DECISION MAKING IN THE INSURANCE INDUSTRY:
A DYNAMIC SIMULATION MODEL AND EXPERIMENTAL RESULTS

by

ALEXANDER ASHER MOISSIS

S.B. Electrical Engineering, MIT
(1984)
S.M. Electrical Engineering & Computer Science, MIT
(1987)
Electrical Engineer, MIT
(1987)

Submitted to the Sloan School of Management
in Partial Fulfillment of
the Requirements of the Degree of
Master of Science in Management

at the
Massachusetts Institute of Technology
January 1989

© Alexander A. Moissis 1989. All rights reserved
The author hereby grants to MIT permission to reproduce and to
distribute copies of this thesis document in whole or in part.

Signature of Author________________________

Sloan School of Management
December 23, 1988

Certified by________________________

John D. Sterman
Associate Professor
Thesis Supervisor

Accepted by________________________

Jeffrey A. Barks
Associate Dean, Master's and Bachelor's Program

ARCHIVES

MAY 10 1989
DECISION MAKING IN THE INSURANCE INDUSTRY: A SIMULATION MODEL AND EXPERIMENTAL RESULTS

by

ALEXANDER ASHER MOISSIS

Submitted to the Sloan School of Management on January 20, 1989 in partial fulfillment of the requirements for the Degree of Master of Science in Management

ABSTRACT

This thesis presents a study of the decision making processes of insurance claim managers. A computerized simulation game of an insurance company claims department (the Claims Game, developed by the System Dynamics Group at MIT) was given to claims managers acting as subjects for this study. Their decisions and performance in the game are analyzed with the following goals in mind:

1. The identification of managerial misperceptions of the feedbacks and time delays associated with the claims adjustment process.

2. The formulation of simple decision rules which capture the essential attributes of the managers' performance.

3. The comparison of the managers' performance in the game to a benchmark test.

4. The identification of areas where the Claims Game can be improved as a learning tool.

The analysis of the decision making process of individual claims managers may provide information linking the behavior of these actors to the aggregate dynamic behavior of the insurance industry. Furthermore, the results of this study may help claims managers improve their mental models of industry dynamics, and thus enhance their learning process and their performance.

The benchmark test included in this thesis suggests that with the use of simple, consistent decision rules, claims managers could improve their company's profitability by as much as 30%. The magnitude of the potential improvement justifies recent efforts to enhance managerial understanding of claims adjustment dynamics.

Thesis Supervisor: Dr. John D. Sterman
# CONTENTS

Acknowledgement  
Table of Illustrations  
INTRODUCTION  

## CHAPTER 1: AN INTRODUCTION TO BEHAVIORAL DECISION THEORY  

## CHAPTER 2: INSURANCE INDUSTRY BACKGROUND  
- Introduction  
- Functions of an Insurance Company  
- Major Competitors  
- Entry Barriers and Economies of Scale  
- Rate Setting  
- Claims Adjustment  
- Performance Measures  
- Chronological Survey of the Insurance Industry  
- Hanover Insurance Company  

## CHAPTER 3: A DYNAMIC MODEL OF CLAIMS MANAGEMENT AT HANOVER INSURANCE  
- Claims Workflow Subsystem  
- Adjuster Capacity Subsystem  
- Work Intensity Subsystem  
- Quality Subsystem  
- The Complete Model  

## CHAPTER 4: THE CLAIMS GAME  

## CHAPTER 5: THE EXPERIMENTAL SESSIONS  

## CHAPTER 6: DATA ANALYSIS  
- Exogenous Input-Features  
- Initial Conditions  
- Strategies Suggested by Players  
- Two Case Studies  
- Distribution of Final Scores  

## CHAPTER 7: A BENCHMARK FOR THE CLAIMS GAME  

## CHAPTER 8: MODEL OF THE SUBJECTS' DECISIONS  
- Desired Settlement Size  
- Desired Productivity  
- Hiring  

- 3 -
LIST OF TABLES AND ILLUSTRATIONS

Tables

<table>
<thead>
<tr>
<th></th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30 Leading Prop./Liab. Companies and Groups</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>List of reports or graphs available to the players of the game</td>
<td>57</td>
</tr>
<tr>
<td>3</td>
<td>Age Statistics of Insurance Managers</td>
<td>59</td>
</tr>
<tr>
<td>4</td>
<td>Manager Industry/Company Experience Data</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>Feature Descriptive Statistics</td>
<td>67</td>
</tr>
<tr>
<td>6</td>
<td>Initial conditions for Claims Game</td>
<td>68</td>
</tr>
<tr>
<td>7</td>
<td>Retrospective Data on Managers Decisions</td>
<td>72</td>
</tr>
<tr>
<td>8</td>
<td>Table of Cumulative Scores</td>
<td>108</td>
</tr>
<tr>
<td>9</td>
<td>Regressions for Desired Settlement Size Decision</td>
<td>115</td>
</tr>
<tr>
<td>10</td>
<td>Desired Productivity Decision: Descriptive statistics</td>
<td>117</td>
</tr>
<tr>
<td>11</td>
<td>Regressions for Hiring (2 variable model)</td>
<td>119</td>
</tr>
<tr>
<td>12</td>
<td>Regressions for Hiring (3 variable model)</td>
<td>121</td>
</tr>
</tbody>
</table>

Figures

<table>
<thead>
<tr>
<th></th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Property/Liability Loss Expense ratios (1976-1986)</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>Property/Liability Industry Performance (1967-1987)</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>Claims Workflow Subsystem Diagrams</td>
<td>31</td>
</tr>
<tr>
<td>4</td>
<td>Relation of average age of claims pending to Time Effectiveness</td>
<td>34</td>
</tr>
<tr>
<td>5</td>
<td>Adjuster Subsystem Diagrams</td>
<td>37</td>
</tr>
<tr>
<td>6</td>
<td>Work Intensity Subsystem Diagrams</td>
<td>42</td>
</tr>
<tr>
<td>7</td>
<td>Quality Subsystem Diagrams</td>
<td>45</td>
</tr>
<tr>
<td>8A</td>
<td>Effect of Time Pressure on Quality Standard</td>
<td>47</td>
</tr>
<tr>
<td>8B</td>
<td>Effect of Quality Standard on Settlement Size</td>
<td>48</td>
</tr>
<tr>
<td>9</td>
<td>Causal Loop Diagram of Insurance Claims Model</td>
<td>51</td>
</tr>
<tr>
<td>10</td>
<td>Claims Game Board</td>
<td>54</td>
</tr>
<tr>
<td>11</td>
<td>Step Feature Input</td>
<td>62</td>
</tr>
<tr>
<td>12</td>
<td>Random Feature Input</td>
<td>63</td>
</tr>
<tr>
<td>13</td>
<td>Sample Insurance Manager Comment Sheets</td>
<td>69</td>
</tr>
<tr>
<td>14-27</td>
<td>Team 10 Data</td>
<td>74</td>
</tr>
<tr>
<td>28-41</td>
<td>Team 5 Data</td>
<td>93</td>
</tr>
</tbody>
</table>

Appendix

- A1-A10: STELLA Model Tables  137
- A11-A21: Benchmark Model Simulation Runs  148
INTRODUCTION:

In this thesis, I study the decision making process of Insurance Claims Managers. A computerized simulation game of the claims requirements of an insurance company (the Claims Game, developed by the System Dynamics Group at MIT) was given to Claims Managers acting as subjects for this study. Their decisions and performance in the game are analyzed with the following goals in mind:

1. The formulation of decision rules which capture the essential attributes of the managers' performance.

2. The comparison of the managers' performance to a benchmark for the game.

3. The identification of areas where the Claims Game could be improved as a learning tool.

The data analysis is followed by recommendations for the improvement of the design of the game and for the improvement of the managers' performance.

Data for this thesis were collected at the three-day Hanover Insurance Claims Learning Laboratory in June and July of 1988. The Claims Learning Laboratory was introduced at Hanover in 1988 as a part of the company's 'Systems Thinking Program.' The program uses System Dynamics models along with the Claims Game to enhance insurance managers' learning and to improve their understanding of insurance business dynamics.

The program was introduced by top management at Hanover in the early Eighties. Contrary to most other insurance managers, Hanover managers believed that insurance industry crises were to some extent self-inflicted. The insurance industry has traditionally blamed high litigation costs, outdated legal statures as well as changing social values for its crises. [11] The people at Hanover suspected that Claims department practices were, to a degree, responsible for the industry's escalating settlement costs.

The Claims Game was designed as a learning tool for Claims managers. The dynamic model of the claims department, which is incorporated in the game, was developed by
members of the Hanover Insurance management, under the guidance of Peter Senge and Nathan Forrester of MIT. The model was tested and refined at MIT by members of the System Dynamics Group.

The game can assist claims managers in their training much like a flight simulator assists airline pilots. Subjects in the game are provided with a simulation environment of an insurance claims department. They are asked to make hiring and production decisions and to set goals for average settlement size. In a real firm, it can take months, maybe years, before the consequences of such decisions become apparent. Managers are thus often encouraged to focus on short term performance rather than long term profitability. The simulation game allows insurance managers to see the potential long term consequences of their decisions in a few minutes. The managers are thus able to obtain a better understanding of the time delays and feedbacks associated with their decisions, and to observe first-hand how their effort to look good in the short run can hurt their company in the longer run. Clearly, the model used for the Claims Game is not a perfect representation of the Claims department. Yet, by simply incorporating the important features of the department's operations, it can provide useful guidelines to managers.

The data presented in this thesis provide yet another illustration of the fact that managerial decisions can be explained by a model of bounded rationality. Managers do not use all the information which is available to them for their decisions. Furthermore, they tend to adopt simplifying strategies using rules of thumb and heuristics. They tend to look for satisfactory choices rather than optimal ones.

I begin this thesis with an introduction to Behavioral Decision Theory and to the Bounded Rationality Model (Chapter 1). I then provide some background information on the insurance industry, on insurance companies and their operations, and on Hanover Insurance (Chapter 2).

Following this introduction, I present the dynamic model of the Claims department at Hanover which was incorporated in the Claims Game (Chapter 3). I then describe the game itself (Chapter 4), and the experimental sessions where the data for the thesis were collected.
(Chapter 5).

The thesis continues with a presentation of the experimental data and with a description of some of the strategies that players pursued (Chapter 6). To evaluate the players' final scores I then establish a benchmark for the game using a simple decision rule (Chapter 7). Finally, the data from the game are used to formulate models which describe the subjects' decisions (Chapter 8).

The thesis is wrapped up with concluding remarks on the data analysis and with recommendations for the Claims Learning Laboratory.
CHAPTER 1
AN INTRODUCTION TO BEHAVIORAL DECISION THEORY

The concept of 'comprehensive rationality,' found in most theories of individual and organizational choice, assumes that when making a decision, 'individuals and organizations choose the best alternative, taking account of consequences, their probabilities, and utilities.' Such a choice by the decision maker, requires:

1) The generation of all possible alternatives.

2) The assessment of the probabilities of all consequences of each.

3) The evaluation of each set of consequences for all relevant goals. [18]

These underlying assumptions of neoclassical theory have often come under attack. An extended bibliography exists of case studies demonstrating human behavior contrary to these assumptions.[1-5],[19] Most widely cited is the work of Kahneman and Tversky. Through experiments they successfully demonstrate that 'people rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations.' They also claim that 'deviations of the actual behavior from the normative model are too widespread to be ignored, too systematic to be dismissed as random error, and too fundamental to be accommodated by relaxing the normative system.' [1]

As Nobel laureate Herbert Simon puts it, 'there can no longer be any doubt that the micro assumptions of the theory—the assumptions of the perfect rationality—are contrary to fact. It is not a question of approximation, they do not even remotely describe the processes that human beings use for making decisions in complex situations.' [2]

Proponents of the neoclassical model assume that, 'if the stakes are large enough people will get it right,' or, 'in the real world people will learn to get it right,' or even, 'in the aggregate, errors will cancel.' These assertions, however, have not been supported convincingly by experimental data. Furthermore, studies have demonstrated that behavior can
be 'purposeful, regular, and yet systematically different from the axioms of economic theory.' [3]

A model of 'bounded rationality' has been promoted as a better description of human decision making behavior. According to H. Simon, 'it may well be that classical theory can be patched up sufficiently to handle a wide range of situations where uncertainty and outguessing phenomena do not play a central role—that is, to handle the behavior of economies that are relatively stable and not too distant from a competitive equilibrium. However, a strong positive case for replacing the classical theory by a model of bounded rationality begins to emerge when we examine situations involving decision making under uncertainty and imperfect competition. These situations classical theory was never designed to handle, and has never handled satisfactorily.' [2] The body of theory and data describing the 'cognitive limitations in the perception and processing of information' of decision makers, as well as the 'organizational strategies people devise to deal with these limitations,' is known as Behavioral Decision Theory (BDT). [4]

BDT researchers have demonstrated that as the tasks and goals of an organization become more complex, decision makers adopt simplifying strategies. One such strategy is to look for satisfactory choices instead of optimal ones. Thus, decision problems are simplified using rules of thumb or 'heuristics'. A second strategy is to 'replace abstract, global goals with tangible subgoals, whose achievement can be obtained and measured.' There is no unique determination of these subgoals; their formulation depends on the knowledge and experience of the decision maker. A third strategy is to 'divide up the decision making task among many specialists, coordinating their work by means of a structure of communications and authority relations.' [2]

Behavioral Decision Theorists thus distinguish between the real world and the decision makers' perception, or 'mental model', of it. Their theory includes not only the 'reasoning processes but also the processes that generate the actor's subjective representation of the decision problem, his or her frame.' [5]
One approach for obtaining information from case studies is computer simulation. Empirical evidence on a particular decision making process is encoded in a computer program. The program is then used to test its adequacy in explaining existing data and in order to discover 'the key features that account, qualitatively, for the interesting and important characteristics of its behavior.' [2] Furthermore, the assumption of bounded rationality, 'enables one to decompose a complex simulation model into pieces to reveal intuitively clear, plausible behavior.' [6] The model also provides information on the effects of feedbacks, time lags and nonlinearity. It is such a 'direct experiment' approach that we use in this study in order to obtain empirical evidence of the decision making process of Insurance company managers. The analysis consists of four basic steps: [4]

1. Identifying the decision to be tested.

2. Converting the original simulation model into an interactive 'game' in which subjects play the roles of decision makers.

3. Running the experiment.

4. Analyzing experimental results and comparing them to the original model assumptions.

As H. Simon pointed out, 'even if there were no present evidence of such relevance (of the behavioral theory of the firm to the macroeconomy), human behavior in business firms constitutes a highly interesting body of empirical phenomena that calls out for explanation as do all bodies of phenomena. And if we may extrapolate from the history of other sciences, there is every reason to expect that as explanations emerge, relevance for important areas of practical application will not be long delayed.'[2]
CHAPTER 2
INSURANCE INDUSTRY BACKGROUND

2.1 INTRODUCTION

In this section we examine the nature and functions of the multiple line property and liability insurance industry, as well as its recent history. Such an overview will help the reader understand the complexity of the tasks that insurance managers face and the competitive environment in which they operate.

Although the concept of insurance and insurance policies can be traced back to antiquity, one need only go back to the year 1949 in order to obtain an understanding of the development of the U.S. property-liability insurance industry to its current form. In that year, a bill was passed by the New York State Legislature which effectively enabled a single insurer to operate in the entire field of insurance outside of life insurance. Before 1949, the following eighteen kinds of insurance (apart from Life Insurance) had been authorized by the New York Insurance Law:

Accident and health
Fire
Miscellaneous property
Water damage
Burglary and theft
Glass
Boiler and machinery
Elevator
Animal
Collision
Personal (bodily) injury liability
Property damage liability

Workmen's compensation and employer's liability

Fidelity and surety

Credit

Motor vehicle and aircraft

Marine

Marine protection and indemnity.

A single insurer was not permitted to operate in all eighteen fields; insurers were allowed to operate only in specified areas such as casualty and surety insurance, or fire and marine insurance. With the new law, which was soon emulated by the other states, the compartmentalization of insurers, which had existed before 1949 under the so-called 'American System', and which had been unique in the United States, was abolished. [7] Multiple insurance companies appeared which could now operate in all eighteen fields.

The introduction of multiple line insurance expanded the opportunities of insurers or producers of insurance, as they often refer to themselves. The radical change in the scope of the companies brought about drastic changes in their operations, their marketing methods, their servicing of clients as well as their financing policies. Furthermore, the state supervisory practices also had to be modified. The overhaul of the U.S. insurance industry which followed the 1949 legislation brought the industry to its current form. Two major types of insurance companies exist today, life insurance firms and multiple line property-liability (or property-casualty) firms.

2.2 FUNCTIONS OF AN INSURANCE COMPANY

The products sold by the insurance industry are purchased by almost all economic agents in the United States. Insurance provides a method for spreading risk of property loss and liability among several groups. For a small premium the insured party trades the risk and uncertainty of
an accidental loss to the certainty of insurance.

Property-liability insurance provides protection against losses resulting from liability claims and damage to property such as buildings, vehicles, ships, or materials. Property insurance may involve automobile physical damage, fire, inland and ocean marine, and associated risks. Under the heading of liability insurance we find automobile and other liability lines, as well as workmen's compensation.

In describing the functions of an insurance company we may distinguish four general areas of operation; these areas are all closely linked to the primary function of an insurance company, which is underwriting: [7]

1. Marketing and distribution. This area involves the development of products and premium systems which will lead to profitable business for the company. The premium is the price of the insurance for an individual risk.

2. Service. Prompt writing of policies and more important, handling of claims, are some aspects of insurance company service. Service is one important form of non-price competition in the insurance industry, since it can lead to a differentiation in the perception of the consumer among seemingly identical insurance products.

3. Investments. Earnings from investments of premium reserves constitute a significant source of the companies' annual income. Moreover, the investment performance of a firm may influence its underwriting operations. Thus, investments constitute an important part of the multi-line insurer's operations.

4. General administrative duties. These include functions such as accounting and finance, human resource management, etc.

The underwriting functions of multiple line insurance companies are often segregated between personal and commercial lines. Automobile insurance accounts for almost half of the insurance industry business.
2.3 MAJOR COMPETITORS

The multiple line insurance industry includes a large number of firms. In 1986 the industry numbered about 2,000 firms which varied enormously in size from $16 billion in total annual premiums to less than $1 million. Total industry net premium volume exceeded $130 billion.

The industry concentration ratio is low, with the top 20 companies controlling only about 50% of the market. However, the concentration ratio has been increasing steadily. A list of the main competitors in the property/liability insurance industry appears in Table 1.

Insurance companies can be found as stock companies, mutual companies, or reciprocal exchanges. Stock companies are owned by individuals who have invested in them. Mutual companies are owned by their policy holders. Reciprocal exchanges are cooperative organizations which allow their members to share specified risks. In 1987 stock companies accounted for about three quarters of all property liability premiums and industry assets.

Sales have traditionally been carried out by non-exclusive agents, under the so-called American Agency System. In recent years, many companies have developed their own direct distribution networks. Due in part to their lower costs, direct writer market share has been slowly increasing over the past quarter century. The increase in the market share has been slowed down however by significant switching costs. Since customers are nominally represented by agents, switching of an insurance company from agents to direct writing implies the abandoning of existing customers and is therefore costly. [8,9,10]

2.4 ENTRY BARRIERS AND ECONOMIES OF SCALE

There seem to be no significant barriers to entry in the multiple line insurance industry. Licensing requirements include a provision of a minimum amount of paid-in capital and surplus, which varies by state between $15,000 and about $500,000 per line. The requirements are relatively low and do not constitute an important barrier to entry.
# TABLE 1

## 30 LEADING PROPERTY-LIABILITY COMPANIES & GROUPS

Ranked by Net Premiums as of 12/31/86  
(Thousands of dollars)

<table>
<thead>
<tr>
<th>Rank</th>
<th>Company</th>
<th>Net Premiums</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>State Farm Mutual Auto</td>
<td>16,248,188</td>
</tr>
<tr>
<td>2</td>
<td>Allstate</td>
<td>9,185,053</td>
</tr>
<tr>
<td>3</td>
<td>Aetna Life &amp; Casualty</td>
<td>6,933,732</td>
</tr>
<tr>
<td>4</td>
<td>Nationwide Mutual</td>
<td>5,799,900</td>
</tr>
<tr>
<td>5</td>
<td>Liberty Mutual</td>
<td>5,473,629</td>
</tr>
<tr>
<td>6</td>
<td>Travelers</td>
<td>5,050,163</td>
</tr>
<tr>
<td>7</td>
<td>Farmers Insurance</td>
<td>4,981,194</td>
</tr>
<tr>
<td>8</td>
<td>Hartford Fire</td>
<td>4,921,780</td>
</tr>
<tr>
<td>9</td>
<td>American International</td>
<td>4,906,757</td>
</tr>
<tr>
<td>10</td>
<td>CIGNA</td>
<td>4,531,885</td>
</tr>
<tr>
<td>11</td>
<td>CNA</td>
<td>3,986,183</td>
</tr>
<tr>
<td>12</td>
<td>Continental Insurance</td>
<td>3,944,040</td>
</tr>
<tr>
<td>13</td>
<td>USF&amp;G</td>
<td>3,562,035</td>
</tr>
<tr>
<td>14</td>
<td>Fireman's Fund</td>
<td>3,351,923</td>
</tr>
<tr>
<td>15</td>
<td>Crum &amp; Forster</td>
<td>3,022,407</td>
</tr>
<tr>
<td>16</td>
<td>Chubb</td>
<td>2,840,379</td>
</tr>
<tr>
<td>17</td>
<td>Kemper</td>
<td>2,816,366</td>
</tr>
<tr>
<td>18</td>
<td>St. Paul</td>
<td>2,554,446</td>
</tr>
<tr>
<td>19</td>
<td>General Re</td>
<td>2,478,577</td>
</tr>
<tr>
<td>20</td>
<td>Royal Insurance</td>
<td>2,203,175</td>
</tr>
<tr>
<td>21</td>
<td>Lincoln National</td>
<td>2,026,886</td>
</tr>
<tr>
<td>22</td>
<td>Prudential of America</td>
<td>1,937,176</td>
</tr>
<tr>
<td>23</td>
<td>USAA</td>
<td>1,926,928</td>
</tr>
<tr>
<td>24</td>
<td>Home Insurance</td>
<td>1,877,260</td>
</tr>
<tr>
<td>25</td>
<td>American Financial</td>
<td>1,686,727</td>
</tr>
<tr>
<td>26</td>
<td>American General</td>
<td>1,523,462</td>
</tr>
<tr>
<td>27</td>
<td>Reliance</td>
<td>1,514,675</td>
</tr>
<tr>
<td>28</td>
<td>Employers Re</td>
<td>1,371,446</td>
</tr>
<tr>
<td>29</td>
<td>SAFECO</td>
<td>1,370,782</td>
</tr>
<tr>
<td>30</td>
<td>GEICO</td>
<td>1,368,222</td>
</tr>
</tbody>
</table>

Note: The America Group, which includes Hanover Insurance, ranks 36th with $1,158,994,000 in net premiums.

Due to the complexity of the underwriting operations and of the claim adjustment process, a supply of trained and experienced personnel is essential to an insurance company. While there may exist short run trained personnel shortages, there is no reason to suspect shortages over the long run. Consequently, there appear to be no significant barriers to entry due to trained personnel shortages.

With the products (policies) offered by insurance firms very similar from company to company, the price offered by each firm for its policies becomes an important factor. Due to its visibility to the customer, price can play an important role in the policyholder's choice of insurer. Thus relative costs of insurers become important. Substantial cost advantages of large firms over smaller ones would constitute barriers to entry in the industry. Paul Joskow of MIT as well as others have argued that there appear to be no such economies of scale. [8,9] Large companies do not appear to provide insurance at a lower cost.

Due to the nature of the product, company identification becomes very important, particularly for direct underwriters. Mass advertising is common in the industry as a means for establishing such identification. The high costs of national advertising campaigns do constitute a barrier to entry. However, the continuous and substantial entry in the property and liability insurance industry over the past years supports the argument that entry in the industry is relatively easy.

The rates, financial organization as well as quality of service of the property liability insurance industry are supervised by state insurance commissions. The regulatory authorities have in general adopted policies aimed at preventing 'destructive competition' among insurance companies. Their primary concern had been (at least until the nineteen Seventies) to guard against rates that were too low. [8]

2.5 RATE SETTING

Insurance policies are usually classified in various risk groups; individual rates are set for each classification. The establishment of classifications is an ongoing and challenging part of
the operations of an insurance company. The determination of rates for various insurance policies is particularly difficult, as it involves in effect a prediction of the likelihood of occurrence of an accident for which the policy is established. While personal judgment of experienced underwriters may be used in cases of uncommon insurance policies, rates for most policy lines are set using past performance statistics. The numbers that are most commonly used in such rate setting calculations are past loss data, forecasts of future losses (which would also take into account trends in past loss data), as well expense data. Furthermore, provisions are allowed for underwriting profit, and catastrophes. Expenses are usually expressed as a percentage of premiums (loss adjustment expenses a percentage of earned premiums, while commissions and premium taxes a percentage of written premiums). [7]

2.6 CLAIMS ADJUSTMENT

Once a loss is reported to an insurance company, it is forwarded to the claims department. This department plays a very important role in the operations of the firm. Claim adjusters are the people who come into contact with the insured and thus play a critical role in establishing the company's reputation. Furthermore, their prompt and efficient investigation of claims can help significantly the operating performance of the firm. It can take several years for a company to develop skilled claims adjustment specialists. Their operation requires discretion as well as informed judgment. In the words of an insurance manager, the claim adjuster must have a broad understanding of human nature, must be firm but at the same time compassionate, and must also have a good understanding of the law. [7] Peter Senge describes a claims adjuster as, 'one who is capable of conducting thorough professional investigations, has excellent communication skills, keeps neat and complete file records, and is able to educate claimants regarding the fair value of their claims, while at the same time being able to detect those with the slightest fraudulent inclinations.' [10]
Claims are subdivided into first party claims (where the claim is made by the insured to the insurer), third party claims (where the claim is made against the insured by a third party), and workmen's compensation claims. It is the operations of the claims department of insurance companies that we examine in this study, particularly the decision making process of insurance claim managers. More information on the claims department is thus presented later in the thesis.

2.7 PERFORMANCE MEASURES

The operating performance of an insurance company is usually described in terms of the so-called underwriting ratios. These ratios are:

**Loss ratio**: The ratio of losses and loss expenses to premiums earned.

**Expense ratio**: The ratio of underwriting expenses to premiums written.

**Combined ratio**: The combination of the loss ratio and the expense ratio.

**Dividend ratio**: The ratio of dividends to premiums earned.

**Combined ratio including dividends**: The combination of the loss ratio, the expense ratio, and the dividend ratio.

These ratios provide an indication of the relation between the operating cash inflows of the firm (premiums) and the cash outflows due to losses, general expenses, and dividends. Thus a combined ratio which is less than 100% would suggest that operating cash inflows exceeded outflows. Typical combined ratios for the industry in the year 1983 ranged from 96% to 114%. For 1987 the average combined ratio for the industry was estimated at 104.5%, indicating a net loss on operations. As illustrated in Figure 1, loss ratios have been rising steadily from about 67% in 1978 to 86% in 1986. Figure 2 presents combined ratios for the Property Liability Insurance Industry for the period from 1965 to 1987. [12]

The above ratios, widely used by insurance companies to describe their performance, but also by regulators in order to approve insurance rates, have come under heavy criticism.
Figure 1: Insurance industry performance: Loss ratios for 14 Casualty Companies (1976-1986)
PROPERTY-LIABILITY INSURANCE:  
INDUSTRY PERFORMANCE (1965-87)

Figure 2: Source: Hanover Insurance, Annual reports 1983-1987
Analysts argue that these ratios provide a poor description of the performance and profitability of insurance firms because they omit investment income. They argue that all sources of income, underwriting profits, interest and dividends from investments, as well as capital gains should be used in order to evaluate the operating performance of insurance firms. [8,9]

2.8 CHRONOLOGICAL SURVEY OF THE INSURANCE INDUSTRY

A first look at the performance of the insurance industry since 1949 indicates business cycles which are particularly pronounced (with an apparent cycle of about 6 years). [13,14] This cyclicality is attributed by some to the so-called 'insurance cycle'. Under this scenario, during bullish times insurance companies seek cash in order to increase their investments. To attract investment capital they compete by cutting their rates sharply. This rate decrease is based on the notion that investment income will more than make up for any underwriting losses. When the returns from investments decline the trends are reversed: Insurers are forced to put more weight on their underwriting performance and are thus led to rate increases. Furthermore, shortages in the availability of insurance are likely to appear during this second half of the insurance cycle, as insurers become more and more cautious and selective in their choice of business.

The Fifties and Sixties

Following the legislative changes of 1949 the industry experienced profitable underwriting results through 1955. In the years 1956-1958 the insurance underwriting operations were unprofitable, but recovered in 1959. Four profitable years (1959-1962) were followed by three years of heavy losses (1963-1965). The 1965 loss, which was due in part to the catastrophe brought about by Hurricane Betsy, was, at the time, the heaviest on record. With increased premium rates and fewer catastrophes, the industry recovered in 1966 and 1967. The recovery, however, was not sustained. The years 1968 and 1969 were once again unprofitable for the industry. Hurricane Camille of 1969 was one of the reasons for the bad year.
Increased losses due to an escalating inflation also became a concern towards the end of the decade. [15]

The Seventies

While severe natural catastrophes were the main cause of the industry's underwriting losses in the sixties, it is inflation that brought about the 'insurance crisis' of the mid seventies. The decade began with losses expanding faster than premiums. With investment income in a decline in the late sixties and early seventies, companies were less volume conscious and more selective. By 1973, however, investment income had improved and, according to the Standard and Poor's Industry survey, was the dominant source of earnings. Seeking to increase their investment income, insurers increased their premium volume. According to the S&P report, 'temptation to sacrifice high underwriting profits through rate cutting in an attempt to obtain additional premium income to boost investment income was apparently irresistible.' [15]

With the abrupt price increases of 1973 and the stock market collapse of 1974 the trend reversed: Accelerated inflation and (according to insurance companies) generous court settlements (social inflation) reduced reserves and consequently investment income. Regulatory lags prevented the insurers from boosting their rates soon enough to avoid losses. Thus, the loss ratio reached an all time high as premiums proved inadequate in the face of unprecedented inflation combined with a high rate of catastrophes. As the S&P report states, 'for too long, property-liability insurers had depended on stock market appreciation as a means for expanding their capital and surplus, providing them with the means to increase their premium volume.' [15]

The crisis of 1974 and 1975 proved to be devastating for the industry. Insurance companies experienced unprecedented losses, despite the absence of extraordinary catastrophes. By the end of 1974 and in 1975, substantial rate increases were approved, as the insurance crisis appeared to lessen the resistance of insurance commissioners to insurance rate boosts. The commissioners approved huge rate increases for medical malpractice insurance
rates, a fact which prompted extended debates between the insurance industry, physicians, insurance commissioners, and the public. As a result of these rate increases insurance industry profits skyrocketed in 1976-1977.

The Eighties

With interest rates climbing to unprecedented levels in the late Seventies and early Eighties, insurers once again began to cut prices and to insure poor risks, so as to obtain as many premium dollars as possible for their investments. Thus, the period from 1979 to 1984 corresponds to sharp drops in insurance rates and to a decline in profits due to undercutting.

As a result of the sharp cuts in insurance rates of the early Eighties, the industry was once again faced with an underwriting crisis in 1985-1986. With interest rates falling, insurers sought once again to boost their rates. As in the mid-Seventies, the 'insurance crisis' of the mid-Eighties was blamed by the insurers themselves on the so-called social inflation, the sharply higher jury awards for damages. However, the coincidence of the 'jury award crises' with downturns in the insurance industry performance suggests that the industry has sought repeatedly to use higher jury awards as an excuse for its poor performance. There seems to be little doubt that the downturns in the insurance industry performance are at least to some extent self inflicted.

As the stock market surged in the first half of 1987, there were once again indications that insurance rates were beginning to drop; the stock market plunge of the latter half of 1987 somewhat slowed this undercutting trend. However, by the third quarter of 1988 insurance rates were still declining.

This brief outline of the performance of the insurance industry over the past forty years should give the reader an idea of the cyclicality that it faces. Some analysts have argued that this cyclicality is becoming more pronounced as deregulation allows insurers to set rates without approval from state regulatory bodies, thus eliminating the so-called regulatory lags.
2.9 HANOVER INSURANCE COMPANY

The Hanover Insurance Company, now based in Worcester Massachusetts, was founded in 1852 in New York. In its first century of operations the company grew in a 'consistent though conservative manner.' [17] The radical changes that took place in the property-liability insurance industry in the Fifties, following the end of the compartmentalization of the industry, found Hanover unprepared to respond. As John Adam, president of Hanover in the Seventies, wrote in 1974, the company 'rested and slept while the other companies were becoming multiple line (property, casualty and marine) insurers.' [17] In fact, it was not until 1958 that the company moved to change its name from 'The Hanover Fire Insurance Company.' By the end of the Sixties, Hanover 'lagged behind the industry in introducing new products, in improving its current products and in taking advantage of the dawning data processing revolution to improve service to policyholders, claimants and agents.' [25] Needless to say, the company also lagged behind the industry in its financial performance.

The year 1969 was a pivotal one for the company. In that year State Mutual purchased a 50% interest in Hanover, thus providing much needed financial backing. The company thus became affiliated with the America Group Companies, a group of insurance companies 'sharing their resources in their mission to fulfill the financial needs of the consumer.' With John Adam taking over as president, the company began its recovery with a move of the company's headquarters from New York to Worcester Massachusetts. In the following decade, Adam established along with his marketing vice president and eventual successor, William O'Brien, a set of guiding ideas for the company which were aimed at solving the past inflexibility of The Hanover. The five main issues addressed were Purpose, Merit, Openness, Localness, and Vision.

The notion of Purpose addressed the weak sense of common direction that had existed in the company. Merit was designed to be an "antidote to rampant politics and bureaucracy." Openness addressed the problem of "wide-spread gamesplaying through hoarding information or operating from private agendas." Localness was aimed at expanding the decision making
authority of peripheral managers and thus enhancing their morale and their professional development. The idea of Vision was provided as an "antidote to low self image and difficulties in communicating the scale of the firm's activities." [11]

For the first time in over a decade, Hanover posted underwriting profits for the year 1971. The improvement continued and in 1972 the company's financial performance bettered the industry average for the first time in many years. The company was not able however to avoid frustrating losses during the 'insurance crisis' which affected the industry in the mid Seventies; it did however continue to outperform the industry average.

The changes in Hanover's management style were not implemented without cost: in the early Seventies management turnover at the company was high, as several managers appeared unprepared to adjust. Furthermore, a level of regional management was eliminated in order to enhance local autonomy.

Despite the costs, the structural changes were successful and by the mid-Eighties Hanover had emerged as a leader in the property and liability industry, consistently outperforming the industry average in profitability and growth. In 1987 the company had total assets amounting to about $2.3 billion, earned $1.3 billion in premiums with a net investment income of $116 million and a net income of $112 million. Losses and loss expenses amounted to $916 million (a loss ratio of 72%), while selling and administrative expenses were $385 million (an expense ratio of 28.9%). With dividends paid to policyholders amounting to 1% of premiums earned, the combined ratio for the company was 101.9%. With this combined ratio, Hanover bettered the industry average of 104.7% by 2.8 percentage points. It was the ninth year in a row that Hanover had outperformed the industry in terms of profitability. [12] As was noted in Table 1, in 1986 Hanover ranked 36th in size (in terms of premiums earned) in the Property-Liability Industry.

I should point out here, in view of the benchmark model which will be presented in Chapter 7, that a 2% reduction in Hanover's total costs would amount to about $20 million.
With earnings at around $100 million, a 2% reduction in total expenses would thus correspond to a 20% boost in the company's net income.
CHAPTER 3
A DYNAMIC MODEL OF CLAIMS MANAGEMENT AT HANOVER

As has often been stated, no model is perfect. A model provides a representation of the interactions and interrelations within a particular system. The model can be very general, providing only a 'big picture' of the system, or it can be very specific, encompassing interrelations in great detail; by its definition however, it provides a simplification of an actual system.

Despite the fact that some level of detail is always omitted, by encompassing the important aspects of the actual system, a model can provide useful insight to an analyst, a designer, a policy maker, or a student. It is in the context of improving managers' mental models of the operations of the company that Hanover began developing a Systems Thinking program in 1982. Managers at Hanover felt that practices within the insurance industry were contributing to the problems that insurers faced. In particular, they felt that the escalating industry loss expenses (which are illustrated in Figure 1A) were due in part to claims adjustment practices. By enhancing the thinking process of claims managers, it was hoped that the program could help improve the company's operating performance.

The first part of Hanover's Systems Thinking Program involved the development of a model that would encompass the major aspects of the claims operations of the company. The model presented here was developed by Peter Senge of the System Dynamics Group at MIT under the 'Program in Systems Thinking and the New Management Style' [11]. Senge, along with other members of the System Dynamics Group at MIT, assisted managers at Hanover in modeling the major elements of the claims management system. Particular attention was paid to interconnections between strategic objectives, as well as to the possible adverse effects of a particular policy choice due to feedback loops within the system. The model was developed
during 50 two-hour meetings extending from 1986 to 1988 between Peter Senge, Nathan Forrester, and claims managers at Hanover.

The meetings between Senge and Hanover, which have been described in more detail elsewhere [11], resulted in a dynamic model of claims operations at Hanover. The model was formulated using the STELLA software [20]. STELLA has simplified the modeling task by allowing the development of dynamic simulation models on the screen of a personal computer. Model developers are able to view on their computer screens graphical representations of the stocks and flows involved in the model.

The modeling process with STELLA is an iterative one: Using the software tool kit, assumptions about the system are mapped out on a structural diagram, such as the ones appearing in Figure 3B. The structural diagram appears on the screen of the computer. Once the assumptions are mapped out, designers can review the screen display of their model and modify it where necessary. When this initial mapping process is completed, STELLA allows them to simulate the model. Simulation enables the modelers to trace the results of their assumptions and to examine the dynamic behavior of their system. If the dynamic behavior observed is different from a desired pattern, the designers can go back and modify their original structural diagram. The iteration is repeated until a satisfactory model is obtained, or until the original assumptions about the structure of a system are rejected. In either case the modelers can improve their mental models of a particular dynamic system.

The behavior that Hanover managers sought to model was the trend of escalating insurance settlement costs. Hanover managers felt that the escalating costs were due to increased production pressures within the industry, rather than due to exogenous factors such as higher court settlements. With an escalating number of claims pending (and no additional staff hired by the firm), adjusters will tend to speed up their investigations. By speeding up their investigations they can then increase their production (settlement) rate and reduce the pile of claims pending on their desks. A hasty investigation, however, is likely to result in escalating settlement costs. Consequently, claims adjusters can look good in the short run by meeting
their production goals, at the expense of their company's long term profitability. It is this concept that Hanover managers sought to model.

The model of the claims operations at Hanover, developed by Senge and the Hanover managers, is presented below. To describe the model, I subdivide it into four general subsystems: Claims Workflow, Adjuster Capacity, Work Intensity, and Quality Standards. The Claims Workflow subsystem examines the link between incoming claims (features), pending claims, adequacy of staff employed (adjuster capacity adequacy), and settlement rate. The Adjuster Capacity subsystem focuses on the human resources of the firm, modeling the effects of increased production pressures on employee turnover, as well as the effects of increased hiring on production rates. The Work Intensity subsystem indicates how increased production pressures may motivate employees to work more intensively in order to reduce their workload. Finally, the Quality Standard subsystem illustrates how increased production pressures can also lead to an erosion of the performance quality of employees and thus to escalating settlement costs.

3.1 CLAIMS WORKFLOW SUBSYSTEM

Figures 3A and 3B present a causal loop and a STELLA stock and flow diagram of the model for the claims workflow of the company. Incoming claims or features accumulate in a pending pool until they are eventually settled. The STELLA diagram introduces the production measures most commonly used in the industry, namely the production ratio, the pending ratio, and the average age of the pending claims (which is closely related to another measure, the average settlement time). The production ratio relates settlements to incoming claims. The pending ratio relates claims pending to incoming claims. An approximate measure of the age of claims pending is found by dividing pending claims by the average number of incoming features; it provides an indication of the average time that has to elapse before a claim is investigated and settled (Appendix 1 presents in detail the equations incorporated in the model).
Figure 3A: Claims Workflow Subsystem Model - Causal Loop Diagram
Figure 3B: Claims Workflow Subsystem Model - STELLA Diagram
The model assumes that incoming features are determined by market forces exogenous to the operations of the company. Due to possible adjuster capacity constraints, the number of settlements per month will in general be only a fraction of the desired production level. The desired production level is the product of incoming features and desired productivity. Desired productivity is assumed to be a constant (as illustrated in Figure A1 of the Appendix). In the Claims Game to be described later, desired productivity is set by the subjects playing the game.

In the presentation of the dynamic model of the claims operations, it is important to identify closed feedback loops. By observing these loops we may be able to trace the effects of a particular policy throughout the system and to identify processes which may lead to side effects or which may thwart the intended effects of policies.

The claims workflow model provides an illustration: As features increase, claims pending go up. Consequently, the desired production level goes up. As desired production goes up, settlements also increase and claims pending drop. The loop described above (which appears on the top left portion of the causal loop diagram of Figure 3A) is a negative feedback loop, as the system responds in such a way as to counteract the initial increase in claims pending.

We may similarly identify a positive feedback loop (on the top right portion of the causal loop diagram of Figure 3A): With an increase in features and in claims pending, and consequently an increase in the desired production level, the adequacy of the existing employee workforce (capacity adequacy) of the company drops (the interconnection between desired production and capacity adequacy is presented in more detail later). A drop in capacity adequacy results in a drop in settlements (the fraction of the desired production that actually takes place decreases). The decline in settlements results in an increase in pending claims. Thus an increase in features can have the adverse effect of further increasing claims pending via the positive loop identified above.

There is one other loop in the claims workflow model that we should identify. This loop appears in the bottom half of the causal loop diagram of Figure 3A. As claims pending increase, so does the average age of these claims. As illustrated in Figure 4, the effective time
Figure 4: Relationship of average age of claims pending to the Time Effectiveness of an investigation (for an average age in excess of 1.5 months, as the average age increases, each settlement requires more and more time).
required to investigate a claim is assumed to depend on the age of the claim. For claims investigated too soon (with an average age in our model of less than forty five days, i.e. before all relevant evidence has been collected) the effective time required for an investigation drops with increasing age. On the other hand, for claims investigated late (after 45 days), the effective time required for the investigation grows with age, as some evidence may be lost etc. Thus, for the typical case of claims with an age of over forty five days, an increase in claims pending leads to an increased age of claims and consequently to an decrease in the effective time available to settle claims. As effective time available to settle the claims goes down, capacity adequacy also drops. With a drop in capacity adequacy, settlements drop (the fraction of desired claims settled drops) and consequently pending claims increase. This loop is then a positive one when the average age of claims pending exceeds 45 days. For an age less than a month and a half, the loop is in fact a stabilizing negative one.

To summarize, we have identified three important feedback loops related to the claims workflow subsystem. There is one negative loop due to the increase in desired production which results from an increase in incoming features; there is a positive loop due to the effect of a change in desired production on the capacity adequacy of the firm: With an increase in desired production, a lower fraction of the desired settlements will be achieved. Finally, we identified a third loop, which, depending on the average age of the claims, could be either stabilizing or destabilizing: For claims older than 45 days the loop is positive: As the average age of claims increases, so does the effective time required to investigate them; consequently, the capacity adequacy of the firm is reduced and fewer claims are settled. For an average age less than a month and a half, the loop provides a stabilizing rather than destabilizing force.

To illustrate the implications of the model described above, let's look at the case where the rate of incoming claims suddenly increases abruptly. We thus assume that the system is originally at a steady state with incoming features at a constant monthly rate (with an equal settlement rate and a constant level of claims pending). Suddenly, and unexpectedly, the monthly rate of incoming claims goes up and remains at this new higher level. How does the
system respond to this step increase in incoming features? If we look at the causal loop diagram of Figure 3A we see the following effects: As features instantaneously increase, with settlements slower to adjust, the stock of claims pending rises. With this rise in features and claims pending, the desired production level increases; if the company is able to meet the increased production requirements, settlements rise thereby bringing the stock of claims pending to their desired level. There are two important side effects however: With the increased production requirements the adequacy of the employee force is reduced. The inadequate adjuster workforce is unable to meet the increased production requirements and so the reduction in claims pending slows down. Furthermore, as the average age of claims pending gradually increases (at levels above 45 days), more and more time is required for each settlement. The adjuster workforce inadequacy is thus further aggravated and the reduction in the pending stock is delayed.

3.2 ADJUSTER CAPACITY SUBSYSTEM

We now focus on the human resources of the claims operations of the company. A model for adjuster stocks and flows appears in Figures 4A and 4B. The STELLA model separates the total number of adjusters employed by the company into two major groups, senior adjusters and new adjusters. Claims adjustment is a highly skilled activity. Expertise is developed slowly, after a period of apprenticeship. To account for the long training delays the model assumes that an average of about two years are required before a new adjuster acquires 'senior' status.

The complexity of the investigation required for the settlement of a claim varies. Some claims, such as the theft of a car radio, are standard, frequent and easy to process. Such claims are typically handled by the new adjusters. Others claims can be rare and more complex; such claims require an experienced adjuster for an effective investigation. Since complex claims can only be handled by experienced adjusters, the fraction of senior adjusters handling incoming claims becomes important. The model introduces the concept of 'adjuster capacity'.

-36-
Figure 5A: Adjuster Capacity Subsystem - Causal Loop Diagram
Figure 5B: Adjuster Capacity Subsystem - STELLA Diagram
The modelers assumed that new adjusters contribute about half as much as senior adjusters. Furthermore, senior adjusters must spend part of their time training new adjusters. As more new adjusters are hired, more senior adjuster time is needed for training sessions and for helping the rookies solve problems. Thus, adjuster capacity is defined as the total number of senior adjusters plus one half of the total number of new adjusters (the weight of one half accounting for the lower performance of new adjusters relative to senior adjusters) minus an adjustment for the training time required. Each new adjuster is assumed to require training equivalent to 10% of the time of a senior adjuster. Thus,

\[
\text{Adjuster Capacity} = 0.5 \times \text{New Adjusters} + \text{SrAdjusters} - 0.1 \times \text{New Adjusters}
\]

Adjuster capacity measures the number of full time equivalent senior adjusters. It accounts for the lower productivity and experience of the rookie adjusters and the time senior adjusters spend training the rookies. The stock of new adjusters is increased by hiring and is reduced either by promotions or by new adjuster turnover. In the Claims Game described in Chapter 4, the hiring decision is made by the players of the game.

Turnover out of the adjuster pool is relatively high. In our model, roughly 4% of the adjuster workforce are assumed to leave each month. The percentage of new and senior adjusters who leave the firm is assumed to depend on the time pressure the employees face. Figure A7 of Appendix 1 illustrates the assumed effect of time pressure on turnover: As time pressure increases, turnover is assumed to increase. Turnover generally drops with decreasing time pressure. However, Figure A7 indicates that turnover is assumed to increase as time pressure becomes too low. The increase corresponds to departures of highly motivated employees who seek a more challenging occupation when their jobs become boring.

There are several important feedback loops that we should identify in the adjuster subsystem diagram of Figures 5A and 5B. First, there is a destabilizing positive loop due to increased pressure. A drop in the stock of new and senior adjusters due to turnover (or
equivalently a reduction in adjuster capacity) reduces the total employee time (effective time) available for the settlement of claims (effective time being the total employee-hours spent on claims adjustment). With reduced adjuster time devoted to settlements, time pressure on the employees increases. Higher time pressure results in the departure of more employees. The turnover rate goes up and consequently the total number of adjusters falls even more. The loop appears on the right hand side of the causal loop diagram of Figure 5A.

With turnovers constant as a percentage of the adjuster workforce, an increase in the number of adjusters also leads to an increase in the number of departing employees. The increase in departures counteracts the original increase in adjusters. This negative loop appears in the center of the causal loop diagram of Figure 5A.

Even though the hiring decision is left to the players of the Claims Game (the Game is introduced in the next Chapter), it is interesting to observe the consequences of one particular hiring policy: Hiring at the turnover rate. If we assume that adjusters are only hired to replace turnovers, so that the total number of adjusters employed remains constant, two other feedback loops emerge. First, as the number of adjusters increases, turnover increases; with increased turnover, hiring is increased; consequently, the number of new claims adjusters further increases (a positive loop).

An increase in time pressure on the adjusters can also lead, however (due to the decision rule assumed), to increased hiring. We may thus identify a negative 'stabilizing' loop: As adjusters increase due to hiring, the effective time available by adjusters for the settlement of claims goes up. Consequently, time pressure on the workforce decreases, turnover drops, and the number of adjusters hired is reduced.

To understand the implications of the dynamic structure assumed for this subsystem, let's examine the effects of a step increase in the rate of incoming claims (features). We begin with the system in equilibrium: The hiring rate is initially constant and equal to the turnover rate. Adjuster capacity and time pressure are also initially constant. As incoming features jump to a new level, the production pressure on the workforce suddenly increases. The increased
pressure causes higher turnover and thus leads to a drop in the total number of adjusters (and adjuster capacity). The drop in the number of adjusters further aggravates the time pressure on the workforce; consequently turnover increases even more.

The destabilizing process described above is slowed down by two forces: First, as the total number of adjusters drops (with turnover constant as a percentage of the workforce), the number of adjusters who depart is also reduced. Furthermore, increased turnover will typically (though not necessarily) result in increased hiring. In the Claims Game described in Chapter 4, the hiring decision is left to the players of the game. As we will illustrate, however, all the players paid close attention to turnover when making their hiring decision.

3.3 WORK INTENSITY SUBSYSTEM

In this section and the next we focus on what the firm calls the 'fuzzy' standards of the claims department. Fuzzy standards refer to aspects of the operations of claims departments which are difficult to measure and to quantify. While difficult to measure, insurance managers agree that these fuzzies play an important role in the company's operations. Such standards include the work intensity of claims adjusters, as well as the quality of their work.

In this section we examine the subsystem which relates to the work intensity of claims adjusters. Causal loop as well as STELLA diagrams of the assumed interrelations appear in Figures 6A and 6B. It is assumed that the harder the adjusters work, the higher the effective time available for settlements. An increase in the department's effective time corresponds to an increase in capacity adequacy. Work intensity is assumed to adjust according to time pressure with a lag. A four month adjustment period is assumed, so that an increase in time pressure eventually results in an increase in work intensity. We may identify one important negative feedback loop in Figure 6A. An increase in work intensity results in an increased effective time available for settlements. Consequently, the department's time pressure decreases, thereby reducing the desired (indicated) intensity. A drop in the work intensity desired will result in lower work intensity after a 4 month delay.
Figure 6A: Work Intensity Subsystem - Causal Loop Diagram
Figure 6B: Work Intensity Subsystem - STELLA Diagram
With the subsystem initially in equilibrium, a step increase in incoming claims will result in an instantaneous increase in the time pressure faced by the workforce. In this subsystem, employees will meet the rise in pressure by slowly increasing their work intensity. As work intensity gradually goes up gradually, time pressure is reduced.

### 3.4 Quality Standard Subsystem

One other 'fuzzy' standard introduced in the model is the quality standard. The quality standard refers to the quality of claim investigations, file quality, as well as service quality. The model includes the possibility of an erosion in the company's quality standard and consequent rise in average settlement size as a result of increased pressure. Figures 6A and 6B present causal loop and STELLA diagrams of the portion of the model related to quality. It is assumed that increased time pressure will result in a decrease in the company's quality standard. However, change in quality can take time; a six month quality drift time is assumed. The quality of the adjustment process will affect settlement size. As quality drops, settlement size generally increases. The rise in settlement costs is due to less thorough investigations.

We now describe this process in more detail: First, we identify a negative feedback loop, illustrated on the right hand side of the causal loop diagram of Figure 7A: A rise in time pressure leads to a reduction in the quality standard (the effect of pressure on the quality standard is illustrated in Figure 8 and also in Figure A8 of Appendix 1). With a lower quality standard, the effective time required for the settlement of claims decreases (as adjusters do a sloppy and hasty job). A drop in the effective time required for settlements leads to a decline in the time pressure that adjusters face.

There is one other loop (which can be positive or negative, depending on the level of the quality standard) that we should note. This loop appears on the left hand side of Figure 6A: When the quality standard is less than its normal level of one, a drop in the quality of the claim adjusters' work will lead to increasing settlement size (the direct effect of quality standard on settlement size is plotted in Figure 8 and also Figure A6 of Appendix 1). As claims adjusters
Figure 7A: Quality Standard Subsystem - Causal Loop Diagram
Figure 7B: Quality Standard Subsystem - STELLA Diagram
Figure 8A: Effect of time pressure on quality standard as was incorporated in the STELLA model.
Figure 8B: Effect of quality standard on settlement size as was incorporated in the STELLA model.
reduce the quality of their work, they tend to rush their investigations. The easiest way for an adjuster to reduce pressure is to settle claims over the phone without careful investigation. Faced with a pile of claims pending, the claims adjuster has the option to simply accept customer claims without a thorough investigation. The customers will probably be happy since they will receive their claims. Furthermore, the pressure on the claims adjusters will drop as the pile of claims pending on their desks will decrease. Inevitably, however, such a practice will lead to higher settlement costs.

If, on the other hand, the quality standard is significantly higher than one (i.e. three or higher), with overzealous adjusters already spending too much time for each settlement, settlement costs will drop as the quality index decreases.

An increase in settlement size will eventually lead to an increase in the quality standard (with a relatively long delay of 18 months). As costs escalate, claims adjusters are eventually forced to pay more attention to the level of their settlements. However, it can take months before the escalating cost effects of hasty investigations become apparent to the company. It is for this reason that the delay time is as high as 18 months.

When the quality standard is less than its normal level of one, the feedback loop on the left hand side of the causal loop diagram of Figure 6A is a negative one: a drop in the quality standard leads to increased settlement size and consequently to an eventual increase in the quality standard. On the other hand, when the quality standard is high (over three) a drop in quality standard leads to a drop in settlement size and consequently to a further drop in quality.

Finally, there is another weak positive loop due to the effect of adjuster capacity adequacy on settlement size: An increase in quality standard increases the effective time required for settlements. This higher time requirement corresponds to a drop in adjuster capacity adequacy. Lower capacity adequacy results in increased settlement size and, with a long delay (18 months) improved quality standards.

With the subsystem originally in equilibrium with incoming claims at a constant rate and with settlement size and quality standard also constant, an unexpected step increase in features
will have the following effect: As features rise, with increased production requirements, the total (effective) time required for the settlement of claims rise. As a result time pressure on the employees also rise instantaneously. Higher time pressure will bring, after a delay, an erosion in quality standards. By lowering the quality of their work, adjusters are able to reduce their production pressure. The above sequence of events describes the effects of the right hand side negative causal loop of Figure 6A.

As quality standards erode, however, the left hand causal loop (which is usually negative) of Figure 6A kicks in. With the quality standard dropping below its normal level of one, and claims adjusters rushing their investigations, the average claim settlement size begins to climb. Higher settlement costs will eventually, though after a long delay, slow down the erosion in quality standards.

3.5 THE COMPLETE MODEL

The complete model of the operations of the Claims Department combines the subsystems presented in sections 3.1 to 3.5. Figure 9 presents a causal loop diagram of the complete system. The complete STELLA model is included in Appendix 1. We now identify the four most important loops of the model.

a) An increase in claims pending results in increased time pressure to the employees. Higher time pressure can lead to an eventual increase in the work intensity of the adjusters and consequently to increased settlements (and a reduced number of claims pending-a negative loop).

b) At the same time, an increase in pressure can lead to an erosion of the company's quality standards; with lower standards (and less time spent on each claim) settlements will increase bringing the number of claims pending down (a negative loop).

c) An increase in claims pending and in time pressure can also lead, however, to increased employee turnover. Higher turnover further aggravates the pressures in the firm and lowers settlements (positive loop).
Figure 9: Causal Loop Diagram for the complete model which was used in the Claims Game. Outlined characters illustrate the three decisions which were made by the players of the game.
d) As the pile of claims pending increases, so does their average age. As claims get older, investigations take more time, since the collection of data becomes more difficult. This increase in the time required for the investigation of claims reduces the total number of settlements and further increases the number of claims pending (a positive loop).

Furthermore, with quality standards eroding, as a result of increased production pressures, settlement costs escalate.

The complete model confirms the original suspicion of the Hanover managers that escalating settlement costs may be due to endogenous production pressures rather than due to exogenous forces. We see that the company can 'look good without being good': Increased pressure from a rising volume of claims may be met by additional hiring or by increasing work intensity. However, these pressures can also be reduced with the erosion of quality standards. The adjusters can easily manage to 'look good' by hastily reducing the pile of claims pending on their desks. Hasty settlements will eventually, however, result in the escalation of loss expenses, thus damaging the company's long term profitability.

The model presented above was incorporated in a simulation game which was given to insurance claim managers. By observing the managers' performance in the game, I examine whether a change in their decision making process could reduce the insurance company's costs and therefore improve the firm's profitability.
CHAPTER 4
THE CLAIMS GAME

Following the development of a model for the Claims operations at Hanover, the company began developing a Claims Learning Laboratory, where managers could 'come for dialogue and debate key issues they face, stimulated by a variety of learning tools' such as the model discussed above. In order to assist managers in discovering the 'fundamental pitfalls and leverage points in managing capacity growth, settlements, and investigation quality,' the Claims Game was designed. [11] The game was developed by members of the System Dynamics Group at MIT. It is based on the STELLA model presented in the previous section, with one important difference: The hiring, desired settlement size, and desired production decisions are made by the players of the game. In the STELLA model presented earlier, assumptions for these decisions were incorporated in the design.

For the development of the game, the general purpose gaming and simulation programs MicroWorlds Creator and MicroWorlds Explorer were used. The programs were developed by Ernst Diehl, a doctoral student at MIT. In this thesis I examine the performance of insurance claims managers on the game. I use their performance to determine how insurance managers make their decisions.

The game can be played on a personal computer (Macintosh). The screen displays the 'game board' illustrated in Figure 10. Subjects assume the role of Manager for the entire claims department of an insurance firm. Each period in the game represents one month. The screen display is divided into four parts:

On the left hand side of the screen, current month claims and adjuster stock and flow information is provided. The claims stock and flow information displayed includes the following: production ratio, pending ratio, features (incoming claims), claims pending, and claims settled. The adjuster stock and flow information displayed includes the total number

-53-
Figure 10: The Claims Game Board as it appears on the screen of a personal computer
of adjusters, the number of adjusters hired in the previous month, as well as the number of turnovers in the past month. Finally, the current month's ratio of claims pending per adjuster and of claims settled per adjuster are also displayed.

A narrow vertical section in the middle of the screen displays current settlement cost information, such as the average settlement size in the past month, the total (administrative) expense per settlement, and finally the total cost per claim (which is the sum of settlement size and expense per settlement).

The information provided on the screen at the start of every round displays the current status of the claims department. In addition to current month information, players have the possibility to view numerical reports or graphs of their performance in previous months. To access this information, players use the menu on the bottom right hand side of the screen and 'click' on the appropriate title. The available information is presented in Table 2. Although not immediately visible, this information was easily accessible.

The top right part of the screen was the region allocated for the input of the players' decisions. The players are required to make three decisions each month: They had to decide on hiring, desired productivity, and desired settlement size.

With the hiring decision, players can choose how many new adjusters will join the company each month. By boosting the hiring rate, the players can thus reduce increasing production pressures on the existing workforce. The desired productivity decision allows players to control the monthly production rate required by the workforce. As illustrated in the STELLA model for the game, desired productivity (or more accurately desired production ratio) is the ratio of desired production rate to incoming feature rate. By increasing desired productivity, and consequently the desired production rate, managers are able to put more pressure on their workforce in order to increase the settlement rate. The desired settlement size (average cost for each settlement) decision allows managers to control the department's quality standard. By requesting a low settlement size, managers are effectively requesting a high quality of investigation for each claim.
Each player's goal is to minimize cumulative costs over a four year period. Lower 'scores' therefore reflect better performances in the game.
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TABLE 2</strong></td>
<td>List of reports or graphs available to the players of the claims game.</td>
</tr>
<tr>
<td>Change in Operating Ratio</td>
<td>Comparison of current month total settlement cost to initial month cost</td>
</tr>
<tr>
<td>Score: Cumulative Total Cost</td>
<td>Individual performance measure</td>
</tr>
<tr>
<td>Cost per Claim</td>
<td>Settlement size, administrative expense per settlement and their sum, total cost</td>
</tr>
<tr>
<td>Settlement Size</td>
<td>Settlement Size, Desired Settlement Size</td>
</tr>
<tr>
<td>Production Decisions</td>
<td>Desired Productivity, Production Ratio, Quality Standard</td>
</tr>
<tr>
<td>Claims: Stock</td>
<td>Stock of claims pending</td>
</tr>
<tr>
<td>Claims: Flows</td>
<td>Claims Settled, Incoming Features</td>
</tr>
<tr>
<td>Claims: Performance</td>
<td>Pending Ratio, Production Ratio, Quality Standard</td>
</tr>
<tr>
<td>Claims: Age</td>
<td>Average age of claims pending</td>
</tr>
<tr>
<td>Adjusters: Stock</td>
<td>Total number of adjusters</td>
</tr>
<tr>
<td>Adjusters: Flows</td>
<td>Hiring rate, Turnover rate</td>
</tr>
<tr>
<td>Adjusters: Workload</td>
<td>Claims Pending per Adjuster, Claims Settled per Adjuster</td>
</tr>
<tr>
<td>Adjusters: Pressures</td>
<td>Time Pressure, Work Intensity, Ratio of Actual to Desired Settlement Size</td>
</tr>
<tr>
<td>Adjusters: Mix</td>
<td>Fraction of Senior Adjusters</td>
</tr>
<tr>
<td>Workload</td>
<td>Required Time, Effective Time</td>
</tr>
<tr>
<td>Time Effectiveness</td>
<td>Indicates the effect of the average age of claims pending on settlement time</td>
</tr>
</tbody>
</table>
CHAPTER 5
THE EXPERIMENTAL SESSIONS

The data analyzed in this thesis were collected during two three-day sessions of the Claims Learning Laboratory which took place at Hanover Headquarters in June and July of 1988. A total of 24 claims managers participated in the sessions. Tables 2 and 3 present data on the age and experience of the subjects. All of the subjects held B.A. degrees; one subject had an S.M. degree, and one other a J.D. degree. Their average age was about 43 years. They had an average experience of about 16 years in the insurance industry. Their average experience with Hanover was around 7 years.

The first day of the sessions began with an introduction to the Claims Learning Laboratory and to the Systems Thinking approach at Hanover. Participating managers were asked to express their thoughts, concerns, or reservations about the program. The discussion then focused on an identification of the major operational objectives, strategies, as well as obstacles of claims management. The listing of the major issues (qualms) which concerned claims managers was followed by a formulation of causal loop diagrams which related these concepts (the conceptualization stage).

The development of causal loop diagrams began with a very general 'big picture' description of the system. This 'big picture' was then followed by more specific causal loop models describing specific aspects of the organization (and their interdependencies). Subsystem models, such as the ones presented in Chapter 3 were developed.

In the second day of the Claims Learning Laboratory session, the participating managers were presented with a causal loop of the complete system, similar to the one presented in Figure 9. Having discussed this model, managers were introduced to the Claims Game itself.

Participants were divided into teams of two and were allowed several practice runs in order to familiarize themselves with the game environment. The practice runs involved several
TABLE 3: Age of managers who participated in the Claims Learning Laboratory sessions of June and July 1988.
TABLE 4: Industry and company experience of managers who participated in the Claims Learning Laboratory sessions of June and July 1988.
'Learning Scenaria'. For example, in the first practice run, participants were asked to manage the production ratio only. The purpose of this exercise was to illustrate to the claims managers that it is relatively easy to maintain the production ratio at 100 percent, by allowing quality standards to erode and settlement size to increase. Another learning scenario asked participants to push on quality. The players were asked to improve the company's quality standard. With this learning scenario, players were able to see the increased pressures on the claims department which will result from a push on quality. Several such scenarios were tested.

For the practice runs, a step input in features was used: Incoming claims began at a constant level in the first and second months and then increased by 20% in the third month (Figure 11). Features remained at this higher level throughout the remainder of the game (50 months). Before the introduction of the step increase in the incoming rate of claims, the system was at a stationary equilibrium, with the stock of claims pending constant and the settlement rate equal to the arrival rate of claims. Work intensity, quality standard as well as time pressure were all initially at their normal level of one. The step in features was not announced to the players in advance.

Before the practice runs participants were asked to describe their strategies and to plot their expected performance (in terms of production, settlement size, quality, pending claims etc). The comparison of actual outcome in the practice runs to expected results helped the managers improve their understanding of the system. By comparing the actual performance of the system to the expected one, players could identify their misperceptions about the claims department model. The trial runs thus gave the managers a better understanding of the forces and delays incorporated in the model.

It was in the third day of the Learning sessions that the data analyzed in this thesis were collected. After the two day discussion of the managers mental models of the claims operations of the company and their introduction to the game, the managers were once again asked to play the game. In this 'free play' session, incoming features did not display the simple step form of the trial run but fluctuated randomly about their first period value (as illustrated in Figure 12).
Figure 11: Step input in incoming features. Features begin at a 2,500 monthly rate and climb in month 3 to a constant rate of 3,000 claims per month.
Figure 12: Incoming features fluctuating randomly about a 2,500 monthly rate (the distribution plotted here, which was actually the input to the game of Team 1, had a mean of 2,499 claims per month and a standard deviation of 152)
Subjects expected the game to last for 50 months, but the game was halted after 44 months to avoid horizon effects. Each simulation game typically lasted about 30-60 minutes.

The 'free play' session was followed by a debriefing session. In this session, which was characterized by some as the most useful part of the exercise, participants described their strategies and performance during the game. The Claims Learning Laboratory sessions were concluded with feedback from the participants on the usefulness of the program and on the insights they may have gained.
CHAPTER 6
DATA ANALYSIS

The Claims Learning Laboratory was introduced at Hanover in order to enhance the learning process of the company's Claim managers. The Claims Game plays an important role in helping managers improve their mental models of the Claims adjustment process. The simulation game allows managers to observe in a few minutes the long term as well as short term consequences of various policy decision choices. With this simulation experience, it is hoped that managers will be able to improve their decision making process.

In this section I examine the game performance of experienced Claims managers. In my analysis of their performance, I look for clues as to how they reach their monthly decisions. With an improved understanding of the Claims management decision making process, I then present suggestions as to how insurance managers may be able to improve their company's profitability (by as much as 30%). Furthermore, using the conclusions of my analysis, I provide specific recommendations on how the design of the Claims Game and of the learning laboratory could be improved.

I begin my analysis with a presentation of the free play data that were collected on the last day of each learning laboratory, in June and July of 1988. The data are presented in the following order: First, I describe the exogenous inputs as well as the initial game conditions. I then present some of the strategies that the participants suggested before playing the game. I continue with a detailed qualitative presentation of two individual teams which pursued different strategies. Finally, I present data on the performance of all 12 participating teams.

The data analysis will be followed in the next chapters by the calculation of a benchmark score for the game and by the derivation of decision rules that explain the managers' decisions.
6.1 EXOGENOUS INPUT-FEATURES

The Claims Game is designed so that only the flow of incoming features is exogenous. Players are not told beforehand what the input features will be like. A typical input for a free play session was illustrated in Figure 12. For the twelve games analyzed here, features fluctuated randomly about their initial value of 2,500 claims per month. Monthly features for the 44 months of the game were drawn from a normal distribution of random numbers with a mean of 2,500 and a standard deviation of about 135. Due to the finite sample size, the mean and standard deviation of incoming features varied from game to game. Table 5 presents detailed information on the incoming features for the twelve teams.

6.2 INITIAL CONDITIONS

For the various 'stocks' of the insurance model, initial conditions had to be established. The initial levels that the managers were faced with are listed in Table 6. Initial conditions were set so that the system began at a stationary equilibrium: For the case where incoming features did not depart from their initial level throughout the game, all system variables remained unchanged for the duration of the game. The initial levels were also chosen so that they would appear realistic to the Hanover claims managers. The game began with 54 claims adjusters employed (22 senior adjusters, 28 new adjusters), 5,000 claims pending (a pending ratio of 2, or 200%), a quality standard of 1 (or 100% of its normal level), and a work intensity of 1 (100%). The initial average settlement size per claim was $2,000. Normal time spent on each claim was set at 34/2,500 (Initial Effective Time/Initial Features).

6.3 STRATEGIES SUGGESTED BY PLAYERS

Before playing the game, participants were asked to describe the strategy they planned to follow during the game. The comments presented in Figures 13A and 13B illustrate two different strategies. Team 5 intended to pay more attention to the experience level of the
Table 5
Incoming features: Descriptive statistics for the 12 teams.

<table>
<thead>
<tr>
<th>TEAM</th>
<th>MEAN</th>
<th>ST. DEV.</th>
<th>MIN.</th>
<th>MAX.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team 1</td>
<td>2,499</td>
<td>152</td>
<td>2,250</td>
<td>2,725</td>
</tr>
<tr>
<td>Team 2</td>
<td>2,502</td>
<td>154</td>
<td>2,250</td>
<td>2,725</td>
</tr>
<tr>
<td>Team 3</td>
<td>2,431</td>
<td>112</td>
<td>2,250</td>
<td>2,725</td>
</tr>
<tr>
<td>Team 4</td>
<td>2,509</td>
<td>133</td>
<td>2,250</td>
<td>2,725</td>
</tr>
<tr>
<td>Team 5</td>
<td>2,496</td>
<td>151</td>
<td>2,250</td>
<td>2,725</td>
</tr>
<tr>
<td>Team 6</td>
<td>2,495</td>
<td>152</td>
<td>2,250</td>
<td>2,725</td>
</tr>
<tr>
<td>Team 7</td>
<td>2,520</td>
<td>131</td>
<td>2,250</td>
<td>2,725</td>
</tr>
<tr>
<td>Team 8</td>
<td>2,513</td>
<td>135</td>
<td>2,250</td>
<td>2,725</td>
</tr>
<tr>
<td>Team 9</td>
<td>2,522</td>
<td>132</td>
<td>2,250</td>
<td>2,725</td>
</tr>
<tr>
<td>Team 10</td>
<td>2,520</td>
<td>131</td>
<td>2,250</td>
<td>2,725</td>
</tr>
<tr>
<td>Team 11</td>
<td>2,537</td>
<td>127</td>
<td>2,250</td>
<td>2,725</td>
</tr>
<tr>
<td>Team 12</td>
<td>2,520</td>
<td>131</td>
<td>2,250</td>
<td>2,725</td>
</tr>
</tbody>
</table>

Note: Monthly features were drawn from a normal distribution of random numbers with a mean of 2,500 and a standard deviation of about 135. Due to the finite sample size, individual means and standard deviations vary slightly from game to game.
### TABLE 6

**Initial Conditions for the Claims Game**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Initial Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Settlement Size ($/feature)</td>
<td>2,000</td>
</tr>
<tr>
<td>Admin. Expense per Settlement ($/feature)</td>
<td>259</td>
</tr>
<tr>
<td>Total Expense per Settlement ($/feature)</td>
<td>2,259</td>
</tr>
<tr>
<td>Desired Settlement Size ($/feature)</td>
<td>2,000</td>
</tr>
<tr>
<td>Desired Productivity (x100)</td>
<td>100</td>
</tr>
<tr>
<td>Quality Standard (% of normal level of 1)</td>
<td>100</td>
</tr>
<tr>
<td>Claims Pending (features)</td>
<td>5,000</td>
</tr>
<tr>
<td>Claims Settled (features per month)</td>
<td>2,500</td>
</tr>
<tr>
<td>Incoming Features (features per month)</td>
<td>2,500</td>
</tr>
<tr>
<td>Pending Ratio (100*months)</td>
<td>200</td>
</tr>
<tr>
<td>Production Ratio (100*months)</td>
<td>100</td>
</tr>
<tr>
<td>Age of Claims Pending (days)</td>
<td>60</td>
</tr>
<tr>
<td>Adjusters</td>
<td>54</td>
</tr>
<tr>
<td>Claims Pending per Adjuster</td>
<td>93</td>
</tr>
<tr>
<td>Claims Settled per Adjuster</td>
<td>46</td>
</tr>
<tr>
<td>Time Pressure (% of normal level of 1)</td>
<td>100</td>
</tr>
<tr>
<td>Work Intensity (% of normal level of 1)</td>
<td>100</td>
</tr>
<tr>
<td>Ratio of Actual to Desired Settlement Size (%)</td>
<td>100</td>
</tr>
<tr>
<td>Fraction of Senior Adjusters</td>
<td>44</td>
</tr>
<tr>
<td>Required Time</td>
<td>34</td>
</tr>
<tr>
<td>Effective Time</td>
<td>34</td>
</tr>
<tr>
<td>Time Effectiveness (% of normal level)</td>
<td>93</td>
</tr>
</tbody>
</table>
CLAIMS LEARNING LABOPATORY

June 27 - 29, 1988

PLAY: FREE PLAY 50 persons 5500 per month
5000 pending

TEAM/GAME: 

DATE: 6/29/88

STRATEGY: DEVELOP STAFF - Staff up from
Maintain high 1800/OVERALL 1700
BUILD EXPERIENCE LEVEL STAFF
MAINTAIN-PENDING UNDER 5000 - PRODUCTION

PROJECTED OUTCOME: EXPENSES HIGHER
Production STABLE
SETTLEMENT VALUE (QUALITY) LOWER

Strategy @ the 34 month level
we tried to increase staff to
handle the rising pending & costs
Also to ease settlement value pressure

Figure 13A. Sample comment sheet of Team 5.
PLAY: FREEPLAY

TEAM/GAME: ________________________________

DATE: 6/29/88

STRATEGY:
- STAFF EARLY
- OVERSTAFF
- EMPHASIZE QUALITY - TREND DOWN AVERAGE PAYOUT - EARLY ON 1750
- PRODUCTION 100% - 101%
- TRACK AND CONTROL TURNOVER CAREFULLY - DON'T EXCEED 2%

PROJECTED OUTCOME: ________________________________

Figure 13B. Sample comment sheet of Team 1.
adjuster force and to the pending ratio than to quality and settlement size (lower quality implied higher settlement size). Team 1 intended to pursue a different strategy with more emphasis on quality and settlement size than on experience and pending ratio. Other teams' strategies varied between these two. Some teams proposed strategies that would clearly impose severe strains on the Claims department. For example, one team intended to reduce pending ratio to 1, reduce settlement size, and maintain turnover at low levels. The production pressures which would result from a desired pending ratio of 1 (which corresponded to a cut in half of claims pending) would lead to a drop in quality (and an increase in settlement size) or to higher turnover (or both). Other teams were very vague in their descriptions of strategies (i.e. 'we will hire as appropriate'). Finally, one team was planning a wait-and-see strategy: They planned to wait for a few periods in order to get an idea of what the incoming 'trend' in the features would be like and then to act (in the free play game where incoming features were random, this strategy was bound to fail). Table 7 presents retrospective results on the variables that teams suggested they monitored during the game. We note that each team claims to have focused on a limited set of information for its decisions. The set of information that each team used was not necessarily the appropriate one. Team 6, for example, monitored the stock of claims pending, the adjuster mix (experience level), the time pressure on the workforce, as well as the production rate. The team appears to have ignored quality, since they do not mention monitoring settlement size or quality standard. Furthermore, the team does not mention monitoring adjuster capacity. The quality subsystem, which was introduced in Chapter 3 plays an important role in the overall outcome of the Claims game. By ignoring the effects of this subsystem, Team 6 was clearly pursuing a strategy which was not optimal.

The retrospective reports therefore confirm the concept of a Bounded Rationality, which was introduced in Chapter 1. Decision makers seem to be simplifying their task by using a limited set of information. The omission of important sources of information suggests that the individual strategies used were not optimal.

There is one other thing that we should point out in Figure 13A. The bottom paragraph
Table 7
Retrospective Data on Managers Decisions.

Xs indicate variables that teams claimed they monitored during the game. The information was recorded by the teams on retrospective reports.

<table>
<thead>
<tr>
<th>Team</th>
<th>Pending</th>
<th>Experience</th>
<th>Quality</th>
<th>Settl. Size</th>
<th>Pressure</th>
<th>Production</th>
<th>Adjuster Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2</td>
<td>(N/A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>10</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>12</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Note: Pending includes Pending Ratio and Claims Pending

Adjuster Experience refers to the Fraction of Senior Adjusters

Quality refers to the Quality standard

Pressure includes Time Pressure as well as Workload
illustrates that Team 5 was forced to abandon its original plan when faced with rising pending claims and costs in the 34th month. This was in fact not uncommon in the data that we studied. Teams did not appear to follow a consistent long-term strategy, but rather appeared to change their original plan, when faced with unexpected problems. This made interpretation of the data more difficult. Frequent changes in the teams strategies within the 44 month duration of the game also suggested a misperception by the managers of the time delays involved in the model. When some effects of a particular policy would take more than a year to become apparent, a change of policy after only three or four months did not make sense.

6.4 TWO CASE STUDIES

I now present the game results for two teams that, based on their performance, appeared to pursue different strategies in their games. Team 10 appeared to be staffing-up early in order to reduce claims pending, without any attention to quality. On the other hand, Team 1 was aggressive in trying to improve quality. We now look at the performance of these two teams in more detail.

Figures 14-27 present the performance of Team 10. Their final score (cumulative cost) was $256.963 million. With this score they ranked eleventh out of twelve teams. We observe the following:

Desired Settlement Size (one of the three monthly decisions each team had to make) was not changed from its original level of $2,000 per feature (Figure 15). Since settlement costs are related to quality (as illustrated in the model of Chapter 3), we may conclude that the team did not pay any attention to quality (the team's retrospective report of variables monitored, which is included in Table 7 agrees with this).

Desired productivity, the other monthly decision was also left constant at 100 (Figure 16). In fact, most teams left the desired productivity level unchanged, or changed it only once in the duration of the game. Not one of the teams changed desired productivity often.
Figure 14: Team 10 performance
Figure 15: Team 10 performance
Figure 16: Team 10 performance
Figure 17: Team 10 performance
Figure 18: Team 10 performance
Figure 19: Team 10 performance
Figure 20: Team 10 performance
Figure 21: Team 10 performance
Figure 22: Team 10 performance
Figure 23: Team 10 performance
Figure 24: Team 10 performance
Figure 25: Team 10 performance
Figure 26: Team 10 performance
Figure 27: Team 10 performance
All of the action in the decisions of Team 10 therefore concentrated on hiring. As Figure 22 illustrates, Team 10 hired aggressively in the first 5 months. After the fifth month hiring was reduced to approximately the turnover rate, with occasional small adjustments. As a consequence of the aggressive initial hiring, the number of adjusters climbed in the first few months (Figure 21). The number of adjusters was then allowed to slip. The drop in hiring after the first few months resulted in an increasing fraction of senior adjusters (Figure 25).

This original aggression in hiring may reflect a desire to bring down the number of claims pending. Over the duration of the game, the team was successful in cutting its stock of claims pending (Figure 17) by about half, the goal stated in their strategy. With features and settlements at a fairly constant rate on average, (despite short term fluctuations- Figure 18), the pending ratio was also cut in half (Figure 19). Since the average age of the claims pending is directly related to the pending pool, it too was halved (Figure 20).

I now present an example of a period-by-period account of the game by Team 10. The examination focuses on the decisions that the team made and on their apparent results. The examination of the overall behavior of the system in response to team decisions illustrates the effects of the various feedback loops of the system described in Chapter 3. Furthermore, the team's decisions provide a clue as to how well the dynamic system was understood by the players and as to how effective their strategy was.

**Team 10 Performance**

Before beginning the game, Team 10 included five goals for the game in its comment sheet (sample comment sheets were presented in Figure 13). The five goals were to maintain the pending ratio at its initial level of 2, to increase the ratio of senior adjusters employed in the firm, to reduce the average settlement cost, to maintain the settlement rate at the level of incoming features (100% productivity), and to maintain turnover ratio at its initial level. As the account below illustrates, most of these goals were indeed accomplished. Comparison of the team's cumulative cost to the performance of other teams will show, however, that despite the
success in meeting their goals, the team could have done better by paying more attention to quality and settlement costs.

**Months 0-10**

Team 10 began the game by hiring very aggressively. As a result of the aggressive hiring of the first five periods, the effective time available for the settlement of claims initially increased (Figure 26). With no aggressive desired productivity or settlement size decisions, time pressure (Figure 24) dropped and remained low for the first 11 months. With pressure low, quality began to improve (Figure 16). Higher quality, with a fixed desired settlement size, resulted in decreasing settlement costs for the first 18 months (Figure 15). With time pressure down, work intensity also decreased for the first ten months. Following positive results for the first 5 months, Team 10 stopped hiring. Only in the tenth month, when things appeared to be deteriorating, did they begin to hire again.

**Months 11-20**

Up until the tenth month, the average age of claims pending had been over 45 days and decreasing, so time effectiveness had been going up (the relationship between age and time effectiveness is illustrated in Figure 4). The rise in time effectiveness (Figure 27), combined with the aggressive hiring of the first months resulted in an increase in the effective time available to adjusters for the settlement of claims. However, in the tenth month, the age of the claims dropped below the optimal level of 45 days. As claims pending and their age continued to drop, time effectiveness began to slip.

Lower time effectiveness and decreasing adjuster capacity resulted in a drop in the effective time available for claims adjustment. The drop lasted from month 11 until month 18. As effective time dropped, time pressure climbed. The rise in time pressure boosted work intensity (with a delay). Also, due to the higher time pressure, the rise in the quality standard began to slow down. By month 20 quality had levelled off.

Settlement size began to deteriorate, with a delay. Furthermore, capacity adequacy, which is the ratio of effective time available by the employee workforce for the settlement of claims to
the time required for these settlements, dropped. With capacity adequacy decreasing the fraction of desired claims settled dropped leading to a deterioration in the production ratio (Figure 16).

**Months 21-30**

The sharp drop in claims pending that took place in the first twenty months slowed to a halt in months 21-30. Consequently, the average age of pending claims also levelled off. Time effectiveness was fairly constant (until month 28). I should point out that for an age of less than 45 days, time effectiveness was very sensitive to changes in the age of the pending claims (this can be understood by referring to Figure 4).

As the quality standard also levelled off, the effective time available for the settlement of claims continued to drop (until month 25). Decreasing effective time available (with the effective time required remaining fairly constant) resulted in a further increase in time pressure. By the second half of this time period (months 25-30), both time pressure and work intensity had also levelled off.

While most other variables levelled off, reaching some sort of a steady state, settlement size did not. As capacity adequacy dropped to levels below 1 (see Figure A9 of the Appendix), the settlement size became particularly sensitive to changes in capacity. This becomes apparent in month 24, when settlement size shoots up. The sharp increase is due to a sudden rise in incoming features during that month. The rise in features resulted in an instantaneous increase in desired production and in the effective time required for the settlement of claims. With the rise in time required came (again instantaneously in our model) a sharp drop in capacity adequacy. Thus, for months 21-30 settlement size climbed and began to display sharp fluctuations. Despite its increasing value, the actual settlement size remained below the desired settlement size, which was fixed at 2,000. The total number of adjusters and their average experience level were also fairly flat during this time period.

Team 10 set its hiring rate at a level equal to the turnover rate, until the 28th and 29th month. It is not clear why the team chose to boost hiring at this point. The hiring increase may
be due to the fact that by period 28 the levelling off of claims pending would have become apparent to the managers. They may have hired to bring the pending ratio down even further.

**Months 31-44**

Two variables display the most interest in this final phase of the game. Settlement size fluctuates very sharply and increases in its average value. The fluctuations are again due to the sharp swings in capacity adequacy. The quality standard also begins to fall, presumably as a result of the earlier increases in time pressure (there is a 6 month quality drift time due to time pressure assumed in the model).

The other variables remained fairly constant during this period. Changes in time effectiveness were due to the high sensitivity of the variable to small changes in the age of pending claims.

Hiring was again mostly done to replace turnover. The reasons for the drop in hiring in months 33 and 34 below turnover are not clear.

To summarize the performance of Team 10, we note that they succeeded in meeting their original objectives: The pending ratio, rather than being maintained at its initial level as desired, was actually improved; the senior adjuster ratio was increased with success; average settlement costs were reduced; the production rate often exceeded 100%; finally, the turnover rate was fairly constant. Despite this success however, their score was relatively low. The low score is due to the fact that the team ignored the effects of the quality standard subsystem of Chapter 3. By allowing desired settlement size to exceed actual settlement size, the team was eventually (though after a long delay) faced with eroding quality standards and rising settlement costs. Rising settlement payments in the second half of the game damaged the team's overall performance. Much like in real life, this game illustrates that emphasis on production without attention to quality will eventually lead to escalating costs and will damage an insurance firm's long term profitability.
Team 5 Performance

Team 10, observed above, did not put any pressure on quality and settlement size. The strategy of Team 5 was markedly different. This team paid close attention to quality and settlement size. The comment sheet of Team 5 is included in Figure 13A. The team planned on increasing the number of adjusters employed, increasing the staff experience level, reducing the stock of claims pending, and reducing the average settlement size. Administrative expenses were expected to rise.

Figures 28 to 41 present the performance of Team 5. Figure 29, in particular, illustrates the team's emphasis on quality: While actual settlement size was improving, the team continued to demand even further reductions in costs. In fact, the team kept increasing the margin between actual and desired settlement size (with actual always exceeding desired). As a result of the pressure on settlement size, quality standards grew continuously for the first three quarters of the game.

By the third year of the simulation game, however, sharply increasing pending claims (Figure 31) eventually forced the team to reduce its pressure on quality. The margin between actual and desired settlement size was cut, bringing about an eventual drop in the quality standard around month 34.

Hiring was aggressive for the first five months and then was set at the turnover rate up until month 28, when the team apparently decided to address the climbing pending cost issue. As of that year, the team became once again aggressive in its hiring (Figure 36); the team remained aggressive throughout the remainder of the game.

Desired productivity was left at its initial 100% level for most of the game (Figure 27).

Of all twelve teams that participated in the Claims Learning Laboratory of June and July 1988, Team 5 achieved the lowest cumulative cost at $237.670 million. While it is difficult to identify exactly the decisions responsible for their success, a comparison of their game to the performance of Team 10 suggests the following:
Figure 28: Team 5 performance
Figure 29: Team 5 performance
Figure 30: Team 5 performance
Figure 31: Team 5 performance
Figure 32: Team 5 performance
Figure 33: Team 5 performance
Figure 34: Team 5 performance
Figure 35: Team 5 performance
Figure 36: Team 5 performance
Figure 37: Team 5 performance
Figure 38: Team 5 performance
Figure 39: Team 5 performance
Figure 40: Team 5 performance
Figure 41: Team 5 performance
The emphasis on quality and on settlement costs was essential for a successful score. Teams (like Team 10) that ignored quality and focused instead only on decreasing the pending ratio ended up with higher cumulative costs.

The flexibility that Team 5 displayed, trading the high quality goal for a reduction in pending ratio when pending costs became too high seemed to be effective. As our benchmark analysis will illustrate, keeping the average age of claims pending around 45 days maximized time effectiveness and reduced costs.

The reader should note that with turnover at a level of 1 employee per month and with hiring ranging for most of the game between 0 and 2 per month, visual interpretation of the reasoning for the hiring level is difficult. A game with different initial conditions, where the initial number of adjusters had been 500 instead of 50, and turnover 10 instead of 1 would provide a better resolution for our analysis.

6.5 DISTRIBUTION OF FINAL SCORES

As was mentioned above, twelve teams (24 managers) participated in the game sessions of June and July 1988. The scores (cumulative total costs, including penalty for claims pending) cover a narrow range from a minimum of $237.6 million (the best score) to a maximum of $255.7. A complete list of the final scores appears in Table 7. The mean score is $249.815 million with a standard deviation of $6.308 million.
Table 7

Cumulative scores for the twelve teams

<table>
<thead>
<tr>
<th>Team</th>
<th>Score ($ thousand)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>244,693</td>
</tr>
<tr>
<td>2</td>
<td>246,647</td>
</tr>
<tr>
<td>3</td>
<td>241,450</td>
</tr>
<tr>
<td>4</td>
<td>247,520</td>
</tr>
<tr>
<td>5</td>
<td>237,670</td>
</tr>
<tr>
<td>6</td>
<td>256,917</td>
</tr>
<tr>
<td>7</td>
<td>252,729</td>
</tr>
<tr>
<td>8</td>
<td>253,082</td>
</tr>
<tr>
<td>9</td>
<td>254,565</td>
</tr>
<tr>
<td>10</td>
<td>256,963</td>
</tr>
<tr>
<td>11</td>
<td>249,786</td>
</tr>
<tr>
<td>12</td>
<td>255,762</td>
</tr>
</tbody>
</table>

Mean score 249,815
Std. deviation 6,308
Benchmark 243,053 (see Chapter 7)

Note: The null hypothesis that the mean score of the twelve teams is equal to the benchmark score was tested and rejected. The t-statistic for the test is 3.71 (with 11 degrees of freedom the significance of the test is 0.003). We thus conclude that the benchmark score is significantly different from the team score distribution.
CHAPTER 7
A BENCHMARK FOR THE CLAIMS GAME

To evaluate the performance of the claims managers, we develop here a benchmark through computer simulation. The complexity of the claims model makes direct calculation of the optimal behavior intractable. The STELLA model of Chapter 3 is used. The model corresponds exactly to the one used in the game. The decision rules used to determine a benchmark are the following:

Hiring is assumed to take an anchoring and adjustment form. Anchoring and adjustment is a common strategy used for the estimation of an unknown quantity. A known reference point (the anchor) is first used to approximate the unknown quantity. The reference level is then adjusted for the effects of other factors which may be less salient. The anchoring and adjustment model has been shown to apply to many decision making processes [4,21-23].

The monthly rate of turnovers is used as the anchor for hiring, while the adjustment is prompted by the discrepancy between the desired stock of adjusters and the actual stock. A stock adjustment time of T₁ is assumed:

\[ \text{Hiring} = \text{MAX} \left[ 0, \text{Turnover} + \left( \text{Desired Adjusters} - \text{Actual Adjusters} \right) / T₁ \right] \] \hspace{1cm} (1)

By taking the maximum of zero and the adjusted turnover rate we ensure that the hiring level is always positive. The game does not allow for a layoff policy.

The level of desired adjusters is modeled as the ratio of desired settlement rate divided by desired productivity per adjuster. Furthermore, the model for the desired level of adjusters includes a second factor; this factor will allow an increase in the desired level of adjusters due to a discrepancy between actual settlement size and desired settlement size:
Desired Adjusters = (Desired Settlements/Desired Productivity per Adjuster) +
+ a₁ [ A (S-S*) / S*] 

(2)

where A denotes the actual level of Adjusters, S denotes the actual Settlement Size, S* the Desired Settlement Size and a₁ is a constant.

Desired Settlements are also modeled using an anchoring and adjustment model. Expected Features (the expected rate of incoming features) are used as the anchor, with an adjustment due to the discrepancy between the actual and the desired stock of claims pending. The stock adjustment time is denoted as T₂:

Desired Settlements = Expected Features + (Pending - Desired Pending)/T₂

(3)

The desired stock of claims pending is assumed to depend on the expected incoming rate of features as well as on the desired age of pending claims:

Desired Pending = Desired Age * Expected Features

(4)

Since no long term growth trend in incoming features was included in the games analyzed, we use the average level of features to model expected features:

Expected Features = Average Features

(5)

Average features are computed in the STELLA model using a first order exponential smoothing with a time constant of 3 months.

Desired productivity per adjuster was originally set at the ratio of initial settlement rate to initial adjuster stock:
Desired Productivity per Adjuster = Initial Settlements/Initial Adjusters \hspace{1cm} (6)

The desired age of claims pending is set at a parameter value AG

Desired Age = AG \hspace{1cm} (7)

Desired Settlement Size is modeled as a stock. It is assumed to decrease with an adjustment time \( T_3 \) when the difference between actual settlement size and an adjustment margin drops below the desired settlement size level:

\[ \frac{dS^*}{dt} = \frac{[(S - \text{Adjustment Margin}) - S^*]}{T_3} \] \hspace{1cm} (8)

The benchmark model of equations (1)-(8) thus include six parameters, namely:

\( T_1 \) : Hiring adjustment time for discrepancy between desired and actual adjuster levels (Equation 1).
\( a_1 \) : Coefficient of contribution to desired adjuster level due to discrepancy between actual and desired settlement size (Eq. 2).
\( T_2 \) : Desired Settlement adjustment time due to discrepancy between actual and desired pending levels (Equation 3).
AG: Desired age of claims pending (Equation 5).
Adjustment Margin for desired settlements (Equation 8).
\( T_3 \) : Desired Settlement Adjustment time (Equation 8).

A set of parameters which produce low total costs were determined by simulating the game over a wide range of plausible parameter values. Trial-and-error was used to determine the benchmark costs. The best parameters were:

\( T_1 = 0.74 \) (months)
\( a_1 = 0 \)
\( T_2 = 2 \) (months)

Desired age of claims pending = 45 days (\( AG = 1.5 \) months)

Adjustment Margin for desired settlements = 240 ($ per feature)

\( T_3 = 1.1 \) (months)

No further improvement was accomplished by varying desired productivity per adjuster.

We note that the benchmark solution involves a desired age of claims pending of 45 days, just as Figure 4 would suggest. At that age time effectiveness is maximized. Furthermore, no adjustment in the desired level of adjusters due to a gap between desired and actual settlement size was necessary (\( a_1 = 0 \)). The response of the STELLA model with the above decision rules is included in Appendix. We should note that for this 'optimal' policy, hiring is aggressive in the first few periods and then drops off.

The 'optimal' score from the benchmark model presented above is $243.053 million. This score surpasses ten out of the twelve teams that participated in the claims learning laboratory (Table 7). Furthermore, this score betters the teams' mean score of $249.815 million by over one standard deviation.

To determine whether the difference between the benchmark score and the team performances is statistically significant, we test the null hypothesis that the benchmark score is equal to the sample mean, versus the hypothesis that the benchmark score is not equal to the sample mean. With a t-statistic of 3.7 in a test with 11 degrees of freedom, the null hypothesis is rejected (the significance level is 0.003).

The benchmark score is therefore statistically superior to the performance of skilled, experienced managers. While the difference in costs is only 2.7%, the magnitude of this improvement is large. For a typical insurance company, a 2.7% reduction in claims costs would raise net income by about 27%. Hanover Insurance, for example, paid out $1.3 billion in loss and administrative expenses in 1987. A 2.7% reduction in costs would amount to about $35 million, or 31% of the $112 million net income.
CHAPTER 8
MODELS OF THE SUBJECTS' DECISIONS

The benchmark test shows that reasonable heuristics for hiring and settlement size can outperform experienced managers. What heuristics, then, were the subjects using? Why do they fail to produce minimum costs? In our effort to understand the claims managers' decisions, we attempted to formulate decision rules that would explain their behavior. Several decision rules were tested, with varying degrees of success.

8.1 DESIRED SETTLEMENT SIZE

Most of the desired settlement size plots (with the exception of teams that paid no attention to quality) resembled that of Team 5 (shown in Figure 26): Desired settlement size was revised (usually downward, as actual settlement size dropped) incrementally and only after several periods had elapsed from a previous change. From the managers' own accounts we can conclude that that revisions in the desired settlement size level were not made every month for the following two reasons:

Some managers adopted a wait-and-see strategy, where they set the desired decision variable at a certain level and then waited to see the results.

Other managers claimed in their prospective reports that they would review settlement size only after a certain number of months had elapsed, i.e. three months or six months.

The simple continuous decision rule that we used in order to model the managers' settlement size decision abstracted from the discrete character of the game. As we illustrate below, however, the decision rule did explain their decisions fairly well.

The decision rule that we used for $S^*$ assumed that managers included an adjustment margin when setting their desired settlement size. This adjustment margin was modified as the actual settlement size departed from a reference level.
Adjustment margin \( = S - S^* = b (S_R - S) \) \( \tag{9} \)

where \( S \) is the actual settlement size, \( S^* \) is the desired settlement size, \( S_R \) is a reference level for settlement size, and \( b \) is a constant. To test this model, we performed linear regressions using the data of each individual team. The values for the parameters \( b, S_R \), as well as the \( R^2 \) for each team are presented in Table 9. We note that the decision rule explained fairly well the managers performance: the values of \( R^2 \) ranged from a high of .90 to a low of .15.

The above parameters suggest that managers tended to increase the adjustment margin as settlement size decreased below their reference. The negative sign of the constant regression coefficient \( a \) (for most of the teams) indicates that players tended to set desired settlement size below actual settlement size. The values of a range between -3.971 and +491. The positive sign of \( b \) indicates that teams became more and more aggressive as their settlement costs improved. Coefficient \( b \) ranges from a low of 0.74 to a high of 2.99, with higher values reflecting more aggressive teams. By using the values of \( a \) and \( b \) we are able to calculate the reference settlement cost for each team, \( S_R \). For most teams this reference settlement size is close to the initial cost of $2,000 per feature.

Most of the teams appeared to pursue a fairly aggressive policy with regard to quality. This aggressiveness seems to contradict recent industry experience of escalating costs and eroding quality standards. I suspect that the strong emphasis on raising quality reflects the lessons of the two days of the learning laboratory before these trials.

8.2 DESIREd PRODUCTIVITY

Desired productivity was modified very infrequently by the subjects. The decision allowed players to control the desired monthly production level of the claims department. The product of desired productivity times the number of incoming claims per month determines the desired monthly rate of claim settlements. An increase in desired productivity can boost monthly
Table 9: Desired Settlement Size Decision Model—Regressions

The following decision rule was tested:

\[ S - S^* = b \left( S_R - S \right) \]

with the linear regression \( S^* = a + (1+b) S \)

so that \( S-S^* = -a - bS \)

and \(-a = bS_R\)

(\(S^*\) and \(S\) are the desired and actual settlement size, \(S_R\) a reference settlement size, and \(a\) and \(b\) regression coefficients)

<table>
<thead>
<tr>
<th>Team</th>
<th>(a) ((13))</th>
<th>(1+b) ((11))</th>
<th>(b) ((8))</th>
<th>(S_R) ((10))</th>
<th>(R^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-3239 (-13)</td>
<td>2.58 (20)</td>
<td>1.58</td>
<td>2,050</td>
<td>.90</td>
</tr>
<tr>
<td>2</td>
<td>-2983 (-7)</td>
<td>2.44 (11)</td>
<td>1.44</td>
<td>2,072</td>
<td>.75</td>
</tr>
<tr>
<td>3</td>
<td>-2630 (-5)</td>
<td>2.30 (8)</td>
<td>1.30</td>
<td>2,023</td>
<td>.59</td>
</tr>
<tr>
<td>4</td>
<td>-3040 (-4)</td>
<td>2.45 (7)</td>
<td>1.45</td>
<td>2,102</td>
<td>.51</td>
</tr>
<tr>
<td>5</td>
<td>-1682 (-6)</td>
<td>1.62 (10)</td>
<td>0.62</td>
<td>2,712</td>
<td>.71</td>
</tr>
<tr>
<td>6</td>
<td>-1262 (-4)</td>
<td>1.65 (11)</td>
<td>0.65</td>
<td>1,941</td>
<td>.73</td>
</tr>
<tr>
<td>7</td>
<td>-405 (-.8)</td>
<td>1.19 (4)</td>
<td>0.19</td>
<td>2,132</td>
<td>.30</td>
</tr>
<tr>
<td>8</td>
<td>+491 (1)</td>
<td>0.74 (3)</td>
<td>-0.26</td>
<td>663</td>
<td>.15</td>
</tr>
<tr>
<td>9</td>
<td>-2220 (-4)</td>
<td>2.11 (7)</td>
<td>1.11</td>
<td>2,000</td>
<td>.54</td>
</tr>
<tr>
<td>10</td>
<td>N/A (desired settlement size fixed at 2,000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>-1145 (-4)</td>
<td>1.56 (11)</td>
<td>0.56</td>
<td>2,044</td>
<td>.74</td>
</tr>
<tr>
<td>12</td>
<td>-3971 (-6)</td>
<td>2.99 (8)</td>
<td>1.99</td>
<td>1,995</td>
<td>.62</td>
</tr>
</tbody>
</table>

Note: Numbers in parentheses represent t-statistics
settlements in the short run. However, increased production demands will decrease the adequacy of the number of adjusters employed (adjuster capacity) and increase time pressures in the claims department. With a delay, increased pressures will limit the monthly settlement rate.

Table 10 presents the mean and variance for the desired productivity rate of various teams. For most teams it was left at or close to 100%. In fact, half of the teams chose not to change the level of desired productivity at all. Consequently, with so few changes in the desired productivity level, it was pointless to try to explain the data with regressions.

Managers did not pay much attention to the desired productivity level during their games. Once again, this may be due to the fact that little emphasis was put on this decision during the managers' introduction to the game, rather than due to the managers' everyday practices. Alternatively, the reason for the infrequent use may be due to a poor understanding of the effects of the desired productivity decision. Finally, trial-and-error suggests that the costs of not using the desired productivity decision are small. The players may have ignored the decision due to its apparent insignificance. If poor understanding of the decision is the cause of the infrequent use, a modified game with a desired production level decision (rather than a desired productivity decision) may provide more interesting results.

8.3 Hiring

Most of the interesting 'action' in the managers' decisions appears in the hiring option. To model the managers' decision we assume that they used an anchoring and adjustment heuristic: hiring is assumed to equal turnover, with an adjustment for a discrepancy between the desired and actual adjuster level:

\[
\text{Hiring} = \text{Turnover} + \frac{[A^* - A]}{T}
\]  

(10)
Table 10

Desired Productivity Decision: Descriptive Statistics

<table>
<thead>
<tr>
<th>Team</th>
<th>Mean</th>
<th>St. dev.</th>
<th>min.</th>
<th>max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>101.3</td>
<td>1.6</td>
<td>100</td>
<td>104</td>
</tr>
<tr>
<td>3</td>
<td>100.9</td>
<td>1.0</td>
<td>100</td>
<td>102</td>
</tr>
<tr>
<td>4</td>
<td>102.5</td>
<td>1.1</td>
<td>100</td>
<td>103</td>
</tr>
<tr>
<td>5</td>
<td>100.4</td>
<td>1.0</td>
<td>100</td>
<td>103</td>
</tr>
<tr>
<td>6</td>
<td>99.9</td>
<td>.54</td>
<td>98</td>
<td>101</td>
</tr>
<tr>
<td>7</td>
<td>98.1</td>
<td>2.4</td>
<td>95</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>9</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>11</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>12</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>
where $A^*$ denotes desired adjusters, $A$ the actual adjusters, and $T$ the adjustment time.

The desired adjuster level was assumed to depend on a sum of cues each with its own weight $w$.

$$A^* = w_0 + w_1 (c\text{ue})_1 + w_2 (c\text{ue})_2 + \ldots + w_i (c\text{ue})_i$$  \hspace{1cm} (11)$$

where $w_0$ is a constant, while $w_1, w_2, \ldots, w_i$, denote the constant weights for each cue. First, using linear regressions, we tested a model where desired adjuster level depended on settlement size and pending ratio only:

$$A^* = w_0 + w_1 (\text{normalized settlement size}) + w_2 (\text{normalized pending ratio})$$  \hspace{1cm} (12)$$

To normalize settlement size and pending ratio we divide them by their mean values for each game. We would expect the weights $w_1$, and $w_2$, to have positive values: As settlement size and pending ratio increase, we would expect that managers would want to hire more adjusters.

The results of our regressions appear in Table 10. We note that the $R^2$ for the regressions ranged from a high of .72 to a low of .08. The $t$-statistics for $w_2$, the coefficient of the pending ratio, indicate that all but two of the coefficients are statistically significant. For $w_1$, the coefficient of the settlement size, all twelve coefficients are not statistically different from zero. The model therefore does a fairly good job in explaining some teams' decisions and a poor job in explaining others. For three cases (with low $R^2$) the regressions gives us parameter values for the adjustment time which are negative, indicating a problem with our model. Furthermore, only two of the regressions gave us values for the coefficients $w_1$ and $w_2$ which were both positive.

One problem is that equation 10 does not distinguish between new and experienced adjusters. Yet Table 7 shows that most teams did attend to the experience level. To correct this
Table 11: The Hiring Decision: Two-variable linear regression testing the decision:

\[
\text{Hiring} = \text{Turnover} + \frac{[A^* - A]}{T}
\]

\[
A^* = w_0 + w_1 \text{ (normalized settlement size)} + w_2 \text{ (normalized pending ratio)}
\]

<table>
<thead>
<tr>
<th>Team</th>
<th>T (months)</th>
<th>(w_0)</th>
<th>(w_1)</th>
<th>(w_2)</th>
<th>(R^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.5 (5.8)</td>
<td>34.6 (3.0)</td>
<td>15.3 (1.0)</td>
<td>11.6 (5.8)</td>
<td>.70</td>
</tr>
<tr>
<td>2</td>
<td>-15 (0.4)</td>
<td>-304 (-0.9)</td>
<td>413 (1.4)</td>
<td>-56 (-2.5)</td>
<td>.14</td>
</tr>
<tr>
<td>3</td>
<td>12.4 (0.9)</td>
<td>200 (0.6)</td>
<td>-183 (-0.6)</td>
<td>40 (2.2)</td>
<td>.20</td>
</tr>
<tr>
<td>4</td>
<td>9.3 (1.1)</td>
<td>247 (1.5)</td>
<td>-226 (-1.4)</td>
<td>38 (3.3)</td>
<td>.19</td>
</tr>
<tr>
<td>5</td>
<td>2.9 (7.5)</td>
<td>58 (5.0)</td>
<td>-211 (-1.4)</td>
<td>22 (7.5)</td>
<td>.62</td>
</tr>
<tr>
<td>6</td>
<td>24 (1.0)</td>
<td>536 (0.6)</td>
<td>-526 (-0.6)</td>
<td>75 (3.0)</td>
<td>.60</td>
</tr>
<tr>
<td>7</td>
<td>-15 (-0.4)</td>
<td>81.9 (0.2)</td>
<td>39 (0.1)</td>
<td>-73 (-2.5)</td>
<td>.16</td>
</tr>
<tr>
<td>8</td>
<td>2 (6.4)</td>
<td>-49 (-1.0)</td>
<td>82 (1.6)</td>
<td>18 (5.8)</td>
<td>.58</td>
</tr>
<tr>
<td>9</td>
<td>22 (0.2)</td>
<td>398 (0.4)</td>
<td>-402 (-0.4)</td>
<td>58 (1.1)</td>
<td>.03</td>
</tr>
<tr>
<td>10</td>
<td>1.9 (4.8)</td>
<td>97 (1.9)</td>
<td>-55 (-1.2)</td>
<td>12 (5.7)</td>
<td>.48</td>
</tr>
<tr>
<td>11</td>
<td>-13 (-0.5)</td>
<td>-369 (-0.6)</td>
<td>457 (0.7)</td>
<td>-38 (-1.0)</td>
<td>.02</td>
</tr>
<tr>
<td>12</td>
<td>5.1 (2.6)</td>
<td>78 (0.8)</td>
<td>-47 (-0.5)</td>
<td>28 (4.5)</td>
<td>.39</td>
</tr>
</tbody>
</table>

Note: Numbers in parentheses represent t-statistics
we include adjuster experience as a cue in our model:

\[ A^* = w_0 + w_1 \text{ (normalize settlement size)} + w_2 \text{ (normalized pending ratio)} + \\
+ w_3 \text{ (normalized fraction of senior adjusters)} \tag{13} \]

Like settlement size and pending ratio, to normalize the fraction of senior adjusters we divide it by its mean value over each game. We would expect the sign of \( w_3 \) to be negative: As the employee experience level increases, fewer adjusters would be desired. The results of our regressions are presented in Table 11. We note that the adjustment times \( T \) are all positive ranging from a low of 1.2 months to a high of 9.1 months. For most teams the adjustment time is close to two months. The values of \( R^2 \) (adjusted for the additional variable) are higher, ranging from a high of .8 to a low of .24.

We should also point out that for each team the weights \( w_0, w_1, w_2 \) and \( w_3 \) add up to a value close to the level of adjusters (which ranged between 50 to 100 adjusters) as we would have expected.

It is difficult however to interpret the individual weights due to the high correlation between some of the independent variables of our regression. For example, for Team 2, settlement size is found to be negatively correlated to the fraction of senior adjusters (by a factor of minus .92).

The anchor and adjustment model allows us to explain a significant part of the managers' decisions. However, high correlations between cues make interpretation of the weights for the cues difficult. Thus, it is not possible for example to separate the exact weight that managers placed on pending ratio when they made their hiring decision, because it is highly correlated to settlement size (a high number of claims pending leads to a deterioration in settlement size due to lower quality).

Most teams appear to hire aggressively in the first few months, building their capacity and eventually their experience level. While this aggressive initial hiring seems to be correct and
Table 12: The Hiring Decision: Three-variable linear regression testing the decision:

\[
\text{Hiring} = \text{Turnover} + \frac{[A^* - A]}{T}
\]

\[
A^* = w_0 + w_1 \text{ (norm. settlement size)} + w_2 \text{ (normalized pending ratio)} + w_3 \text{ (normalized fraction of senior adjusters)}
\]

<table>
<thead>
<tr>
<th>Team</th>
<th>T (months)</th>
<th>w₀</th>
<th>w₁</th>
<th>w₂</th>
<th>w₃</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.2 (7.4)</td>
<td>115</td>
<td>-8.9 (0.7)</td>
<td>3.3 (1.5)</td>
<td>-48 (3.7)</td>
<td>.77</td>
</tr>
<tr>
<td>2</td>
<td>9.1 (0.6)</td>
<td>560</td>
<td>-256 (-1.5)</td>
<td>-13 (-0.5)</td>
<td>-232 (-2.5)</td>
<td>.24</td>
</tr>
<tr>
<td>3</td>
<td>2.9 (3.0)</td>
<td>133</td>
<td>-6.7 (0.1)</td>
<td>-6.7 (-1.0)</td>
<td>-77 (-3.1)</td>
<td>.34</td>
</tr>
<tr>
<td>4</td>
<td>1.2 (9.9)</td>
<td>138</td>
<td>29 (-2.8)</td>
<td>-2.1 (-2.2)</td>
<td>-51 (-9.0)</td>
<td>.80</td>
</tr>
<tr>
<td>5</td>
<td>2.9 (5.1)</td>
<td>57</td>
<td>-8.1 (-1.3)</td>
<td>22 (7.4)</td>
<td>1 (0.1)</td>
<td>.61</td>
</tr>
<tr>
<td>6</td>
<td>3.5 (5.0)</td>
<td>278</td>
<td>-117 (-1.1)</td>
<td>-9.4 (-1.9)</td>
<td>-76 (-5.0)</td>
<td>.75</td>
</tr>
<tr>
<td>7</td>
<td>2.7 (1.9)</td>
<td>436</td>
<td>-273 (-2.3)</td>
<td>0.3 (0.1)</td>
<td>-110 (-3.2)</td>
<td>.32</td>
</tr>
<tr>
<td>8</td>
<td>1.4 (8.1)</td>
<td>77</td>
<td>3.7 (0.1)</td>
<td>4.3 (1.4)</td>
<td>-34 (-3.7)</td>
<td>.68</td>
</tr>
<tr>
<td>9</td>
<td>1.6 (2.6)</td>
<td>193</td>
<td>-89 (-1.4)</td>
<td>-1.2 (-0.3)</td>
<td>-51 (-4.2)</td>
<td>.31</td>
</tr>
<tr>
<td>10</td>
<td>1.8 (5.9)</td>
<td>145</td>
<td>-47 (-1.2)</td>
<td>0.9 (0.3)</td>
<td>-43 (-3.5)</td>
<td>.59</td>
</tr>
<tr>
<td>11</td>
<td>2.2 (2.5)</td>
<td>200</td>
<td>-92 (-0.9)</td>
<td>-0.7 (-0.1)</td>
<td>-55 (-4.2)</td>
<td>.30</td>
</tr>
<tr>
<td>12</td>
<td>2.0 (5.5)</td>
<td>176</td>
<td>-50 (-1.4)</td>
<td>-0.4 (-0.1)</td>
<td>-67 (-4.6)</td>
<td>.59</td>
</tr>
</tbody>
</table>

Note: Numbers in parentheses represent t-statistics
effective, as the benchmark analysis of the previous chapter illustrated, it may have been implemented simply as a result of the nature of the trial runs. All trial runs that the managers faced before their free-play session included a step increase in features. For a step in features aggressive early hiring was essential. The managers may have thus developed the habit of hiring early from their practice runs.

There is no presumption that the subjects used the rules presented in the equations above to calculate their decisions. However, the rules provide a good model of the heuristics that the managers used. In addition to turnover, managers used other information (such as the number of claims pending, or the adjuster experience level) to make their monthly hiring decision. The managers' decisions were not optimal. While misperceptions about the effects of feedback loops and about time lags were apparent in the teams' performance, the above regressions indicate that the managers understood, at least in part, the system incorporated in the game. The extent of their learning during the training session is not clear. It would be interesting to give the game to insurance managers who have not participated in the Claims learning laboratory, or even to outsiders, in order to compare their scores to the performances presented above. The comparison would provide clues on the effectiveness of the learning lab.
CHAPTER 9
CONCLUSIONS AND RECOMMENDATIONS

The analysis of the performance of insurance claim managers on the Claims Game provides useful information at the micro level about their decision making process. While the managers demonstrated some understanding of the dynamic model, their performance also indicates misperceptions of the feedback loops and of the time lags in the system. The managers' performance was not optimal. Furthermore, their retrospective reports suggest that they used a limited set of information for their decisions. The prospective reports indicate that players set simplified subgoals, such as reducing settlement costs or reducing the pending ratio, in order to achieve their objective. The conclusions of this thesis therefore support the concept of Bounded Rationality in decision making.

The benchmark analysis of Chapter 7 suggests that by using a simple decision rule consistently, insurance claims managers may be able to cut costs by 2.7% and improve their company's profitability by as much as 30%. The magnitude of the potential improvement in profitability supports the attempts made by Hanover Insurance to enhance managerial understanding of the dynamic system of claims management.

The success of the learning laboratory will depend on the extent of the assimilation of the System Thinking ideas in the everyday decisions of insurance claim managers. An assessment of the success of the program will be based primarily on subjective accounts of the managers involved. In their accounts, the managers would discuss the extent of the incorporation of system model ideas in their decisions. The claims game itself can also be used, however, to provide statistical information on the effectiveness of the learning laboratory.
Bounded rationality of managers' decisions

In Chapter 1 the concept of Bounded Rationality was introduced. According to the theory, decision makers adopt simplifying strategies in order to reach their goals, using a limited subset of the information available. Abstract global goals are replaced by measurable subgoals.

The data analyzed in this thesis support the Bounded Rationality Theory. The insurance managers simplify their task of minimizing cumulative costs by setting measurable subgoals, such as the reduction in the stock of claims pending, the reduction in monthly average settlement costs, or the increase in the experience level of the employed adjuster force. These subgoals are stated by the players in their prospective reports. Even though several teams succeed in meeting their subgoals, their performance is not optimal.

Retrospective reports suggest that the managers used a limited set of information to make their decisions. The apparent lack of attention on information essential to determine the state of the dynamic system, suggests that the information sets used by the managers were not necessarily the appropriate ones.

Misperceptions of feedbacks

The managers performance, as well as their prospective and retrospective reports suggest misperceptions of the feedbacks in the system. The discrepancy between the projected outcomes that some teams included in their comment sheets, to the actual results, indicate a limited understanding of the feedbacks in the system. The lack of attention on entire subsystems of the model, such as the quality subsystem of Chapter 3, also suggests a misperception of the feedbacks in the system.

Misperceptions of time delays

Some managers adopted a wait-and-see strategy in their decisions. After implementing a decision, managers would wait several months in order to assess the effects of the decision. The period of their wait, which was about 2-6 months indicates a misperception of the time lags involved in the system. Some of the adjustment times of the system are in the order of 18
to 24 months. The complete effects of a particular policy can therefore not be assessed in 2-6 months.

Potential improvements: The benchmark test

By using a simple decision rule consistently, I was able to surpass most teams' performance. The benchmark test improves the managers' mean cumulative cost by a statistically significant 2.7%. For Hanover insurance a 2.7% improvement in costs corresponds to about $35 million, or 31% of net income. The magnitude of the potential benefits is such that even small improvements in the managers mental models of the claims adjustment process and in their decisions may enhance the company's profitability.

Managerial decision pattern

My analysis indicates that managers are aggressive on quality, pay little attention to the desired productivity decision, and hire aggressively in the early months of the game. The emphasis on quality, which contradicts recent industry experience, may reflect the lessons of the two days of the learning laboratory which preceded the game trials. The infrequent use to the desired productivity decision raises questions about its usefulness in the game. Finally, the aggressive early hiring, despite the fact that it seems to be correct and effective, according to the benchmark test, may also be due to the managers' practice runs. In their practice runs, managers faced step inputs in incoming claims. Aggressive hiring is important for such a sudden increase in features. The managers may have developed the habit of aggressive early hiring in their practice runs.

Models of the managers' decisions

With the use of simple decision rules, I was able to explain fairly well the managers performance. The simple anchoring and adjustment rules work because they encompass the essential attributes of the managers decisions. However, it was difficult to identify the effect of each individual piece of information on the managers' decision. The claims adjustment model which was used in the game is a realistic one. Like other complicated real life models, the cues in the system are correlated. High correlations make interpretation of the effects of each
cue on the decision impossible. In the claims adjustment system, for example, settlement size was found to be highly correlated to the pending ratio.

The relative success of the simple decision rules in explaining the managers' decisions does indicate however that the managers did understand, at least in part, the claims adjustment system.

9.1 RECOMMENDATIONS

The analysis presented in this thesis motivates specific suggestions for the assessment and improvement of the learning laboratory, as well as for the design of the game:

1) The game should be used for an assessment of the effectiveness of the learning laboratory. As I mentioned above the success of the learning lab will depend on the degree of the assimilation of the ideas introduced in the sessions on the managers' everyday decisions. An assessment of the success of the program will therefore depend to a great extent on the subjective assessments of the managers involved. However, statistical information on the effectiveness of the laboratory can be obtained from the game itself.

The game should be introduced to a group of claims managers who have not participated in the learning laboratory, as well as to a group of outsiders. Comparison of the performance of the two groups to the performance of the group studied in this thesis may provide useful information about the effectiveness of the learning laboratory.

Specifically, it would be interesting to see whether the game performance of experienced claims adjusters who have gone through the learning laboratory betters the performance of experienced adjusters who have not gone through the program. Furthermore, comparison of the performance of both groups to a group of outsiders, would indicate how important previous experience in claims adjustment is for a good performance in the game.

2) The benchmark test may be used in the learning laboratory as an illustration of the potential benefits of improved decision making. The success of the learning laboratory depends to a great extent on the commitment of the managers involved. The benchmark test of Chapter 7 indicates that with the consistent use of a simple
decision rule cumulative costs could be improved by as much as 2.7%. A 2.7% improvement in costs at Hanover would correspond to about 31% of net income. The potential improvements from a consistent decision rule may increase the managers' interest in the program. The results of the benchmark test should therefore be presented to managers during the learning laboratory sessions.

The benchmark model could also be used to demonstrate the effects of different strategies, such as a strategy which emphasizes quality, or a strategy which focuses only on production. By varying the various parameters (cue weights) in the decision rules of the benchmark model, the effects of such different strategies can be illustrated.

3) **Suggestions for improvements in the design of the Claims Game.**

**Simulation time.** A look at the graphs displaying the performance of each team illustrates that the simulation time of 44 months is not sufficient for the system to reach a 'steady state'. Since some of the time lags in the model are as high as 18 (settlement size adjustment time) or 24 months (promotions), it seems that at least 60 months are required before all the dynamics of the system become apparent. This is particularly true if we look at the pending ratio. For most teams the pending ratio dropped off monotonically, without increasing or levelling off until the very last months of the game.

An extension of the simulation time to 60 months would be helpful. Managers may then be able to bring the system 'under control' at an effective steady state, and may thus learn more about the nature of the system. The current time span of 44 months makes this 'control' very difficult to accomplish.

**Resolution of hiring decision.** The Claims Game began with an initial number of 54 adjusters employed. Due to this initial level of adjusters, monthly hiring and turnover rates were close to one or two adjusters. These low integer numbers provided a poor resolution for the analysis of the hiring decision.

For the purpose of future analyses, the initial level of adjusters should be increased tenfold to about 500. The higher numbers will reduce the grain size imposed by the integer constraint
on the hiring decision. Observers of the hiring decision will then have a better resolution for their analysis. The higher number of adjusters is still a fairly realistic one.

**Effect of Adjuster Capacity on Settlement Size.** The designers of the model should examine whether the sharp sensitivity of settlement size to adjuster capacity (which is illustrated in section 6.4 for months 31-44 of Team 10) is realistic. When adjuster capacity is low, settlement size becomes extremely sensitive to the rate of incoming features and fluctuates sharply.
REFERENCES


[20] STELLA is a trademark of High Performance Systems, Inc. of Lyme, New Hampshire


BIBLIOGRAPHY


APPENDIX 1

EQUATIONS FOR STELLA MODEL OF THE CLAIMS GAME

In this section I include the complete listing of the equations for the STELLA model of the game, which was introduced in Chapter 3.
AvgFeatures = AvgFeatures + dt * (ChgAvgFeatures)
INIT(AvgFeatures) = Features

CumTotalCost = CumTotalCost + dt * (TotalCost)
INIT(CumTotalCost) = 0

InAdj = InAdj
INIT(InAdj) = Adjusters

InEffTime = InEffTime
INIT(InEffTime) = EffectiveTime

InitialTotal = InitialTotal
INIT(InitialTotal) = Total

NewAdjusters = NewAdjusters + dt * (Hiring - Promotions - NewTO)
INIT(NewAdjusters) = 50*(.02*24)/(.02*24+1)

NormalTimePerClaim = NormalTimePerClaim
INIT(NormalTimePerClaim) = EffectiveTime/INT(PIF*.05+.5)

Pending = Pending + dt * (-Settlements + Features)
INIT(Pending) = Features*2

PIF = PIF + dt * (ChgPIF)
INIT(PIF) = 50000

QualityStandard = QualityStandard + dt * (ChgQualStd)
INIT(QualityStandard) = 1

SrAdjusters = SrAdjusters + dt * (Promotions - SrTO)
INIT(SrAdjusters) = 50*(1/(/.02*24))/1/(/.02*24+1)

WorkIntensity = WorkIntensity + dt * (ChgIntensity)
INIT(WorkIntensity) = 1

AdjusterCapacity = (NewAdjusters*.5+SrAdjusters-NewAdjusters*.1)

Adjusters = NewAdjusters+SrAdjusters

Age = Pending/AvgFeatures

CapacityAdedcacy = EffectiveTime/EfftvTimeRequired

ChangeInOpRatio = (.7*(Total-InitialTotal)/InitialTotal)/(-.05)

ChgAvgFeatures = (Features-AvgFeatures)/3

ChgIntensity = (IndicatedIntensity-WorkIntensity)/IntensAdjTime

ChgPIF = 0

ChgQualStd = MAX(-.5*QualityStandard,MIN(QualityStandard*.2,QualityStandard*(-EffPressOnQual-1)/QualityDriftTime+(RatioActDesSettle-1)/SettleSizeAdjTime))

CostPerAdjuster = 12000

DesiredProduction = MIN(Features*DesiredProductivity,Features+Pending)

DesSettleSize = IF DesrdSettSizeTable=-1000 THEN SettleSize ELSE MAX(0,MIN(DesrdSettSizeTable,10000))

EffectiveTime = AdjusterCapacity*WorkIntensity*TimeEffctvnsFrmPen

EfftvTimeRequired = DesiredProduction*QualityStandard*NormalTimePerClaim

Expenses = Adjusters*CostPerAdjuster

ExpPerSettl = Expenses/Settlements

Features = INT(PIF*.05+0*PULSE(0,5,1000)+.5)*(1+0*NORMAL)

FractSrAdj = SrAdjusters/Adjusters

Hiring = IF HiringTable=-1000 THEN Turnover ELSE INT(MAX(-NewAdjusters*.5,MIN(}
HiringTable, Adjusters*0.1)) + 0.5)
O IntensAdjTime = 4
O NewTO = INT(NewAdjusters*0.02*EffPressOnTO + 0.5)
O PendPerAdj = Pending/Adjusters
O PendRatio = Pending/Features
O PotenQualIndex = (EffectiveTime/Settlements)/NormalTimePerClaim
O ProdRatio = Settlements/Features
O Promotions = INT(NewAdjusters/24 + 0.5)
O QualityDriftTime = 6
O RatioActDesSette = SettleSize/DesSettleSize
O Score = CumTotalCost + Pending*Total
O SettCost = Settlements*SettleSize
O SettledPerAdj = Settlements/Adjusters
O Settlements = MAX(0, MIN(INT(DesiredProduction*FractDesSettled + 0.5), Pending + Features))
O SettleSize = 2000*EffQualStdOnSettSze*EffCapOnSettSze
O SettleSizeAdjTime = 18
O SrTO = INT(SrAdjusters*0.02*EffPressOnTO + 0.5)
O t = TIME
O TimeEffctvns = (EffectiveTime/Adjusters)/(InEffTime/InAdj)
O TimePressure = EfftvTimeRequired/EffectiveTime
O Total = TotalCost/Settlements
O TotalCost = SettCost + Expenses
O Turnover = SrTO + NewTO
\[\text{DesiredProductivity} = \text{graph}(t)\]
\[= (0.0, 1.0000), (4.8000, 1.0000), (9.6000, 1.0000), (14.4000, 1.0000), (19.2000, 1.0000),
(24.0000, 1.0000), (28.8000, 1.0000), (33.6000, 1.0000), (38.4000, 1.0000),
(43.2000, 1.0000), (48.0000, 1.0000)\]
\[\text{DesrdSettSizeTable} = \text{graph}(t)\]
\[= (0.0, -1000.0000), (4.0000, -1000.0000), (8.0000, -1000.0000), (12.0000, -1000.0000),
(16.0000, -1000.0000), (20.0000, -1000.0000), (24.0000, -1000.0000), (28.0000, -1000.0000),
(32.0000, -1000.0000), (36.0000, -1000.0000), (40.0000, -1000.0000),
(44.0000, -1000.0000), (48.0000, -1000.0000)\]
\[\text{EffCapOnSettSze} = \text{graph}(\text{CapacityAdequacy})\]
\[= (0.0, 2.0000), (0.1000, 1.7250), (0.2000, 1.5500), (0.3000, 1.4050), (0.4000, 1.3000),
(0.5000, 1.2250), (0.6000, 1.1600), (0.7000, 1.1000), (0.8000, 1.0600), (0.9000, 1.0300),
(1.0000, 1.0000)\]
\[\text{EffPressOnQual} = \text{graph}(\text{TimePressure})\]
\[= (0.0, 1.1000), (0.2000, 1.1000), (0.4000, 1.0900), (0.6000, 1.0800), (0.8000, 1.0500),
(1.0000, 1.0000), (1.2000, 0.8700), (1.4000, 0.7300), (1.6000, 0.6000), (1.8000, 0.5300),
(2.0000, 0.5000)\]
\[\text{EffPressOnTO} = \text{graph}(\text{TimePressure})\]
\[= (0.0, 1.4000), (0.2000, 1.2500), (0.4000, 0.9250), (0.6000, 0.8125), (0.8000, 0.8125),
(1.0000, 1.0000), (1.2000, 1.4000), (1.4000, 1.7250), (1.6000, 1.9375), (1.8000, 2.0500),
(2.0000, 2.0875)\]
EffQualStdOnSettSze = \text{graph}(\text{QualityStandard})
(0.0, 2.0000), (1.0000, 1.0000), (2.0000, 0.7000), (3.0000, 0.7000), (4.0000, 0.7800),
(5.0000, 0.9400), (6.0000, 1.0800), (7.0000, 1.1900), (8.0000, 1.2900), (9.0000, 1.3700),
(10.0000, 1.4000)

FractDesSettled = \text{graph}(\text{CapacityAdequacy})
(0.0, 0.0), (0.2000, 0.3300), (0.4000, 0.5800), (0.6000, 0.7700), (0.8000, 0.9000),
(1.0000, 1.0000), (1.2000, 1.0700), (1.4000, 1.1200), (1.6000, 1.1400), (1.8000, 1.1500),
(2.0000, 1.1500)

HiringTable = \text{graph}(t)
(0.0, -1000.000), (5.0000, -1000.000), (10.0000, -1000.000), (15.0000, -1000.000),
(20.0000, -1000.000), (25.0000, -1000.000), (30.0000, -1000.000), (35.0000, -1000.000),
(40.0000, -1000.000), (45.0000, -1000.000), (50.0000, -1000.000)

IndicatedIntensity = \text{graph}(\text{TimePressure})
0.0 \rightarrow 0.0
0.2000 \rightarrow 0.2700
0.4000 \rightarrow 0.5100
0.6000 \rightarrow 0.7100
0.8000 \rightarrow 0.8700
1.0000 \rightarrow 1.0000
1.2000 \rightarrow 1.0800
1.4000 \rightarrow 1.1500
1.6000 \rightarrow 1.1800
1.8000 \rightarrow 1.2000
2.0000 \rightarrow 1.2000

TimeEffectvnessFrmPen = \text{graph}(\text{Age})
0.0 \rightarrow 0.0100
0.5000 \rightarrow 0.6700
1.0000 \rightarrow 1.0400
1.5000 \rightarrow 1.2000
2.0000 \rightarrow 1.0000
2.5000 \rightarrow 0.7400
3.0000 \rightarrow 0.5400
3.5000 \rightarrow 0.4600
4.0000 \rightarrow 0.3900
4.5000 \rightarrow 0.3500
5.0000 \rightarrow 0.3300
<table>
<thead>
<tr>
<th>DesiredProductivity</th>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>4.8000</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>9.6000</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>14.4000</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>19.2000</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>24.0000</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>28.8000</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>33.6000</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>38.4000</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>43.2000</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>48.0000</td>
<td>1.0000</td>
<td></td>
</tr>
</tbody>
</table>

Figure A1: STELLA model table
Figure A2: STELLA model table
Figure A3: STELLA model table
Figure A4: STELLA model table
Figure A5: STELLA model table
Figure A6: STELLA model table
Figure A7: STELLA model table
Figure A8: STELLA model table
Figure A9: STELLA model table
Figure A10: STELLA model table
APPENDIX 2

BENCHMARK MODEL SIMULATION RUNS

In this section I include the results of the STELLA simulation run, using the benchmark decision model of Chapter 7.
Figure A11: Benchmark model simulation run
Figure A12: Benchmark model simulation run
Figure A13: Benchmark model simulation run
Figure A14: Benchmark model simulation run
Figure A15: Benchmark model simulation run
Figure A16: Benchmark model simulation run
Figure A17: Benchmark model simulation run
Figure A18: Benchmark model simulation run
Figure A19: Benchmark model simulation run
Figure A20: Benchmark model simulation run
Figure A21: Benchmark model simulation run
ACKNOWLEDGEMENT

As Goscinny and Uderzo, creators of Asterix would have said, the completion of my studies at MIT would not have been possible without the following:

14 liters of inc.
62 pencils.
38 kilos of paper.
6,500 hours of school work.
Three pairs of soccer cleats.
Several hundred telephone calls to Greece.

The love and support (moral and financial) of my family.

The inspiration and guidance of Prof. Elias Gytopoulos and his family, who have provided me throughout these years with a loving home away from home.

The countless extended lunch breaks, evenings in front of the TV set (the idiot box), and weekend beach or ski trips, with the Ossipee Club at TUFTS, the ATO Slugs at 405 Memorial Drive, and the Greek Mafia at MIT.

At the Sloan School, I would like to thank Prof. Peter Senge and Dan Kim of the System Dynamics Group for their help. Finally, I am grateful to my thesis supervisor, Prof. John Sterman, for his insightful advice and direction throughout this project.