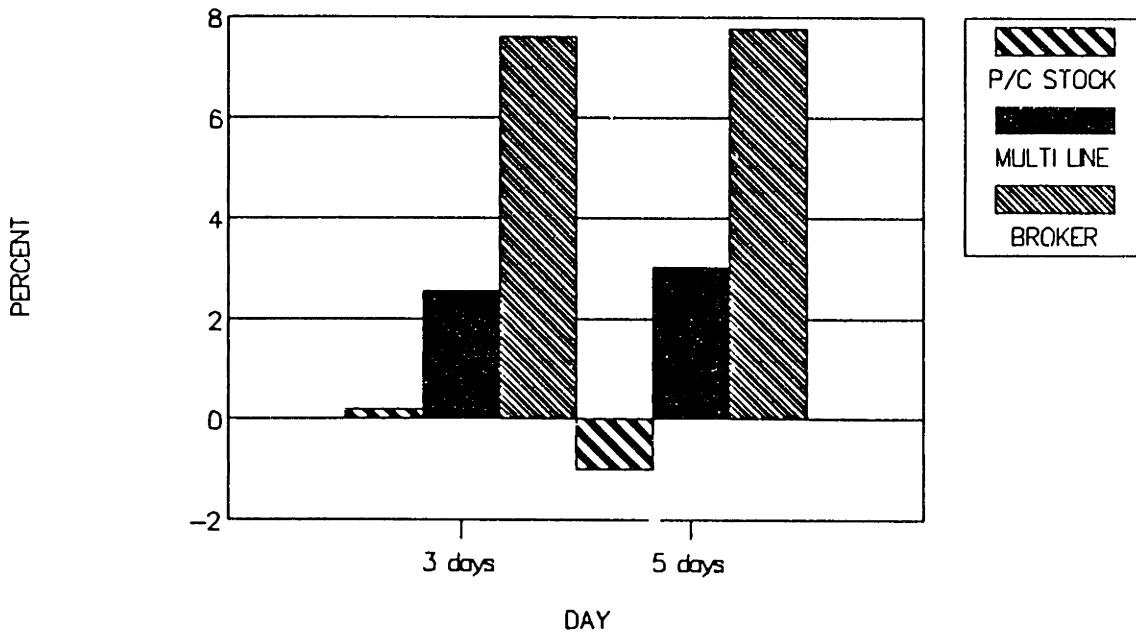
SAN FRANCISCO EARTHQUAKE

NET OF MARKET RETURNS			
	P/C STOCKS	MULTILINE STOCKS	BROKER STOCKS
day	daily	daily	daily
-4	.06	.00	.01
-3	.26	.16	1.80
-2	1.45	1.62	2.08
-1	-1.54	-1.51	-1.59
0	.33	-.45	-2.08
1	-.33 (-.83)	1.37 (2.35)	5.31 (7.42)
2	.53 (1.47)	-.15 (-.20)	1.80 (2.51)
3	.003 (.00)	1.32 (2.28)	.51 (.71)
4	-.89 (-2.49)	.53 (.91)	.36 (.50)
5	-.29 (.80)	-.03 (-.05)	-.19 (-.27)
3 day cum.	.20	2.54	7.61
5 day cum.	-.98	3.03	7.78

day 0 is tuesday, Oct. 17. The earthquake occurred just after 8:00 p.m. EST.

t statistics are in parentheses

FIGURE 7
THREE AND FIVE DAY CUM. RETURN
SAN FRANCISCO EARTHQUAKE



Broker stocks have a cumulative excess return of 7.6 and 7.8% for the three and five days after the earthquake. The value weighted average three day cumulative excess return for property-casualty, multiline and broker stocks are 3.08, 2.52 and 8.51 percent, respectively.⁴⁹ Property-casualty insurers had an average increase in value of \$108 million, multiline insurers had an average increase of \$111 million, and brokers had an average increase of \$198 million for the same period. The results are consistent with the capacity constraint hypothesis. For property-casualty insurers, the losses from the earthquake are just offset by the expected increase in future profits from the decrease in capacity using the simple weighted average return. The increase in expected profits more than offsets the loss for property-casualty insurers if calculated using value weighted averages, and for multiline insurers using either measure. The average excess return for broker companies, which participate in premium increases but do not pay claims, is over twice as large as that of multiline companies, both in terms of simple and value weighted averages.

Hurricane Hugo occurred at a time when some insurers had already begun to increase premiums for some lines of insurance, suggesting that, for some insurers, capacity was less abundant.⁵⁰ Under these conditions, the significant capital loss associated with Hugo is expected to have substantially increased the expected return to capital still in the

⁴⁹The five day value-weighted cumulative return is 2.61 percent for property-casualty insurers, 2.51 percent for multiline insurers and 9.55 percent for brokers.

⁵⁰See Wall Street Journal Nov. 6, 1989, p. A7A.

industry.⁵¹ The market reaction to the San Francisco earthquake reflects the additional increase in the expected return to net worth from a significant loss of capital to the industry at a time when capacity is already substantially reduced from the hurricane.

An alternative explanation is that the two events caused a change in insurance demand. If insurance consumers have difficulty in estimating the probability of very low frequency events, then large, well-publicized events such as Hugo and the earthquake, may cause them to change their subjective probabilities of these events. This will produce a shift in demand for insurance covering catastrophic events. If firms have some reputational capital, those firms supplying insurance for catastrophic events will realize a positive effect from the increase in demand. The positive effect from the demand shift may more than offset the losses to firms in the sample especially since those with the greatest losses are not included.

The data analyzed here cannot distinguish between these two hypotheses. It may well be that both capacity and demand factors played a role in the outcome. Combined with the evidence from section IV, however, the conclusion that changes in capacity play a primary role in the observed change in market value is quite reasonable. Most of the transitions to high premiums and accounting profitability covered in the time series analysis are not associated with highly publicized catastrophic events which might cause a sudden and substantial shift in

⁵¹ Even if net worth in the industry is expected to have a higher return when there are capacity shortages, Hugo increased the probability that a shortage would occur in the immediate future.

demand.

VI. SUMMARY AND CONCLUSION

The aggregate time series regression analysis relating accounting profitability to variations in capacity and other important variables demonstrates that variations in capacity have a significant effect on movements in property-casualty profitability. The evidence is consistent with the capacity constraint hypothesis where net worth adjustment costs prevent the industry from rapidly returning to long run equilibrium following shocks to losses or asset values.

The evidence from the stock market reaction to Hurricane Hugo and the San Francisco earthquake are consistent with the capacity constraint analysis. The stock market reaction shows that large reductions in net worth have two offsetting effects on the market value of insurance firms: one is the capital loss associated with event itself and the other is the expected increase in future returns to the remaining net worth in the industry. The stock market evidence supports the hypothesis that real returns to industry net worth increase when industry capacity is significantly reduced.

The mechanism underlying property-casualty insurance cycles is of particular interest to insurance regulators. Regulators are interested in controlling the large premium swings and reduction in available coverage that accompany the transition from low premiums and accounting profitability to high premiums and accounting profitability. These abrupt market changes can have real effects on economic activity as observed during the 1985-6 "insurance crisis" when large increases in

premiums and reductions in quantity caused some businesses to withdraw products and services and some municipalities suspended certain activities and services. As a consequence of this experience numerous states have enacted tort reform measures aimed at reducing liability costs. While these actions are likely to reduce insurance losses and therefore premiums, the results of this chapter indicate the measures will not prevent future insurance cycles as some regulators and legislators imply.

Another recent trend in insurance regulation is the movement toward increased price regulation. Whereas earlier price regulation was aimed at establishing minimum premium levels, the current regulation is directed at keeping rates down.⁵² Under the capacity constraint hypothesis, short run variation in the expected return to net worth is part of the adjustment process. When capacity is below the long run equilibrium level, higher expected returns to capital generate retained earnings and attracts additional capital to the industry. This process increases net worth and brings the industry toward long run equilibrium. If the new forms of premium regulation prevent or dampen variations in expected returns to net worth, there may be substantial impact on the dynamics of property-casualty insurance markets.

Another implication of the results is that the current form of guaranty funds may exacerbate the effects of a capacity shortage.

⁵²The most controversial example of this is California's proposition 103 which mandates a rate rollback on all property-casualty premiums to 20% below 1987 levels and places substantial restrictions on firms' ability to cancel or refuse to renew policies as well as restricting future rate adjustments.

Insurance buyers are protected from insurer insolvency through state guaranty funds. These "funds" generally operate as follows. Once an insurer is declared insolvent, the state insurance department assesses all licensed insurers (in that state) to contribute to the fund.⁵³ While there is a maximum annual assessment that can be made, insurers may be assessed several years in a row. Policyholders of the insolvent insurer will have their claims paid up to a state-set maximum-- usually \$300,000.⁵⁴

Rules governing how the fund operates vary by state. Guaranty funds differ in the lines of insurance they cover and the limit per claim. Some states separate insurance into "accounts," such as workman's compensation, automobile and other lines, and assess only those insurers licensed to write the same lines as the insolvent company. Many states provide for a mechanism by which insurers operating within a state can police one another. Organizations called state guaranty associations, made up of insurers licensed in the state, were intended to prepare reports on the causes of insolvencies and "alert the regulators about financially troubled companies." The associations have not operated in this manner, due to potential antitrust allegations.

If the property-casualty insurance cycle is associated with variations in capacity, as supported by the results, this form of

⁵³Only New York's guaranty fund operates with a prefunded system.

⁵⁴The primary source for this discussion is Government Accounting Office, July 1987.

guaranty funds will exaggerate any capacity shortages. Figures 1 and 2 show that larger numbers of insolvencies are correlated with the abrupt transition between low and high premiums and profitability. A regulation which further reduces net worth by assessing solvent insurers will further reduce industry capacity. While guaranty funds do protect a few consumers from insurer insolvencies, they may increase premium level throughout the industry by reducing net worth at a time when the industry is already capacity constrained.

In closing, I would like to indicate some areas for further research. The capacity constraint hypothesis relies on the assumption that there is a cost differential between externally and internally generated capital. The evidence here along with the results from chapter two provide support for the theory which is based on this assumption. Further evidence of this cost differential might be observable in firm behavior, specifically in decisions to issue new equity.

The regression results indicate that there is a difference between capacity for auto liability and the two predominantly property coverages, homeowners' and auto physical damage. They also suggest there has been substantial change in other liability insurance over the time period examined. Additional research on the components of liability capacity and the factors which influence other liability accounting profitability is clearly needed.

APPENDIX A

This appendix details the assumptions and calculations underlying the characterization of incurred losses as developed in section III of the text. Incurred losses reflect revisions to expected claims from expired policies, unexpected changes in payouts made during the accounting period and the undiscounted estimate of future payouts from premiums earned during the accounting period. In order to arrive at a specification for estimation which controls for the effects of these factors on accounting profitability, it is instructive to describe the separate components of incurred losses in some detail.

The loss reserve in year t is a current estimate of all outstanding liabilities at the end of the year. In the discussion so far, all timing has been based on the beginning of the period. To keep the notation consistent, the information available to insurers at the time the year t loss reserve is reported will be Ω_{t+1} (the same information set insurers have at the beginning of the next period).

The loss reserve for year t is equal to the loss reserve at the end of the previous year, plus the estimated liabilities from policies earned in year t net of any payments made to those policies, less the funds allocated for payments in year t associated with premiums earned prior to period t , plus any revisions to future liabilities from premiums earned prior to t .

$$\begin{aligned}
 \text{LSR}_t - \text{LSR}_{t-1} &= \sum_{p=0}^{t-1} c_{t-p} E[\hat{L}_p | \Omega_t] + (E[\hat{L}_t | \Omega_{t+1}] - c_0 E[\hat{L}_t | \Omega_{t+1}]) \\
 &+ \sum_{p=1}^t \sum_{i=1}^{\infty} c_{t-p+i} (\epsilon_t) \tag{A1}
 \end{aligned}$$

where p is the subscript for premiums earned prior to year t , i indexes years (beyond t). Note that estimated mean losses for year t is $\hat{\bar{L}}_t$, not \bar{L}_t as in equation 2 in the text. The estimated future liabilities from year t correspond to premiums earned in year t which are not necessarily the same as the losses associated with the premiums written in year t .

Change in the loss reserve, ΔLSR , is easily found by subtracting LSR_{t-1} from both sides of equation A1. Incurred losses (IL), the numerator for the loss ratio, is then found by adding paid losses to the ΔLSR .

$$\begin{aligned}
 \text{IL}_t &= c_0 \left\{ \hat{\bar{L}}_t + \epsilon_t + \nu_t \right\} + \sum_{p=0}^{t-1} c_{t-p} \left\{ \hat{\bar{L}}_p + \epsilon_t + \nu_t \right\} - \sum_{p=0}^{t-1} c_{t-p} E[\hat{L}_p | \Omega_t] \\
 &+ (\hat{\bar{L}}_t - c_0 \bar{L}_t) + \sum_{p=1}^t \sum_{i=1}^{\infty} c_{t-p+i} (\epsilon_t) \tag{A2}^{55}
 \end{aligned}$$

Canceling terms and simplifying yields equation A3. The ϵ_t is the total effect on nominal losses of new information about cost inflation, loss frequency, etc. learned in year t ; the ν_t term is the effect of random

⁵⁵All expectations are as of the beginning of period t ($\Omega = \Omega_t$) unless explicitly written otherwise. The last term now includes the information revealed in year t that changes estimated losses for premiums earned in t .

shocks to the losses (i.e. sampling errors, large idiosyncratic losses such as a large hurricane).

$$IL_t = \sum_{p=0}^t c_{t-p} \left\{ \epsilon_t + \nu_t \right\} + \hat{\bar{L}}_t + \sum_{p=1}^t \sum_{i=1}^{\infty} c_{t-p+i} (\epsilon_t) \quad (A3)$$

APPENDIX B

TABLE B1

OTHER LIABILITY ESTIMATION

coeff. on	<u>PERIOD 52-68</u>		<u>PERIOD 69-86</u>		<u>PERIOD 52-86</u>	
	OLS	IV	OLS	IV	OLS	IV
CAP_t	.05 (.05)	.04 (.06)	.55 (.22)	.46 (.38)	.22 (.12)	.03 (.11)
CAP_{t-1}	.04 (.05)	.05 (.05)	-.32 (.22)	-.26 (.32)	.02 (.08)	.16 (.08)
$\Delta E[INFL]_t$.66 (1.07)	.67 (1.07)	-9.58 (6.22)	-7.71 (9.00)	.56 (1.46)	2.28 (1.35)
$ERRINF_t$	-2.58 (.67)	-2.62 (.68)	8.07 (5.48)	6.42 (7.91)	-.37 (1.69)	-2.03 (.98)
$E[INFL]_t$	-4.36 (1.99)	-4.41 (1.99)	-2.86 (1.30)	-2.69 (1.43)	-2.04 (.69)	-2.11 (.80)
ΔPV_t	1.92 (.99)	1.92 (.99)	3.07 (1.90)	3.41 (2.23)	4.60 (.99)	4.79 (1.20)
ΔPV_{t-1}	-.08 (.64)	-.08 (.64)	3.87 (1.35)	3.65 (1.57)	2.18 (.65)	2.05 (.76)
N	17	17	18	18	35	35
\bar{R}^2	.77	.77	.74	.73	.83	.81
SEE	2.48	2.48	8.71	8.79	8.31	8.70

standard errors are in parentheses

SEE: standard error of estimate

APPENDIX C

TABLE C1

REGRESSION RESULTS

(using other capacity measure)

coeff. on	AUTO.PHYS.DAM.		HOMEOWNERS'		AUTO LIABILITY	
	OLS	IV	OLS	IV	OLS	IV
CAP	-.07 (.07)	-.10 (.07)	-.16 (.06)	-.18 (.04)	-.01 (.05)	-.06 (.06)
$\Delta E[\text{INFL}]$	-3.93 (1.84)	-4.34 (1.91)	.86 (.98)	1.16 (1.01)	-.33 (1.52)	1.72 (1.92)
ERRINF_t	1.14 (.88)	1.15 (.88)	-.56 (.59)	-.74 (.61)	.12 (1.26)	-1.69 (1.62)
$E[\text{INFL}]_t$	-1.73 (.48)	-1.85 (.50)	-1.21 (.33)	-1.18 (.34)	-1.42 (.42)	-1.29 (.46)
ΔPV	.39 (.43)	.36 (.44)	.65 (.32)	.65 (.32)	2.66 (.48)	2.60 (.53)
CATAS_t			-3.1×10^{-6} (1.4×10^{-6})	-3.3×10^{-6} (1.45×10^{-6})		
N	28	28	27	27	28	28
\bar{R}^2	.43	.39	.73	.72	.71	.67
SEE	4.56	4.65	2.90	2.93	3.88	4.26
D-W	1.60	1.46	1.59	1.70	1.42	1.26

standard errors are in parentheses

SEE: standard error of estimate

APPENDIX D

All the stock insurers in the market reaction analysis are listed on the New York Stock Exchange. The classification of the companies was based on Best's Review's listing of property casualty stocks and Business Insurance's "BI Industry Stock Report."

Property-Casualty Companies:

AmBase
AIG (American International Group)
Chubb
Continental Corp.
Fireman's Fund
GEICO (Government Employees Insurance Company)
Kemper
Orion Capital Corp.
USF&G (United States Fidelity and Guaranty)

Multiline Insurers:

Aetna
CIGNA
CNA Financial
Lincoln National
Transamerica
Travelers

Brokers:

Alexander and Alexander
Coroon and Black
Marsh and McLennan

Stock data were collected from issues of the Wall Street Journal and New York Times. Number of outstanding shares was reported in Moody's Financial vol. 2.

REFERENCES

- A.M. Best Company Best's Aggregates and Averages Property/Casualty edition
- , Best's Trend Reports.
-----, Best's Insurance Reports.
-----, Best's Convention Statements.
- Best's Review P-C Edition, "Insurance Stock Trends," March 1989, p.22.
- Berger, Lawrence A. "A Model of the Underwriting Cycle in the Property-Liability Insurance Industry," Journal of Risk and Insurance June 1988 p. 298-306.
- Bloom, Thomas S. "Cycles and Solutions," Best's Review (p-c edition) October 1987 p. 20-24.
- Clarke, Richard N., Frederick Warren-Boulton, David D. Smith and Marilyn J. Simon, "Sources of the Crisis in Liability Insurance: An Economic Analysis," Yale Journal on Regulation, Summer 1988.
- Cummins, J. David and J. François Outreville "An International Analysis of Underwriting Cycles in Property-Liability Insurance," Journal of Risk and Insurance 1986 pp. 246-262.
- Cummins, J. David and Scott Harrington, Fair Rate of Return in Property-Liability Insurance, Boston, 1987.
- Doherty, Neil and Han B. Kang "Interest Rates and Insurance Cycles," Journal of Banking and Finance vol 12, 1988, pp 199-214.
- Fairley, William B. "Investment Income and Profit Margins in Property-Liability Insurance: Theory and Empirical Results," Bell Journal of Economics, spring 1979 p. 192-210.
- Fields, Joseph A. and Emilio C. Venezian " Interest Rates and Profit Cycles: A Disaggregated Approach," Journal of Risk and Insurance June 1989 p. 312-319.
- Garcia, Beatrice E., "Several Big Insurers are Raising Rates....," Wall Street Journal, Nov. 6, 1989, p. A7A.
- Greenwald, Judy. "'Bad news' drives up insurer stock prices," Business Insurance Oct, 23, 1989 p. 1, 4.

- Gottheimer, George M. Jr. "Crisis in Confidence," Crisis Avoidance: Insurance Responsibilities CPCU 1988.
- Harrington, Scott E. "Prices and Profits in the Liability Insurance Market," in Robert Litan and Clifford Winston, eds. Liability: Perspectives and Policy. Brookings. 1988.
- Harrington, Scott E. and Robert Litan "Causes of the Liability Insurance Crisis" Science February 1988 p. 737-741.
- Hill, Raymond D. "Profit Regulation in Property Liability Insurance," Bell Journal of Economics, spring 1979 p. 172-191.
- Insurance Information Institute, Insurance Facts (insurance fact book) various years.
- "Insurance Stock Trends," Best's Review P/C edition, March 1989 p. 18-22.
- Insurance Services Office Insurer Profitability : The Facts February 1986.
- The Coming Capacity Shortage February 1985.
- Insurer Failures: P/C Insurer Insolvencies and State Guaranty Funds, GAO report, General Accounting Office , July 1987.
- Kleinfield, N. R., "A Double Whammy for the Insurers," New York Times October 22, 1989, section 3 pp. 1, 12.
- Masterson, Norton, E. "Property/Casualty Insurance Inflation Indexes," Best's Review, P-C edition July 1981, p. 40..52.
- McGee, Robert T. "The Cycle in Property-Casualty Insurance," Federal Reserve Bank of New York Quarterly Review Autumn 1986.
- McIntyre, Kathryn J., "Renewals to Remain Competitive," Business Insurance, Oct. 9, 1989 p. 1+.
- Myers, Stuart C. and R A. Cohn, " Insurance Rate of Return Regulation and the Capital Asset Pricing Model,"
- Myers, S. and N. Majluf, "Corporate Financing and Investment Decisions when Firms Have Information That Investors Do Not Have," Journal of Financial Economics, vol 52, 1984.
- Nash, Nathaniel C., "As more Insurance Concerns Fail, Fears Mount of an Industry Crisis," New York Times April 5, 1989 p. A1 and D5.
- National Association of Insurance Commissioners Using the Insurance

Regulatory Information System Property and Liability Edition, 1985.

Priest, George L. "The Current Insurance Crisis and Modern Tort Law," Yale Law Journal, vol 96 June 1987 p. 1521-90.

Stewart, Barbara D. "Profit Cycles in Property-Liability Insurance," John D. Long and Everett D. Randall (eds) Issues in Insurance vol 2 1984 American Institute for Property and Liability Underwriters, Inc. USA.

Strain, Robert, W. (ed.) Property-Liability Accounting, 1986.

"The Cost of the Earthquake," The Economist Nov. 4, 1989 p. 102.

Tort Policy Working Group, Report on the Causes, Extent and Policy Implications of the Current Crisis in Insurance Availability and Affordability, Department of Justice, Inter-Agency report, Feb. 1986.

-----, An Update on the Liability Crisis, Department of Justice, March 1987.

Troxel, Terrie and Cormick Breslin, Property-Liability Insurance Accounting and Finance, 1983.

Venezian, Emilio "Ratemaking Methods and Profit cycles in Property and Liability Insurance," Journal of Risk and Insurance, 1985 p. 477-500.

Weiss, Mary "A Multivariate Analysis of Loss Reserving Estimates and Property-Liability Insurers," Journal of Risk and Finance vol 52, June 1985 p. 199-221.

Winter, Ralph A., "The Liability Crisis and the Dynamics of Competitive Insurance Markets," Yale Journal on Regulation vol 5, 1988, p. 455-500.

-----, "The Dynamics of Competitive Insurance Markets," mimeo, University of Toronto, July 1989.

Wooldridge, Jeffrey. "Regression-Based Inference in Linear Time-Series Models With Incomplete Dynamics," MIT mimeo, August 1989.