# Casualty Actuarial Society E-Forum, Fall 2014-Volume 1

# The CAS *E-Forum*, Fall 2014-Volume 1

The Fall 2014-Volume 1 Edition of the CAS *E-Forum* is a cooperative effort between the CAS *E-Forum* Committee and various other CAS committees, task forces, or working parties.

This *E-Forum* contains a report of the CAS Automated Vehicles Task Force and four independent research papers.

## CAS Automated Vehicles Task Force

# CAS *E-Forum*, Fall 2014-Volume 1

## Table of Contents

# *E-Forum* Committee



For information on submitting a paper to the *E-Forum*, visit http://www.casact.org/pubs/forum/.

# Restating the National Highway Transportation Safety Administration's National Motor Vehicle Crash Causation Survey for Automated Vehicles

Casualty Actuarial Society Automated Vehicles Task Force

## EXECUTIVE SUMMARY

The National Highway Transportation Safety Administration (NHTSA) concluded its National Motor Vehicle Crash Causation Survey (NMVCCS) in 2008. The NMVCCS analyzed the events leading up to a motor vehicle crash to determine what was causing automobile accidents. This study, which found that 93% of accidents are caused by human error, is often referenced to justify and quantify automated vehicles' accident reduction potential. However, this study was never intended to be applied to automated vehicles.

Currently celebrating its 100th year, the Casualty Actuarial Society fulfills its mission to advance actuarial science through a singular focus on research and education for property/casualty actuarial practice. Among its 6,200 members are experts in property-casualty insurance, reinsurance, finance, risk management, and enterprise risk management. The Casualty Actuarial Society has created an Automated Vehicles Task Force (CAS AVTF) to research the technology's risks and their implications for insurance and risk management. To this end, the Task Force has re-evaluated the NMVCCS in the context of an automated vehicle world. It found that 49% of accidents contain at least one limiting factor that could disable the technology or reduce its effectiveness. The safety of automated vehicles should not be determined by today's standards; things that cause accidents today may or may not cause accidents in an automated vehicle era. Rather, things like the vehicle's failure rate (after accounting for any fail-safes, infrastructure investments, and driver interactions) and unavoidable accidents (e.g., falling rocks) should be the gauge by which they should be measured. Safety metrics should also consider additional criteria that would not be part of today's standards and safety concerns, as automation introduces additional risks to consider.

This report details the CAS AVTF's re-evaluation of the NMVCCS and notes areas for future research.

## 1. INTRODUCTION

John Capp, director of electrical and control systems at General Motors R&D recently stated, "Someone has to get this story straight—that it's going to take a long time before we see the true autonomous vehicle. There's so much to do before we can tell a customer as he leaves the dealer's lot, 'Just close your eyes. It's good to go.[1]'" The problem with Mr. Capp's statement is that no clear safety benchmark has been or is being established. Company XYZ may conclude that their product is safe for consumers while Company ABC may believe more testing is required. This may be because Company XYZ has more advanced technology than ABC, has performed different tests, or has a different safety standard. If the technology is safe, a delay in its implementation can result in accidents, including fatal ones, which could have been avoided with the technology. If it's unsafe, its introduction not only puts lives at risk, it also risks delaying future technological advancements.

The absence of any clearly established safety benchmark has led interested parties to rely on the National Highway Transportation Safety Administration (NHTSA) National Motor Vehicle Crash Causation Survey's (NMVCCS) conclusions. The NMVCCS found that 93% of accidents are caused by human error. Publications such as *The New York Times* and many of the witnesses testifying before the U.S. Senate and House have suggested that automated vehicles have the potential to reduce all of these accidents. However, the NMVCCS's conclusions have no bearing on automated vehicles' potential. The study analyzed the risks of today's driving environment. A future of automated vehicles will look much different and involve different risks. As drivers rely more on the vehicle's technology and less on themselves, the accident causation variables will change. An accurate baseline requires a new analysis that looks at these variables through a lens of automated vehicle performance.

The Casualty Actuarial Society's Automated Vehicles Task Force (CAS AVTF) has re-evaluated the NMVCCS data through just such a lens. The new benchmark illuminates and quantifies a broad array of risks that need to be overcome before the technology can reach its potential; however, this benchmark is merely the first step in the creation of a process that will ensure the product's safety. The quantification of the risks allows stakeholders to identify and prioritize areas of concern. A more robust analysis also allows them to make more accurate cost-benefit decisions, thereby optimizing their investments. Ultimately, increasing the risks' transparency also increases Company XYZ's ability to prove the technology's safety and benefits. In turn, this can inform decisions on how to support the development of a revolutionary product while balancing safety and innovation concerns.

[1] http://www.nempa.org/events/2014/engineering-safer-drivers/

## 2. BACKGROUND AND OBJECTIVE

For many people, getting into a motor vehicle is their greatest everyday risk. Each year, automobile accidents result in 50 million injuries and 1.2 million deaths worldwide. Automated vehicles have the potential to make our roads dramatically safer. However, as the technology reduces some risks, others may be increased and new risks may be introduced. The Casualty Actuarial Society has formed an Automated Vehicles Task Force to study these risks and identify opportunities for stakeholders and the insurance industry to improve the product.

### Automated Vehicle Background

Advancements over the past decade have brought autonomous cars out of the realm of science fiction and onto our public roads. Google, Audi, and Continental have received licenses in Nevada to test their automated vehicles on public roadways. Volvo is conducting road tests in Sweden and Nissan is doing the same in Japan. Mercedes Benz has tested its technology in Germany. Great Britain will allow automated vehicles to be tested on public roads in January 2015.

Regulators have tried to keep pace with the technological advancements. In the United States, California, Florida, Michigan, and Washington D.C. have passed bills regarding automated vehicle testing. The NHTSA has issued a preliminary policy statement on automated vehicles. As the technology advances, there will be further legislative and regulatory developments.

### National Motor Vehicle Crash Causation Survey

The U.S. Congress authorized the NHTSA to conduct a National Motor Vehicle Crash Causation Survey to understand the events leading up to a motor vehicle crash. In 2008, the NHTSA reported its findings to Congress. These findings were specific to the current, non-autonomous transportation environment. There are a number of reasons why these conclusions are not applicable to automated vehicles.

First, the NHTSA did not collect information on all accidents. Minor accidents and accidents occurring between 12:00 a.m. and 5:59 a.m. were not analyzed; care needs to be taken when extrapolating the results to a set of accidents that were not analyzed. Second, automobiles have undergone a number of changes since the evaluation period. Approximately 40% of vehicles on today's roads have electronic stability control, up from 10% in 2005 when the study began. Forward collision avoidance systems, like Volvo's autonomous braking system City Safe, have become much more prevalent in recent years. Third, the study left out variables important to determining the success of automated vehicles. More specifically, the location was not tracked in the dataset; therefore, we cannot develop as granular or as accurate a baseline as will be required to make actual cost-benefit decisions. For example, a countrywide estimate of snow's impact on accidents is not applicable to San Diego. Lastly, automated vehicles represent a fundamental change in transportation risk: one in which the driver depends on his car rather than himself. This requires a complete paradigm shift in the way the data is analyzed.

### Goal

There is a need for a clear, accurate, and applicable baseline to estimate the technology's potential benefits, and more importantly, what actions can help the technology reach its potential. The CAS AVTF has chosen to use the NMVCCS to establish this baseline for two reasons. First, the NMVCCS contains over 600 data elements from 5,471 accidents. This allows us to analyze a wide range of causation variables. Second, it allows us to demonstrate the impact from simply changing the study's focus. Ultimately, it represents merely the first step on the path towards creating the optimal testing approach and risk management structure.

## 3. SCOPE

The NHTSA collected over 600 data elements on 6,950 crashes from January 1, 2005, through December 31, 2007. The Report to Congress used the data on the 5,471 crashes that occurred over the two-and-a-half year period, July 3, 2005, through December 31, 2007. The NHTSA collected information on accidents with the following characteristics:

- The crash must have resulted in a harmful event associated with a vehicle in transport on a traffic way.
- Emergency medical services must have been dispatched to the crash scene.
- At least one of the first three crash-involved vehicles must be present at the crash scene when the NMVCCS researcher arrives.
- The police must be present at the scene of the crash when the NMVCCS researcher arrives.
- At least one of the first three vehicles involved in the crash must be a light passenger vehicle that was towed or will be towed due to damage.
- A completed police accident report for this crash must be available.

The Report to Congress further states: "To make the NMVCCS sample representative of all similar types of crashes for the whole of the United States each of the 5,471 investigated crashes has been assigned a certain weight based on the sample design used in this survey." The recalculation of these weights falls outside the scope of this project. Therefore, the report will show both the unweighted (observed) and weighted (extrapolated) frequencies.

The CAS AVTF also restricted its analysis to these accidents. Since this report is interested in quantifying potential risks to automated vehicles' accident reduction, the Task Force focused on identifying instances in which the technology would be inoperable, disengaged, or involve a risky behavior that requires additional research. In order to determine whether the accident would be prevented, the Task Force asked two questions: would the technology have been operable and would the technology have been operated safely? Specifically, the following variables were identified as potential hurdles:

| Variable | Dataset | Description |
|---|---|---|
| **Traffic way** | ENV | Used to determine when the technology would be disabled due to environmental conditions<br>• Values of Roadway Immersed, Heavy Snow, Heavy Rain, and Dust Storm assumed to disable technology |
| **Weather** | ENV | Used to determine when the technology would be disabled due to environmental conditions<br>• Values of Snow, Rain, Sleet, Blowing Snow assumed to disable technology |
| **Vehicle Condition** | PCAEXT | Used as a proxy for vehicle failures. If the vehicle or technology fails and the driver is not properly engaged, the accident could still occur. The frequency of errors is likely to increase as the vehicle shoulders more of the driving responsibility. |
| **Traffic Control Device** | TCD | Tells us if a traffic control device (TCD) was present and operable at the accident.<br>• If the TCD was inoperable, the automated vehicle technology may not have been able to correctly interpret its environment to prevent the accident. |
| **Alcohol** | PCAEXT | Used to determine if alcohol was present in any of the drivers in the accident |
| **Drug** | DRUGS | Used to determine what other drugs were present in any of the drivers in the accident |
| **Critical Event** | PCA | NHTSA's variable, identifying what they believed to be the critical reason for the accident<br>• Values of Sleeping, Heart Attack/Other Physical Impairment, External Distraction, Internal Distraction, and Inattention were identified as potential risks for automated vehicle use |

The new baseline is limited to the data elements collected in the NVMCCS. However, additional variables will be required to calculate the true baseline. A few of the variables that will be needed are:

**Location:** As noted above, location is a key variable for establishing a reasonable cost-benefit analysis. It is likely that separate baselines will be needed for city driving, highway driving, and country driving. Cities with different environmental risks will also have different cost-benefit relationships to consider.

**Animal accidents:** The majority of the 1.2 million annual animal hits occur in the Midwest. The dataset does not contain a variable detailed enough to break out the unavoidable animal hits from the avoidable ones. The lack of a location identifier also reduces the value of any observed variable.

Volvo is reportedly working on advancing its automated braking system to recognize and stop for deer, which may further reduce the risk.

**Other risks:** The engineering and coding risks are obviously one of the largest hurdles that need to be overcome. The process risk involved in the tests—the risk that the tests differ from reality—should also be quantified and adjusted for.

### Differences from NMVCCS

The CAS AVTF's accident level evaluation differs slightly from the NMVCCS for two reasons:

- The data file used to produce the statistics in the Report to Congress was compiled at April 30, 2008. The final file available at the NMVCCS website is a later version (compiled on October 28, 2008) and is slightly different from the data file used for the report to Congress.
- The CAS AVTF's focus caused some accidents to be categorized differently. For example, if any vehicle in the accident had a pre-identified risk feature then the entire accident was segmented into that risk bucket.

## 4. RESULTS

The results have been divided into two sections, technological issues and behavioral issues. These groups of risks are then further segmented by the actual risk:

- Technological issues: Weather, Vehicle Condition/Error, Inoperable Traffic Control Devices
- Behavioral issues: Aggressive Driving/Driver Disables, Alcohol & Illicit drugs, Sleeping, Physical Issues/Heart Attack, Distraction

### 4.1 Technological Issues

Technological issues are ones where the technology may be inoperable or may inaccurately interpret the environment. Therefore, the technology's presence would not have prevented these accidents.

### 4.1a Weather

Based on publicly available information at the time of this writing, automated vehicles are inoperable in inclement weather. Therefore, no accident could be prevented if it occurred in such weather, regardless of the crash's actual cause. "Inoperable weather" is defined as snow, sleet, rain, or blowing snow as well as times when the road was immersed in water, heavy snow, heavy rain, or a dust storm.

| Weather | Unwtd Freq | Wtd Freq | Unwtd Freq | Wtd Freq |
|---|---|---|---|---|
| No weather issues | 4,868 | 1,921,312 | 89.0% | 87.8% |
| Inoperable weather | 602 | 267,657 | 11.0% | 12.2% |
| **Grand Total** | **5,470** | **2,188,970** | **100%** | **100%** |

Accounting for weather in the analysis reduces the number of preventable accidents by over 12%. However, the value to overcoming this hurdle varies dramatically depending where you live. For example, San Diego and San Francisco average less than 65 days of precipitation a year. However, Seattle averages approximately 150 days of precipitation a year.

Automakers and technology companies may develop an in-vehicle solution to inclement weather, reducing the vehicles' risk and increasing their benefits. In the meantime, a number of tests are being conducted that may offer alternative ways to overcome this hurdle. Volvo's experiments with road magnets indicate that certain infrastructure investments might neutralize this risk. The University of Michigan is building a fake city in Ann Arbor to test automated vehicles. These tests may yield additional insights to help overcome the inclement weather issue.

### 4.1b Vehicle Conditions

NHTSA tracked if a vehicle error or deficiency occurred. For example, NHTSA noted if there was an issue with the braking system even if it was not the main cause of the crash. While some of these conditions may be easily overcome, others could be dramatically worse as the technology becomes more advanced and human involvement decreases. In a fully autonomous world, the driver is free to read, text, or even sleep. Currently, an engaged driver may be able to overcome a braking system error or a blown tire. However, in the automated vehicle world, the driver may not be sufficiently engaged to overcome such issues. These accidents will require additional research to ensure that the automated vehicle will be able to interpret the incident and make any necessary adjustments to avoid a crash. Rather than make any assumption about the unknown accidents, we will exclude all of the accidents where we do not know whether the technology could have prevented the accident.

| Vehicle Condition Present | Unwtd Freq | Wtd Freq | Unwtd Freq | Wtd Freq |
|---|---|---|---|---|
| Not Present | 4,351 | 1,769,134 | 79.5% | 80.8% |
| Present* | 681 | 254,948 | 12.4% | 11.6%* |
| Unknown if present | 438 | 164,888 | 8.0% | 7.5% |
| **Grand Total*** | **5,470** | **2,188,970** | **100%** | **100%** |

*Removing the Unknowns from the subset decreases the total number of accidents from 5,470 to 5,032 and increases the percentage of accidents with a vehicle condition present to 12.6%.

### 4.1c Traffic Control Devices (TCD)

For a small number of accidents, the traffic control device operating the intersection was not working properly. Automated vehicle technology communicates with the traffic control devices to determine the correct action. In these scenarios, it is unknown whether the technology will correctly interpret the inoperable TCD and respond correctly.

| Traffic Signal | Unwtd Freq | Wtd Freq | Unwtd Freq | Wtd Freq |
|---|---|---|---|---|
| Not operating properly | 22 | 7,933 | 0.4% | 0.4% |
| Not present/Operating properly | 5,447 | 2,180,736 | 99.6% | 99.6% |
| Unknown | 1 | 300 | 0.0% | 0.0% |
| **Grand Total** | **5,470** | **2,188,970** | **100%** | **100%** |

These numbers, though tiny, may underestimate the impact inoperable traffic control devices will have on automated vehicles. First, an "inoperable traffic control device" may be defined differently for an automated vehicle than for a human driver. Google's automated vehicles use extremely detailed maps to navigate safely. If a TCD is not at the regulated height or is moved slightly, it may now be "inoperable" from a technological perspective while it works just fine for a human driver. Second, if humans do a better job of identifying and adjusting to inoperable TCDs, then NHTSA's numbers will underweight these accidents in the future state.

The University of Michigan's study might also provide more insights into this risk. Maintaining the roadways and the maps might decrease the amount of technology the vehicles require while also increasing their safety.

### 4.1d Summary of Technology Issues

If automated vehicle technology cannot overcome the weather, vehicle errors, and inoperable traffic control devices, it will only be able to address 78% of the accidents on the roads. While human error accounts for an even greater part of these remaining crashes, it means that the upper bound is limited to 78% accident reduction and not the 93% that is often stated. The difference between a 78% accident reduction and a 93% accident reduction is 830,000 accidents or $45 billion, using the Eno Center for Transportation's accident and cost estimates.[2]

Note: the total does not match the sum of the pieces. This is because the same accident could have more than one disqualifying identifier on it. An accident that occurs in inoperable weather and has a vehicle issue present will be counted in each variable's numbers, but, as it is a single accident, it will only be counted once when it is rolled up to "Total AV Inoperable Accidents."

[2] Eno Center assumes 5.5 million crashes per year with an average economic cost of $55K per crash.

| Disabling Factor | UnWtd Freq | Wtd Freq | UnWtd Freq | Wtd Freq |
|---|---|---|---|---|
| Inoperable Weather | 602 | 267,657 | 11.0% | 12.2% |
| Vehicle Issue Present | 681 | 254,948 | 12.4% | 11.6% |
| TCD Not Operating Properly | 22 | 7,933 | 0.4% | 0.4% |
| **AV Inoperable Accidents** | **1,183** | **466,269** | **21.6%** | **21.3%** |
| **All Accidents*** | **5,470** | **2,188,970** | *100%* | *100%* |

*Removing the accidents where Vehicle Issues and TCD Operability are unknown decreases the subset of accidents from 5,470 to 5,031. It also decreases the UnWtd (unweighted) Frequency of Inoperable Weather accidents to 543 and TCDs to 20. AVs are inoperable in 1,183 of UnWtd accidents or 22.1% of Wtd accidents.

## 4.2 Behavioral (Driver) Issues

Simply because the technology works does not mean it will be used or used effectively. Seat belts have not only been proven to reduce the risk of severe to fatal crashes by up to 50%, their use is also required by law in most states. In spite of seat belts' availability, safety, and legality, one in seven adults still refuses to buckle up.

While it is impossible to perfectly predict how people will react to and use automated vehicles, it is possible to identify risky scenarios and develop policies to minimize that risk. For example, automated vehicles that do not speed may not only encourage their drivers to disengage the system but may also be a risk to other drivers on the road (e.g., Chicago's highways, where the speed limit is 55mph but the average non-congestion speed is typically closer to 70 mph-80 mph). It remains to be seen how this may be addressed in practice: for example, via "automated vehicle only" lanes, allowing automated vehicles to speed, or removing the driver from the loop entirely (e.g., Google's new automated vehicle). Different solutions could be developed in different environments depending on the relative risks. For example, drivers may be less inclined to take control back from a law-abiding automated vehicle while in the city versus one that is on a highway.

While Google's approach appears to eliminate this risk by removing the driver from the equation, its solution does not prevent other automakers from producing a system that allows or requires the individual to maintain some control over the driving function. Using NHTSA's NMVCCS, we can identify scenarios where the driver's involvement may prevent the technology from eliminating the accident.

### 4.2a Aggressive Driving

With approximately 41 million people receiving speeding tickets each year, drivers regularly prioritize speed over safety. While a fleet of automated vehicles may create a more efficient transportation system, a law-abiding automated vehicle's restrictions may encourage the driver to take over in situations in which speed is the main concern.

The NMVCCS report breaks out accidents where aggressive driving was present and lists the reason for the aggressive driving. The reasons included the following: anger, frustration, always drive this way, other, in a hurry/late, fleeing, and racing. Of these, it seems plausible that drivers who were in a hurry, fleeing, or racing would not engage an automated vehicle that follows the speed limit. Drivers who "always drive this way" are also problematic, as they could believe that they will drive better than the technology or that their driving will lead to a better result. If it seems unlikely that anyone will believe that, remember that 64% of Americans believe they are "excellent" or "very good" drivers.

The 3.1% estimate below may be too conservative as some of the "other" aggressive driving actions could also lead a driver to disengage the system. Removing these 808 accidents from the subset slightly increases the estimate to 3.6%. While this may not seem like a large number, with 5.5 million crashes occurring each year, it represents approximately 170,000-200,000 accidents and $9.3 billion to $10.8 billion in economic costs, using the Eno Center's estimates.

| Aggressive Driving Reason | UnWtd Freq | Wtd Freq | UnWtd Freq | Wtd Freq |
|---|---|---|---|---|
| Always drive this way | 95 | 52,155 | 1.7% | 2.4% |
| Racing | 12 | 2,472 | 0.2% | 0.1% |
| Fleeing | 26 | 6,666 | 0.5% | 0.3% |
| In a hurry | 27 | 7,254 | 0.5% | 0.3% |
| **Total – Driver Disables*** | **152** | **67,304** | **2.8%** | **3.1%** |
| **All Accidents** | **5,470** | **2,188,970** | | |

*Note, the total does not equal the sum of the pieces as more than one aggressive driving may be given per accident.

### 4.2b-f Driver Engagement Issues

In the event the technology follows NHTSA's levels of automation, from level 0 (no automation) to level 4 (fully autonomous car),[3] the driver will remain an integral part of the equation throughout much of its development. The driver could still be an integral piece to the driving solution in a level 4 vehicle if BMW's or Mercedes Benz's vision, which will always allow the individual take over the driving, is implemented. Therefore, having a sober and engaged driver who is able to take over control when needed remains an important piece to the accident reduction equation. The driver may not be able to perform this task if he has been drinking or was already distracted.

[3] See Appendix A for a more detailed explanation of NHTSA's levels of automation.

### 4.2b Alcohol

Alcohol was found to be present in approximately 8% of all accidents. However, the risk increases dramatically at night, where it was present in approximately one out of five accidents. We have assumed that if alcohol wasn't tested for (an unknown accident) then it wasn't involved. Therefore, we group these accidents with the "no alcohol" accidents.

| Alcohol Present | Unwtd Freq | Wtd Freq | Unwtd | Wtd Freq |
|---|---|---|---|---|
| No/Unknown | 5,115 | 2,016,577 | 93.5% | 92.1% |
| Yes | 355 | 172,393 | 6.5% | 7.9% |
| **Grand Total** | **5,470** | **2,188,970** | **100%** | **100%** |

Partially automated vehicles may decrease the risk of drunk drivers by transferring some of the driving responsibility from the impaired driver to the unimpaired system. However, it may also increase the risk in two key ways. First, it may increase the pass-off risk of a partially automated vehicle. An inebriated driver may not be as capable of taking over the driving task as a sober person. Second, it may increase the incidence of alcohol in the driver's seat. An individual may be less likely to have a designated driver or call a cab if he believes he will not have to drive the automated vehicle. While this may decrease the accident frequency of each drinking and driving occurrence, it may expand the subset of inebriated drivers on the road.

The chart below shows how much the presence of alcohol changes based on the time of day. The large discrepancy between alcohol presence in daytime accidents and nighttime accidents calls into question extrapolating our results to accidents occurring between 12:00 a.m. - 5:59 a.m., which are not included in the NMVCCS dataset, as noted above.

| Day/Time | Alcohol present | Unwtd Freq | Wtd Freq | Unwtd | Wtd Freq |
|---|---|---|---|---|---|
| **Daytime** | no/unknown | 3,394 | 1,301,946 | 97% | 97% |
| | Yes | 95 | 44,723 | 3% | 3% |
| **Weekday Night*** | no/unknown | 557 | 237,227 | 82% | 79% |
| | Yes | 120 | 61,253 | 18% | 21% |
| **Weekend Day** | no/unknown | 1,059 | 426,213 | 90% | 88% |
| | Yes | 113 | 56,011 | 10% | 12% |
| **Weekend Night**** | no/unknown | 105 | 51,191 | 80% | 83% |
| | Yes | 27 | 10,405 | 20% | 17% |
| **Grand Total** | | **5,470** | **2,188,970** | | |

***Weekday night: 6:30 p.m. – 11:59 p.m.**
**Weekend night: 9:00 p.m. – 11:59 p.m.

### 4.2c Illicit drugs

In addition to alcohol, NHTSA tracked if the driver had taken any other drugs. The types of drugs ranged from ones as innocuous as Lipitor to illegal drugs like cocaine. For a number of legal drugs, like Nyquil, it is recommended not to drive while taking them. We have broken out the illegal drugs from the drugs that may cause drowsiness to estimate a range of impacts.

| Drug Use | UnWtd Freq | Wtd Freq | UnWtd Freq | Wtd Freq |
|---|---|---|---|---|
| Illegal | 103 | 45,132 | 1.9% | 2.1% |
| Drowsy | 78 | 44,788 | 1.4% | 2.0% |
| Drowsy or Illegal | 181 | 89,920 | 3.3% | 4.1% |
| Alcohol or Illegal* | 431 | 205,920 | 7.9% | 9.4% |
| Alcohol, Illegal or Drowsy* | 502 | 241,596 | 9.2% | 11.0% |
| **All Accidents** | **5,470** | **2,188,970** | *100.0%* | *100.0%* |

*The total won't equal the sum of the pieces as more than one drug may be present in each accident.

The following drugs were defined to be "drowsy" and "illegal."

| Drowsy | | |
|---|---|---|
| Amphetamine | Codeine | Opiate |
| Baclofen | Morphine | Oxycodone |
| Barbiturates | Nyquil | Percocet |

| Illegal | |
|---|---|
| Cocaine | Marijuana |
| Crack cocaine | PCP |
| Heroin | |

### 4.2d Heart attack or other physical impairment of the ability to act

A driver who is suffering a heart attack or other physical impairment is similarly unable to take control of a partially automated vehicle. Observing that 2% of accidents are caused when some sort of physical impairment, such as a heart attack, inhibits the driver's ability to effectively control his vehicle may suggest some important risk management controls. Depending on the trip and the automated vehicle's response, the technology could produce either a better or worse result for the car's passenger. If the heart attack causes only a minor accident or forces the driver to pull off the road, having the driver in control may actually save his life. In these instances, the driver may get medical attention more quickly than if he were to continue all the way to his destination, assuming, of course, that his destination is not a hospital.

On the other hand, an automated vehicle will likely provide better protection to the other drivers. Additionally, if the driver's destination is close at hand, delivering the driver safely to his destination may allow for medical attention to be delivered in a much safer way than causing a minor accident or forcing the driver to pull over.

| Critical Reason | Unwtd Freq | Wtd Freq | Unwtd Freq | Wtd Freq |
|---|---|---|---|---|
| Heart attack or other physical impairment | 138 | 49,868 | 2.5% | 2.3% |
| Other | 5,332 | 2,139,101 | 97.5% | 97.7% |
| **Grand Total** | **5,470** | **2,188,970** | **100%** | **100%** |

### 4.2e Sleeping

A sleeping driver was the cause of almost 3% of the accidents studied. Disengaging the driver further may increase the frequency of this occurrence. Volvo's Driver Alert system could be used to minimize this risk, but only if the driver is given enough time to become alert and fully engaged.

| Critical Reason | Unwtd Freq | Wtd Freq | Unwtd Freq | Wtd Freq |
|---|---|---|---|---|
| Sleeping, that is, actually asleep | 159 | 62,974 | 2.9% | 2.9% |
| Other | 5,311 | 2,125,996 | 97.1% | 97.1% |
| **Grand Total** | **5,470** | **2,188,970** | **100%** | **100%** |

### 4.2f Driver distraction

The ability to divert your attention away from the task of driving to something more interesting or productive is one of automated vehicles' key attractions. However, it can also compound the driver inattention problem if the driver needs to become an active participant at random times through the vehicle's trip. Almost 17% of all accidents are caused by distractions or driver inattention. Removing the times when alcohol or illicit drugs are involved decreases this to 15.3% of all accidents. For partially automated vehicles, where the driver is required to remain part of the driving loop, these are the accidents where a successful pass-off is paramount to the technology's success.

| Critical reason | Unwtd | Wtd | Unwtd Freq | Wtd Freq |
|---|---|---|---|---|
| External distraction | 235 | 75,917 | 4.3% | 3.5% |
| Inattention | 217 | 73,059 | 4.0% | 3.3% |
| Internal distraction | 477 | 216,460 | 8.7% | 9.9% |
| **Total** | **929** | **365,436** | **17.0%** | **16.7%** |
| Total Excluding Alcohol and Drugs | 866 | 334,314 | 15.8% | 15.3% |
| **Total** | **5,470** | **2,188,970** | **100%** | **100%** |

Excluding alcohol and drugs from the subset of distracted drivers allows us to focus on its prevalence among sober individuals. Over 15% of accidents are caused by distraction even though the driver is not under any influence of alcohol or drugs.

### 4.2g Summary of Behavioral Issues

Initial research into today's technology supports the claim that continued technological advancements will reduce the impact of human error. The Highway Loss Data Institute (HLDI) found Volvo's autonomous braking system, City Safe, to reduce the accident risk and insurance claims by 20%.

However, there is a large difference between today's technology, which supports the driver and leaves the driver in full control of the vehicle, and Level 3 and Level 4 automated vehicles, which shift the burden of driving and decision making from the driver to the vehicle. Part of automated vehicles' potential value comes from their predictability. Vehicles are able to travel very closely together at high speeds, enabling them to increase highway capacity and fuel efficiency. However, the more involved the driver is, the less predictable the driving becomes. In our current environment, over 30% of accidents involve a behavioral characteristic that may cause the automated vehicle to be used incorrectly.

| Disabling Factor | UnWtd Freq | Wtd Freq | UnWtd Freq | Wtd Freq |
|---|---|---|---|---|
| Driver disables* | 152 | 67,304 | 2.8% | 3.1% |
| Alcohol or Illicit Drugs | 502 | 241,596 | 9.2% | 11.0% |
| Heart Attack/Physical Impairment | 138 | 49,868 | 2.5% | 2.3% |
| Sleeping | 159 | 62,974 | 2.9% | 2.9% |
| Distraction | 929 | 365,436 | 17.0% | 16.7% |
| **AV Usage Questioned**** | **1,742** | **709,153** | **31.8%** | **32.4%** |
| **All Accidents***** | **5,470** | **2,188,970** | *100%* | *100%* |

* Driver is in a hurry/late, fleeing, racing, or "always drives aggressively"
** The sum of the pieces won't equal the total as an accident could involve many of the risk characteristics.
***If we remove the accidents where the aggressive driving reason was unknown, the AV Usage Questioned Totals increase to 34.3% of accidents.

### 4.3 Overall Results

Contrary to statements such as those made in *The New York Times*, *Forbes*, *The Economist*, and even those made by professionals in front of the U.S. Senate and House, the NMVCCS's results do not conclusively determine the number of accidents automated vehicles will eliminate. The NMVCCS data can, however, provide insight into the risks this technology faces and hurdles that must be overcome before it can reach its full safety potential.

Restating the NVMCCS allows us to identify and quantify potential risks that could limit automated vehicles' benefits. Based on the new benchmark we conclude that:

- Technological advances are required to address 21.3% of accidents.
- Some issues, like inclement weather risk, will need to be measured and addressed on a local level.
- The driver remains a vital part of the accident-reduction equation.
- Success is not only dependent on the technology's operation but also on the circumstances surrounding its use. Driver behavioral issues may interfere with optimal implementation of the technology in over 30% of the accidents.

| Category | Disabling Factor | UnWtd Freq | Wtd Freq | UnWtd Freq | Wtd Freq |
|---|---|---|---|---|---|
| **Technology Issues** | Inoperable Weather | 602 | 267,657 | 11.0% | 12.2% |
| | Vehicle Issue Present | 681 | 254,948 | 12.4% | 11.6% |
| | Inoperable Traffic Control Device | 22 | 7,933 | 0.4% | 0.4% |
| | **Total Technology Issues** | **1,183** | **466,269** | **21.6%** | **21.3%** |
| **Behavioral (Driver) Issues** | Driver Disables | 152 | 67,304 | 2.8% | 3.1% |
| | Alcohol/Illicit Drugs | 502 | 241,596 | 9.2% | 11.0% |
| | Physical Impairment (heart attack) | 138 | 49,868 | 2.5% | 2.3% |
| | Sleeping | 159 | 62,974 | 2.9% | 2.9% |
| | Distraction | 929 | 365,436 | 17.0% | 16.7% |
| | **Total Usage Issues** | **1,742** | **709,153** | **31.8%** | **32.4%** |
| | **Total AV Issues** | **2,644** | **1,070,757** | **48.3%** | **48.9%** |
| | **Total Accidents** | **5,470** | **2,188,970** | **100.0%** | **100.0%** |

## 5. IMPLICATIONS FOR TESTING

The new baseline indicates that a more robust, transparent, and collaborative effort is needed to optimize automated vehicles' safety. The technology's safety depends not only on its engineering and coding but also consumers' use of it. Our work suggests two issues worthy of consideration with respect to test data collection.

**More comprehensive collection of real-world test data:** Some states[4] and the District of Columbia have already passed laws governing the testing of these vehicles on public roads. California's Department of Motor Vehicles requires testers to report all instances when the automated vehicle technology was disengaged.[5] In order to understand the technology's risks, the entire set of miles driven must be analyzed. More disengagements do not necessarily equate to a faultier product if Company XYZ is performing more tests or more difficult tests than Company ABC.

**Limitations of simulator tests:** Testing approaches that rely significantly on computer simulations involve a great deal of model risk, i.e., the risk that the test model does not accurately reflect reality. Over 30% of accidents involve a behavior risk that is very difficult to test in a simulator. An inebriated individual may react very differently in a simulator test than in real life. Testers are also unlikely to be in a hurry or race when using a computer simulation, while they may in real life. Removing responsibility from the driver altogether, as Google is doing, is one way to overcome this risk.

[4] California, Florida, Michigan, and Nevada
[5] http://apps.dmv.ca.gov/about/lad/pdfs/auto_veh2/adopted_txt.pdf section 227.46

## 6. CONCLUSION

Automated vehicles have the potential to transform our world, making transportation safer, cheaper, quicker, and greener. The technology could also transform transportation risk by shifting the driving responsibility and decision making from the individual to the technology. This will reduce or eliminate some risks while others will be increased or introduced.

In order to properly assess the risk, accurate risk measures need to be established. While NHTSA's NMVCCS found that human error is the main cause for over 90% of accidents using current technology, automated vehicles have a significantly different risk profile. Inclement weather, technological errors, and infrastructure issues can reduce the technology's benefit. Risky driver behaviors may further deteriorate automated vehicles' safety. Understanding and quantifying these risks will allow all stakeholders to make better decisions.

The Casualty Actuarial Society's Automated Vehicles Task Force believes the best way to balance the competing aims of safety, cost, and innovation is through robust, transparent testing and the application of statistical analysis appropriate to the data. Actuarial insights can support all stakeholders in making decisions to help the technology reach its potential and ensure it is brought to market as safely and efficiently as possible.

## REFERENCES

[1.] Allstate. "New Allstate Survey Shows Americans Think They Are Great Drivers—Habits Tell a Different Story." News release, August 2, 2011. Allstate Newsroom. http://www.allstatenewsroom.com/channels/News-Releases/releases/new-allstate-survey-shows-americans-think-they-are-great-drivers-habits-tell-a-different-story

[2.] California Department of Motor Vehicles. "Article 3.7—Autonomous Vehicles." Adopted Regulations for Testing of Autonomous Vehicles by Manufacturers. Section 227.46. May 19, 2014.

[3.] "Driving Citation Statistics—Statistic Brain." 2013 Statistic Brain Research Institute, Publishing as Statistic Brain. July 8, 2014. http://www.statisticbrain.com/driving-citation-statistics/

[4.] Fagnant, Daniel J., and Kara M. Kockelman. *Preparing a Nation for Autonomous Vehicles*. Eno Center for Transportation. October 2013. http://www.enotrans.org/wp-content/uploads/wpsc/downloadables/AV-paper.pdf

[5.] George, Alexander. "Volvo Thinks Magnetic Roads Will Guide Tomorrow's Autonomous Cars." *Wired*. March 17, 2014. http://www.wired.com/2014/03/volvo-magnets-autonomous/

[6.] Griffith, Bill. "Engineering Safer Drivers." New England Motor Press Association. June 10, 2014. http://www.nempa.org/events/2014/engineering-safer-drivers/

[7.] Kameda, Masaaki. "Nissan Road-tests Self-driving Vehicle." *Japan Times*. November 25, 2013. http://www.japantimes.co.jp/news/2013/11/25/business/nissan-road-tests-self-driving-vehicle/#.VCLFl_ldWao

[8.] Moore, Nicole C. "Driverless, Connected Cars: Ann Arbor as an Early Adopter." University of Michigan Engineering. November 5, 2013. http://www.engin.umich.edu/college/about/news/stories/2013/november/driverless-connected-carsOsborn, Liz. "Average Annual Precipitation by City in the United States." Current Results.

[9.] Ottley, Stephen. "Car Drives Itself across Germany." *Drive*. September 10, 2013. http://www.drive.com.au/motor-news/car-drives-itself-across-germany-20130910-2thn8.html

[10.] "U.K. to Allow Driverless Cars on Roads." BBC News. July 30, 2014. http://www.bbc.com/news/technology-28551069

[11.] U.S. Department of Transportation National Highway Transportation Safety Administration, National Motor Vehicle Crash Causation Survey Report to Congress, July 2008, http://www.thenewspaper.com/rlc/docs/2008/us-crashcause.pdf.

[12.] U.S. Department of Transportation's National Highway Traffic Safety Administration. "U.S. Department of Transportation Releases Policy on Automated Vehicle Development." News release, May 30, 2013. http://www.nhtsa.gov/About+NHTSA/Press+Releases/U.S.+Department+of+Transportation+Releases+Policy+on+Automated+Vehicle+Development

[13.] Volvo Car Group. Global Newsroom. "Volvo Car Group's First Self-driving Autopilot Cars Test on Public Roads around Gothenburg." News release, April 29, 2014. Volvo Car Group. https://www.media.volvocars.com/global/en-gb/media/pressreleases/145619/volvo-car-groups-first-self-driving-autopilot-cars-test-on-public-roads-around-gothenburg

[14.] Weiner, Gabriel, and Bryant W. Smith. *The Center for Internet and Society at Stanford Law School.* cyberlaw.stanford.edu/wiki/index.php/Automated_Driving:_Legislative_and_Regulatory_Action

## APPENDIX A—NHTSA'S LEVELS OF AUTOMATION[6]

### Level 0—No Automation

The driver is in complete and sole control of the primary vehicle controls (brake, steering, throttle, and motive power) at all times, and is solely responsible for monitoring the roadway and for safe operation of all vehicle controls. Vehicles that have certain driver support/convenience systems but do not have control authority over steering, braking, or throttle would still be considered "level 0" vehicles. Examples include systems that provide only warnings (e.g., forward collision warning, lane departure warning, blind spot monitoring) as well as systems providing automated secondary controls such as wipers, headlights, turn signals, hazard lights, etc. Although a vehicle with V2V warning technology alone would be at this level, that technology could significantly augment, and could be necessary to fully implement, many of the technologies described below, and is capable of providing warnings in several scenarios where sensors and cameras cannot (e.g., vehicles approaching each other at intersections).

### Level 1—Function-Specific Automation

Automation at this level involves one or more specific control functions; if multiple functions are automated, they operate independently from each other. The driver has overall control, and is solely responsible for safe operation, but can choose to cede limited authority over a primary control (as in adaptive cruise control), the vehicle can automatically assume limited authority over a primary control (as in electronic stability control), or the automated system can provide added control to aid the driver in certain normal driving or crash-imminent situations (e.g., dynamic brake support in emergencies). The vehicle may have multiple capabilities combining individual driver support and crash avoidance technologies, but does not replace driver vigilance and does not assume driving responsibility from the driver. The vehicle's automated system may assist or augment the driver in operating one of the primary controls—either steering or braking/throttle controls (but not both). As a result, there is no combination of vehicle control systems working in unison that enables the driver to be disengaged from physically operating the vehicle by having his or her hands off the steering wheel AND feet off the pedals at the same time. Examples of function-specific automation systems include cruise control, automatic braking, and lane keeping.

### Level 2—Combined Function Automation

This level involves automation of at least two primary control functions designed to work in unison to relieve the driver of control of those functions. Vehicles at this level of automation can

[6] http://www.nhtsa.gov/About+NHTSA/Press+Releases/U.S.+Department+of+Transportation+Releases+Policy+on+Automated+Vehicle+Development

utilize shared authority when the driver cedes active primary control in certain limited driving situations. The driver is still responsible for monitoring the roadway and safe operation and is expected to be available for control at all times and on short notice. The system can relinquish control with no advance warning and the driver must be ready to control the vehicle safely. An example of combined functions enabling a Level 2 system is adaptive cruise control in combination with lane centering. The major distinction between Level 1 and Level 2 is that, at Level 2 in the specific operating conditions for which the system is designed, an automated operating mode is enabled such that the driver is disengaged from physically operating the vehicle by having his or her hands off the steering wheel AND foot off pedal at the same time.

### Level 3—Limited Self-Driving Automation

Vehicles at this level of automation enable the driver to cede full control of all safety-critical functions under certain traffic or environmental conditions and in those conditions to rely heavily on the vehicle to monitor for changes in those conditions requiring transition back to driver control. The driver is expected to be available for occasional control, but with sufficiently comfortable transition time. The vehicle is designed to ensure safe operation during the automated driving mode. An example would be an automated or self-driving car that can determine when the system is no longer able to support automation, such as from an oncoming construction area, and then signals to the driver to reengage in the driving task, providing the driver with an appropriate amount of transition time to safely regain manual control. The major distinction between Level 2 and Level 3 is that at Level 3, the vehicle is designed so that the driver is not expected to constantly monitor the roadway while driving.

### Level 4—Full Self-Driving Automation

The vehicle is designed to perform all safety-critical driving functions and monitor roadway conditions for an entire trip. Such a design anticipates that the driver will provide destination or navigation input, but is not expected to be available for control at any time during the trip. This includes both occupied and unoccupied vehicles. By design, safe operation rests solely on the automated vehicle system.

# Tail Factor Convergence in Sherman's Inverse Power Curve Loss Development Factor Model

Jon Evans

**Abstract**

The infinite product of the age-to-age development factors in Sherman's inverse power curve model is proven to converge to a finite number when the power parameter is less than -1, and alternatively to diverge to infinity when the power parameter is -1 or greater. For the convergent parameter values, a simple formula is derived, in terms of any finite product of age-to-age factors, for the endpoints of an interval containing the limit of the infinite product. These endpoints converge to the limit as the finite time cutoff point increases. For any finite time cutoff, the product of age-to-age factors lies below the interval, and thus the lower endpoint of the interval is always a better estimate of the limit than the finite product itself. Several numerical examples are included for illustration.

**Keywords**. Tail Factor, Inverse Power Curve.

## 1. BACKGROUND AND INTRODUCTION

In 1984, Sherman [5] found that an inverse power curve of the form $1+a\left(t+c\right)^b$ fit empirical age-to-age loss development factors better than several other basic functional forms he tested. Lowe and Mohrman in 1985 [3], expressed concern about the convergence of the product of the age-to-age factors. Boor in 2006 [1, p. 373], and the CAS Tail Factor Working Party in 2013 [2, p. 52] noted that there has been no known closed form expression that approximates the tail generated by the inverse power curve.

In practice, the age-to-age development factors produced by the curve are multiplied together out to some finite age cutoff, such as $t$ = 80, to produce a cumulative development factor. The impact of factors beyond that age to ultimate, or the tail factor beyond the cutoff, in this case $t$ = 81…, is assumed to be negligible. Alternatively, if the impact of the tail factor is not negligible, then some other modeling consideration must inform the selection of the cutoff time.

The potential danger in the assumption of negligible tail factor impact is illustrated in Table 1. Two different sets of parameters share the same initial age-to-age factor of 1.01 at $t$ = 1 and the same cumulative factor of 1.30 from $t$ = 1 to 100. However, while the cumulative factor for Example 1, using power parameter $b$ = –4.0, grows only a little past $t$ = 100, Example 2, using $b$ = –0.5, appears to zoom toward infinity in the tail.

**Table 1:** Examples of Apparently Convergent and Divergent Tail Factors for the Inverse Power Curve Model

Parameter Values

| Parameters | Example 1 | Example 2 |
|---|---|---|
| *a* | 545540.243359093 | 0.0150014750112457 |
| *b* | -4.0 | -0.5 |
| *c* | 84.9422458022239 | 1.25044252421429 |

Cumulative Development Factors From 1 to *n*

| *n* | Example 1 | Example 2 |
|---|---|---|
| 1 | 1.010 | 1.010 |
| 10 | 1.085 | 1.065 |
| 100 | 1.300 | 1.300 |
| 1,000 | 1.337 | 2.482 |
| 10,000 | 1.338 | 19.293 |
| 100,000 | 1.338 | 1.27E+04 |
| 1,000,000 | 1.338 | 1.03E+13 |
| 10,000,000 | 1.338 | 1.54E+41 |

This paper uses basic real analysis ([4] being a standard textbook reference) to prove that the infinite product of the age-to-age factors converges to a finite number when the power parameter $b$ is less than -1, and diverges to $+\infty$ when $b \geq -1$. Note, in this paper we refer to a sequence that increases without any upper bound as diverging to $+\infty$, or having a limit of $+\infty$. Furthermore, when $b < -1$, for any finite product of the age-to-age factors up to a specific age $n$, there is a simple formula for an interval containing the limit of the infinite product. As $n$ increases, the interval becomes tighter and the endpoints each converge to the limit of the infinite product. The lower endpoint of this interval is always a better estimate of the infinite product than the finite product of the age-to-age factors which is always less than the lower endpoint.

It is worth noting again that tail divergence does not necessarily mean the model is invalid, but simply that any specific finite cutoff point should be otherwise justified. For a convergent tail, either a cutoff point must still be justified by some other consideration or care must be taken that the tail factor past the cutoff is reasonably close to 1. The interval estimate derived in this paper can help answer the latter question.

The proof of convergence/divergence is laid out in Section 2.1, with the proof of several useful lemmas in Appendix A. The interval estimate is derived in Section 2.2. Numerical examples of the progressive convergence/divergence of the finite product and the interval estimate of the infinite product for several sets of parameters are shown in Section 2.3.

## 2. CONVERGENCE THEOREM AND LIMIT ESTIMATION

Following the notational conventions of the recent CAS working party [2], in the remainder of this paper, *d*, instead of *t*, is used for age or time.

### 2.1 Statement and Proof of Primary Theorem

First we will set up a definition for the finite product of the age-to-age factors in the inverse power curve model.

**Definition:** $F_n(a,b,c) = \prod_{d=1}^{n} \left(1 + a(d+c)^b\right)$ where $a > 0$, $b$, and $c \geq 0$ are real numbers and $n$ is a positive integer.

Note, this definition includes cases where *d* begins at a higher value than 1 as the *c* parameter can be increased to handle such cases. It is also worth noting that $a\,(d+c)^b > 0$, a key fact that will be used in subsequent derivations.

### Theorem 1

(i) If $b \geq -1$ then $\lim_{n\to\infty} F_n(a,b,c) = +\infty$.

(ii) If $b$ < -1 then $\lim_{n\to\infty} F_n(a,b,c) = F(a,b,c) < +\infty$ exists.

Proof:

(i) For any sequence of numbers $x_i > 0$ where $i = 1,\ldots,n$ and $n \geq 2$ the inequality

$\prod_{i=1}^{n}(1+x_i) > 1+\sum_{i=1}^{n} x_i$ holds according to Lemma A.3. Applying this we have

$$F_n(a,b,c) = \prod_{d=1}^{n}\left(1+a\,(d+c)^b\right) > 1+\sum_{d=1}^{n} a\,(d+c)^b = 1+a\sum_{d=c+1}^{n+c} d^b \ .$$

If $b \geq$ -1 then $\lim_{n\to\infty} \sum_{d=c+1}^{n+c} d^b = +\infty$ according to Lemma A.1, and consequently

$\lim_{n\to\infty} F_n(a,b,c) = +\infty$ .

(ii) By Lemma A.2, $\log(1+x) < x$ for any $x > 0$, so $\log\left(1+a\,(d+c)^b\right) < a\,(d+c)^b$ .

Summing over $d$ gives

$$\log F_n(a,b,c) = \sum_{d=1}^{n} \log\left(1+a(d+c)^b\right) < \sum_{d=1}^{n} a(d+c)^b = a\sum_{d=c+1}^{n+c} d^b \ .$$

If $b$ < -1 then $L = a\left(\lim_{n\to\infty}\sum_{d=c+1}^{n+c} d^b\right)$ exists and is less than +∞ according to Lemma A.1. Now note that $\log F_n(a,b,c)$ is an increasing sequence, because $1+a(d+c)^b > 1$ implies that $\log\left(1+a(d+c)^b\right) > 0$, and is bounded by $L$. Consequently, $\lim_{n\to\infty} \log F_n(a,b,c)$ exists and is less than +∞. So $\lim_{n\to\infty} F_n(a,b,c) = F(a,b,c)$ exists and is less than +∞.

## 2.2 An Interval Estimate for the Infinite Product Limit

For the convergent case of $b < -1$, it is possible to construct a useful interval estimate for the infinite product.

**Definition:** The *tail upper bound factor* is $U_n(a,b,c) = \exp\left(-a\frac{(n+c)^{b+1}}{b+1}\right)$.

**Definition:** The *tail lower bound factor* is $L_n(a,b,c) = 1 - a\frac{(n+c+1)^{b+1}}{b+1}$.

### Theorem 2

Let $F(a,b,c) = \lim_{n\to\infty} F_n(a,b,c)$. If $b < -1$ then:

(i) $\lim_{n\to\infty} U_n(a,b,c) = 1$.

(ii) $\lim_{n\to\infty} L_n(a,b,c) = 1$.

(iii) $F(a,b,c) \in \left( L_n(a,b,c)F_n(a,b,c),\ U_n(a,b,c)F_n(a,b,c) \right)$.

Proof:

(i) $b + 1 < 0$ implies that $\lim_{n\to\infty}\left(-a\frac{(n+c)^{b+1}}{b+1}\right) = 0$ and consequently

$$\lim_{n\to\infty} \exp\left(-a\frac{(n+c)^{b+1}}{b+1}\right) = 1.$$

(ii) $b + 1 < 0$ implies $\lim_{n\to\infty}\left(1 - a\frac{(n+c+1)^{b+1}}{b+1}\right) = 1 - \lim_{n\to\infty}\left(a\frac{(n+c+1)^{b+1}}{b+1}\right) = 1.$

(iii) $F(a,b,c) = F_n(a,b,c)\prod_{d=n+1}^{\infty}\left(1 + a\,(d+c)^b\right).$ Taking the logarithm of the tail factor and applying bounding techniques described in Lemmas A.1 and A.2

$$\log\left(\prod_{d=n+1}^{\infty}\left(1 + a\,(d+c)^b\right)\right) < \sum_{d=n+1}^{\infty} a\,(d+c)^b = a\sum_{d=n+c+1}^{\infty} d^b < -a\frac{(n+c)^{b+1}}{b+1}.$$

Exponentiating produces $\prod_{d=n+1}^{\infty}\left(1 + a\,(d+c)^b\right) < \exp\left(-a\frac{(n+c)^{b+1}}{b+1}\right).$ Consequently,

$F(a,b,c) \quad < \quad U_n(a,b,c)F_n(a,b,c).$

Similarly, using techniques from Lemmas A.1 and A.3 produces

$$\prod_{d=n+1}^{\infty}\left(1 + a\,(d+c)^b\right) > 1 + \sum_{d=n+1}^{\infty} a\,(d+c)^b = 1 + a\sum_{d=n+c+1}^{\infty} d^b > 1 - a\frac{(n+c+1)^{b+1}}{b+1}.$$

Consequently, $F(a,b,c) \quad > \quad L_n(a,b,c)F_n(a,b,c).$ This completes the proof of Theorem 2.

The lower endpoint of the estimation interval is always a better estimate of the infinite product since $L_n(a,b,c) \quad > \quad 1$ and consequently $F_n(a,b,c) \quad < \quad L_n(a,b,c)F_n(a,b,c).$

The tail bound factors are computationally simple even for large values of *n* and give a measure of the relative width of the estimation interval prior to doing the computationally intense calculation of the finite product. For example, to achieve a certain target *U* for the upper bound requires $n \approx -c + \left( -\frac{(1+b)\log(U)}{a} \right)^{\frac{1}{1+b}}$. A more relevant measure of relative error, but without any simple formula for *n* that the author is aware of, is the ratio of the tail upper bound factor to the tail lower bound factor

$$U_n(a,b,c)/L_n(a,b,c) = \exp\left( -a\frac{(n+c)^{b+1}}{b+1} \right) \left( 1 - a\frac{(n+c+1)^{b+1}}{b+1} \right)^{-1}.$$

**Example 1:** An upper bound factor target set at $U$ = 1.01 for the parameter values $a$ = 545540, $b$ = -4.0, and $c$ = 84.9422 requires $n \approx 178$. However, by $n = 29$ the ratio of the tail upper bound factor to the tail lower bound factor is about 1.01.

## 2.3 More Numerical Examples

Table 2 shows six different sets of parameters, each of which produces an age-to-age factor at $d$ = 1 of 1.01 and a cumulative factor from $d$ = 1 to 100 of 1.30. The parameter sets are indexed by a set of values {-2.0, -1.5, -1.1, -1.0, -0.9, -0.6} for the power parameter *b*. For $b$ = -1 the divergence happens very slowly, but for $b$ = -1.1 the convergence happens remarkably slowly. To achieve $U_n(a,b,c) \approx 1.01$ for the $b$ = -1.1 parameter set would require $n \approx 2.7 \times 10^{22}$, although by $n \approx 5.3 \times 10^{10}$ $U_n(a,b,c)/L_n(a,b,c) \approx 1.01$, still an astronomically slow rate of convergence.

**Table 2:** Examples of Finite Development Factor Products and Interval Estimates For Infinite Development Factor Products

| Parameters | Parameter Values<br>Example 3 | Example 4 | Example 5 |
|---|---|---|---|
| *a* | 12.1209748535112 | 1.07747300550919 | 0.174451676891596 |
| *b* | -2.0 | -1.5 | -1.1 |
| *c* | 33.815190439679 | 21.6431893821624 | 12.4522704340826 |

| *n* | Cumulative Development Factor Product (Infinite Product Lower Bound, Infinite Product Upper Bound)<br>Example 3 | Example 4 | Example 5 |
|---|---|---|---|
| 1 | 1.010 (1.352, 1.431) | 1.010 (1.458, 1.589) | 1.010 (2.359, 3.877) |
| 10 | 1.083 (1.375, 1.428) | 1.081 (1.488, 1.585) | 1.078 (2.449, 3.868) |
| 100 | 1.300 (1.417, 1.423) | 1.300 (1.553, 1.581) | 1.300 (2.713, 3.858) |
| 1,000 | 1.406 (1.423, 1.423) | 1.477 (1.576, 1.580) | 1.610 (3.017, 3.856) |
| 10,000 | 1.421 (1.423, 1.423) | 1.546 (1.579, 1.580) | 1.926 (3.263, 3.856) |
| 100,000 | 1.423 (1.423, 1.423) | 1.569 (1.580, 1.580) | 2.221 (3.447, 3.856) |
| 1,000,000 | 1.423 (1.423, 1.423) | 1.576 (1.580, 1.580) | 2.488 (3.578, 3.856) |
| 10,000,000 | 1.423 (1.423, 1.423) | 1.579 (1.580, 1.580) | 2.723 (3.670, 3.856) |
| 100,000,000 | 1.423 (1.423, 1.423) | 1.580 (1.580, 1.580) | 2.925 (3.733, 3.856) |
| 1,000,000,000 | 1.423 (1.423, 1.423) | 1.580 (1.580, 1.580) | 3.096 (3.776, 3.856) |

| Parameters | Parameter Values<br>Example 6 | Example 7 | Example 8 |
|---|---|---|---|
| *a* | 0.112891979103701 | 0.0737384147594275 | 0.0219230164116958 |
| *b* | -1.0 | -0.9 | -0.6 |
| *c* | 10.2891979090266 | 8.20670480785112 | 2.69970572509898 |

| *n* | Cumulative Development Factor Product<br>Example 6 | Example 7 | Example 8 |
|---|---|---|---|
| 1 | 1.010 | 1.010 | 1.010 |
| 10 | 1.077 | 1.075 | 1.069 |
| 100 | 1.300 | 1.300 | 1.300 |
| 1,000 | 1.668 | 1.744 | 2.185 |
| 10,000 | 2.161 | 2.550 | 8.118 |
| 100,000 | 2.803 | 4.119 | 219.782 |
| 1,000,000 | 3.635 | 7.534 | 8.72E+05 |
| 10,000,000 | 4.714 | 16.111 | 9.55E+14 |
| 100,000,000 | 6.113 | 41.946 | 4.86E+37 |
| 1,000,000,000 | 7.928 | 139.919 | 5.27E+94 |

**Acknowledgment**

The author is greatly thankful to John Robertson, Dan Corro, and Len Herk for review and comments on this paper.

## Appendix A - Lemmas

### Lemma A.1

Let *n* be a positive integer and $l > 0$.

(i) If $b \geq -1$ then $\lim_{n\to\infty}\sum_{k=l}^{l+n} k^b = +\infty$.

(ii) If $b < -1$ then $\lim_{n\to\infty}\sum_{k=l}^{l+n} k^b < +\infty$ exists.

Proof:

It suffices to show convergence or divergence for $\lim_{n\to\infty}\sum_{k=l+1}^{l+n} k^b$ since $l^b$ is a finite number.

For $k > 1$ and $b \geq 0$, $k^b \geq 1$, and therefore $\lim_{n\to\infty}\sum_{k=l+1}^{l+n} k^b = +\infty$.

For $k > 1$ and $b < 0$, $k^b$ is a strictly decreasing function of *k*, and therefore there is a sandwich inequality

$$\int_k^{k+1} t^b dt < k^b < \int_{k-1}^{k} t^b dt \quad \text{and consequently} \quad \int_{l+1}^{l+n+1} t^b dt < \sum_{k=l+1}^{l+n} k^b < \int_{l}^{l+n} t^b dt .$$

Solving the integrals when $b \neq -1$

$$\frac{(l+n+1)^{b+1}-(l+1)^{b+1}}{b+1} < \sum_{k=l+1}^{l+n} k^b < \frac{(l+n)^{b+1}-l^{b+1}}{b+1}.$$

For $b < -1$, taking limits produces

$$\frac{-(l+1)^{b+1}}{b+1} < \lim_{n\to\infty} \sum_{k=l+1}^{l+n} k^b < \frac{-l^{b+1}}{b+1}$$

In this case, the upper bound of the inequality is a finite number and implies convergence to a finite number since the sequence of partial sums in the middle is non-decreasing.

For $-1 < b < 0$, taking limits results in $\lim_{n\to\infty} \sum_{k=l+1}^{l+n} k^b = +\infty$, since in this case the lower bound of the earlier inequality diverges $\lim_{n\to\infty} \frac{(l+n+1)^{b+1}-(l+1)^{b+1}}{b+1} = +\infty$.

For the case $b = -1$, integration of the earlier inequality leads to

$$\log\left(\frac{l+n+1}{l+1}\right) < \sum_{k=l+1}^{l+n} k^b < \log\left(\frac{l+n}{l}\right).$$

Once again taking limits leads to $\lim_{n\to\infty} \sum_{k=l+1}^{l+n} k^b = +\infty$ from the lower bound of the inequality diverging $\lim_{n\to\infty} \log\left(\frac{l+n+1}{l+1}\right) = +\infty$.

**Lemma A.2**

If x > 0, then $\log(1+x) < x$.

Proof:

If $t > 1$ then $1/t - 1 < 0$, and consequently $\int_1^{1+x} (1/t - 1)\, dt < 0$.

So, $\int_1^{1+x} dt/t - \int_1^{1+x} dt < 0$ and solving the integrals produces $\log(1+x) - x < 0$.

**Lemma A.3**

$\prod_{i=1}^{n}(1+x_i) > 1 + \sum_{i=1}^{n} x_i$ for any sequence of numbers $x_i > 0$, $i = 1,\ldots,n$, and integer $n \geq 2$.

Proof:

We proceed by induction.

For $n = 2$, since $x_1 x_2 > 0$, it follows that $1 + x_1 + x_2 + x_1 x_2 > 1 + x_1 + x_2$.

Assume the conclusion of the lemma is true for *n* where $n \geq 2$. We will show that the lemma is then true for *n* + 1.

In general

$$\prod_{i=1}^{n+1}(1+x_i) = (1+x_{n+1})\prod_{i=1}^{n}(1+x_i) = \prod_{i=1}^{n}(1+x_i) + x_{n+1}\prod_{i=1}^{n}(1+x_i)$$

But $x_{n+1}\prod_{i=1}^{n}(1+x_i) > x_{n+1}$ so

$$\prod_{i=1}^{n}(1+x_i) + x_{n+1}\prod_{i=1}^{n}(1+x_i) > 1 + \sum_{i=1}^{n} x_i + x_{n+1} = 1 + \sum_{i=1}^{n+1} x_i \,,$$

which establishes the lemma.

# REFERENCES


[1] Boor, J., "Estimating Tail Development Factors: What to do When the Triangle Runs Out," *Casualty Actuarial Society Forum*, Winter, 2006.

[2] CAS Tail Factor Working Party, "The Estimation of Loss Development Tail Factors: A Summary Report," *Casualty Actuarial Society Forum*, Fall, 2013.

[3] Lowe, S. P. and Mohrman, D. F., "Discussion of Extrapolating, Smoothing and Interpolating Development Factors," *PCAS* LXXII, 1985.

[4] Rudin, W., *Principles of Mathematical Analysis*, 3rd ed., McGraw-Hill, Inc., 1976.

[5] Sherman, R. E., "Extrapolating, Smoothing and Interpolating Development Factors," *PCAS* LXXI, 1984.


## Biography of the Author

**Jonathan Evans, FCAS, FSA, FCA, CERA, MAAA, WCP** is an actuary at the National Council on Compensation Insurance in Boca Raton, FL. His work primarily involves research and development of NCCI's ratemaking, reserving, and catastrophe modeling procedures.

# A Continuous Version of Sherman's Inverse Power Curve Model with Simple Cumulative Development Factor Formulas

Jon Evans

**Abstract**

A continuous version of Sherman's discrete inverse power curve model for loss development is defined. This continuous version, apparently unlike its discrete counterpart, has simple formulas for cumulative development factors, including tail factors. The continuous version has the same tail convergence conditions and basic analytical properties as the discrete version. Parameter fitting and numerical comparisons between the discrete and continuous model versions are explored.

**Keywords**: Tail Factor, Inverse Power Curve.

## 1. INTRODUCTION

The inverse power curve model for loss development factors was introduced by Sherman in 1984 [6]. Several papers have commented on the lack of a closed form representation and/or tail convergence information [1], [2], and [4]. The conditions for tail factor convergence and estimates of the rate of tail factor convergence have been determined [3]. However, simple or closed form formulas for finite and infinite products of the discrete incremental development factors still appear illusive.

This paper will demonstrate that a continuous version of the inverse power curve model captures the relevant properties of the discrete model. At the same time this continuous model leads to simple cumulative development factor, including tail factor, formulas. Basic real analysis is used throughout this paper, as described in standard textbooks such as [5].

Section 2.1 defines the continuous version of the model. Convergence conditions are proved in Section 2.2. Several basic analytical properties are proven in Section 2.3. Empirical fitting and a comparison with the discrete model is discussed in Section 2.4. Further numerical comparisons are shown in Sections 2.5. Section 3 contains concluding remarks. Appendix A contains the proof of two lemmas and Theorem 2 from Section 2.3.

## 2. ANALYSIS AND RESULTS

In the remainder of this paper $t$ is used for age or time, whether discrete or continuous.

### 2.1 Continuous Inverse Curve Model Definition

The discrete version of the inverse power curve model can be defined in terms of a cumulative development factor $F_t(a,b,c)$ from time 1 to $t$.

**Definition:** For real numbers $a > 0$, $b$, and $c \geq 0$,

(i) If $t \geq 2$ is an integer, then $F_t(a,b,c) = \prod_{k=1}^{t-1}\left(1 + a(k+c)^b\right)$.

(ii) If $t = 1$, then $F_1(a,b,c) = 1$.

$F_t(a,b,c)$ obeys the finite difference equation

$$F_{t+1}(a,b,c) - F_t(a,b,c) = a(t+c)^b F_t(a,b,c).$$

If the boundary value of $F_1(a,b,c) = 1$ is included then this equation, along with the previous parameter restrictions, is an equivalent definition of $F_t(a,b,c)$.

A corresponding continuous version of the inverse power curve model can similarly be defined in terms of a cumulative development factor $F_t^*(a,b,c)$ from time 1 to $t$, but with much simpler closed form expressions.

**Definition:** For $a > 0$, $b$, $c \geq 0$, and $t \geq 0$

(i) $F_t^*(a,b,c) = \exp\left(\dfrac{a(c+t)^{1+b} - a(c+1)^{1+b}}{1+b}\right)$ if $b \neq -1$

(ii) $F_t^*(a,b,c) = \left(\frac{c+t}{c+1}\right)^a$ if $b = -1$

Using an analogous boundary value $F_1^*(a,b,c) = 1$, $F_t^*(a,b,c)$ is the solution to a differential equation in continuous time $t$,

$$\frac{dF_t^*(a,b,c)}{dt} = a(t+c)^b F_t^*(a,b,c),$$

that is analogous to the finite difference equation satisfied by the discrete model.

## 2.2 Tail Convergence

The tail of $F_t^*(a,b,c)$ converges when $b < -1$ and diverges when $b \geq -1$, which are exactly the same as the conditions for convergence of $F_t(a,b,c)$ as shown in [3].

## Theorem 1

(i) If $b \geq -1$, $\lim_{t\to+\infty} F_t^*(a,b,c) = +\infty$ .

(ii) If $b < -1$, $\lim_{t\to+\infty} F_t^*(a,b,c) = \exp\left(\frac{-a(c+1)^{1+b}}{1+b}\right)$.

Proof:

(i) If $b = -1$, clearly $\lim_{t\to+\infty} \frac{c+t}{c+1} = +\infty$ and since $a > 0$ by Lemma **A.1**

$\lim_{t\to+\infty} \left(\frac{c+t}{c+1}\right)^a = +\infty$. If $b > -1$, then $b + 1 > 0$, so again by Lemma **A.1**

$\lim_{t\to+\infty} \frac{a(c+t)^{1+b}}{1+b} = +\infty$ and since $\exp(x)$ is an increasing function of $x$,

$$\lim_{t\to+\infty} \exp\left(\frac{a(c+t)^{1+b} - a(c+1)^{1+b}}{1+b}\right) = +\infty .$$

(ii) If $b < -1$, then $b + 1 < 0$, so by Lemma **A.1** $\lim_{t\to+\infty} \frac{a(c+t)^{1+b}}{1+b} = 0$ and

consequently since exp(*x*) is continuous

$$\lim_{t\to+\infty} \exp\left(\frac{a(c+t)^{1+b} - a(c+1)^{1+b}}{1+b}\right) = \exp\left(\frac{-a(c+1)^{1+b}}{1+b}\right).$$

## 2.3 Some Basic Analytical Properties

For convenience we will first set up notational definitions of the one period development factors, $f_t(a,b,c)$ for the discrete model and $f_t^*(a,b,c)$ for the continuous model.

**Definition:** $f_t(a,b,c) = \frac{F_{t+1}(a,b,c)}{F_t(a,b,c)} = 1 + a(c+t)^b$

**Definition:**

(i) For $b \neq -1$, $f_t^*(a,b,c) = \frac{F_{t+1}^*(a,b,c)}{F_t^*(a,b,c)} = \exp\left(\frac{a(c+t+1)^{1+b} - a(c+t)^{1+b}}{1+b}\right)$

(ii) For $b = -1$, $f_t^*(a,b,c) = \frac{F_{t+1}^*(a,b,c)}{F_t^*(a,b,c)} = \left(\frac{c+t+1}{c+t}\right)^a$

Lowe and Mohrman [4] listed several analytical properties for a curve of one period loss development factors to be "well-behaved".

**Definition:** A curve of one period loss development factors, $f(t)$ for $t \geq 0$, is said to be *well-behaved* if it has all of the following properties:

(i) $f(t) \geq 1$

(ii) $\lim_{t \to +\infty} f(t) = 1$

(iii) $f'(t) < 0$

(iv) $\lim_{t \to +\infty} f'(t) = 0$

(v) $f''(t) > 0$

(vi) $\lim_{t \to +\infty} f''(t) = 0$

## Theorem 2

If $b < 0$ then $f_t(a,b,c)$ and $f_t^*(a,b,c)$ are both well-behaved.

See Appendix A for details of the proof.

## 2.4 Fitting to Empirical Data

Table 1 includes an example from Sherman's original paper [6] of the discrete model fit to empirical data. Also shown are one period development factors from the continuous model, using the same parameter values and then using another set of parameter values refit for the continuous model itself. The continuous model development factors using the same parameter values is fairly close to the discrete model. When the parameters are refit the resulting development factors are very close to the discrete model.

**Table 1:** Comparison of Discrete and Continuous Models Fit to General Liability Data (Actual and discrete fit are from Exhibit 2 in Sherman's original paper [6]. Time convention is shifted by -1 from the original paper.)

| | | Parameter Values | | |
|---|---|---|---|---|
| Parameters | | Discrete Fit | Discrete Fit | Continuous Fit |
| *a* | | 0.88614 | 0.88614 | 1.20154 |
| *b* | | -1.7338 | -1.7338 | -1.8306 |
| *c* | | 0 | 0 | 0 |
| | | One Period Development Factors From *t* to *t+1* | | |
| *t* | Actual | Discrete Model | Continuous Model | Continuous Model |
| 1 | 1.839 | 1.886 | 1.618 | 1.884 |
| 2 | 1.279 | 1.266 | 1.205 | 1.262 |
| 3 | 1.185 | 1.132 | 1.108 | 1.131 |
| 4 | 1.077 | 1.080 | 1.068 | 1.080 |
| 5 | 1.039 | 1.054 | 1.048 | 1.055 |
| 6 | 1.033 | 1.040 | 1.035 | 1.040 |
| 7 | 1.029 | 1.030 | 1.027 | 1.031 |
| 8 | 1.030 | 1.024 | 1.022 | 1.024 |
| 9 | 1.019 | 1.020 | 1.018 | 1.020 |
| 10 | 1.014 | 1.016 | 1.015 | 1.016 |
| 11 | 1.016 | 1.014 | 1.013 | 1.014 |
| 12 | 1.013 | 1.012 | 1.011 | 1.012 |
| 13 | 1.012 | 1.010 | 1.010 | 1.010 |
| 14 | 1.008 | 1.009 | 1.009 | 1.009 |
| Goodness of Fit ($R^2$) | | 98.3% | 97.8% | 98.2% |

**Definition:** $f^a(t)$ will denote an empirical observation of a one period development factor from time $t$ to $t$+1.

The fits and goodness of fit ($R^2$) numbers in Table 1 are determined using the squared error function

$$\sum_{t=1}^{n}\left(\log(f^a(t)-1)-\log(f_t(a,b,c)-1)\right)^2=\sum_{t=1}^{n}\left(\log(f^a(t)-1)-\log(a)-b\log(c+t)\right)^2 .$$

for the discrete model and correspondingly for the continuous model

$$\sum_{t=1}^{n}\left(\log(f^a(t)-1)-\log(f_t^*(a,b,c)-1)\right)^2$$

$$=\sum_{t=1}^{n}\left(\log(f^a(t)-1)-\log\left(\exp\left(\frac{a(c+t+1)^{1+b}-a(c+t)^{1+b}}{1+b}\right)-1\right)\right)^2 .$$

A simpler squared error function for the continuous model would be

$$\sum_{t=1}^{n}\left(\log(f^a(t))-\log(f_t^*(a,b,c))\right)^2$$

$$=\sum_{t=1}^{n}\left(\log(f^a(t))-\left(\frac{a(c+t+1)^{1+b}-a(c+t)^{1+b}}{1+b}\right)\right)^2 .$$

This is still a fairly complicated function, likely not having a simple formulaic solution to minimize the parameters *a*, *b*, and *c*. One of the few apparent advantages of the discrete model is a somewhat simpler error function. It may be advantageous to use a numerical optimization program (like Solver in Excel) to first optimize *a*, *b*, and *c* for the discrete model. Those values can then be used as a starting point, or initial values, for the optimizer to search for values optimal for the continuous model.

## 2.5 More Numerical Comparisons with Discrete Model

Table 2 follows the basic layout of Table 1, except that the fitting targets a one period development factor of 1.01 from time 1 to 2 and a development factor of 1.30 from time 1 to 101. The $b$ parameter runs through the set of values {-2.0, -1.5, -1.1, -1.0, -0.9, -0.6} in these examples. Similar to what happened in Table 1, in Table 2 the continuous model is fairly close to the discrete model using the same parameter values and very close - identical up to 3 digits past the decimal in some examples - when the parameter values are refit.

**Table 2:** Some Numerical Comparisons of Discrete and Continuous Models

| Parameters | Discrete Fit | Discrete Fit | Continuous Fit |
|---|---|---|---|
| | | Parameter Values | |
| *a* | 12.121 | 12.121 | 12.1528 |
| *b* | -2.0 | -2.0 | -2.0 |
| *c* | 33.8152 | 33.8152 | 33.4513 |

Cumulative Development Factors From 1 to *t*

| *t* | Discrete Model | Continuous Model | Continuous Model |
|---|---|---|---|
| 2 | 1.010 | 1.010 | 1.010 |
| 11 | 1.083 | 1.081 | 1.083 |
| 101 | 1.300 | 1.295 | 1.300 |
| 1,001 | 1.406 | 1.400 | 1.406 |
| 10,001 | 1.421 | 1.415 | 1.421 |
| 100,001 | 1.423 | 1.416 | 1.423 |
| 1,000,001 | 1.423 | 1.416 | 1.423 |
| 10,000,001 | 1.423 | 1.416 | 1.423 |
| 100,000,001 | 1.423 | 1.416 | 1.423 |
| 1,000,000,001 | 1.423 | 1.416 | 1.423 |

Parameter Values

| Parameters | Discrete Fit | Discrete Fit | Continuous Fit |
|---|---|---|---|
| *a* | 1.07747 | 1.07747 | 1.07894 |
| *b* | -1.5 | -1.5 | -1.5 |
| *c* | 21.6432 | 21.6432 | 21.2437 |

Cumulative Development Factors From 1 to *t*

| *t* | Discrete Model | Continuous Model | Continuous Model |
|---|---|---|---|
| 2 | 1.010 | 1.010 | 1.010 |
| 11 | 1.081 | 1.079 | 1.081 |
| 101 | 1.300 | 1.295 | 1.300 |
| 1,001 | 1.477 | 1.470 | 1.477 |
| 10,001 | 1.546 | 1.539 | 1.546 |
| 100,001 | 1.569 | 1.562 | 1.569 |
| 1,000,001 | 1.576 | 1.569 | 1.577 |
| 10,000,001 | 1.579 | 1.572 | 1.579 |
| 100,000,001 | 1.580 | 1.572 | 1.580 |
| 1,000,000,001 | 1.580 | 1.573 | 1.580 |

**Table 2 (cont.):** Some Numerical Comparisons of Discrete and Continuous Models

| | Parameter Values | | |
|---|---|---|---|
| Parameters | Discrete Fit | Discrete Fit | Continuous Fit |
| *a* | 0.174452 | 0.174452 | 0.174523 |
| *b* | -1.1 | -1.1 | -1.1 |
| *c* | 12.4523 | 12.4523 | 12.0248 |

| | Cumulative Development Factors From 1 to *t* | | |
|---|---|---|---|
| *t* | Discrete Model | Continuous Model | Continuous Model |
| 2 | 1.010 | 1.010 | 1.010 |
| 11 | 1.078 | 1.075 | 1.078 |
| 101 | 1.300 | 1.295 | 1.300 |
| 1,001 | 1.610 | 1.603 | 1.611 |
| 10,001 | 1.926 | 1.917 | 1.926 |
| 100,001 | 2.221 | 2.211 | 2.222 |
| 1,000,001 | 2.488 | 2.477 | 2.489 |
| 10,000,001 | 2.723 | 2.711 | 2.723 |
| 100,000,001 | 2.925 | 2.912 | 2.926 |
| 1,000,000,001 | 3.096 | 3.082 | 3.097 |

| | Parameter Values | | |
|---|---|---|---|
| Parameters | Discrete Fit | Discrete Fit | Continuous Fit |
| *a* | 0.112892 | 0.112892 | 0.112913 |
| *b* | -1.0 | -1.0 | -1.0 |
| *c* | 10.2892 | 10.2892 | 9.85493 |

| | Cumulative Development Factors From 1 to *t* | | |
|---|---|---|---|
| *t* | Discrete Model | Continuous Model | Continuous Model |
| 2 | 1.010 | 1.010 | 1.010 |
| 11 | 1.077 | 1.074 | 1.077 |
| 101 | 1.300 | 1.295 | 1.300 |
| 1,001 | 1.668 | 1.661 | 1.669 |
| 10,001 | 2.161 | 2.152 | 2.162 |
| 100,001 | 2.803 | 2.790 | 2.803 |
| 1,000,001 | 3.635 | 3.618 | 3.635 |
| 10,000,001 | 4.714 | 4.693 | 4.715 |
| 100,000,001 | 6.113 | 6.086 | 6.115 |
| 1,000,000,001 | 7.928 | 7.892 | 7.930 |

**Table 2 (cont.):** Some Numerical Comparisons of Discrete and Continuous Models

| | Parameter Values | | |
|---|---|---|---|
| Parameters | Discrete Fit | Discrete Fit | Continuous Fit |
| *a* | 0.0737384 | 0.0737384 | 0.0737367 |
| *b* | -0.9 | -0.9 | -0.9 |
| *c* | 8.2067 | 8.2067 | 7.7661 |

| | Cumulative Development Factors From 1 to *t* | | |
|---|---|---|---|
| *t* | Discrete Model | Continuous Model | Continuous Model |
| 2 | 1.010 | 1.010 | 1.010 |
| 11 | 1.075 | 1.073 | 1.075 |
| 101 | 1.300 | 1.295 | 1.300 |
| 1,001 | 1.744 | 1.737 | 1.744 |
| 10,001 | 2.550 | 2.539 | 2.550 |
| 100,001 | 4.119 | 4.101 | 4.119 |
| 1,000,001 | 7.534 | 7.500 | 7.534 |
| 10,000,001 | 16.111 | 16.039 | 16.110 |
| 100,000,001 | 41.946 | 41.759 | 41.944 |
| 1,000,000,001 | 139.919 | 139.293 | 139.906 |

| | Parameter Values | | |
|---|---|---|---|
| Parameters | Discrete Fit | Discrete Fit | Continuous Fit |
| *a* | 0.021923 | 0.021923 | 0.021913 |
| *b* | -0.6 | -0.6 | -0.6 |
| *c* | 2.69971 | 2.69971 | 2.24551 |

| | Cumulative Development Factors From 1 to *t* | | |
|---|---|---|---|
| *t* | Discrete Model | Continuous Model | Continuous Model |
| 2 | 1.010 | 1.009 | 1.010 |
| 11 | 1.069 | 1.066 | 1.068 |
| 101 | 1.300 | 1.295 | 1.300 |
| 1,001 | 2.185 | 2.176 | 2.185 |
| 10,001 | 8.118 | 8.083 | 8.113 |
| 100,001 | 219.782 | 218.839 | 219.322 |
| 1,000,001 | 8.72E+05 | 8.69E+05 | 8.67E+05 |
| 10,000,001 | 9.55E+14 | 9.51E+14 | 9.40E+14 |
| 100,000,001 | 4.86E+37 | 4.84E+37 | 4.67E+37 |
| 1,000,000,001 | 5.27E+94 | 5.24E+94 | 4.76E+94 |

# 3. CONCLUSIONS

The continuous inverse power curve model presented in this paper has the same tail convergence conditions and "well-behaved" analytical properties as the discrete model. Unlike the discrete model it is known to have very simple closed formulas for cumulative development factors, including tail factors. It tends to produce numerical values extremely close to the discrete value when fit to the same data. Squared error functions for fitting the parameters of the continuous model tend to be more complex, but fits borrowed from the discrete model can be used as initial values to facilitate fitting the continuous model.

**Appendix A**

**Lemma A.1**

(i) If p > 0 then $\lim_{x\to+\infty} x^p = +\infty$

(ii) If p < 0 then $\lim_{x\to+\infty} x^p = 0$

**Proof:**

(i) For any $\varepsilon > 0$, choose $x > \varepsilon^{1/p}$ to make $x > \varepsilon$.

(ii) For any $\varepsilon > 0$, choose $x > \varepsilon^{-1/p}$ to make $x < \varepsilon$.

**Lemma A.2**

If p < 1 and q ≥ 0 then $\lim_{x\to+\infty}((x+q)^p - x^p) = 0$

**Proof:**

If $p < 0$ then by **A.1** $\lim_{x\to+\infty}((x+q)^p - x^p) = 0 - 0 = 0$. If $p = 0$ then trivially

$\lim_{x\to+\infty}((x+q)^p - x^p) = 1 - 1 = 0$. For $1 > p > 0$, since $\frac{d(x+q)^p}{dq} = p(x+q)^{p-1} > 0$ but

$\frac{d^2(x+q)^p}{dq^2} = p(p-1)(x+q)^{p-2} < 0$, it follows regarding the tangential approximation from $q$=0 that $(x+q)^p \leq x^p + q\left(\left.\frac{d(x+q)^p}{dq}\right|_{q=0}\right) = x^p + qpx^{p-1}$. So

$(x+q)^p - x^p \leq qpx^{p-1}$ and consequently $0 \leq \lim_{x \to +\infty}((x+q)^p - x^p) \leq \lim_{x \to +\infty} qpx^{p-1} = 0$ by **A.1**

since $p$-1<0.

**Proof of Theorem 2 from Section 2.3:**

For $f_t(a,b,c)$:

(i) $a(c+t)^b > 0$ and consequently $1 + a(c+t)^b > 1$.

(ii) If $b < 0$ then by **A.1** $\lim_{t \to +\infty} a(c+t)^b = 0$ and consequently $\lim_{t \to +\infty}\left(1 + a(c+t)^b\right) = 1$.

(iii) $\frac{df_t(a,b,c)}{dt} = ba(c+t)^{b-1}$. Since $b < 0$, clearly $ba(c+t)^{b-1} < 0$ .

(iv) Since $b - 1 < b < 0$, by **A.1** it follows that $\lim_{t \to +\infty} ba(c+t)^{b-1} = 0$.

(v) $\frac{d^2 f_t(a,b,c)}{dt^2} = b(b-1)a(c+t)^{b-2}$. Since $b - 1 < b < 0$, obviously $b(b-1) > 0$, and consequently. $b(b-1)a(c+t)^{b-2} > 0$.

(vi) Since $b - 2 < b < 0$, by **A.1** it follows that $\lim_{t \to +\infty} b(b-1)a(c+t)^{b-2} = 0$.

For $f_t^*(a,b,c)$:

(i) If $b + 1 < 0$ then $a(c+t+1)^{1+b} \leq a(c+t)^{1+b}$, or if $b + 1 > 0$ then $a(c+t+1)^{1+b} \geq a(c+t)^{1+b}$. Either way it follows when taking the ratio that $\frac{a(c+t+1)^{1+b} - a(c+t)^{1+b}}{1+b} \geq 0$ and consequently that

$\exp\left(\dfrac{a(c+t+1)^{1+b}-a(c+t)^{1+b}}{1+b}\right)\geq 1$. Since $t \geq 1$, $c+t > c+1$, and

consequently $\left(\dfrac{c+t+1}{c+t}\right)^a \geq 1$ for the case of $b+1=0$.

(ii) If $b<0$ then $b+1<1$ and by **A.2** $\lim\limits_{t\to+\infty}\left(a(c+t+1)^{1+b}-a(c+t)^{1+b}\right)=0$. Since

$\exp(x)$ is continuous $\lim\limits_{t\to+\infty}\exp\left(\dfrac{a(c+t+1)^{1+b}-a(c+t)^{1+b}}{1+b}\right)=1$.

(iii) $\dfrac{df_t^*(a,b,c)}{dt}=\left(a(c+t+1)^b-a(c+t)^b\right)f_t^*(a,b,c)$. Since $b<0$,

$a(c+t+1)^b < a(c+t)^b$ and $a(c+t+1)^b - a(c+t)^b < 0$, and since

$f_t^*(a,b,c)>0$ it follows that $\left(a(c+t+1)^b-a(c+t)^b\right)f_t^*(a,b,c)<0$.

(iv) $\lim\limits_{t\to+\infty}\left(a(c+t+1)^b-a(c+t)^b\right)=\lim\limits_{t\to+\infty}a(c+t+1)^b-\lim\limits_{t\to+\infty}a(c+t)^b=0-0=0$ by

applying **A.1**. Therefore, since $\lim\limits_{t\to+\infty}f_t^*(a,b,c)=1$ it follows that

$\lim\limits_{t\to+\infty}\left(\left(a(c+t+1)^b-a(c+t)^b\right)f_t^*(a,b,c)\right)=0$.

(v) $\dfrac{d^2 f_t^*(a,b,c)}{dt^2}=\left(ba(c+t+1)^{b-1}-ba(c+t)^{b-1}\right)f_t^*(a,b,c)$

$+\left(a(c+t+1)^b-a(c+t)^b\right)\dfrac{df_t^*(a,b,c)}{dt}$

$=\left(ba(c+t+1)^{b-1}-ba(c+t)^{b-1}\right)f_t^*(a,b,c)+\left(a(c+t+1)^b-a(c+t)^b\right)^2 f_t^*(a,b,c)$

Clearly $\left(a(c+t+1)^b-a(c+t)^b\right)^2>0$. Since $b<0$ and

$a(c+t+1)^{b-1} < a(c+t)^{b-1}$ it follows that $ba(c+t+1)^{b-1} - ba(c+t)^{b-1} > 0$. Since $f_t^*(a,b,c) \geq 1$, altogether it follows that

$$\left(ba(c+t+1)^{b-1} - ba(c+t)^{b-1}\right) f_t^*(a,b,c) + \left(a(c+t+1)^b - a(c+t)^b\right)^2 f_t^*(a,b,c) > 0$$

.

(vi) Since $b-1 < b < 0$ by **A.2** it follows that

$$\lim_{t\to+\infty}\left(ba(c+t+1)^{b-1} - ba(c+t)^{b-1}\right) = 0 \text{ and } \lim_{t\to+\infty}\left(a(c+t+1)^b - a(c+t)^b\right)^2 = 0.$$

Since $\lim_{t\to+\infty} f_t^*(a,b,c) = 1$ and $\lim_{t\to+\infty} \frac{df_t^*(a,b,c)}{dt} = 0$, it follows that

$$\lim_{t\to+\infty}\left(\left(ba(c+t+1)^{b-1} - ba(c+t)^{b-1}\right) f_t^*(a,b,c) + \left(a(c+t+1)^b - a(c+t)^b\right)^2 f_t^*(a,b,c)\right) = 0$$

.

## 5. REFERENCES


[1] Boor, J., "Estimating Tail Development Factors: What to do When the Triangle Runs Out," *Casualty Actuarial Society Forum*, Winter, 2006.

[2] CAS Tail Factor Working Party, "The Estimation of Loss Development Tail Factors: A Summary Report," *Casualty Actuarial Society Forum*, Fall, 2013.

[3] Evans, J., "Tail Factor Convergence in Sherman's Inverse Power Curve Loss Development Factor Model," submitted to *Casualty Actuarial Society Forum* on July 10, 2014.

[4] Lowe, S. P. and Mohrman, D. F., "Discussion of Extrapolating, Smoothing and Interpolating Development Factors," *PCAS* LXXII, 1985.

[5] Rudin, W., *Principles of Mathematical Analysis*, 3rd ed., McGraw-Hill, Inc., 1976.

[6] Sherman, R. E., "Extrapolating, Smoothing and Interpolating Development Factors," *PCAS* LXXI, 1984.


### Biography of the Author


**Jonathan Evans, FCAS, FSA, FCA, CERA, MAAA, WCP** is a property and casualty consulting actuary with 20 years of experience. He is currently President of Convergent Actuarial Services, Inc.

# An Empirical Investigation of the Dependence between Catastrophe Events and the Performance of Various Asset Classes

Romel G. Salam

**Abstract:** For insurance and reinsurance companies primarily involved in the Catastrophe business, the dependence between losses from catastrophe events and the returns on their asset portfolio can significantly impact their risk capital calculation. This dependence is also of relevance to capital market investors involved in Insurance Linked Securities (ILS) funds. In this paper, we draw on more than 60 years of data to investigate the dependence between insured and economic losses from catastrophe events and the relative performance of several asset classes, commodities, and economic indices in the US. We also look at the association between catastrophes and equities for selected catastrophe prone countries around the world. For US equities, our investigation suggests two correlation effects: one corresponding to the lowest 80$^{th}$ percentile of catastrophe losses, and another corresponding to the highest 20$^{th}$ percentile.

**Keywords:** Enterprise Risk Management, Economic Capital Model, Dependence, Correlations, Assets, Catastrophes, Insurance Linked Securities

***Acknowledgements**: We would like to thank James Norman, Anthony Grandolfo, Christian Bird, Sajad Obaydi, and Fabien Cœur-Uni for their advice, feedback, and assistance with this paper. However, the views expressed herein are ours alone. We are also solely responsible for all errors and infractions committed in this paper.*

## 1. INTRODUCTION

Within the realm of dynamic financial analysis, there is broad recognition of both the importance and the challenges of adequately representing the dependence between risk variables. Indeed, some have partially blamed the 2008 financial crisis on the failure of quantitative analysts across the financial industry to accurately model the dependence between complex financial instruments[1]. Those tasked with building capital models for P&C insurance and reinsurance entities need to account for the dependence between a number of risk variables across multiple dimensions. There is some consensus around modeling the dependence of risk variables that fall within the same risk categories, which, for general insurance companies, are generally defined as insurance, market, credit, and operational. For instance, many practitioners use normal correlation matrices to capture the dependence between the underwriting results for various classes of business (i.e. Marine, Property, Medical Malpractice), or between the performance of different asset classes (i.e. Equities, Mortgage Backed Securities, Treasuries). There is much less agreement around how to represent the dependence between risk variables that fall in different risk categories.

For insurance and reinsurance companies primarily involved in the Catastrophe business and capital market investors involved in Insurance Linked Securities (ILS) funds, the dependence between losses from catastrophe events and the returns on their asset portfolio are particularly relevant. In this paper, we investigate the dependence between insured and economic losses from catastrophe events and the relative performance of several asset classes, commodities, and economic indices in the US. We also investigate the dependence between economic losses from catastrophes and the performance of equities in Australia, Chile, Japan, the Philippines, and Thailand. In section 2, we provide a brief overview of our approach. We present our findings and offer commentary in sections 3 and 4, respectively. We describe the data underlying this study and provide data sources in Appendix A. In Appendix B, we describe the calculations of the P-values and provide the distributions from which they are derived. Finally, we show selected graphs in Appendix C.

## 2. OVERVIEW OF APPROACH

For the purpose of this study, we expressed aggregate catastrophe losses incurred in a calendar year as a percentage of Gross Domestic Product (GDP) in the same year. We believe this provides a more consistent measure of the relative importance of catastrophe losses across time but also across countries. Graph 2.1 below shows annual insured catastrophe losses as a percentage of GDP for the

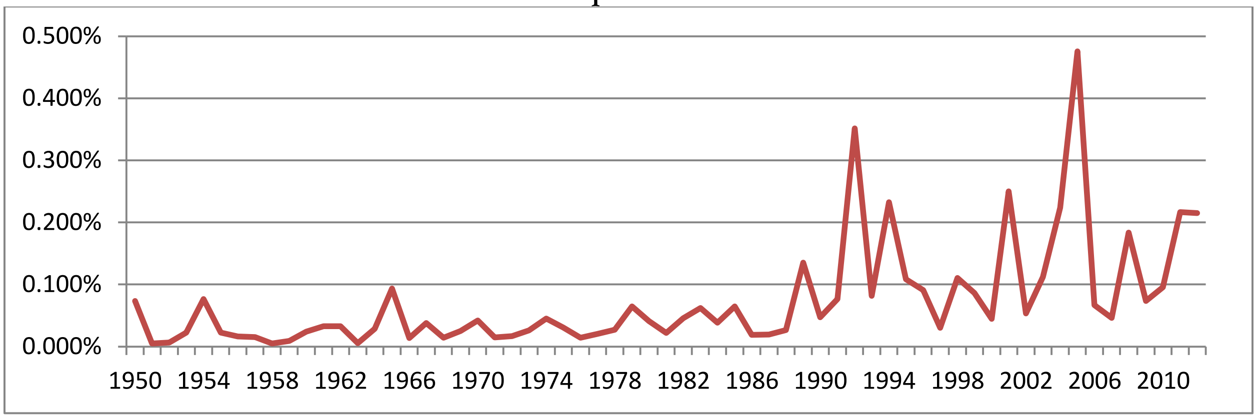
---

[1] See Mackenzie, D. and Spears, T. (2012): "The Formula That Killed Wall Street"? (School of Social & Political Science, University of Edinburgh, Scotland) http://www.sps.ed.ac.uk/__data/assets/pdf_file/0003/84243/Gaussian14.pdf

US from 1950 to 2013. Throughout the remainder of this text, we will use the terms "catastrophe losses" and "catastrophe losses as a percentage of GDP" interchangeably. Annual catastrophe losses are compared to the percentage change in various financial and economic indices over the same calendar year. In the remainder of this text, we will sometimes use the term "return" when referring to the percentage change in the financial indices.

Graph 2.1

**US Insured Catastrophe Losses as a % of GDP**




Our investigation relies on the following tools:

a. Visual representations of the relationships through the use of percentile scatter plots – Each point on the plots represents the percentile value for each pair of observations within their respective sample. These scatter plots represent the empirical copula of each pair of variables. We reviewed these plots to search for trends and other patterns in the data. For brevity, we refer to the percentile scatter plots simply as scatter plots in the remainder of this document. Selected scatter plots are shown throughout the paper and in Appendix C.
b. Rank correlation measurements – We used the Kendall's Tau[2] and the Kendall's Partial Tau[3] statistics as non-parametric measures of rank correlation. We chose non-parametric measures as we did not want to make any assumptions about the distributions underlying the variables we were studying. As Graph 2.1 shows, US insured catastrophe losses show an upward trend over time even after being normalized for GDP. Without controlling for time, some of the correlations we observe may simply be driven by common time dependencies coming across the data for both catastrophes and the financial and economic indices. Hence, we used the Kendall's Partial Tau to provide a measure of correlation between any pair of variables that removes the effect of common time correlations. We assess significance by calculating the P-values associated with the Kendall's Tau and the Kendall's

[2] We reach virtually the same conclusions about the significance of the observed correlations using a Spearman Rho rather than a Kendall's Tau statistic. We prefer the latter statistic as it has a more intuitive interpretation than the Spearman Rho.

[3] Assume we have three variables, X, Y, and Z, the Kendall's Partial Tau correlation coefficient for X and Y after removing the effect of Z is given by: $\tau_{xy.z} = \frac{\tau_{xy} - \tau_{xz}\tau_{yz}}{\sqrt{1-\tau_{xz}^2}\sqrt{1-\tau_{yz}^2}}$ where $\tau_{xy}, \tau_{xz}, \tau_{yz}$ represent the Kendall's Tau correlation coefficients for the pairs XY, XZ, and YZ respectively. See Gibbons, J.D. (1993, p. 49) *Nonparametric measures of association* (Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-091). Newbury Park, CA: Sage.

Partial Tau statistics. The calculations of the P-values, including the underlying null hypotheses, are described in Appendix B.

c. Difference in rank correlation measurements – We measured the differences between the Kendall's Tau and the Kendall's Partial Tau statistics wherever we had an indication that there was a shift in the correlation trends. We assess significance by calculating the P-values associated with the differences in the Kendall's Tau and Kendall's Partial Tau statistics. The calculations of the P-values are described in Appendix B.

d. We determine the significance of the Kendall's Tau, Kendall's Partial Tau, and of the differences in the Kendall's Tau and Kendall's Partial Tau values throughout this paper based on the interpretation of P-values shown in Table 2.1 below. This is perhaps the most subjective and also the most important table in this entire study. Different interpretations of the P-values will likely lead to different conclusions about the statistical significance of the observed correlations.

**Table 2.1**
**P-Value Interpretation**

| One-Tailed P- value Ranges | Reject Null Hypothesis? |
|---|---|
| **P-value ≤ .05** | Yes |
| **P-value > .05** | No |

## 3. FINDINGS

We present our key findings below:

a. We find two correlation trends between US annual catastrophe losses – either insured or economic – and annual changes in US equity prices. We observe a zero or a weak positive correlation when catastrophe losses as a percentage of GDP fall in the first 80th percentile and a negative correlation when they are at or above the 80th percentile. This is shown in Table 3.1.a below. This finding is unchanged when we remove the effect of time on the Kendall's Tau correlations as shown in Table 3.1.b. Tables 3.11.a and 3.11.b show the P-values for the differences in the Kendall's Tau and Kendall's Partial Tau values, respectively. We show the annual returns of the DJIA and DJCA against the highest 20th percentile of annual insured catastrophe losses in Table 3.2. We also show the annual returns of the DJIA against the highest 20th percentile of economic losses due to catastrophe in Table 3.3. Graph 3.1 shows a scatter plot of the annual DJIA returns against annual insured catastrophe losses. Graphs 3.1.a and 3.1.b show separate scatter plots corresponding to the lowest 80th percentile and the highest 20th percentile of annual insured catastrophe losses. We show the corresponding scatter plots for economic losses against the DJIA in Graphs 3.2, 3.2.a, and 3.2.b.

**Table 3.1.a**
**Annual Catastrophe Losses against Annual Changes in Equities – US**

| Catastrophe Losses | Catastrophe Loss Percentile | Index | No. of Observations | Kendall's Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|---|
| **Insured** | <.80 | DJIA | 51 | 17.3% | 0.036 | Yes |
| **Insured** | ≥.80 | DJIA | 13 | -53.8% | 0.005 | Yes |
| **Insured** | <.80 | DJCA | 51 | 14.5% | 0.066 | No |
| **Insured** | ≥.80 | DJCA | 13 | -41.0% | 0.025 | Yes |
| **Insured** | <.80 | S&P 500 | 51 | 12.9% | 0.090 | No |
| **Insured** | ≥.80 | S&P 500 | 13 | -46.2% | 0.014 | Yes |
| **Economic** | <.80 | DJIA | 56 | 18.8% | 0.020 | Yes |
| **Economic** | ≥.80 | DJIA | 15 | -48.6% | 0.006 | Yes |

**Table 3.1.b**
**Annual Catastrophe Losses against Annual Changes in Equities after Removing Effect of Time – US**

| Catastrophe Losses | Catastrophe Loss Percentile | Index | No. of Observations | Kendall's Partial Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|---|
| **Insured** | <.80 | DJIA | 51 | 15.7% | 0.054 | No |
| **Insured** | ≥.80 | DJIA | 13 | -56.2% | 0.003 | Yes |
| **Insured** | <.80 | DJCA | 51 | 13.9% | 0.078 | No |
| **Insured** | ≥.80 | DJCA | 13 | -42.7% | 0.021 | Yes |
| **Insured** | <.80 | S&P 500 | 51 | 12.8% | 0.095 | No |
| **Insured** | ≥.80 | S&P 500 | 13 | -48.7% | 0.009 | Yes |
| **Economic** | <.80 | DJIA | 56 | 18.2% | 0.025 | Yes |
| **Economic** | ≥.80 | DJIA | 15 | -48.5% | 0.006 | Yes |

**Table 3.2**
**Highest 20th Percentile of Annual Insured Catastrophe Losses against Annual Equity Returns – US**

| Year | Insured Cat as % of GDP | Insured Cat(1) (USD MM) | DJIA Return | DJIA Rank (2) | DJCA Return | DJCA Rank (2) |
|---|---|---|---|---|---|---|
| **2010** | 0.096% | 14,315 | 11.0% | 32 | 13.1% | 29 |
| **1995** | 0.109% | 8,325 | 33.5% | 4 | 32.9% | 3 |
| **1998** | 0.111% | 10,070 | 16.1% | 24 | 10.1% | 34 |
| **2003** | 0.112% | 12,885 | 25.3% | 9 | 26.3% | 7 |
| **1989** | 0.135% | 7,642 | 27.0% | 6 | 25.3% | 10 |
| **2008** | 0.184% | 27,045 | -33.8% | 64 | -29.8% | 64 |
| **2012** | 0.215% | 34,960 | 7.3% | 35 | 5.0% | 41 |
| **2011** | 0.217% | 33,640 | 5.5% | 38 | 4.9% | 42 |
| **2004** | 0.224% | 27,490 | 3.1% | 43 | 13.2% | 28 |
| **1994** | 0.233% | 17,010 | 2.1% | 46 | -7.7% | 54 |
| **2001** | 0.250% | 26,549 | -7.1% | 53 | -12.8% | 57 |
| **1992** | 0.351% | 22,970 | 4.2% | 42 | 4.1% | 43 |
| **2005** | 0.476% | 62,301 | -0.6% | 47 | 7.1% | 38 |

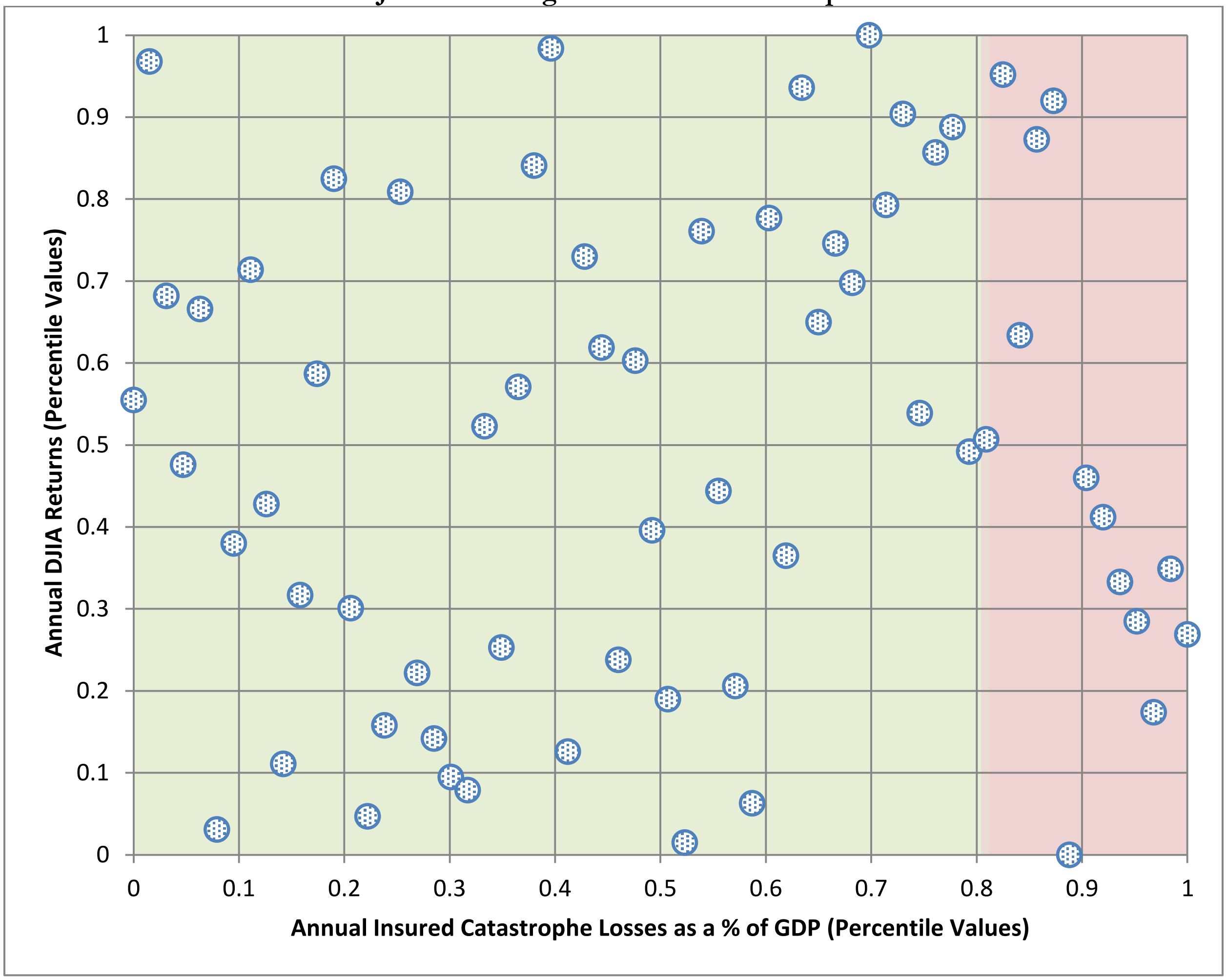
*(1) Source: PCS*
*(2) Rank from best to worst out of 64*

**Table 3.3**
**Highest 20th Percentile of Annual Economic Losses due to Catastrophes against Annual Equity Returns – US**

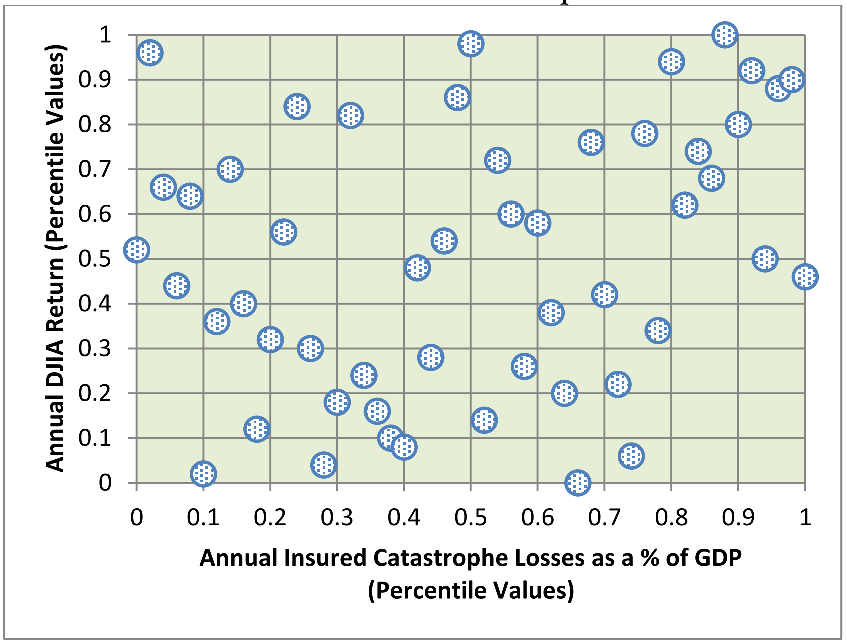
| Year | Economic Losses Due to Cats as % of GDP | Economic Losses Due to Cats(1) (USD MM) | DJIA Return | DJIA Rank (2) |
|---|---|---|---|---|
| **1995** | 0.220% | 16,890 | 33.5% | 5 |
| **1989** | 0.238% | 13,480 | 27.0% | 8 |
| **1993** | 0.268% | 18,423 | 13.7% | 33 |
| **1951** | 0.296% | 1,029 | 14.4% | 31 |
| **1964** | 0.305% | 2,090 | 14.6% | 30 |
| **2011** | 0.331% | 51,433 | 5.5% | 42 |
| **1938** | 0.350% | 306 | 28.1% | 6 |
| **2008** | 0.392% | 57,762 | -33.8% | 71 |
| **1994** | 0.432% | 31,554 | 2.1% | 50 |
| **1943** | 0.443% | 900 | 13.8% | 32 |
| **2004** | 0.454% | 55,692 | 3.1% | 47 |
| **1937** | 0.471% | 438 | -32.8% | 70 |
| **2012** | 0.483% | 78,469 | 7.3% | 39 |
| **1992** | 0.534% | 34,950 | 4.2% | 46 |
| **2005** | 1.215% | 159,060 | -0.6% | 51 |

*(1) Source: EM-Dat; Years with no losses are excluded*
*(2) Rank from best to worst out of 71*

**Graph 3.1**
**Scatter Plot of Annual DJIA Returns against Insured Catastrophe Losses as a % of GDP**

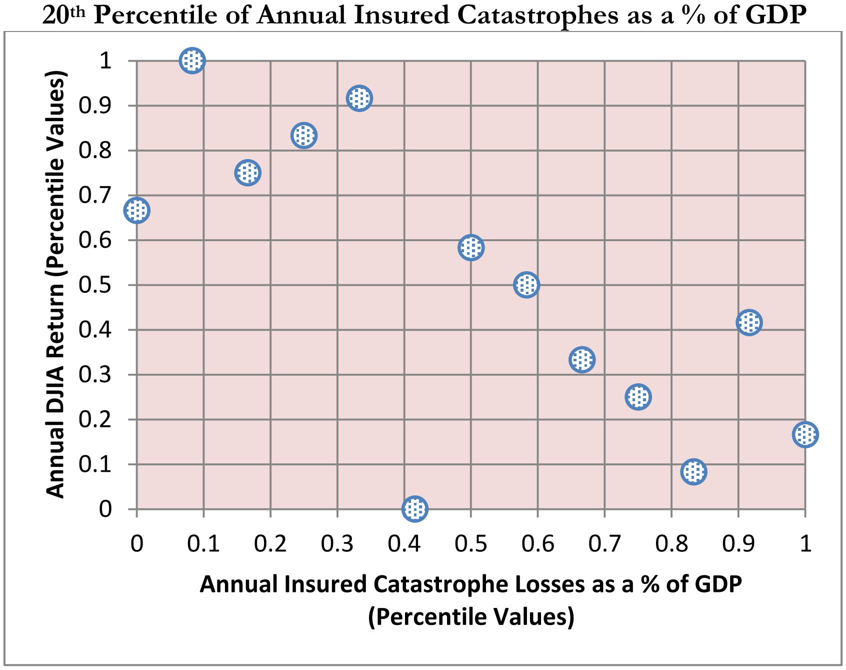


**Graph 3.1.a**
**Scatter Plot of Annual DJIA Returns against Insured Catastrophe Losses as a % of GDP for lowest 80th Percentile of Annual Insured Catastrophes as a % of GDP**



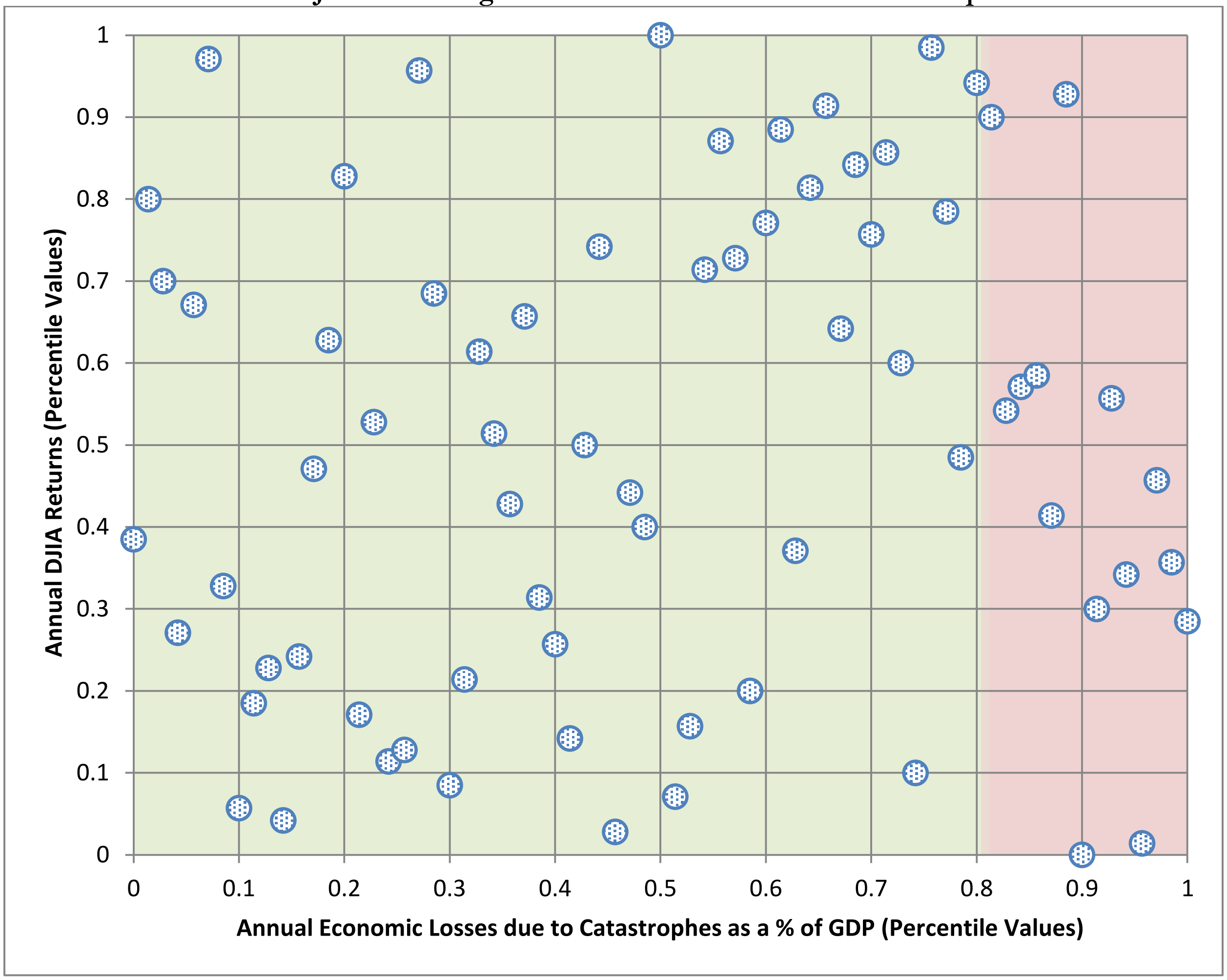

*Please note that the percentiles in the above graph have been recalculated based on the relative rankings of the subset of points that fall in the lowest 80th percentile of annual insured catastrophe losses as a % of GDP. As such, these percentiles range from 0 to 1.*

**Graph 3.1.b**
**Scatter Plot of Annual DJIA Returns against Insured Catastrophe Losses as a % of GDP for highest 20th Percentile of Annual Insured Catastrophes as a % of GDP**

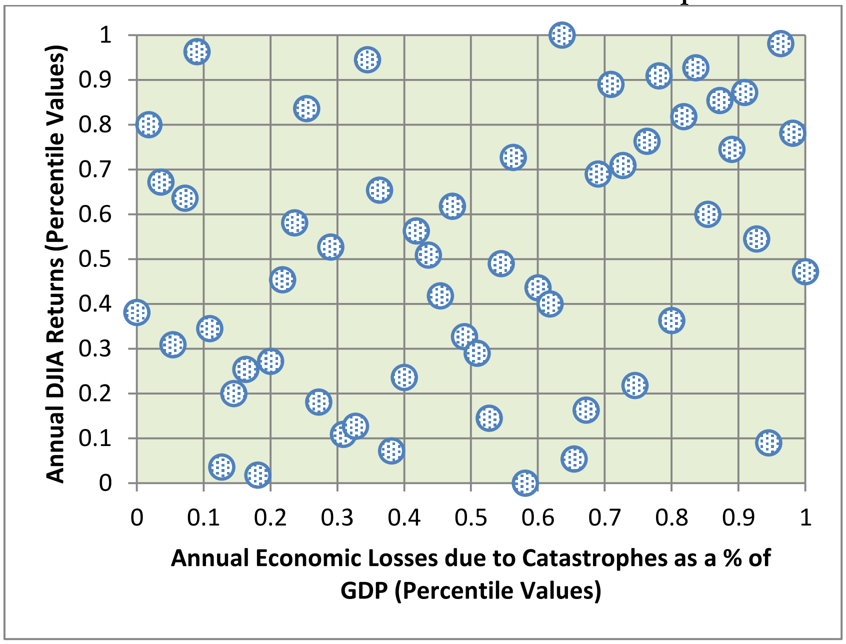



*Please note that the percentiles in the above graph have been recalculated based on the relative rankings of the subset of points that fall in the highest 20th percentile of annual insured catastrophe losses as a % of GDP. As such, these percentiles range from 0 to 1.*

**Graph 3.2**
**Scatter Plot of Annual DJIA Returns against Economic Losses due to Catastrophes as a % of GDP**

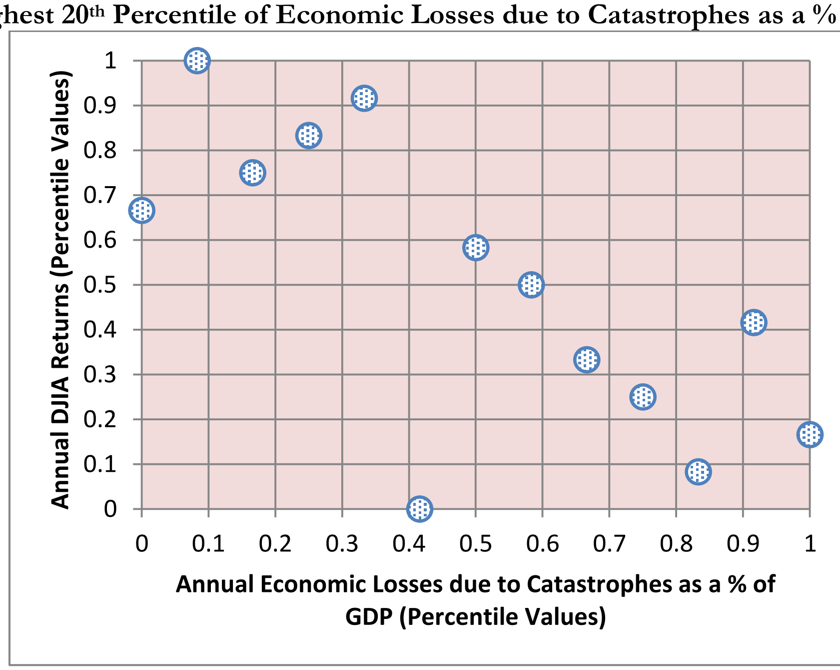


**Graph 3.2.a**
**Scatter Plot of Annual DJIA Returns against Economic Losses due to Catastrophes as a % of GDP for lowest 80th Percentile of Economic Losses due to Catastrophes as a % of GDP**




*Please note that the percentiles in the above graph have been recalculated based on the relative rankings of the subset of points that fall in the lowest 80th percentile of annual insured catastrophe losses as a % of GDP. As such, these percentiles range from 0 to 1.*

**Graph 3.2.b**
**Scatter Plot of Annual DJIA Returns against Economic Losses due to Catastrophes as a % of GDP for highest 20th Percentile of Economic Losses due to Catastrophes as a % of GDP**




*Please note that the percentiles in the above graph have been recalculated based on the relative rankings of the subset of points that fall in the highest 20th percentile of annual insured catastrophe losses as a % of GDP. As such, these percentiles range from 0 to 1.*

b. We find no significant correlation between catastrophe losses and changes in equity prices for Australia, Chile, Japan, and Thailand. We did not find any significant shift in correlation based on the relative size of catastrophe losses. This is shown in Table 3.4.a below. This finding is unchanged when we remove the effect of time on the Kendall's Tau correlations as shown in Table 3.4.b. Please note that we only have economic losses for these countries. Also, the data set for these countries is much sparser compared to the US.

**Table 3.4.a**
**Annual Catastrophe Losses against Annual Changes in Equities – Australia, Japan, Chile, and Thailand**

| Catastrophe Losses | Index | No. of Observations | Kendall's Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|
| **Economic Losses** | All Australia Shares | 44 | -11.2% | 0.140 | No |
| **Economic Losses** | Nikkei 225 | 38 | 1.3% | 0.455 | No |
| **Economic Losses** | IGPA | 30 | 11.3% | 0.191 | No |
| **Economic Losses** | SET Index | 22 | 4.8% | 0.378 | No |

**Table 3.4.b**
**Annual Catastrophe Losses against Annual Changes in Equities after Removing Effect of Time – Australia, Japan, Chile, and Thailand**

| Catastrophe Losses | Index | No. of Observations | Kendall's Partial Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|
| **Economic Losses** | All Australia Shares | 44 | -11.2% | 0.142 | No |
| **Economic Losses** | Nikkei 225 | 38 | 0.4% | 0.488 | No |
| **Economic Losses** | IGPA | 30 | 3.3% | 0.401 | No |
| **Economic Losses** | SET Index | 22 | 3.9% | 0.383 | No |

c. Similar to US equities, we find two correlation trends between annual catastrophe losses in the Philippines and annual changes in the main Philippines stock index, PSEi. We observe a zero correlation when catastrophe losses as a percentage of GDP fall in the first 70$^{th}$ percentile and a negative correlation when they are at or above the 70$^{th}$ percentile. This is shown in Table 3.5.a below. This finding is unchanged when we remove the effect of time on the Kendall's Tau correlations as shown in Table 3.5.b. Tables 3.11.a and 3.11.b show the P-values for the differences in the Kendall's Tau and Kendall's Partial Tau values. We show the annual returns of the PSEi index against the highest 30$^{th}$ percentile of annual economic losses due to catastrophes in Table 3.6. Please note that the number of observations is quite sparse compared to the US data. Graph 3.3 shows a scatter plot of the PSEi return against catastrophe losses. Graphs 3.3.a and 3.3.b show separate scatter plots corresponding to the lowest 70$^{th}$ percentile and the highest 30$^{th}$ percentile of catastrophe losses.

**Table 3.5.a**

**Annual Catastrophe Losses against Annual Changes in Equities – Philippines**

| Catastrophe Losses | Catastrophe Loss Percentile | Index | No. of Observations | Kendall's Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|---|
| **Economic Losses** | <.70 | PSEi | 18 | 3.3% | 0.425 | No |
| **Economic Losses** | ≥.70 | PSEi | 8 | -64.3% | 0.013 | Yes |

**Table 3.5.b**
**Annual Catastrophe Losses against Annual Changes in Equities after Removing Effect of Time – Philippines**

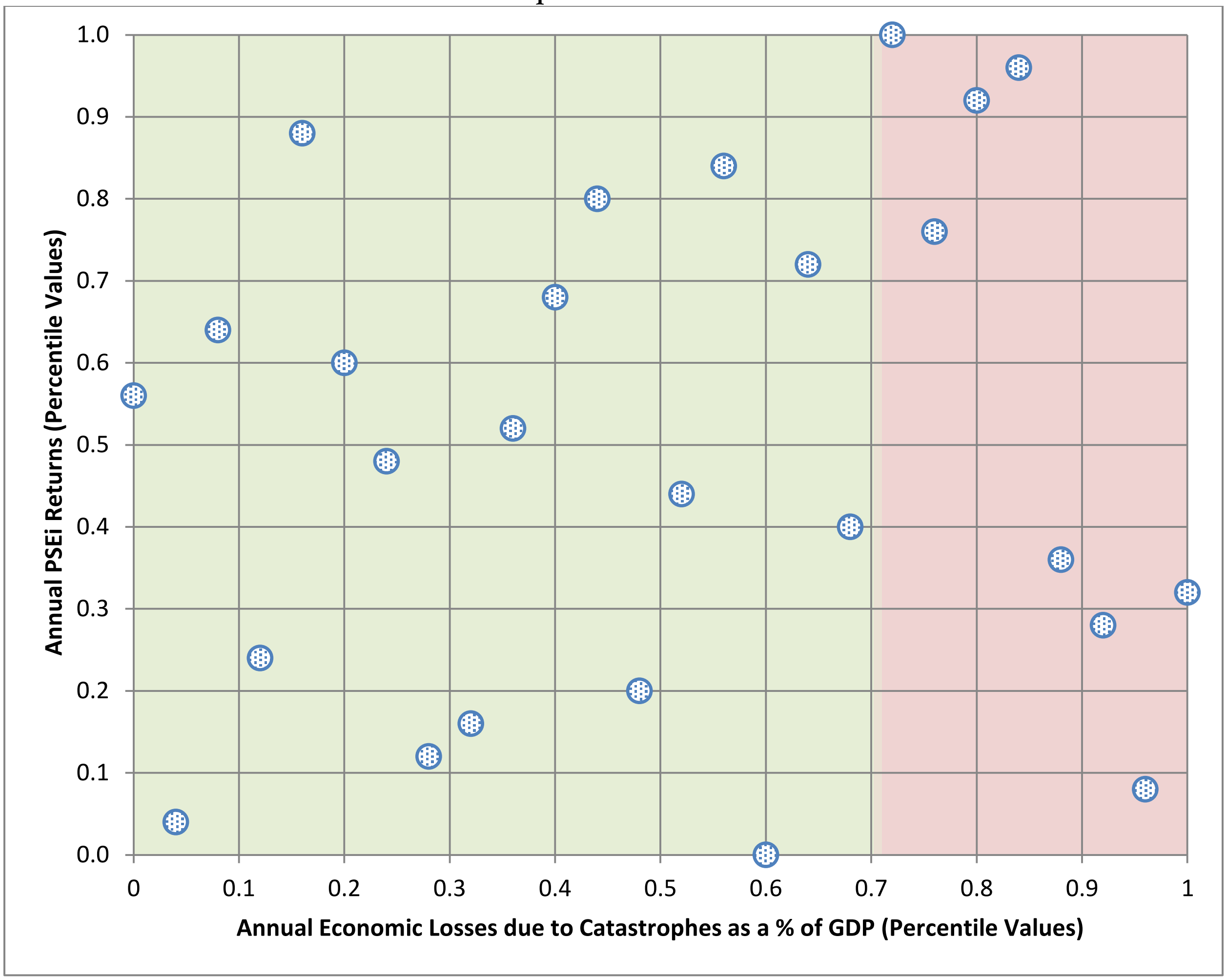
| Catastrophe Losses | Catastrophe Loss Percentile | Index | No. of Observations | Kendall's Partial Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|---|
| **Economic Losses** | <.70 | PSEi | 18 | 2.8% | 0.439 | No |
| **Economic Losses** | ≥.70 | PSEi | 8 | -64.5% | 0.011 | Yes |

**Table 3.6**
**Highest 30th Percentile of Annual Catastrophe Losses against PSEi Returns – Philippines**

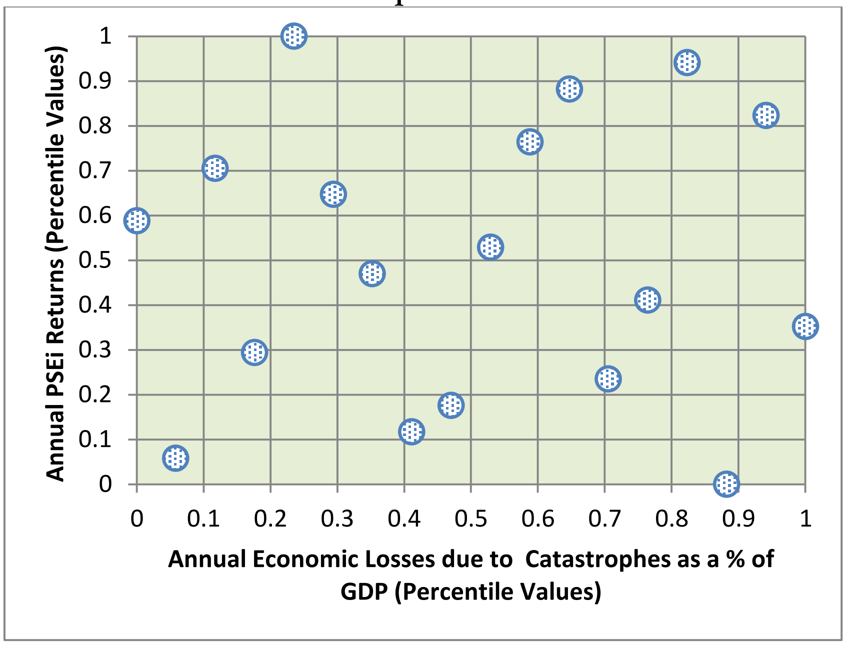
| Year | Economic Losses Due to Cats as % of GDP | Economic Losses Due to Cats(1) (USD MM) | PSEi Return | PSEi Rank (2) |
|---|---|---|---|---|
| **1993** | 0.209% | 456 | 152.0% | 1 |
| **2012** | 0.241% | 855 | 33.6% | 7 |
| **2009** | 0.295% | 876 | 62.9% | 3 |
| **1991** | 0.349% | 699 | 76.7% | 2 |
| **1988** | 0.400% | 673 | 2.8% | 17 |
| **1995** | 0.490% | 1,305 | -7.9% | 19 |
| **1990** | 0.588% | 1,134 | -41.2% | 24 |
| **2013** | 2.742% | 10,413 | 1.3% | 18 |

*(1) Source: EM-Dat; Years with no losses are excluded*
*(2) Rank from best to worst out of 26*

**Graph 3.3**
**Scatter Plot of Annual PSEi (Philippines Stock Index) Returns against Economic Losses due to Catastrophes as a % of GDP**

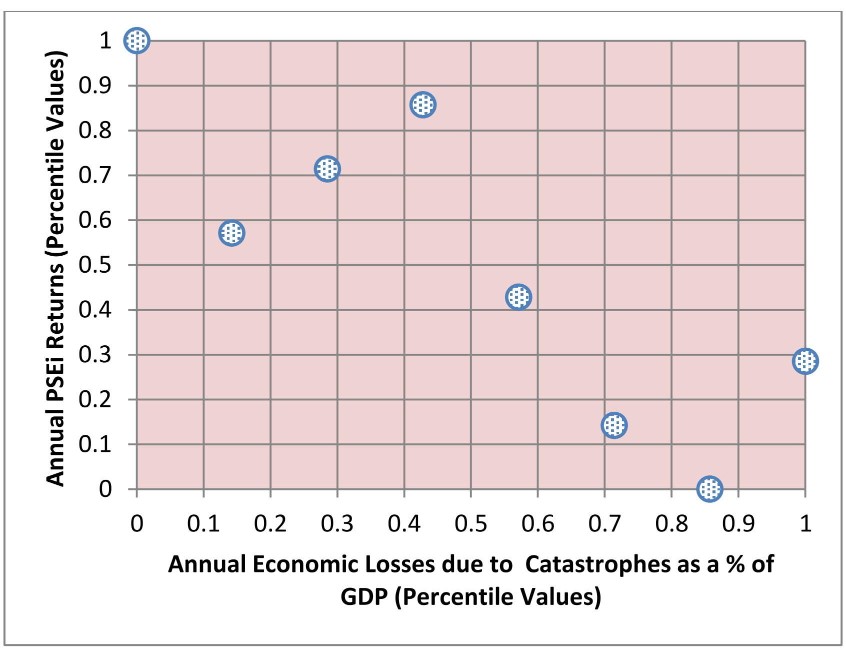


**Graph 3.3.a**
**Scatter Plot of Annual PSEi (Philippines) Returns against Economic Losses due to Catastrophes as a % of GDP for lowest 70th Percentile of Economic Losses due to Catastrophes as a % of GDP**




*Please note that the percentiles in the above graph have been recalculated based on the relative rankings of the subset of points that fall in the lowest 70th percentile of annual insured catastrophe losses as a % of GDP. As such, these percentiles range from 0 to 1.*

**Graph 3.3.b**
**Scatter Plot of Annual PSEi (Philippines) Returns against Economic Losses due to Catastrophe Losses as a % of GDP for highest 30th Percentile of Economic Losses due to Catastrophes as a % of GDP**




*Please note that the percentiles in the above graph have been recalculated based on the relative rankings of the subset of points that fall in the highest 30th percentile of annual insured catastrophe losses as a % of GDP. As such, these percentiles range from 0 to 1.*

d. We find no significant correlation between catastrophe losses and the returns on the Barclays Capital US bond indices shown below. We did not find any significant shift in correlation based on the relative size of catastrophe losses. This is shown in Table 3.7.a below. This finding is unchanged when we remove the effect of time on the Kendall's Tau correlations as shown in Table 3.7.b. Please note that the data for the Barclays Capital indices only goes back to 1973.

**Table 3.7.a**
**Annual Catastrophe Losses against Annual Returns on US Treasury, Agency, and Corporate Bonds**

| Catastrophe Losses | Barclays Capital Index | No. of Observations | Kendall's Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|
| **Insured Losses** | US Treasury | 41 | -9.8% | 0.188 | No |
| **Insured Losses** | US Intermediate Treasury | 41 | -16.6% | 0.066 | No |
| **Insured Losses** | US Long Treasury | 41 | 3.9% | 0.365 | No |
| **Insured Losses** | US Credit | 41 | -6.6% | 0.277 | No |
| **Insured Losses** | US Intermediate Credit | 41 | -10.5% | 0.172 | No |
| **Insured Losses** | US Long Credit | 41 | -0.7% | 0.479 | No |

**Table 3.7.b**
**Annual Catastrophe Losses against Annual Returns on US Treasury, Agency, and Corporate Bonds after Removing Effect of Time**

| Catastrophe Losses | Barclays Capital Index | No. of Observations | Kendall's Partial Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|
| **Insured Losses** | US Treasury | 41 | -4.06% | 0.355 | No |
| **Insured Losses** | US Intermediate Treasury | 41 | -8.10% | 0.229 | No |
| **Insured Losses** | US Long Treasury | 41 | 4.02% | 0.357 | No |
| **Insured Losses** | US Credit | 41 | -4.39% | 0.344 | No |
| **Insured Losses** | US Intermediate Credit | 41 | -6.03% | 0.291 | No |
| **Insured Losses** | US Long Credit | 41 | -1.08% | 0.462 | No |

e. We find no significant correlation between annual catastrophe losses and annual movements in crude oil prices as shown in Table 3.8.a below. We did not find any significant shift in correlation based on the relative size of catastrophe losses. This finding is unchanged when we remove the effect of time on the Kendall's Tau correlations as shown in Table 3.8.b.

**Table 3.8.a**
**Annual Catastrophe Losses against Annual Changes in Crude Oil Price – US**

| Catastrophe Losses | Index | No. of Observations | Kendall's Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|
| **Insured Losses** | Crude Oil | 63 | 9.9% | 0.126 | No |

**Table 3.8.b**

**Annual Catastrophe Losses against Annual Changes in Crude Oil Price after Removing Effect of Time – US**

| Catastrophe Losses | Index | No. of Observations | Kendall's Partial Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|
| **Insured Losses** | Crude Oil | 63 | 6.8% | 0.222 | No |

f. We find no significant correlation between annual catastrophe losses and annual movements in the US CPI as shown in Table 3.9.a below. We did not find any significant shift in correlation based on the relative size of catastrophe losses. This finding is unchanged when we remove the effect of time on the Kendall's Tau correlations as shown in Table 3.9.b.

**Table 3.9.a**
**Annual Catastrophe Losses against Annual Changes in the US CPI**

| Catastrophe Losses | Index | No. of Observations | Kendall's Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|
| **Insured Losses** | CPI | 64 | -6.0% | 0.242 | No |

**Table 3.9.b**
**Annual Catastrophe Losses against Annual Changes in the US CPI after Removing Effect of Time**

| Catastrophe Losses | Index | No. of Observations | Kendall's Partial Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|
| **Insured Losses** | CPI | 64 | -8.6% | 0.158 | No |

g. We find evidence of a negative correlation between annual catastrophe losses and annual changes in both nominal and real GDP as shown in Table 3.10.a below. We did not find any significant shift in correlation based on the relative size of catastrophe losses. However, when we remove the effect of time, the correlation between catastrophe losses and nominal GDP gets weaker while that between catastrophe losses and real GDP becomes statistically insignificant as shown in Table 3.10.b.

**Table 3.10.a**
**Annual Catastrophe Losses against Annual Changes in US GDP**

| Catastrophe Losses | Index | No. of Observations | Kendall's Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|
| **Insured Losses** | Nominal GDP | 64 | -25.6% | 0.001 | Yes |
| **Insured Losses** | Real GDP | 64 | -18.3% | 0.017 | Yes |

**Table 3.10.b**
**Annual Catastrophe Losses against Annual Changes in US GDP after Removing Effect of Time**

| Catastrophe Losses | Index | No. of Observations | Kendall's Partial Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|
| **Insured Losses** | Nominal GDP | 64 | -15.0% | 0.041 | Yes |
| **Insured Losses** | Real GDP | 64 | -6.6% | 0.226 | No |

**Table 3.11.a**
**P-values for Difference in the Kendall's Tau Values**

| Catastrophe Losses | Index | Percentile / # obs | Kendall's τ | Percentile / # obs | Kendall's τ | Kendall's Diff δ | One Tailed P-value | Is Difference Significant? |
|---|---|---|---|---|---|---|---|---|
| **Insured** | DJIA | <.80/51 | 17.3% | ≥.80/13 | -53.8% | -71.2% | - | Yes |
| **Insured** | DJCA | <.80/51 | 14.5% | ≥.80/13 | -41.0% | -55.5% | 0.007 | Yes |
| **Insured** | S&P 500 | <.80/51 | 12.9% | ≥.80/13 | -46.2% | -59.1% | 0.004 | Yes |
| **Economic** | DJIA | <.80/56 | 18.8% | ≥.80/15 | -48.6% | -67.4% | - | Yes |
| **Economic** | PSEi | <.7/18 | 3.3% | ≥.7/8 | -64.3% | -67.6% | 0.021 | Yes |

**Table 3.11.b**
**P-values for Difference in the Kendall's Partial Tau Values**

| Catastrophe Losses | Index | Percentile / # obs | Kendall's Partial τ | Percentile / # obs | Kendall's Partial τ | Kendall's Partial Diff δ | One Tailed P-value | Is Difference Significant? |
|---|---|---|---|---|---|---|---|---|
| **Insured** | DJIA | <.80/51 | 15.7% | ≥.80/13 | -56.2% | -71.9% | 0.001 | Yes |
| **Insured** | DJCA | <.80/51 | 13.9% | ≥.80/13 | -42.7% | -56.6% | 0.007 | Yes |
| **Insured** | S&P 500 | <.80/51 | 12.8% | ≥.80/13 | -48.7% | -61.5% | 0.004 | Yes |
| **Economic** | DJIA | <.80/56 | 18.2% | ≥.80/15 | -48.5% | -66.7% | 0.001 | Yes |
| **Economic** | PSEi | <.70/18 | 2.8% | ≥.70/8 | -64.5% | -67.3% | 0.023 | Yes |

## 4. COMMENTARY

There is a widely held view in the catastrophe insurance space that the capital markets are uncorrelated to catastrophe losses. This seems to hold true for fixed income securities in the US but not for equities according to our investigation. However, we can see how a casual evaluation of the data might tend to validate the conventional wisdom. Had we measured the Kendall's Tau and Kendall's Partial Tau statistics for US equities without taking into account the shifts in correlation, we might come to the conclusion that equities are indeed uncorrelated to catastrophes as shown in Tables 4.1.a and 4.1.b below.

**Table 4.1.a**
**Annual Catastrophe Losses against Annual Changes in Equities – US**

| Country | Catastrophe Losses | Index | No. of Observations | Kendall's Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|---|
| **US** | Insured Losses | DJCA | 64 | 4.9% | 0.285 | No |
| **US** | Insured Losses | DJIA | 64 | 6.8% | 0.212 | No |
| **US** | Insured Losses | S&P 500 | 64 | 6.3% | 0.229 | No |
| **US** | Economic Losses | DJIA | 71 | 7.9% | 0.164 | No |

**Table 4.1.b**
**Annual Catastrophe Losses against Annual Changes in Equities after Removing Effect of Time – US**

| Country | Catastrophe Losses | Index | No. of Observations | Kendall's Partial Tau | One Tailed P-value | Is Correlation Significant? |
|---|---|---|---|---|---|---|
| **US** | Insured Losses | DJCA | 64 | 5.5% | 0.265 | No |
| **US** | Insured Losses | DJIA | 64 | 6.6% | 0.226 | No |
| **US** | Insured Losses | S&P 500 | 64 | 6.9% | 0.215 | No |
| **US** | Economic Losses | DJIA | 71 | 8.0% | 0.164 | No |

When thinking about the relationship between US catastrophe losses and equities, say in terms of 2014 dollars, it helps to split the annual insured losses into two ranges: one below $16.5B and one above that. Similarly, annual economic losses can be split into two ranges: one below $38B and one above that. These thresholds represent approximately .096% and .22% of projected 2014 GDP[4] and correspond to the 80th percentile of annual insured and economic catastrophe losses, respectively. Below these thresholds, the correlation is either neutral or slightly positive. Above, the correlation is negative indicating that equity returns tend to deteriorate as the size of catastrophe losses increases. This deterioration entails weaker but not necessarily negative equity returns. Also, just as importantly, the deterioration is only relative to the 13 to 15 data points that fall in the range of the highest 20th percentile of catastrophe losses. For some institutions, most of the coverage they sell is only triggered for large enough catastrophe events so the correlation in the highest 20th percentile is really the most relevant.

The Kendall's Tau and Kendall's Partial Tau values above the 80th percentile thresholds are approximately -50% as shown in Tables 3.1.a and 3.1.b. These values imply that, once catastrophe losses are above these thresholds, equity returns are approximately three times more likely to deteriorate as the size of catastrophe losses increases than they are to improve[5].

---

[4] 2014 GDP is estimated at $17.3T by applying a growth rate of 2.8% to the 2013 GDP. This growth forecast is taken from the World Economic Outlook Update published on January 21, 2014 by the International Monetary Fund. http://www.imf.org/external/pubs/ft/weo/2014/update/01/pdf/0114.pdf

[5] The Kendall's Tau coefficient τ is equal to
$\frac{2(C-D)}{n(n-1)}$ where C and D represent the number of concordant and discordant pairs, respectively; $\frac{n(n-1)}{2}$

A discussion on how to model the dependence structure between assets and catastrophe losses is beyond the scope of this paper. Whatever the chosen modeling approach, it needs to be complemented by robust sensitivity and scenario testing.

There are many caveats and limitations to this study, some of which are discussed below:

a. Our findings only apply to calendar year data and should not be extrapolated to longer or shorter time periods. We did not necessarily observe the same degree of correlation when we studied some of the data for quarterly periods. A presentation of our findings based on quarterly data is beyond the scope of this paper.
b. This investigation neither demonstrates nor suggests any causation relationship between catastrophes and the various indices we evaluated even where the correlations are significant.
c. The percentages of GDP discussed in the preceding paragraphs are purely a function of the PCS and EM-Dat data sets. They should not be interpreted as some fixed, exact, or universal thresholds above which catastrophe losses and equity returns are negatively correlated.
d. The losses in the observation period are primarily from natural catastrophes with 9/11 being the most notable exception. We would not extrapolate the findings of this study to losses stemming from terrorist acts, especially those that involve nuclear, biological, chemical, or radioactive material.

We hope this paper provides but the beginnings of a robust discussion around the subject of dependence between catastrophes and assets. We have shared our data sources in Table A.1 of Appendix A hoping that others will take a critical look at the data in order to correct or augment our findings. We would be interested in expanding the analysis to more countries and to a broader set of financial and economic indices for the five countries we reviewed outside of the US. Finally, we would like to examine the sensitivity of our findings to the data on which we relied by looking into alternative data sources.

---

represent the total number of pair combinations. $\tau$ can be interpreted as follows: $\tau = \pi_c - \pi_d$, where $\pi_c = \frac{C}{n(n-1)/2}$ and $\pi_d = \frac{D}{n(n-1)/2}$ represent the probability of concordant and discordant pairs C and D, respectively.

Solving for the following equations: $\begin{cases} \pi_c - \pi_d = -.5 \\ \pi_c + \pi_d = 1 \end{cases}$ yields $\pi_c = .25 \; and \; \pi_d = .75$. Hence $\pi_d = 3\pi_c$. For a definition of the Kendall's Tau coefficient, see Gibbons, J.D. (1993, p. 11) *Nonparametric measures of association* (Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-091). Newbury Park, CA: Sage.

# APPENDIX A

# DATA

### A.1. United States

### Insured Losses

For insured losses in the United States, we relied on data from Verisk's PCS and their definition of catastrophe events. Losses are in actual dollars and not adjusted for price levels. The PCS insured loss catalog goes back to 1950. We understand that the threshold above which losses are captured in the PCS database has changed a few times since 1950. We are comfortable that the changes in these thresholds do not create any significant distortion in our analysis.

### Economic Losses

For economic losses in the US, we relied on data from the International Disaster Database (EM-DAT) published by the Center for Research on the Epidemiology of Disasters of the Université Catholique de Louvain in Brussels, Belgium. As with PCS's insured losses, the economic losses are in nominal dollars and are not adjusted for inflation. We only considered catastrophe events associated with windstorm, wildfire, earthquake, and flood. The EM-DAT loss catalog goes back to 1900 and is available on an annual aggregate basis. However, there were a number of years for which the estimated losses from the EM-DAT catalog amount to zero. We think that the prevalence of years with no losses could distort the statistics we use to measure correlation. As such, we performed our analysis excluding years with no losses.

### GDP

We used the historical GDP information available from the US Department of Commerce's Bureau of Economic Analysis. As noted above, we relied on the GDP in nominal (current) dollars unadjusted for price levels. The annual GDP information goes back to 1929.

### Financial and Economic Indices

The following indicators of financial and economic performance were used for the US:

**US Equities** – We analyzed the percentage change in the value of the Dow Jones Composite Average (DJCA), the Dow Jones Industrial Average (DJIA), and the S&P 500 as a proxy for the performance of US equities. We made no attempt to factor in dividend yields. We obtained the value of the indices at the daily market close for both the DJCA and the DJIA from the FRED database, which start from 1/3/1949 and 5/26/1896 for the DJCA and DJIA, respectively. The S&P 500 data was obtained from proprietary sources but is widely available across a number of different sources.

**US Treasury Bonds** – We obtained the annual returns on the Barclays Capital US Treasury, US

Intermediate Treasury, and US Long Treasury indices. The return information spans 41 years dating back to 1973.

According to the website http://etfdb.com/, the Barclays Capital U.S. Treasury Index includes all publicly issued, U.S. Treasury securities that are rated investment grade, and have $250 million or more of outstanding face value. The Barclays Capital US Intermediate Treasury and US Long Treasury Indices have a remaining maturity of between 1 and 10 years, and 10 or more years, respectively.

**US Corporate Bonds** – We obtained the annual returns on the Barclays Capital US Credit, US Credit Intermediate, and US Credit Long indices. The return information spans 41 years dating back to 1973.

According to the website http://etfdb.com/, the Barclays Capital US Credit and US Intermediate Credit indices measure the performance of investment grade corporate debt and agency bonds that are dollar denominated and have a remaining maturity of greater than one year, and between more than one year and ten years, respectively. The Barclays Capital U.S. Long Credit Index measures the performance of the long term sector of the United States investment bond market, which as defined by the Long Credit Index includes investment grade corporate debt and sovereign, supranational, local authority and non-U.S. agency bonds that are dollar denominated and have a remaining maturity of greater than or equal to 10 years.

**Oil** – We analyzed the changes in the spot price for West Texas Intermediate crude oil. We obtained the information at quarterly intervals dating back to 1/1/1946 from the FRED database. This data series has been discontinued as of 7/1/2013.

**GDP Growth** – We analyzed the changes in both nominal and Gross GDP. The GDP information is obtained from the Bureau of Economic Analysis as explained above.

**CPI** – We used the annual average CPI for all urban consumers as published by the Minneapolis Fed. The data goes back to 1913.

### Observation Frequency

We studied the correlation of data observed over annual periods. We paired the annual aggregate catastrophe losses with changes in the index value taken over the same period. For instance, annual losses incurred in 1969 are paired with the changes in asset prices observed from January 1 to December 31 of 1969.

### A.2. Australia, Chile, Japan, Thailand, and the Philippines

Similar to the US data, we expressed annual economic losses incurred in a year as a percentage of Gross Domestic Product (GDP) in the same year. For Chile, Japan, Thailand, and the Philippines, both economic loss and GDP figures are adjusted to 2005 US price levels. For Australia, these figures are adjusted to 2011 US price levels.

**Economic Losses**

For all five countries, we relied exclusively on the economic loss data available from the International Disaster Database (EM-DAT) mentioned above. This loss information is originally provided in nominal US dollars. Because the corresponding GDP information we used for these countries is adjusted to 2011 and 2005 price levels for Australia and the remaining countries, respectively, we brought the economic losses to the price levels corresponding to the GDP using US CPI data obtained from the US Bureau of Labor Statistics. We only considered catastrophe events associated with windstorm, wildfire, earthquake, and flood for Chile and Japan. We also included industrial accidents for Australia, Thailand and the Philippines. The EM-DAT loss catalog goes back to 1900 and is only available on an annual aggregate basis. However, there were a number of years for which the estimated losses from the EM-DAT catalog amounted to zero. For many of those years, it appears that there were significant catastrophes for which the economic loss data was not recorded or not available. We think that the prevalence of years with no losses could distort the statistics we use to measure correlation. As such, we performed our analysis excluding years with no losses. Please note that the standards and sources used by EM-Dat to collect and measure economic losses due to catastrophes may vary significantly by country. Also, the EM-Dat data may not match corresponding statistics collected by other local and international agencies.

**GDP**

We used the historical GDP information available from a database established by the Economic Research Division of the Federal Reserve Bank of St. Louis's (FRED). As noted above, this information is already expressed in 2011 USD for Australia and in 2005 USD for the remaining four countries. The FRED GDP data only goes up to 2011 so we extrapolated the GDP figures for 2012 and 2013 for each country by applying to the 2011 GDP figure the GDP growth rates corresponding to 2012 and 2013 obtained from various sources. While this extrapolation is an oversimplification of the correct calculation of GDP in 2011 or 2005 US prices, we believe it is adequate for our purposes. We also extrapolated the 2005 level GDP for years prior to 1950 for Japan and prior to 1951 for Chile by using the growth rates corresponding to these prior years observed from information available from a database established by the Maddison Project.

**Financial and Economic Indices**

We only looked at equity performance for Australia, Chile, Japan, Thailand, and The Philippines.

**Equities** – We analyzed the price changes for all shares in Australia, for the Nikkei 225 (Japan), IGPA (Chile), SET (Thailand), and PSEi (Philippines) indices as a proxy for the performance of equities in those countries. Similar to US equities, we made no attempt to factor in dividend yields.

The total share prices for all shares for Australia was obtained at annual periods from the FRED database along with the daily market close for the Nikkei 225 index. The IGPA data was obtained from proprietary sources while the SET and PSEi historical data were obtained from Bloomberg.

**Summary of Data Sources**

Table A.1 below provides a comprehensive list of our data sources, most of which are available publicly. We hope that others will use this data to rectify or augment our findings.

**Table A.1**
**Data Source List**

| Data | Source | Web Address |
|---|---|---|
| **Insured Catastrophe Losses – US** | PCS - Verisk | Subscription |
| **Economic Catastrophe Losses – Australia, Chile, Japan, Philippines, Thailand, US** | International Disaster Database | http://www.emdat.be/database |
| **GDP – US** | US Department of Commerce | http://www.bea.gov/ |
| **GDP – Australia, Chile, Japan, Philippines, Thailand** | St Louis Fed | https://research.stlouisfed.org/ |
| **DJCA, DJIA, Nikkei 225, Australia Total Share Price Index** | St Louis Fed | https://research.stlouisfed.org/ |
| **S&P 500** | Proprietary | Subscription |
| **IGPA (Chile Stock Index)** | Proprietary | Subscription |
| **SET: Index (Thailand), PSEi (Philippines)** | Bloomberg | Subscription |
| **Barclays Capital US Treasury and US Credit Indices** | Barclays Capital | Subscription |
| **West Texas Oil Spot Rate** | St Louis Fed | https://research.stlouisfed.org/ |
| **US CPI** | Minneapolis Fed | https://www.minneapolisfed.org/ |

# APPENDIX B

# CALCULATION OF P-VALUES

### B.1 Kendall's Tau

We used the Kendall's Tau τ statistic to measure the correlations. For the null hypothesis, we posit that catastrophe losses and the performance of the financial and economic indices are independent. We used a simulation to generate the distributions, under the null hypothesis, of the Kendall's Tau corresponding to each specific number of observations. For instance, where we have 64 years of observations, we calculated the Kendall's Tau based on the simulation of 64 pairs of independent variables uniformly distributed on [0,1]. We derived the P-values from these distributions, which are summarized in Table B.1.a below for various observation counts. Alternatively, for a large enough number of observations, P-values can be determined by assuming the Kendall's Tau to be approximately normally distributed under the null hypothesis with a mean of zero and variance given by $\frac{2(2n+5)}{9n(n-1)}$.

### B.2 Kendall's Partial Tau

We used the Kendall's Partial Tau τ statistic to measure the correlations after removing the effect of time. For the null hypothesis, we posit that catastrophe losses and the performance of the financial and economic indices are independent after removing the effect of time. We used a simulation to generate the distributions, under the null hypothesis, of the Kendall's Partial Tau corresponding to each specific number of observations. For instance, where we have 64 years of observations, we simulated 64 pairs, say X and Y, of independent variables uniformly distributed on [0,1] for each of the 64 years, say Z. We calculated the Kendall's Partial Tau based on the triplets (X,Y,Z) as described in the Overview of Approach in section 2 above[6]. We derived the P-values from these distributions, which are summarized in Table B.1.b below for various observation counts.

### B.3 Difference in Kendall's Tau and Kendall's Partial Tau

We calculate the differences δ in the Kendall's Tau (Kendall's Partial Tau) statistics for the two different ranges of percentile value under consideration. For the null hypothesis, we posit that the Kendall's Tau (Kendall's Partial Tau) is zero across the entire range of percentile values. We used a simulation to generate the distributions, under the null hypothesis, of the differences in the Kendall's Tau (Kendall's Partial Tau) statistics for the two different ranges of percentile values. For instance, where we have 64 years of observations and want to compare the Kendall's Tau (Kendall's

[6] The P-values as calculated assume X, Y, and Z are mutually independent. Further work is needed to calculate the P-values when the assumption of independence is violated, as is the case for some of the variables we reviewed.

Partial Tau) for the first 80th percentile to the Kendall's Tau (Kendall's Partial Tau) for the last 20th percentile, we calculated the difference in the Kendall's Tau (Kendall's Partial Tau) for the lowest 51st observations ranked by size of catastrophe to the Kendall's Tau (Kendall's Partial Tau) for the highest 13 observations. We derived the P-values from the distributions of δ obtained through the simulations. Tables B.2.a and B.2.b show the distributions of δ for various observation counts.

**Table B.1.a**
**Simulation-based Distribution of Kendall's Tau Statistic under Null Hypothesis by Number of Observations**

| Observations | 8 | 13 | 15 | 18 | 22 | 30 | 38 |
|---|---|---|---|---|---|---|---|
| **Mean** | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| **Std Dev** | 0.288 | 0.210 | 0.192 | 0.173 | 0.153 | 0.129 | 0.113 |
| **1.00%** | (0.643) | (0.487) | (0.448) | (0.399) | (0.359) | (0.301) | (0.263) |
| **5.00%** | (0.500) | (0.359) | (0.314) | (0.281) | (0.255) | (0.214) | (0.186) |
| **10.00%** | (0.357) | (0.282) | (0.257) | (0.229) | (0.195) | (0.168) | (0.147) |
| **15.00%** | (0.286) | (0.231) | (0.200) | (0.176) | (0.160) | (0.136) | (0.118) |
| **20.00%** | (0.286) | (0.179) | (0.162) | (0.150) | (0.134) | (0.108) | (0.095) |
| **25.00%** | (0.214) | (0.154) | (0.124) | (0.124) | (0.108) | (0.090) | (0.078) |
| **30.00%** | (0.143) | (0.103) | (0.105) | (0.098) | (0.082) | (0.067) | (0.061) |
| **35.00%** | (0.143) | (0.077) | (0.067) | (0.072) | (0.056) | (0.053) | (0.044) |
| **40.00%** | (0.071) | (0.051) | (0.048) | (0.046) | (0.039) | (0.034) | (0.030) |
| **45.00%** | (0.071) | (0.026) | (0.029) | (0.020) | (0.022) | (0.016) | (0.016) |
| **50.00%** | - | - | (0.010) | (0.007) | (0.004) | (0.002) | (0.001) |
| **55.00%** | 0.071 | 0.026 | 0.029 | 0.020 | 0.022 | 0.016 | 0.013 |
| **60.00%** | 0.071 | 0.051 | 0.048 | 0.046 | 0.039 | 0.034 | 0.030 |
| **65.00%** | 0.143 | 0.077 | 0.067 | 0.072 | 0.056 | 0.048 | 0.044 |
| **70.00%** | 0.143 | 0.103 | 0.105 | 0.085 | 0.082 | 0.067 | 0.058 |
| **75.00%** | 0.214 | 0.154 | 0.124 | 0.111 | 0.100 | 0.085 | 0.075 |
| **80.00%** | 0.214 | 0.179 | 0.162 | 0.150 | 0.126 | 0.108 | 0.095 |
| **85.00%** | 0.286 | 0.231 | 0.200 | 0.176 | 0.160 | 0.136 | 0.118 |
| **90.00%** | 0.357 | 0.282 | 0.238 | 0.216 | 0.195 | 0.163 | 0.144 |
| **95.00%** | 0.500 | 0.333 | 0.314 | 0.281 | 0.255 | 0.209 | 0.186 |
| **99.00%** | 0.643 | 0.487 | 0.448 | 0.399 | 0.351 | 0.297 | 0.260 |

**Table B.1.a (continues)**
**Simulation-based Distribution of Kendall's Tau Statistic under Null Hypothesis by Number of Observations**

| Observations | 41 | 44 | 51 | 56 | 63 | 64 | 71 |
|---|---|---|---|---|---|---|---|
| **Mean** | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| **Std Dev** | 0.109 | 0.104 | 0.097 | 0.092 | 0.087 | 0.086 | 0.081 |
| **1.00%** | (0.254) | (0.241) | (0.225) | (0.214) | (0.200) | (0.198) | (0.189) |
| **5.00%** | (0.180) | (0.173) | (0.159) | (0.152) | (0.142) | (0.142) | (0.134) |
| **10.00%** | (0.139) | (0.135) | (0.125) | (0.119) | (0.111) | (0.111) | (0.105) |
| **15.00%** | (0.112) | (0.108) | (0.101) | (0.096) | (0.091) | (0.089) | (0.085) |
| **20.00%** | (0.093) | (0.089) | (0.082) | (0.078) | (0.073) | (0.072) | (0.069) |
| **25.00%** | (0.073) | (0.072) | (0.067) | (0.062) | (0.059) | (0.058) | (0.055) |
| **30.00%** | (0.059) | (0.055) | (0.051) | (0.049) | (0.046) | (0.046) | (0.043) |
| **35.00%** | (0.041) | (0.040) | (0.038) | (0.035) | (0.033) | (0.034) | (0.031) |
| **40.00%** | (0.029) | (0.027) | (0.024) | (0.023) | (0.022) | (0.022) | (0.021) |
| **45.00%** | (0.015) | (0.013) | (0.012) | (0.012) | (0.011) | (0.011) | (0.010) |
| **50.00%** | - | - | (0.001) | - | (0.001) | - | 0.000 |
| **55.00%** | 0.012 | 0.013 | 0.012 | 0.012 | 0.011 | 0.011 | 0.010 |
| **60.00%** | 0.027 | 0.025 | 0.024 | 0.023 | 0.022 | 0.022 | 0.021 |
| **65.00%** | 0.041 | 0.040 | 0.037 | 0.035 | 0.033 | 0.033 | 0.031 |
| **70.00%** | 0.056 | 0.055 | 0.051 | 0.048 | 0.046 | 0.045 | 0.042 |
| **75.00%** | 0.073 | 0.070 | 0.065 | 0.062 | 0.058 | 0.058 | 0.054 |
| **80.00%** | 0.090 | 0.087 | 0.081 | 0.077 | 0.072 | 0.071 | 0.068 |
| **85.00%** | 0.112 | 0.108 | 0.100 | 0.095 | 0.090 | 0.088 | 0.084 |
| **90.00%** | 0.139 | 0.133 | 0.123 | 0.118 | 0.111 | 0.109 | 0.104 |
| **95.00%** | 0.178 | 0.171 | 0.159 | 0.151 | 0.142 | 0.141 | 0.133 |
| **99.00%** | 0.251 | 0.241 | 0.225 | 0.214 | 0.201 | 0.199 | 0.190 |

**Table B.1.b**
**Simulation-based Distribution of Kendall's Partial Tau Statistic under Null Hypothesis by Number of Observations**

| Observations | 8 | 13 | 15 | 18 | 22 | 30 | 38 |
|---|---|---|---|---|---|---|---|
| **Mean** | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| **Std Dev** | 0.292 | 0.211 | 0.193 | 0.173 | 0.154 | 0.129 | 0.113 |
| **1.00%** | (0.645) | (0.485) | (0.443) | (0.399) | (0.356) | (0.299) | (0.264) |
| **5.00%** | (0.486) | (0.351) | (0.318) | (0.285) | (0.254) | (0.214) | (0.187) |
| **10.00%** | (0.382) | (0.275) | (0.249) | (0.223) | (0.198) | (0.166) | (0.147) |
| **15.00%** | (0.309) | (0.223) | (0.202) | (0.181) | (0.161) | (0.135) | (0.119) |
| **20.00%** | (0.253) | (0.180) | (0.165) | (0.147) | (0.131) | (0.110) | (0.097) |
| **25.00%** | (0.207) | (0.145) | (0.133) | (0.118) | (0.105) | (0.088) | (0.077) |
| **30.00%** | (0.162) | (0.113) | (0.104) | (0.092) | (0.082) | (0.069) | (0.061) |
| **35.00%** | (0.125) | (0.083) | (0.076) | (0.068) | (0.061) | (0.051) | (0.045) |
| **40.00%** | (0.076) | (0.055) | (0.050) | (0.045) | (0.040) | (0.034) | (0.030) |
| **45.00%** | (0.042) | (0.028) | (0.025) | (0.022) | (0.020) | (0.017) | (0.015) |
| **50.00%** | - | (0.000) | (0.000) | (0.001) | (0.001) | (0.001) | (0.000) |
| **55.00%** | 0.042 | 0.026 | 0.025 | 0.021 | 0.019 | 0.016 | 0.014 |
| **60.00%** | 0.075 | 0.053 | 0.049 | 0.044 | 0.039 | 0.032 | 0.028 |
| **65.00%** | 0.125 | 0.082 | 0.075 | 0.067 | 0.059 | 0.049 | 0.044 |
| **70.00%** | 0.162 | 0.113 | 0.102 | 0.090 | 0.081 | 0.067 | 0.059 |
| **75.00%** | 0.207 | 0.146 | 0.132 | 0.117 | 0.104 | 0.087 | 0.077 |
| **80.00%** | 0.253 | 0.180 | 0.164 | 0.146 | 0.130 | 0.109 | 0.095 |
| **85.00%** | 0.309 | 0.221 | 0.200 | 0.180 | 0.160 | 0.134 | 0.117 |
| **90.00%** | 0.380 | 0.272 | 0.247 | 0.222 | 0.196 | 0.165 | 0.145 |
| **95.00%** | 0.486 | 0.346 | 0.316 | 0.286 | 0.252 | 0.211 | 0.185 |
| **99.00%** | 0.645 | 0.481 | 0.442 | 0.398 | 0.351 | 0.294 | 0.261 |

**Table B.1.b (continues)**
**Simulation-based Distribution of Kendall's Partial Tau Statistic under Null Hypothesis by Number of Observations**

| Observations | 41 | 44 | 51 | 56 | 63 | 64 | 71 |
|---|---|---|---|---|---|---|---|
| **Mean** | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| **Std Dev** | 0.109 | 0.104 | 0.097 | 0.092 | 0.087 | 0.086 | 0.081 |
| **1.00%** | (0.254) | (0.241) | (0.226) | (0.214) | (0.201) | (0.199) | (0.189) |
| **5.00%** | (0.180) | (0.173) | (0.160) | (0.152) | (0.143) | (0.142) | (0.134) |
| **10.00%** | (0.140) | (0.134) | (0.125) | (0.119) | (0.112) | (0.111) | (0.105) |
| **15.00%** | (0.113) | (0.109) | (0.101) | (0.096) | (0.090) | (0.089) | (0.085) |
| **20.00%** | (0.092) | (0.088) | (0.082) | (0.078) | (0.073) | (0.072) | (0.069) |
| **25.00%** | (0.074) | (0.071) | (0.066) | (0.063) | (0.058) | (0.058) | (0.055) |
| **30.00%** | (0.058) | (0.055) | (0.052) | (0.049) | (0.046) | (0.045) | (0.043) |
| **35.00%** | (0.043) | (0.041) | (0.038) | (0.036) | (0.034) | (0.033) | (0.031) |
| **40.00%** | (0.029) | (0.027) | (0.025) | (0.024) | (0.022) | (0.022) | (0.021) |
| **45.00%** | (0.014) | (0.014) | (0.012) | (0.012) | (0.011) | (0.011) | (0.010) |
| **50.00%** | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | 0.000 |
| **55.00%** | 0.013 | 0.013 | 0.012 | 0.011 | 0.011 | 0.011 | 0.010 |
| **60.00%** | 0.027 | 0.026 | 0.025 | 0.023 | 0.022 | 0.022 | 0.021 |
| **65.00%** | 0.042 | 0.039 | 0.037 | 0.035 | 0.033 | 0.033 | 0.031 |
| **70.00%** | 0.057 | 0.054 | 0.051 | 0.048 | 0.045 | 0.045 | 0.042 |
| **75.00%** | 0.073 | 0.070 | 0.065 | 0.062 | 0.058 | 0.058 | 0.054 |
| **80.00%** | 0.092 | 0.088 | 0.081 | 0.077 | 0.072 | 0.072 | 0.068 |
| **85.00%** | 0.113 | 0.108 | 0.100 | 0.095 | 0.089 | 0.089 | 0.084 |
| **90.00%** | 0.139 | 0.134 | 0.124 | 0.118 | 0.111 | 0.110 | 0.104 |
| **95.00%** | 0.178 | 0.172 | 0.159 | 0.151 | 0.142 | 0.141 | 0.133 |
| **99.00%** | 0.250 | 0.241 | 0.225 | 0.214 | 0.201 | 0.200 | 0.189 |

**Table B.2.a**
**Simulation-based Distribution of Difference δ between Kendall's Tau Statistics under Null Hypothesis by Number of Observations**

| Observation Splits | 18/8 | 51/13 | 56/15 |
|---|---|---|---|
| **Mean** | (0.001) | 0.000 | 0.000 |
| **Std Dev** | 0.336 | 0.231 | 0.213 |
| **1.00%** | (0.769) | (0.530) | (0.491) |
| **5.00%** | (0.553) | (0.381) | (0.351) |
| **10.00%** | (0.436) | (0.298) | (0.275) |
| **15.00%** | (0.352) | (0.242) | (0.223) |
| **20.00%** | (0.287) | (0.197) | (0.181) |
| **25.00%** | (0.233) | (0.158) | (0.145) |
| **30.00%** | (0.182) | (0.123) | (0.113) |
| **35.00%** | (0.131) | (0.090) | (0.083) |
| **40.00%** | (0.089) | (0.059) | (0.054) |
| **45.00%** | (0.045) | (0.029) | (0.027) |
| **50.00%** | (0.000) | 0.000 | 0.000 |
| **55.00%** | 0.040 | 0.030 | 0.028 |
| **60.00%** | 0.085 | 0.059 | 0.055 |
| **65.00%** | 0.130 | 0.090 | 0.084 |
| **70.00%** | 0.176 | 0.123 | 0.113 |
| **75.00%** | 0.228 | 0.158 | 0.145 |
| **80.00%** | 0.286 | 0.197 | 0.181 |
| **85.00%** | 0.352 | 0.242 | 0.223 |
| **90.00%** | 0.435 | 0.299 | 0.275 |
| **95.00%** | 0.553 | 0.381 | 0.351 |
| **99.00%** | 0.762 | 0.532 | 0.491 |

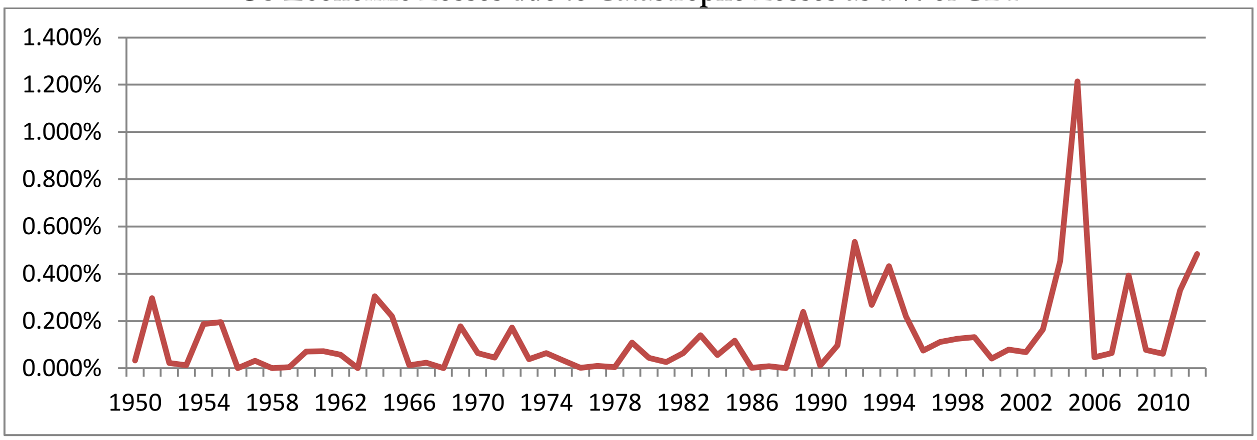
**Table B.2.b**
**Simulation-based Distribution of Difference δ between Kendall's Partial Tau Statistics under Null Hypothesis by Number of Observations**

| Observation Splits | 18/8 | 51/13 | 56/15 |
|---|---|---|---|
| **Mean** | 0.001 | 0.001 | 0.000 |
| **Std Dev** | 0.339 | 0.231 | 0.213 |
| **1.00%** | (0.768) | (0.530) | (0.492) |
| **5.00%** | (0.559) | (0.380) | (0.351) |
| **10.00%** | (0.439) | (0.297) | (0.275) |
| **15.00%** | (0.357) | (0.241) | (0.222) |
| **20.00%** | (0.289) | (0.197) | (0.180) |
| **25.00%** | (0.233) | (0.159) | (0.145) |
| **30.00%** | (0.181) | (0.123) | (0.113) |
| **35.00%** | (0.132) | (0.090) | (0.083) |
| **40.00%** | (0.085) | (0.059) | (0.054) |
| **45.00%** | (0.040) | (0.028) | (0.027) |
| **50.00%** | 0.003 | 0.001 | 0.001 |
| **55.00%** | 0.047 | 0.031 | 0.028 |
| **60.00%** | 0.091 | 0.061 | 0.056 |
| **65.00%** | 0.137 | 0.092 | 0.084 |
| **70.00%** | 0.186 | 0.125 | 0.114 |
| **75.00%** | 0.237 | 0.160 | 0.146 |
| **80.00%** | 0.293 | 0.198 | 0.181 |
| **85.00%** | 0.358 | 0.242 | 0.222 |
| **90.00%** | 0.438 | 0.298 | 0.275 |
| **95.00%** | 0.557 | 0.379 | 0.350 |
| **99.00%** | 0.769 | 0.528 | 0.488 |

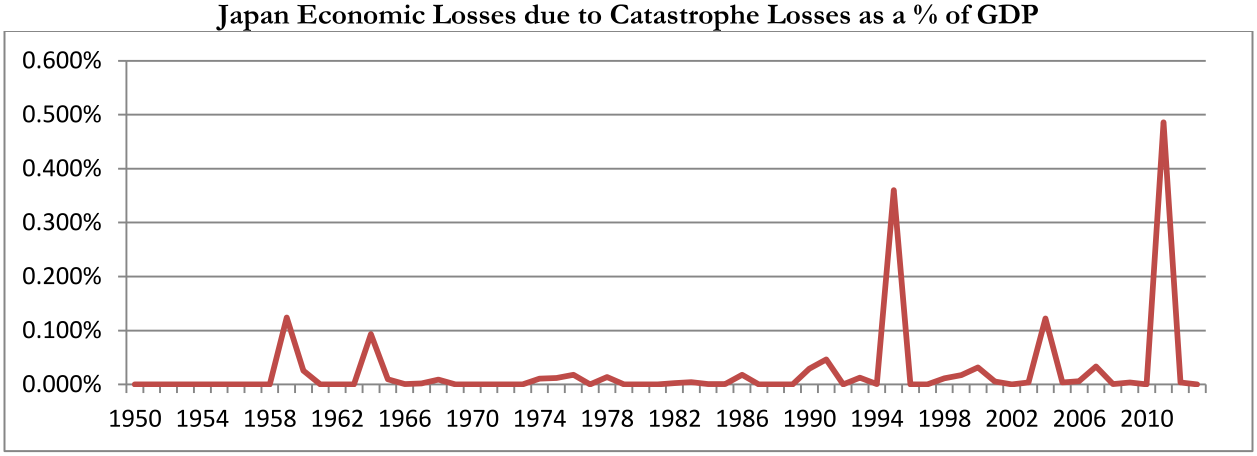
# APPENDIX C

# SELECTED GRAPHS

**Graph C.1**
**US Economic Losses due to Catastrophe Losses as a % of GDP**

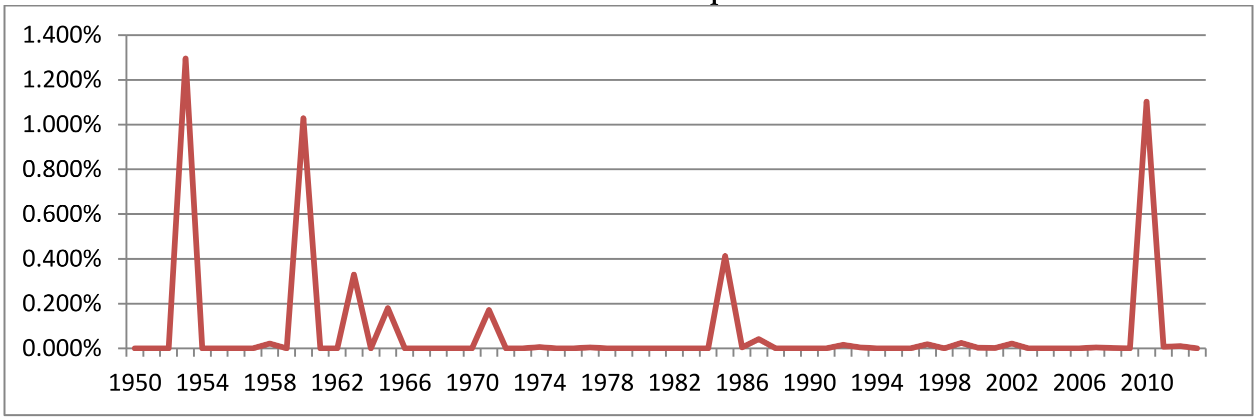



**Graph C.2**
**Japan Economic Losses due to Catastrophe Losses as a % of GDP**

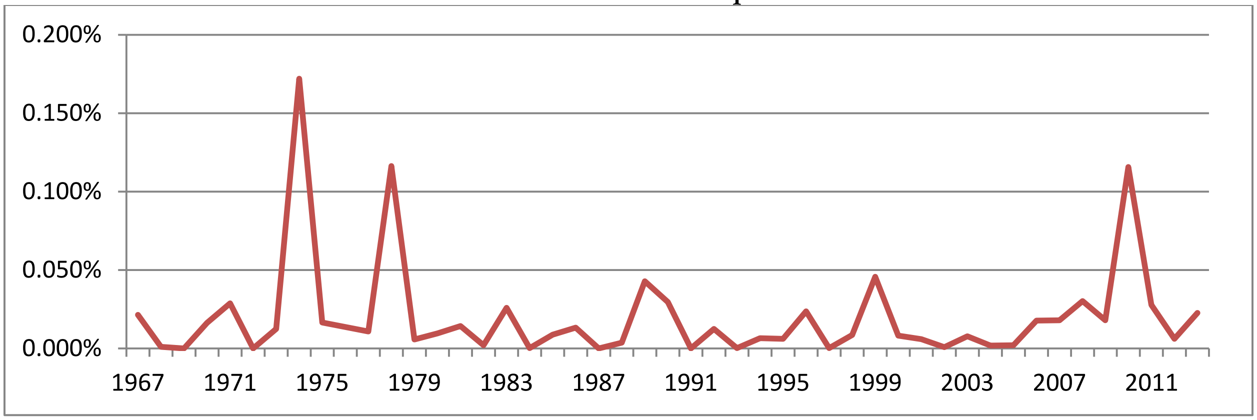



*Note: When comparing graphs across countries, please note that the standards and sources used by EM-Dat to collect and measure economic losses due to catastrophes may vary significantly by country. The data may not match corresponding statistics collected by other local and international agencies.*

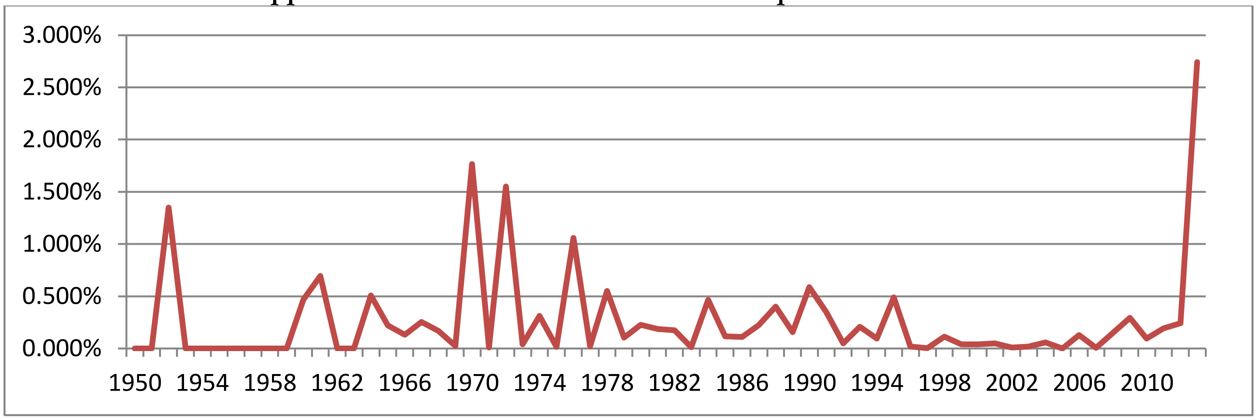
**Graph C.3**
**Chile Economic Losses due to Catastrophe Losses as a % of GDP**




**Graph C.4**
**Australia Economic Losses due to Catastrophe Losses as a % of GDP**

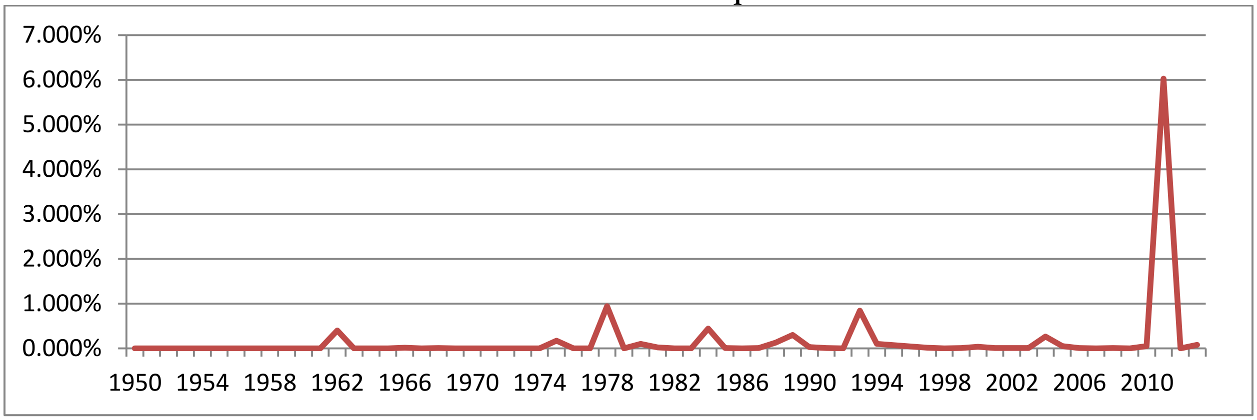



**Graph C.5**
**The Philippines Economic Losses due to Catastrophe Losses as a % of GDP**



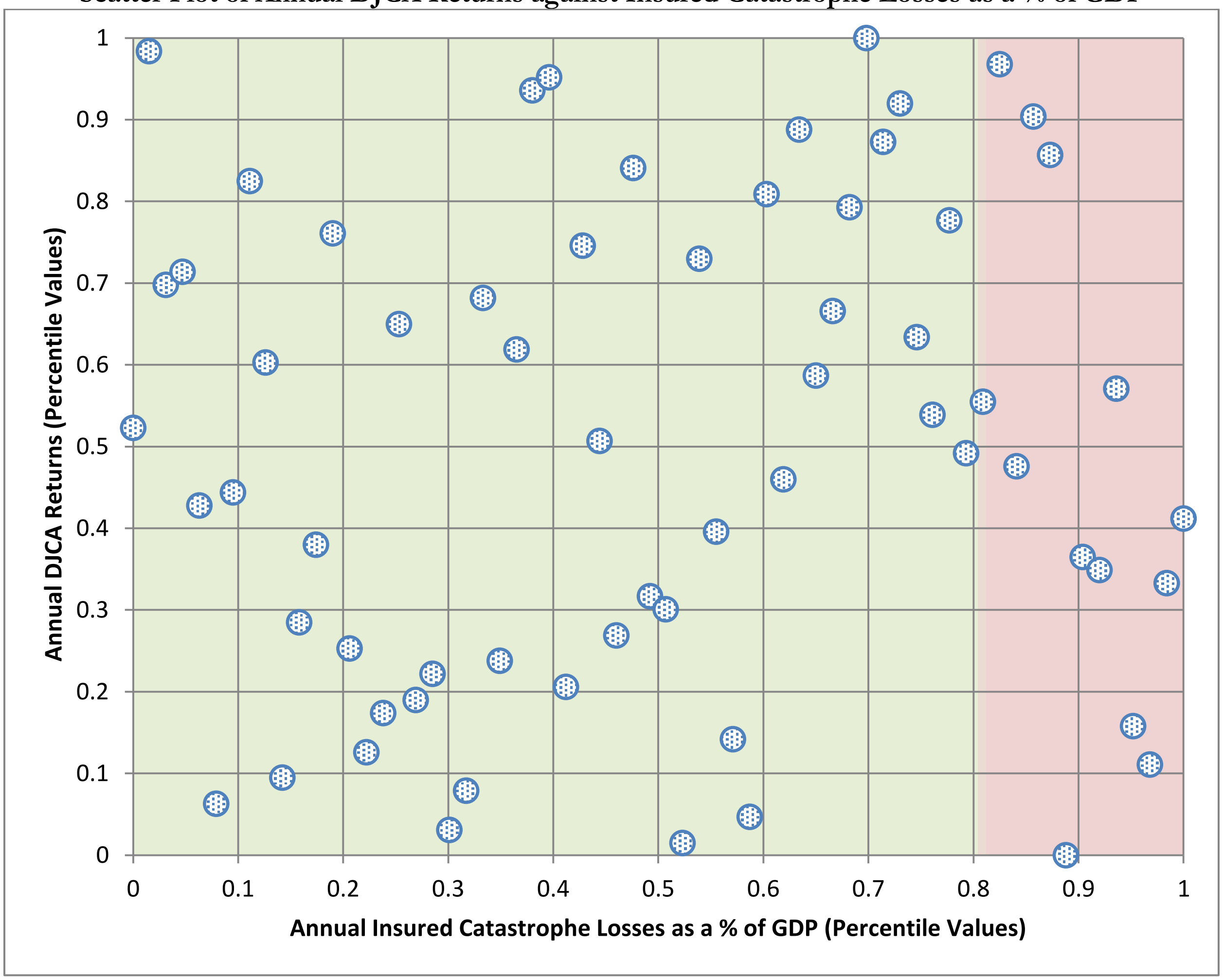

*Note: When comparing graphs across countries, please note that the standards and sources used by EM-Dat to collect and measure economic losses due to catastrophes may vary significantly by country. The data may not match corresponding statistics collected by other local and international agencies.*

**Graph C.6**
**Thailand Economic Losses due to Catastrophe Losses as a % of GDP**

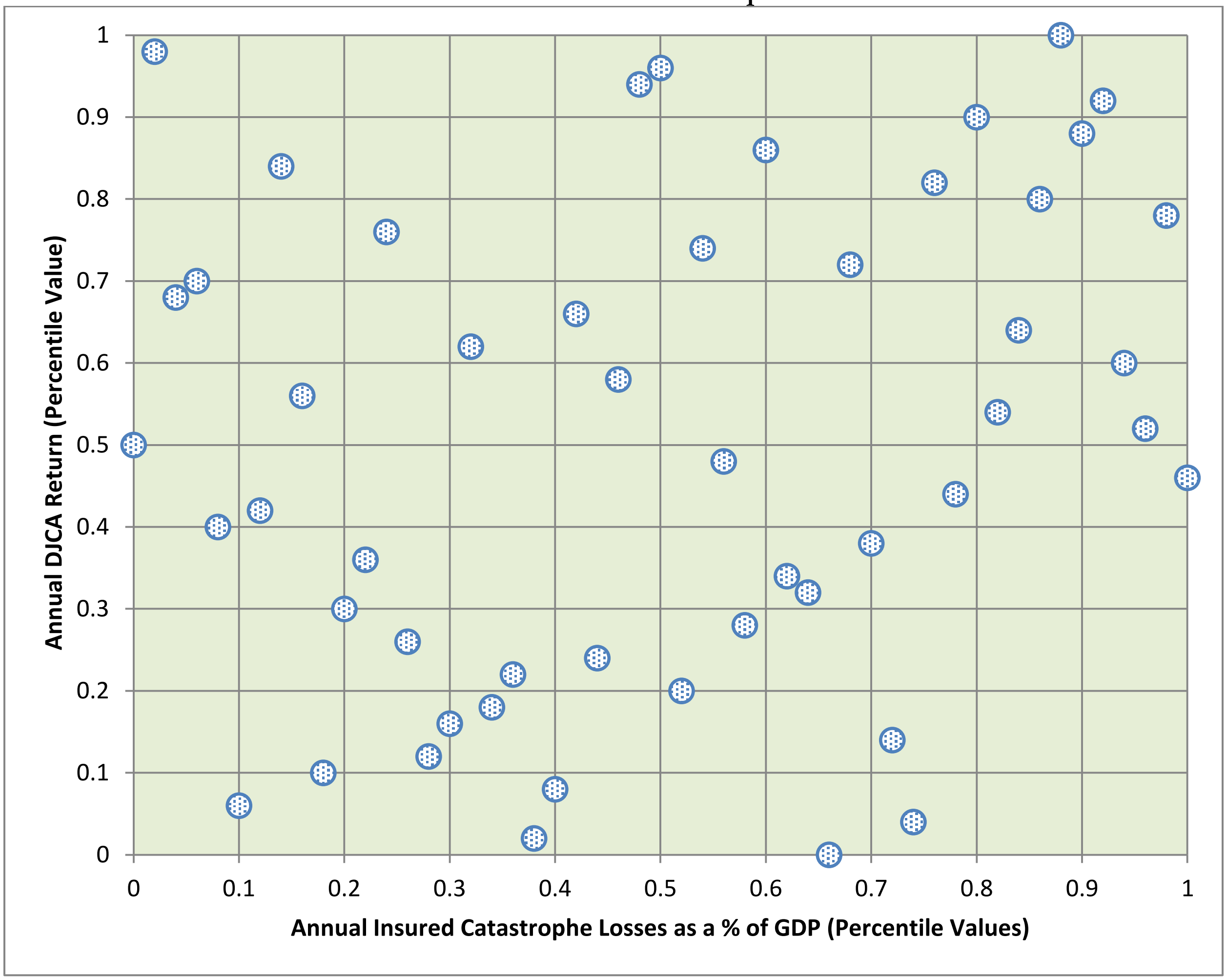



*Note: When comparing graphs across countries, please note that the standards and sources used by EM-Dat to collect and measure economic losses due to catastrophes may vary significantly by country. The data may not match corresponding statistics collected by other local and international agencies.*

**Graph C.7**
**Scatter Plot of Annual DJCA Returns against Insured Catastrophe Losses as a % of GDP**

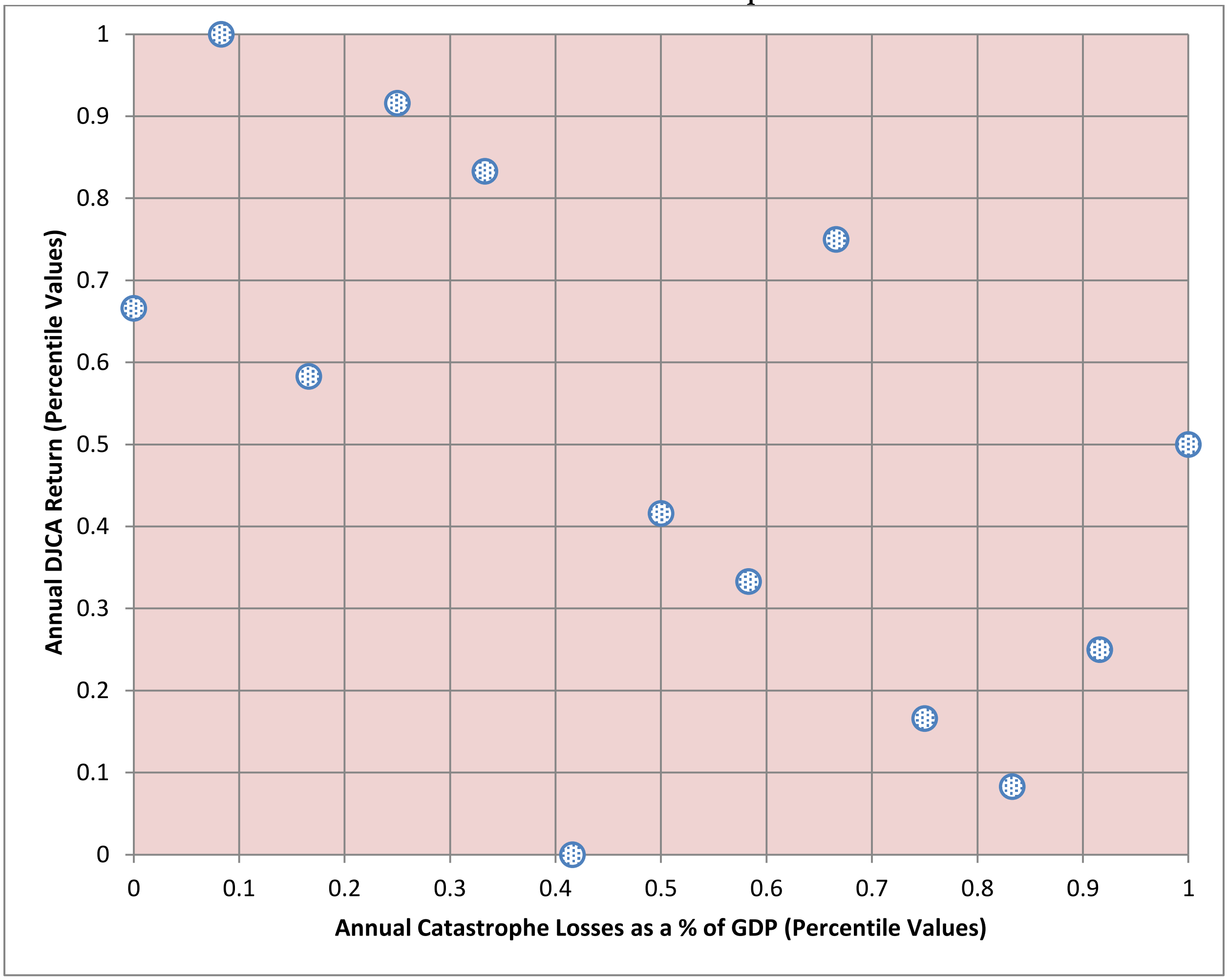


**Graph C.7.a**
**Scatter Plot of Annual DJCA Returns against Insured Catastrophe Losses as a % of GDP for lowest 80th Percentile of Annual Insured Catastrophes as a % of GDP**



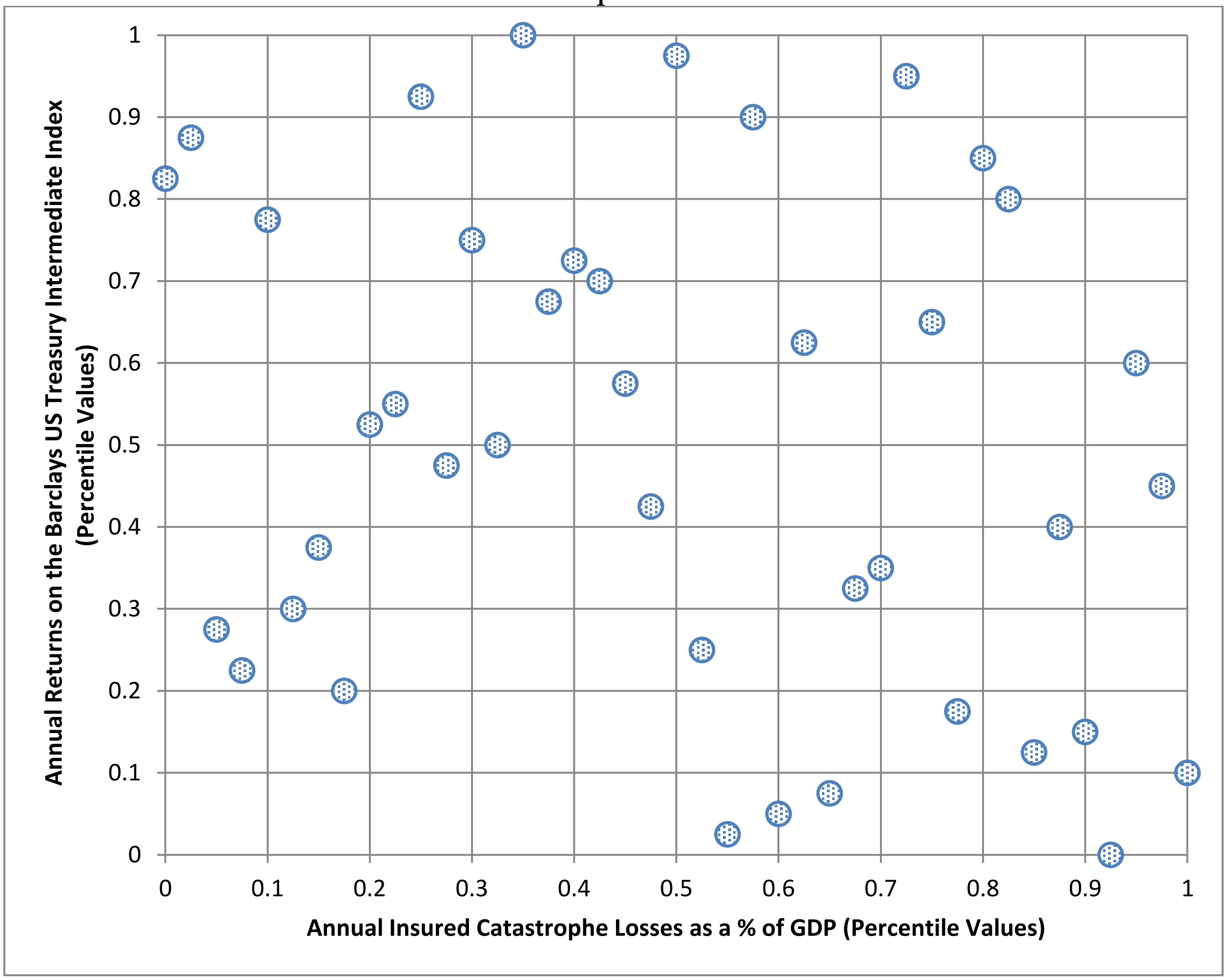

*Please note that the percentiles in the above graph have been recalculated based on the relative rankings of the subset of points that fall in the lowest 80th percentile of annual insured catastrophe losses as a % of GDP. As such, these percentiles range from 0 to 1.*

**Graph C.7.b**
**Scatter Plot of Annual DJCA Returns against Insured Catastrophe Losses as a % of GDP for highest 20th Percentile of Annual Insured Catastrophes as a % of GDP**

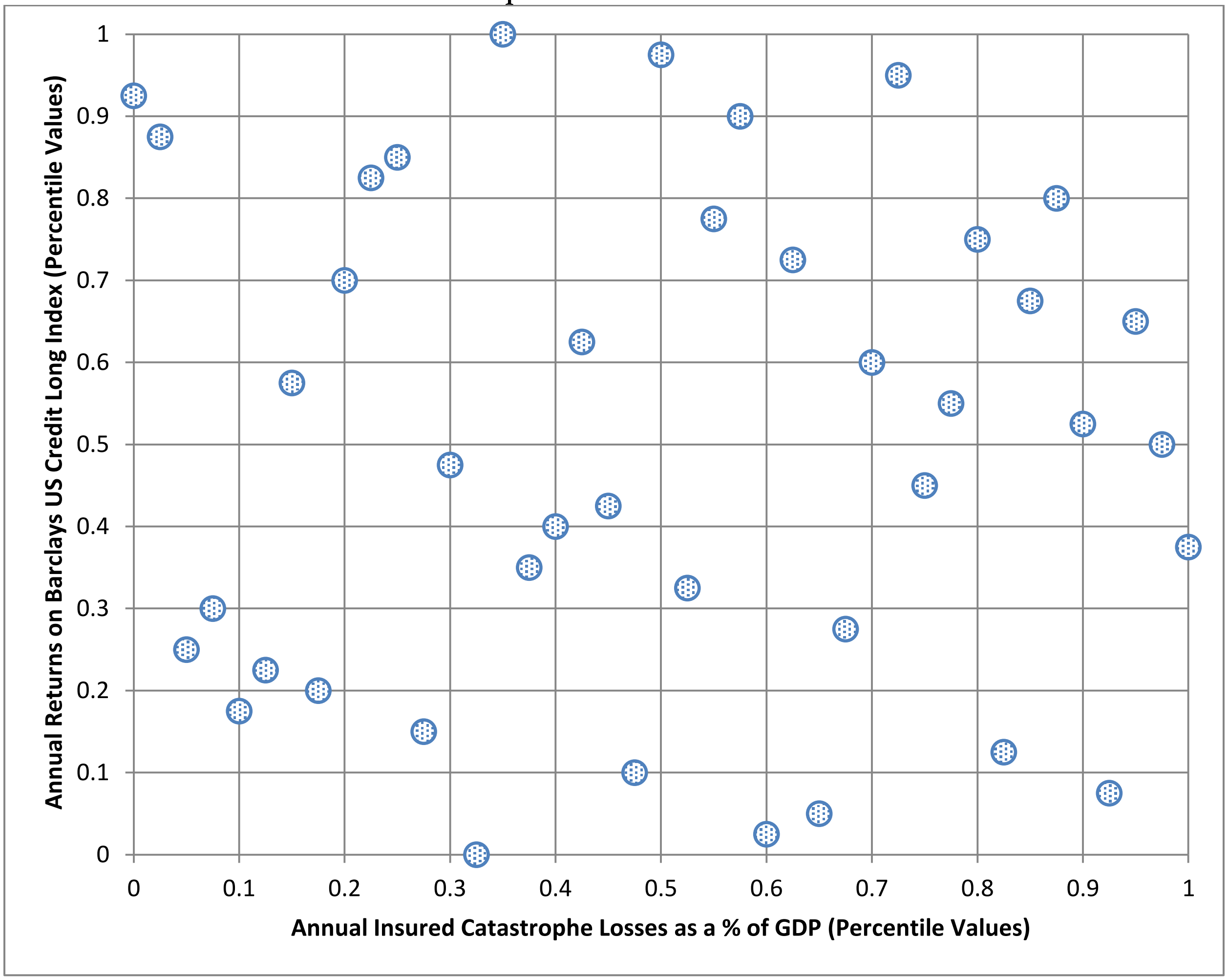



*Please note that the percentiles in the above graph have been recalculated based on the relative rankings of the subset of points that fall in the highest 20th percentile of annual insured catastrophe losses as a % of GDP. As such, these percentiles range from 0 to 1.*

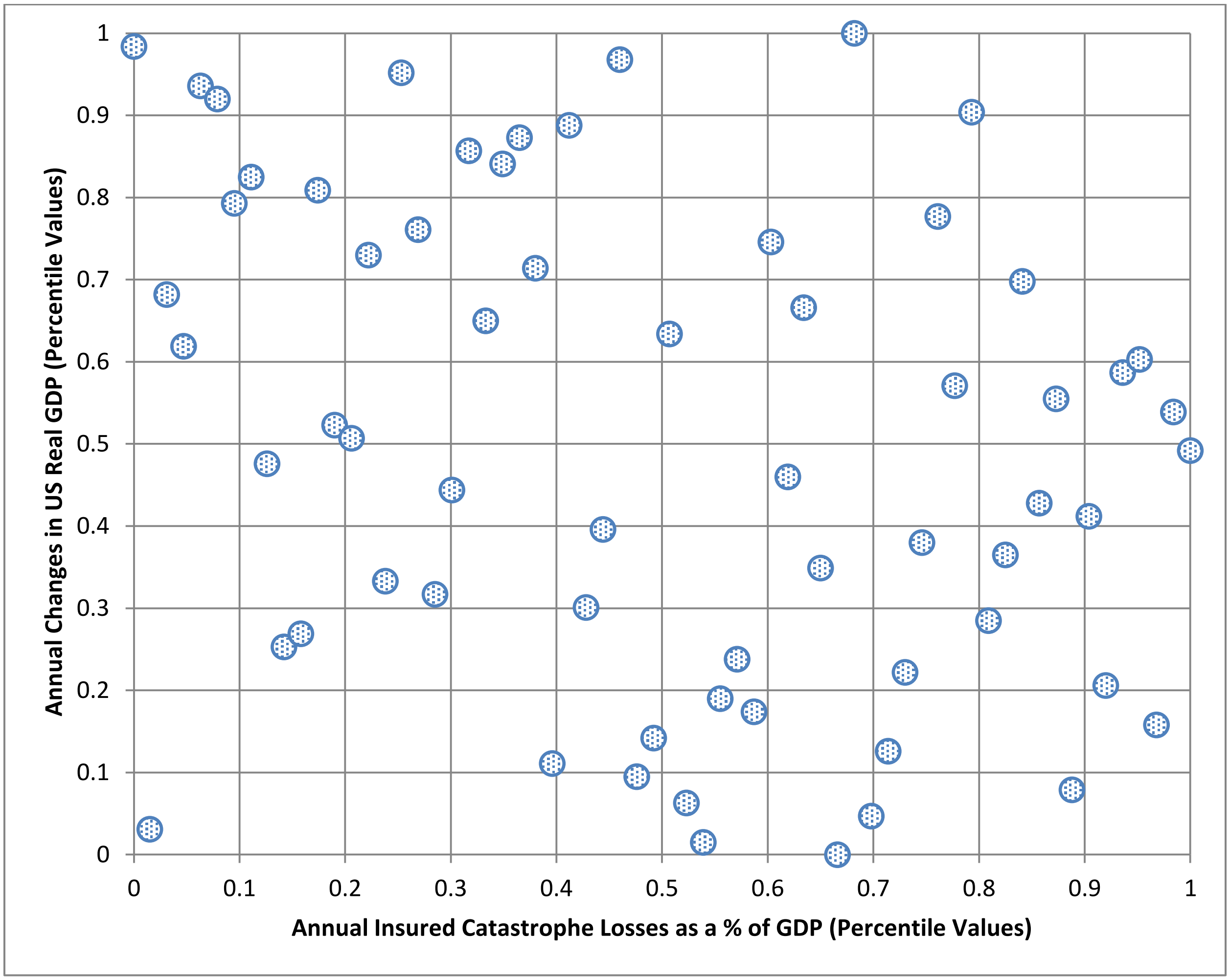
**Graph C.8**
**Scatter Plot of Annual Returns on Barclays Capital US Treasury Intermediate Index against Insured Catastrophe Losses**



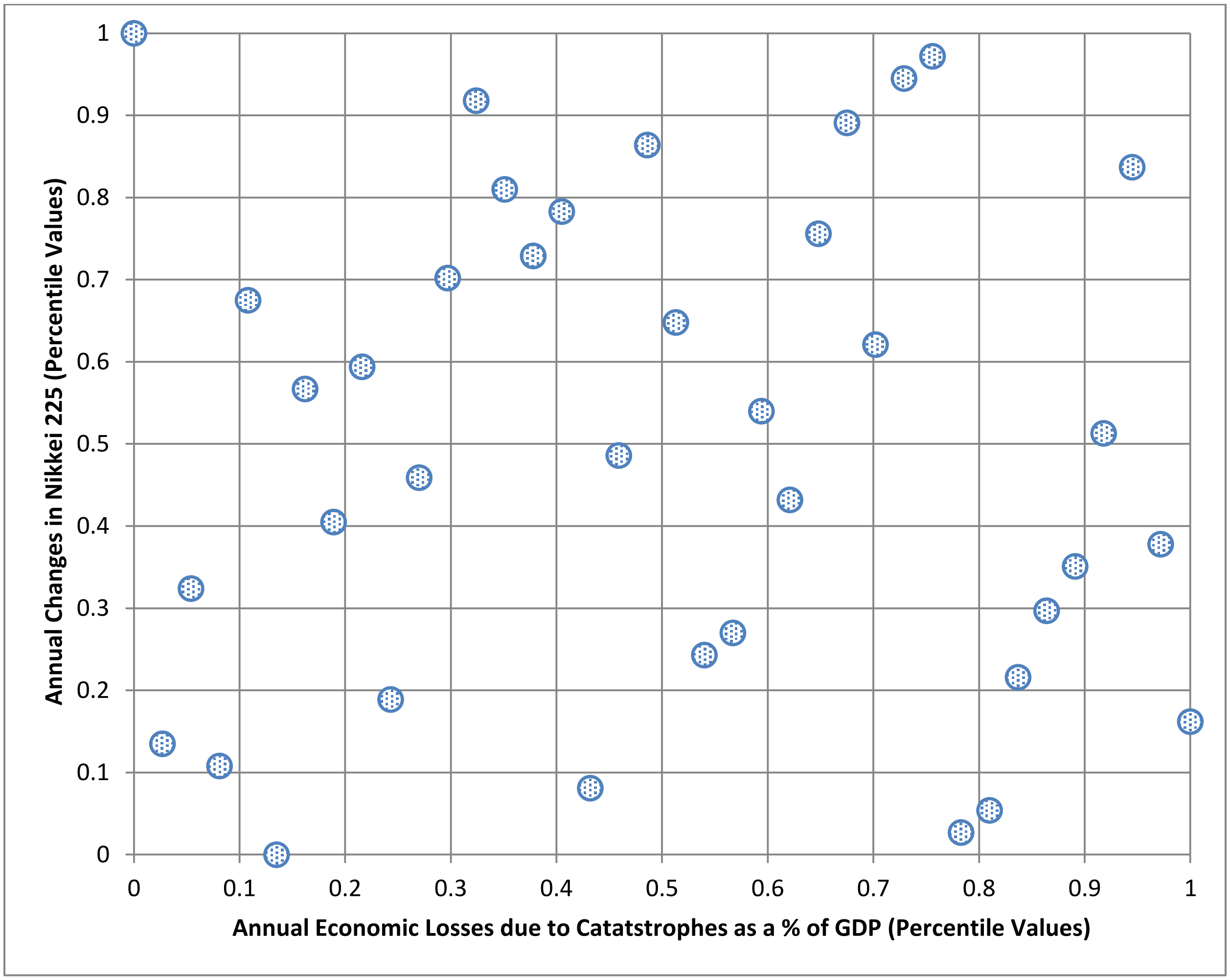
**Graph C.9**
**Scatter Plot of Annual Returns on Barclays Capital US Credit Long Index against Insured Catastrophe Losses as a % of GDP**



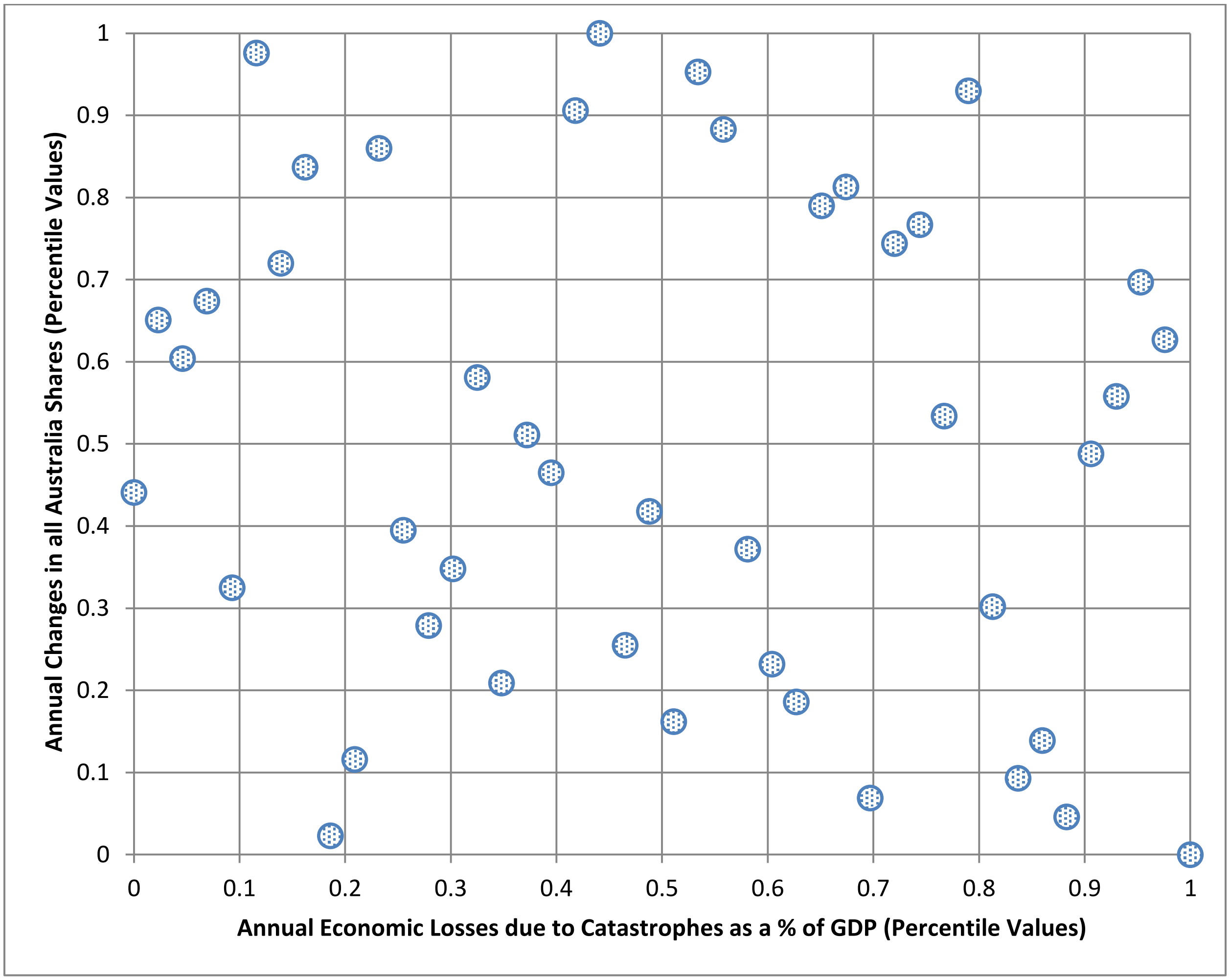
**Graph C.10**
**Scatter Plot of Annual Changes in the US Real GDP against Insured Catastrophe Losses as a % of GDP**



**Graph C.11**
**Scatter Plot of Annual Changes in the Nikkei 225 against Economic Losses due to Catastrophes as a % of GDP**



**Graph C.12**
**Scatter Plot of Annual Changes in all Australia Shares against Economic Losses due to Catastrophe Losses as a % of GDP**



# The Unearned Premium Reserve for Warranty Insurance

Richard L. Vaughan, FCAS, FSA, MAAA

**Abstract.** The Unearned Premium Reserve (UPR) is the largest liability on the balance sheet of most writers of Warranty Insurance. Despite the specialized nature and small size of the line, the NAIC has seen fit in recent years to discuss the UPR for Warranty Insurance in its Regulatory Guidance on Property and Casualty Statutory Statements of Actuarial Opinion.

The UPR is subject to the rules set out for long-duration contracts in Statement of Statutory Accounting Principles 65 (SSAP 65). Because of the high frequency and narrow size-of-loss distribution of Warranty claims, conventional reserve estimators such as Bornhuetter-Ferguson work quite well to estimate the UPR, but to be applied properly they require special modifications. In particular, it is necessary to adjust for unreported losses in recent diagonals of the issue-versus-breakdown lag triangle, to adjust for exposures declining by development month because of cancellations, to estimate appropriate tail factors, to modify expected emergence patterns for coverage of an obligor's failure to perform, and to reserve appropriately for unpaid future refunds; the last two items are not specifically addressed in the regulations.

This paper discusses the purpose and structure of the UPR for Warranty Insurance in general, describes the necessary modifications of conventional actuarial methods in detail, and illustrates them with examples.

## 1. INTRODUCTION

### 1.1 Warranty Insurance

Manufacturers of consumer products are required, either by common law or by the Uniform Commercial Code or similar legislation, to warrant that their products are reasonably fit for their intended use. Making a virtue of necessity, nearly all manufacturers formalize this "implied warranty of merchantability" as an express warranty for a stated period, and often advertise such *factory*, or *manufacturers'*, *warranties* as a guarantee of satisfaction and as evidence of their own confidence in their products.

An extensive market has arisen for contracts that supplement factory warranties by running for longer terms and/or by covering more parts and services. Most commonly such *extended warranties* or *extended service contracts* are purchased at the same time as the underlying product, or shortly thereafter. Their appeal to the individual consumer depends on his or her risk aversion and his or her liquidity relative to the cost of repairing or replacing the particular product should it prove defective.

The classic example of extended warranties with wide consumer appeal are those covering automobiles, with their high initial cost, their many components subject to failure, the high price of replacement parts, and the large labor component of repairs. This paper discusses Warranty reserving mainly from the viewpoint of contracts covering automobiles, because they have special features, such as terms defined in months and miles and separate manufacturers' warranties for power train and for other parts, which, when incorporated in a model, also allow that model to cover simpler warranty contracts as special cases.

While factory warranties attach automatically to every product sold, obligate the manufacturer, and are embedded in the price of the product, extended warranties are usually optional, obligate the retailer or a third party, and are paid for by the consumer with a single premium. Retailers offering such contracts on the products they sell, and assuming such service obligations themselves, are known as *obligors*, and in most states are not regulated as insurers. For tax reasons, many retailers actually issue their extended warranties through affiliated warranty companies. Obligors of this kind are also not regulated as insurers in most states, provided they insure their obligations with a licensed insurer.

Most manufacturers retain the risks associated with their warranties and keep the associated loss data proprietary. Therefore manufacturers' warranties, while they are essential to the design, pricing, and reserving of other Warranty contracts, are not themselves part of the Warranty Insurance marketplace. On the other hand, most warranty companies insure all or some of the risks of their extended service contracts. They may do so to satisfy legal requirements, for the usual benefits of risk transfer, or to obtain the expertise of the insurer in administration, data management, ratemaking, and reserving. Perhaps the most critical specialized expertise of a Warranty underwriter is the ability to calculate the Unearned Premium Reserve (UPR), on which both solvency and the ability to measure rate adequacy depend. The remainder of this paper discusses the UPR in the context of other reserves, of statutory requirements, and of the special adaptations necessary to estimate it for Warranty Insurance.

The extended service contracts we are concerned with in this paper are supplemental to the factory warranty. They may extend it by covering a longer term in months or miles, more components of the product, additional services such as towing, a wider selection of repair facilities, or peripheral contingencies such as those envisioned by Guaranteed Asset Protection (GAP) and Involuntary Unemployment Insurance (IUI) policies. However, the principles of reserving for extended warranties are perfectly applicable to factory warranties and might equally well be used by manufacturers in accounting for their warranty liabilities.

In recent years, Warranty Insurance, despite its relatively small and specialized nature, has drawn the attention of regulators because of the exposure of insurers to the failure of obligors for which they have written Warranty Insurance on a failure-to-perform basis. In 2009, the NAIC introduced a discussion of this issue in its Regulatory Guidance on Property and Casualty Statutory Statements of Actuarial Opinion, which was included as Appendix 9a of the American Academy's Practice Note on the Statements of Actuarial Opinion.

## 1.2 Loss Reserves and the UPR

Like any other insurance product, Warranty Insurance gives rise to obligations for claims incurred but not reported (IBNR) and reported but not paid (RBNP), together making up the reserve for losses incurred but not paid (IBNP). Except for GAP insurance, reporting and settlement lags are short, the size-of-loss distribution is narrow, and frequencies are high, so these loss reserves are both modest in size and straightforward to estimate using conventional triangle-based actuarial techniques.

Unlike many other insurance products, Warranty Insurance policies tend to be of long duration and are usually purchased with a single premium, giving rise to a very significant obligation for losses "paid for" by the single premium but not yet incurred. This obligation is provided for by the Unearned Premium Reserve (UPR). The UPR is by far the largest liability of most Warranty insurers, and is therefore of paramount importance in evaluating their solvency and the solvency of individual Warranty programs. Moreover, the UPR is complementarily related to cumulative earned premium, so it is also of paramount importance when evaluating loss ratios and rate adequacy.

As with other lines of insurance, we may analyze Warranty loss triangles (or similar arrays of more than two dimensions) to establish the UPR and/or loss reserves or to evaluate the adequacy of carried reserves determined in other ways. Typically we analyze the UPR on a policy-month basis (which we shall call *issue month*), the IBNR or IBNP reserves on an accident-month basis (which we shall call *breakdown month*), and the RBNP reserves on a report-month basis.

Many standard actuarial techniques, such as chain-ladder and Bornhuetter-Ferguson loss development, the Cape Cod estimator of expected loss ratios, and the analysis and application of trend and seasonality, may be applied successfully to Warranty Insurance. However, the necessity of issue-month loss development, the importance of cancellations and refunds, and the prevalence of coverage on a failure-to-perform basis in whole or in part, all require certain technical adjustments to the standard techniques, which are a main focus of this paper.

## 2. BACKGROUND AND METHODS

### 2.1 Regulatory Requirements

Most Warranty contracts are issued for terms longer than one year. Such *long-duration* contracts are subject to the requirements of Statement of Statutory Accounting Principles 65 (SSAP 65), which may be paraphrased as follows: the UPR must be at least as great as the greatest of (1) the amount payable if all policyholders surrendered their contracts for refund on the accounting date, (2) the sum over all in-force policies of the gross premium times the expected fraction of ultimate losses not yet incurred as of the accounting date, and (3) the expected present value of future losses, from in-force policies, not yet incurred as of the accounting date. These are called Tests 1, 2, and 3. Test 1 values the surrender option, albeit very conservatively; Test 2 recognizes earnings as risk is borne and services performed; Test 3 addresses claim-paying ability.

All three tests apply prospectively to the portfolio of contracts remaining in force at the valuation date after earlier cancellations. Test 1 assumes that all of these policies cancel immediately, and Test 2 by implication assumes no further cancellations. Test 3 may take into account the effect of future cancellations on expected losses, but it does not measure the expected refunds payable for such cancellations. Therefore Test 3 is an incomplete measure of unpaid future obligations and for that purpose we believe it should be supplemented with an estimate of expected future refunds.

For Warranty business, Test 2 is usually dominant. It is normally greater than Test 1 because the UPR on most policies is greater than the required refund; it is normally greater than Test 3 because most policies are priced to produce a loss ratio less than 1.00. Moreover, the formulas or vectors of monthly factors used by Warranty insurers to calculate the UPR for individual contracts are normally calibrated to match the unincurred fraction of ultimate losses, in the manner of Test 2, since only this may be converted into earned premium suitable for loss ratios or other measures of performance.

A few Warranty contracts are sold for terms of 12 months or less. The UPR for such contracts may simply be taken as pro rata, i.e., gross premium times the unelapsed fraction of the total term, as is usual for other lines. However, since Warranty insurers have in place a mechanism for calculating more precise UPR's satisfying SSAP 65 for their long-duration contracts, they may apply the same technique to shorter contracts, with results that are more accurate and, usually, a bit more conservative.

SSAP 65 applies to an insurer's long-duration business in aggregate and need not be satisfied for any given contract or program. However, it is good practice to attempt to satisfy Test 2 for each program considered separately, for then the aggregate will automatically satisfy Test 2 and the inception-to-date loss ratio for each program will be a reasonable predictor of the ultimate loss ratio. Test 3 may fail for a few programs running loss ratios greater than 100%; correcting this is usually a pricing issue. Test 1 may fail for coverages such as GAP that earn more rapidly at first than pro rata. This will only be a problem in aggregate for companies that write mainly GAP or similar coverages. For them, the system UPR should probably continue to be on a Test 2 basis, but the UPR shown on the books may need to be taken from Test 1.

## 2.2 Reserve Structure

The UPR is chronologically the first component of a reserve structure ultimately designed to recognize all obligations "paid for" but not yet paid; it addresses those obligations not yet incurred. For this reason, in analyzing the UPR, the second, or development, dimension of a loss triangle is usually issue-to-breakdown lag. However, for some sublines of Warranty Insurance the insurer may find it more convenient to analyze the combined UPR and IBNP reserves by issue month versus issue-to-payment lag, or the combined UPR and IBNR reserves by issue month versus report lag, in effect giving rise to several possible reserve structures: (a) UPR, IBNR, and RBNP, or (b) (UPR+IBNR, called UPR) and RBNP, or (c) UPR and (IBNR+RBNP, called IBNR), or (d) (UPR+IBNR+RBNP, called UPR). Because the statutory UPR is governed not only by expected unincurred losses but by the expected emergence pattern of losses *applied to gross premium*, it is usually conservative, so that when UPR principles are applied to some or all of the loss reserves, they in turn become conservative. While this practice is therefore benign, it may be necessary to separate the total reserve into "proper" components for annual statement purposes.

## 2.3 Earnings

The UPR governs the recognition of earnings from an individual policy or cohort of policies, through the general formula (earned premium) = (written premium) – (change in UPR). At the moment of writing, the UPR equals the written premium, having previously been zero, so no earnings are recognized immediately. Thereafter the UPR for a contract is monotonic non-increasing, change in UPR is nonpositive, and earned premium is nonnegative. When the UPR reaches zero, premium earned since inception equals the original written premium and the contract is said to be fully earned.

## 2.4 Cancellations and Refunds

A characteristic feature of Warranty policies is that they are subject to cancellation for refund throughout their term. The amount refunded is specified by law or by contract and is usually different from the UPR carried at the moment of cancellation, giving rise to a gain or loss. Cancellations mean that in-force exposure for a cohort of contracts already issued may decline from month to month. By itself this creates a problem for issue-month loss development, which we address below.

Moreover, the usual accounting treatment of cancellations, which nets refunds against premiums written in the same calendar period, even though the refunds may have arisen from earlier contracts, is awkward and creates difficulties for actuarial analysis, especially when the actuary is attempting to measure the performance of a cohort of written contracts.

We have found it useful to distinguish premium earned by providing coverage from premium earned by cancellation, and we digress here to discuss this in some detail because of its close connection to the UPR. First, we define *UPR released by cancellation* to be the UPR on canceling policies just before cancellation. Then the following definitions relate to premium used to provide coverage:

(pure written premium) = premium for new policies, gross of refunds

(pure earned premium) = (pure written premium) – (UPR released) – (change in UPR)

(pure loss ratio) = (incurred losses) / (pure earned premium)

The corresponding figures in the financial statements are defined somewhat differently:

(statement written premium) = premium for new policies, net of refunds paid in month (statement earned premium) = (statement written premium) – (change in UPR)

(statement loss ratio) = (incurred losses) / (statement earned premium)

The "pure" definitions are very useful when tracking a cohort, for which there is no written premium after the first month. The "statement" definitions are confusing for this purpose, since the actual financial statements would show the initial written premium reduced by refunds paid in the same calendar month on behalf of policies written in that month or earlier months, and would show refunds paid in later months as part of the written premium of later cohorts.

Note that (statement earned premium) = (pure earned premium) + (gain from cancellations), where (gain from cancellations) = (UPR released) – (refunds). For Warranty Insurance on automobiles, where refunds are close to pro rata while UPR declines more slowly, (gain from cancellations) is usually positive, the statement earned premium is greater than the pure earned premium, and the statement loss ratio is less than the pure loss ratio.

The "pure" definitions above involve losses but not refunds. We can create similar definitions treating refunds, or refunds plus losses, in a manner parallel to losses, as follows. Here the refunds are only for policies in the cohort being tracked:

(refund ratio) = (refunds) / (UPR released by cancellation)

(payout ratio) = (incurred losses + refunds) / ((pure WP) - (change in UPR))

Note that UPR released by cancellation is analogous to pure earned premium except that it measures premium earned through cancellation rather than through coverage. Therefore the refund ratio is analogous to the pure loss ratio. Finally, the payout ratio is also similar to the pure loss ratio except that refunds are treated as equivalent to losses; its denominator, (pure WP) – (change in UPR), is the total premium earned either by coverage or by cancellation.

If we make the following two definitions,

(cancelled UPR ratio) = (UPR released by cancellation) / ((pure WP) - (change in UPR))

(earned premium ratio) = (pure earned premium) / ((pure WP) - (change in UPR))

then

(payout ratio) = (cancelled UPR ratio)(refund ratio)+ (earned premium ratio)(loss ratio)

showing how the parallel treatments of losses and refunds fit together. Here the cancelled UPR ratio is the fraction of the total premium earned in *some* manner that is earned by cancellation, and earned premium ratio is the complementary fraction earned by coverage.

These ratios are most useful over the time period from inception to ultimate, for then the initial UPR, the final UPR, and the change in UPR are all zero, the "big" denominators just equal the pure written premium, and the ultimate payout ratio gives a good measure of premium adequacy:

(ultimate payout ratio) = (ultimate losses plus refunds) / (pure written premium)

By contrast,

(ultimate statement loss ratio) = (ultimate losses) / ((pure written premium) – refunds)

which will not, in general, equal the ultimate payout ratio even if the refunds are those paid over time on behalf of the cohort rather than those from any source paid in the initial calendar month.

### 2.5 Carried Reserves

In principle, the management of a Warranty insurer could establish its UPR and loss reserves shortly after the end of each accounting period by developing inception-to-date experience data. But most insurers find it more practical (a) to embed formulas or strings in their administrative systems to generate the UPR automatically from in-force contract data at each month's end, (b) to pull the RBNP reserve directly from reported loss data, and (c) to calculate the IBNR as a factor times RBNP, recent paid losses, or some similar base quantity. These insurers still analyze inception-to-date experience, but do so to validate or modify their strings and formulas, rather than to establish the carried reserves directly.

*Strings* are vectors of UPR factors that are stored with the data for each contract and multiplied by gross written premium for that contract to obtain the UPR at each elapsed month's end from issue to expiration. These factors are normally adjusted so as to apply to an entire month's cohort of similar contracts, on the assumption that contracts are issued uniformly throughout the month. The usual adjustment at lag $n$ is to take an average of the factor that would be held at lag $n$ and the factor that would be held at lag $n$-1, if all policies were written at the start of the month; this is the so-called one-half-month adjustment.

In principle the string could be extended beyond the term in months to accommodate "goodwill" claims paid after expiration, but commonly the length of a string equals the term in months. The factor for the moment of issue is 1, and need not be stored with the string; similarly, the factor for any earlier month's end is 0 (no UPR is needed yet since no premium has been received) and the factor for the end of any month beyond the end of the string is also 0 (contract fully earned).

There may be many strings associated with a Warranty program. The choice of string for a given contract depends on characteristics such as term, manufacturer's warranty, and type of product insured. The string is assigned when the contract is first entered into the administrative system, and is usually fixed thereafter, ensuring a stable accounting treatment through the life of the contract.

*Formulas* are rules for determining UPR factors from contract characteristics and elapsed months. They may be mathematical formulas in the usual sense, such as pro rata or Rule of 78, or they may be lookup tables in every respect analogous to a set of strings, except not stored with the contract. They normally include the one-half-month adjustment. The factors are generated fresh at each valuation; making it possible to change formulas easily and keep UPR patterns responsive to current estimates, for all contracts, old as well as new. The formula UPR for a contract equals its initial written premium multiplied by the formula UPR factor.

Just as formulas may in fact be lookups of strings, strings may have originally been derived from formulas. A set of UPR factors, whether described by string or by formula, may be called a *UPR curve* because of its appearance when displayed graphically.

**2.6 Graphical Representation of UPR and Earnings**

The UPR for a block of contracts of term $T$ months may be plotted against lag $t$ ; it will equal the written premium at lag $t$=0 and will equal zero at full term, $t=T$, or at $t=T+1$ with the one-half month adjustment. Usually we rescale such a graph to represent UPR factors, starting with 1 at $t$=0 and reaching 0 on or before $t=T+1$.

Such a graph need not be continuous; it is possible to conceive of products for which claims, if any, emerge at a few discrete times, creating a step-function UPR curve. We shall ignore such special cases here. Actually, in practice all strings and many formulas *are* step functions with discontinuities at the end of each elapsed month. For our purposes it is convenient to "connect the dots" and treat such strings and formulas as continuous for nonnegative real $t$.

As mentioned above, in the absence of new written premium, earnings are measured by negative change in UPR. Therefore, (a) the complement of the UPR curve is the cumulative earnings factor curve, (b) the negative slope of the UPR curve is proportional to the earnings rate, (c) the cumulative earnings factor curve is an ogive like that of a cumulative distribution function, and (d) the UPR curve is a mirror image of the cumulative earnings curve around the line y=0.5. *Incremental* earnings are often step functions, for example, assuming one level while the full manufacturer's warranty is in effect, another level when only the power-train warranty is in effect, and still another level when both parts of the manufacturer's warranty have expired. This produces changes of slope in the UPR curve but leaves it continuous.

One example of a UPR formula is linear or "pro rata", equivalent to a uniform incidence of losses over an earning period defined by the term and manufacturer's warranty. This is theoretically correct for equipment each of whose components has an exponential distribution of time to first failure, where each failed component is replaced with an identical one, and where there is neither trend nor "breakage" (contract abandonment). Figure 1 shows (to different scales) a uniform earnings pattern over 24 months and the corresponding UPR factors; this graph also illustrates the effect of the one-half-month adjustment for uniform writings throughout the month.

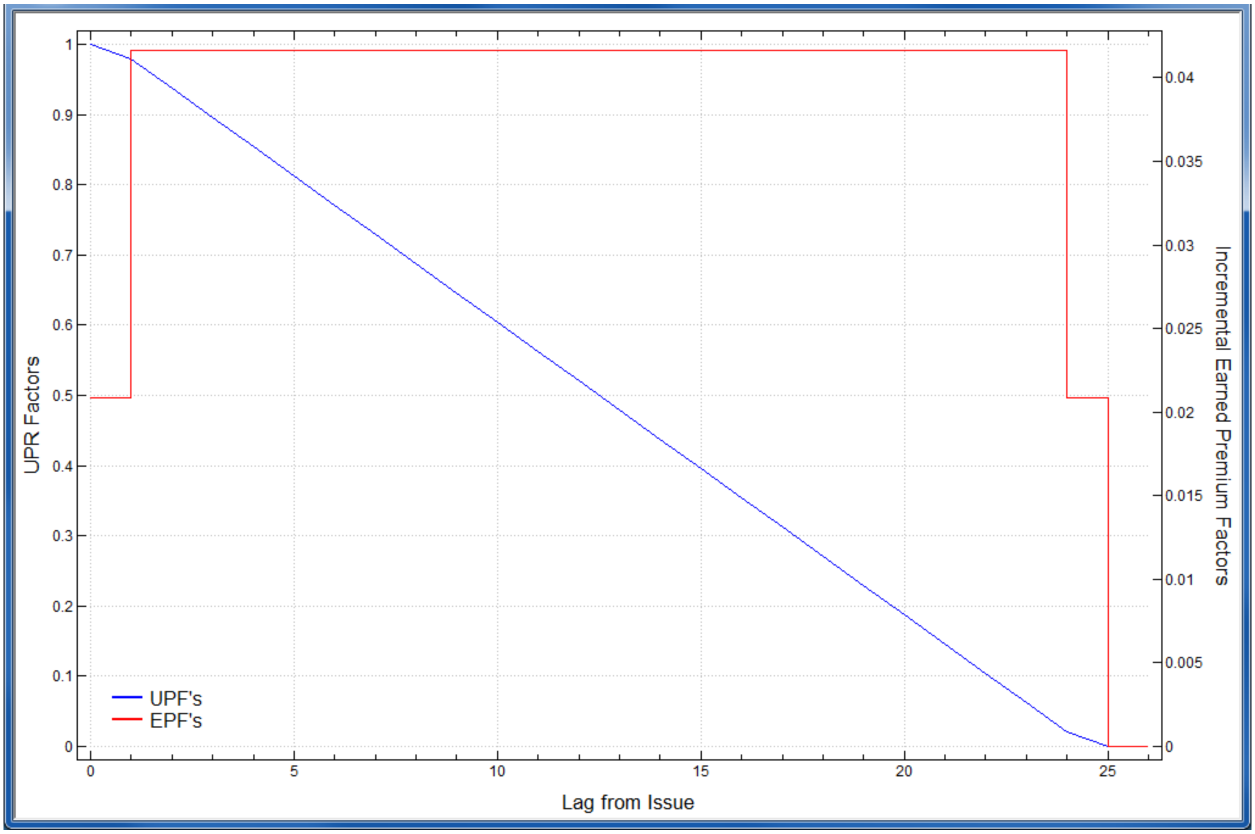



**Fig. 1.** Uniform earnings and pro rata UPR factors for term 24 months

Note that the one-half-month adjustment as graphed is only an approximation to uniform writings *within* the first and the last months; the true earnings pattern would not be constant, and the true UPR pattern within those months would be a second-degree curve.

The pro rata UPR formula is often modified to start at the end of the manufacturer's warranty (MW) instead of at issue. Warranties on so-called "brown and white" goods (electronics and appliances) are often reserved in this manner, with considerable justification from experience. Figure 2 illustrates 24-month pro rata earnings from the end of a 12-month MW.

Another example is sum-of-digits or Rule of 78, a second-degree curve whose first differences (evaluated at successive months' ends) are proportional to the number of months remaining in the term. This is theoretically appropriate for situations where the size of loss decreases linearly to zero at expiration while the probability of a loss remains constant, again with neither trend nor breakage. An example of such coverage would be a warranty covering failure of parts subject to normal wear and tear, where reimbursement reflects the amount of use already received from the part. This pattern is illustrated in Figure 3. Again, earnings have been shown as a step function and the UPR factors as a corresponding stepwise-linear function, reflecting the common practice of calculating UPR factors only at months' ends.

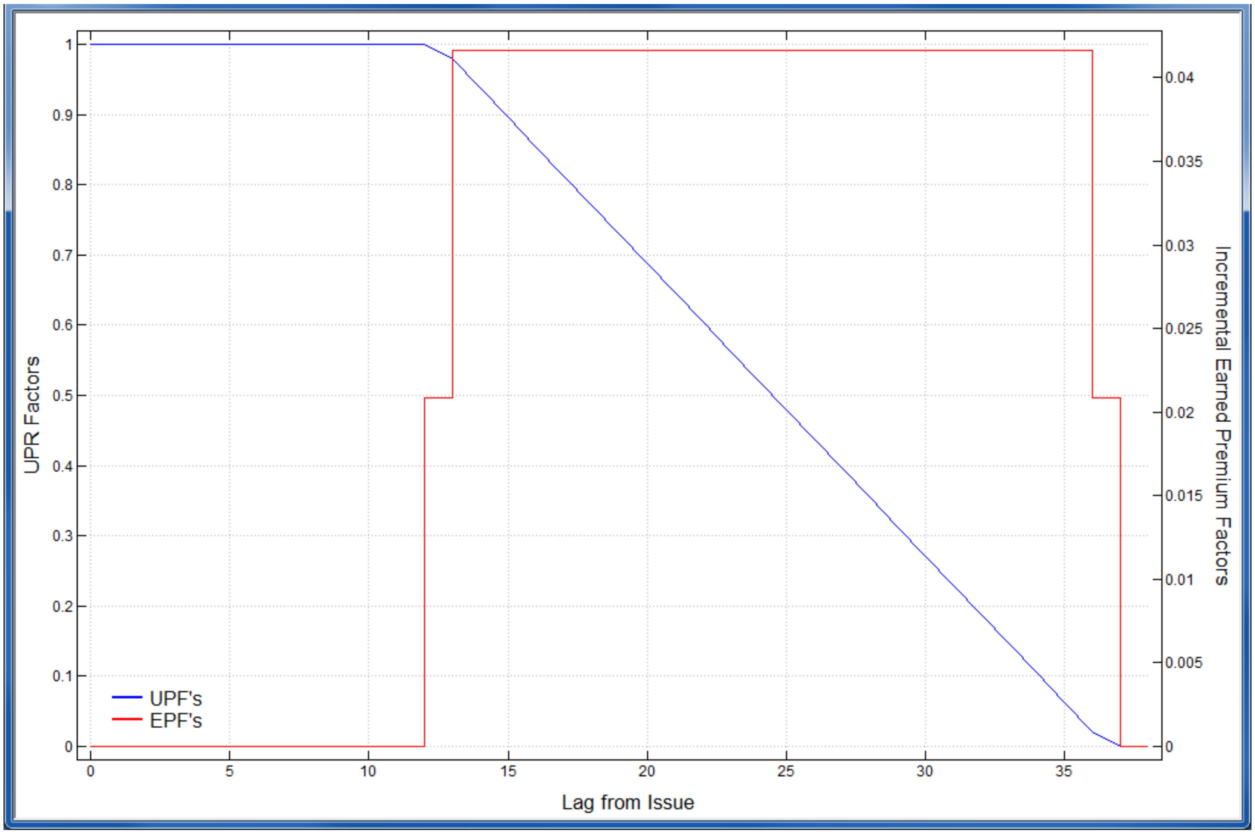



**Fig. 2.** Pro rata from end of manufacturer's warranty, with one-half-month adjustment.

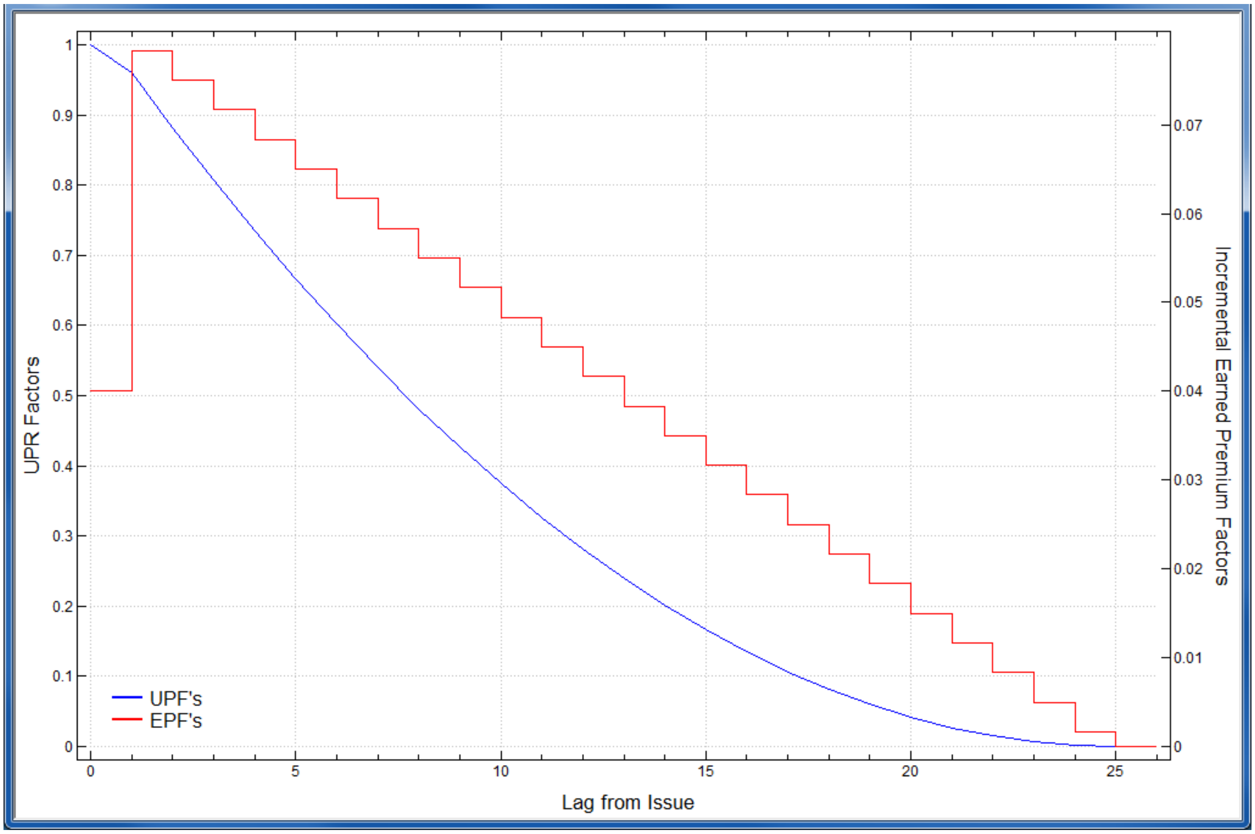



**Fig. 3.** Rule of 78, with one-half-month adjustment.

Here, with a term of 24 months, the "Rule of 78" should more properly be called "Rule of 300", or just "sum of digits". The Reverse Rule of 78 – first differences proportional to number of months elapsed – is sometimes used as a UPR formula, with little theoretical justification, other than that its slow early earnings produce a "shoulder" in the UPR curve that resembles that produced by the manufacturer's warranty for automobile business. This is illustrated in Figure 4. Interestingly, the Reverse Rule of 78 *is* theoretically appropriate as a UPR formula for Warranty Insurance providing contractual liability coverage for an obligor, without coverage of the underlying contracts unless the obligor fails to perform, where the underlying contracts earn uniformly, and where the probability of failure to perform (i.e., bankruptcy) approaches zero.

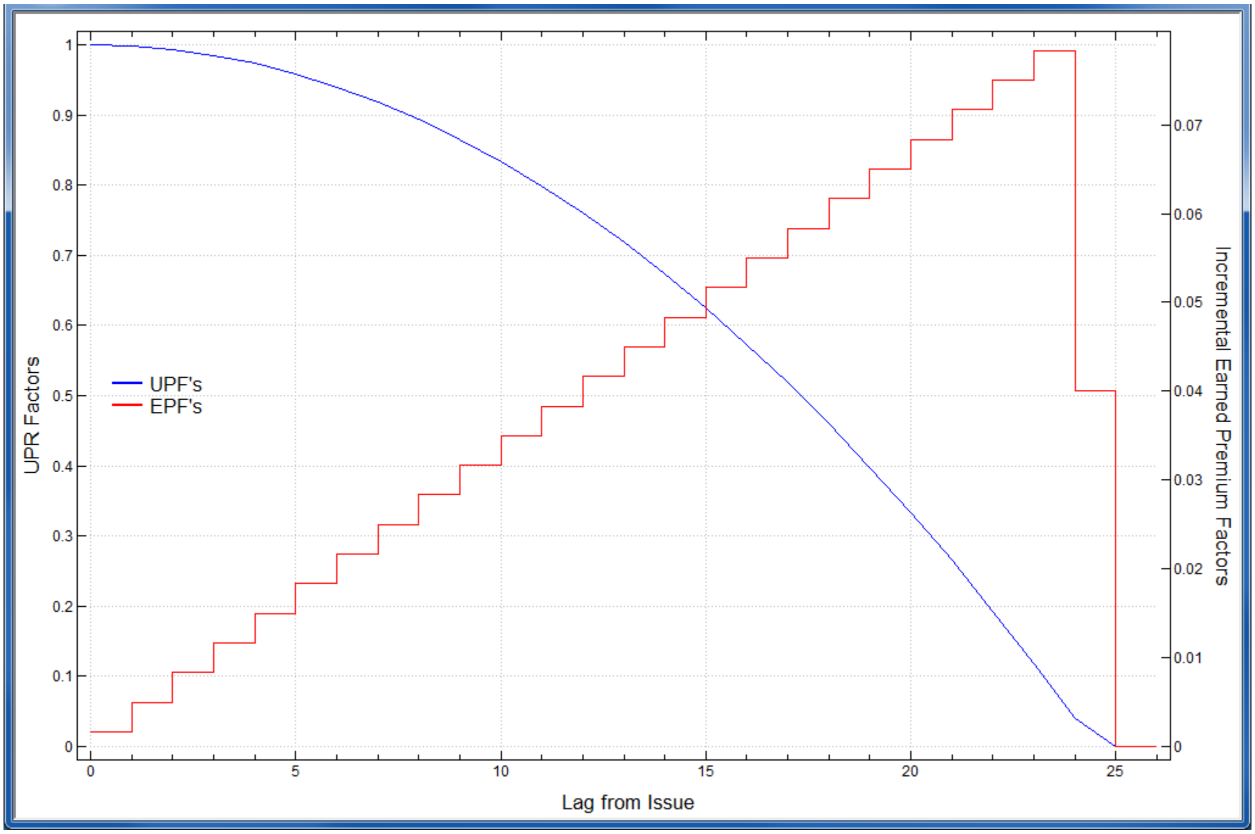



**Fig. 4.** Reverse Rule of 78, for term 24 months, with one-half-month adjustment.

Some Warranty administrators are known to operate with just three possible choices for UPR formula: pro rata from end of MW, Rule of 78, and Reverse Rule of 78.

## 2.7 Indicated Reserves

By an *indicated reserve* we mean an estimate of a liability based on analysis of loss experience. The reliability of such analysis is enhanced when summed over a collection of subdivisions of the data, each large enough to be credible but also as homogeneous as practical. Usually this means grouping contracts by term, although finer subdivisions, involving other contract or product details, may sometimes be necessary. For purposes of this discussion, we assume we have already subdivided the data into one or more reasonably homogeneous and reasonably large collections of contracts and associated claims.

Case reserves are usually good estimators of the RBNP reserve since pending Warranty payments are often known quite accurately at the report date. So loss reserving amounts to estimating either the IBNR reserve from breakdown dates versus report lags or estimating the IBNP reserve from breakdown dates versus payment lag and then subtracting the RBNP to obtain the IBNR reserve. Either estimate may be done with the chain-ladder estimator, not requiring any measure of exposure by breakdown month, or with the Bornhuetter-Ferguson estimator, using the results of a UPR analysis to obtain earned exposure. The actual calculations are straightforward and familiar and we do not discuss them further.

The indicated UPR is more complex and involves choices that may depend on the purpose of the analysis. When evaluating the adequacy of the carried UPR for a single program, a SSAP 65 Test 2 estimate may be appropriate; when evaluating rate adequacy, Test 3; when evaluating compliance with SSAP 65 in aggregate, Tests 1, 2, and 3.

We may evaluate SSAP 65 Test 1 directly by applying the refund formula to each contract in-force at the valuation date and summing the results. The typical refund formula is a flat 100% for 30 or 60 days, thereafter pro rata between issue and expiration (meaning there is a discontinuity at the end of the flat refund period), less a small surrender charge capped as a multiple of premium. This is a routine non-actuarial calculation and we do not discuss it further.

From an actuarial perspective the key to evaluating Test 2 is that we must have a sound technique for estimating issue-to-breakdown lag patterns *after* removing the effect of cancellations. Part of the solution involves modifying Chain-Ladder loss development to accommodate exposure declining across each row of the triangle, part involves adjusting for the deficiency of later diagonals of the issue-breakdown lag triangle because of unreported losses, and part involves special procedures for estimating tail factors. These techniques are discussed in sections 2.11, 2.12, and 2.13, below.

We mention in passing another estimator of loss emergence patterns, which Bühlmann [2] called Complementary Loss Ratio and Stanard [5] called Additive, but which we prefer to call Partial Loss Ratios. This method fits an additive model with only column effects to the triangle of partial loss ratios. Stanard's simulations found it efficient and unbiased. For our purposes, it appears to be affected more than the adjusted Chain Ladder when there is trend in the historical losses that is not well matched with trend in the premium or other measure of exposure.

Evaluating Test 3 requires all of the above plus sound techniques for estimating loss ratios and projecting future in-force exposures. In our work we usually project in-force premiums using a survivorship model analogous to chain-ladder loss development, but applied to the premium surviving cancellations at each lag. Next we use loss development to estimate issue-to-breakdown and breakdown-to-payment lag patterns. With these in hand, we use variations of the Cape Cod technique to estimate expected loss ratios, and then project future losses using a Bornhuetter-Ferguson model. These techniques are discussed in 2.14 below. In some ways they are more easily understood in the Warranty context than for other lines of business; for example, when the first dimension of our triangles is issue month, the Bornhuetter-Ferguson expected emerged exposure is simply earned exposure and the Cape Cod ELR is simply the ITD loss ratio.

Often we are not so much interested in the actual UPR for a block of business as in the average string of UPR factors, one for each lag, that would produce the correct UPR, not only at the given valuation date but at any date. Or we may be interested in the string of UPR factors appropriate for each contract in the data, or for proposed contracts yet to be written. We call such strings, derived from experience, *indicated UPR factors*.

The estimation of SSAP 65 Test 2 or Test 3 produces as an intermediate step an average set of indicated UPR factors for each subdivision of the data. If the subdivision is perfectly homogeneous, the indicated average UPR factors will also be appropriate for each contract in the subdivision. But subdivisions large enough to be credible are seldom perfectly homogeneous, and the indicated average UPR factors may not be quite right for any given contract.

To get factors for individual contracts in the data, or for proposed contracts different from any of those in the data, we describe below a technique which we call “All-Terms Factors”: this is based on an exposure definition similar to that of Kerper and Bowron [4], together with an algorithm for using the experience of all contracts in one set of data to obtain a UPR curve for any contract whether in the data or proposed.

### 2.8 Comparison of Carried and Indicated UPR

At any given valuation date the indicated UPR is likely to differ from the carried UPR; if this difference is large enough the actuary will want to determine its cause and possibly revise the strings or formula used to establish the carried UPR. However, the difference between carried and indicated UPR at a single date gives little information as to what changes are needed, and moreover is highly dependent on the maturity distribution of the contracts that happen to be present in the data.

For example, if the carried UPR is less than the indicated, it may be because the contracts are clustered at the early lags and the early carried UPR factors are inadequate, or because the contracts are clustered at later lags and the later carried UPR factors are inadequate, or because the maturities of the contracts are evenly spread and the entire set of carried UPR factors is inadequate.

We have found it more useful to compare entire average carried UPR curve with entire indicated UPR curves, as shown in Figure 5 below. If the two sets of factors are not close to each other, we can conclude that the carried strings for some or all of the contracts need to be revised. If the carried and indicated factors are close, then we have evidence, though not proof, that the carried strings are indeed satisfactory contract by contract.

Our comparison procedure also generates single statistics that quantify the adequacy of the carried UPR factors independently of the maturity distribution of the data at hand.

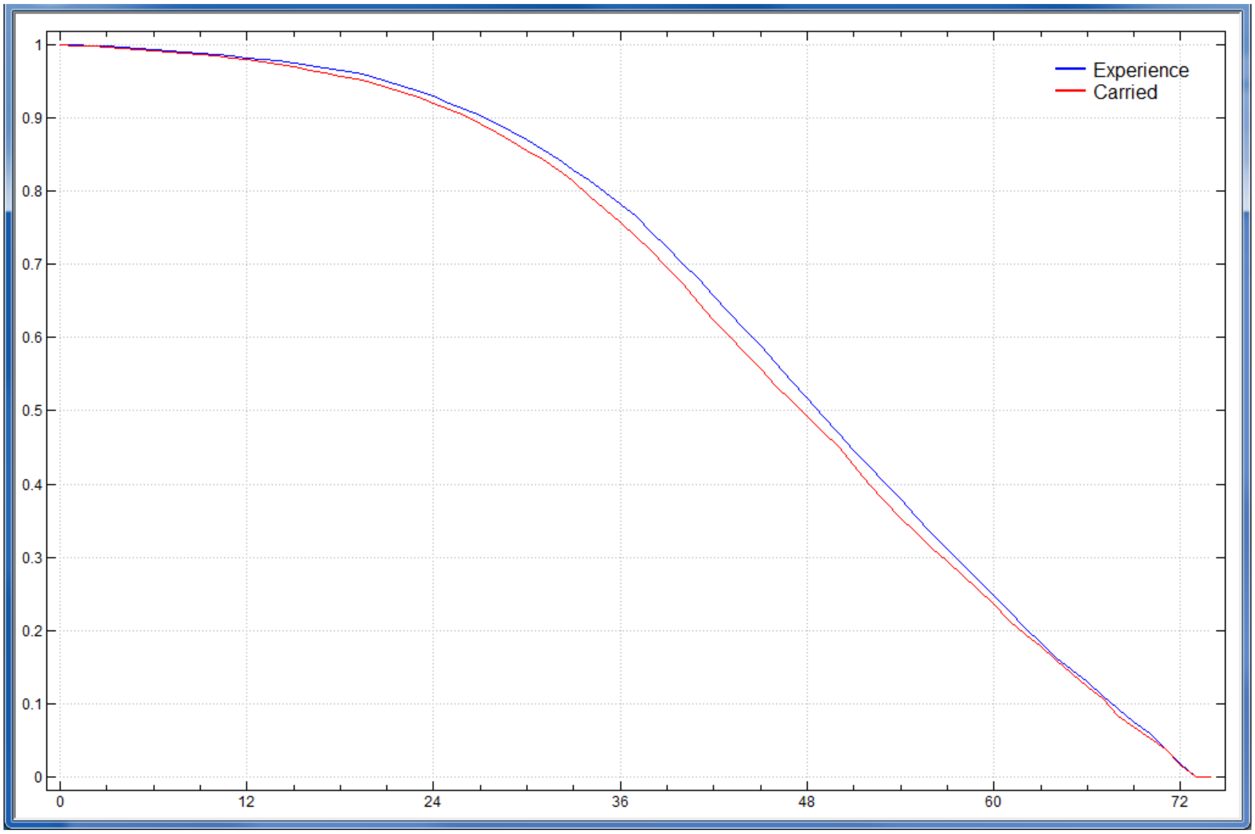



**Fig. 5**. Comparison of carried UPR factors with Test 2 UPR factors indicated by experience

The comparison in Figure 5 is based on about 272,000 contracts, and 55,000 claims, with nominal term 72 months, from a program insured by a large Warranty underwriter. Like all UPR factor curves, these start with the value 1 at issue and end with the value 0 when all contracts are fully earned, in this case by 73 months. We may read cumulative earnings as the complement of the UPR curve, and the instantaneous earnings rate as proportional to its absolute slope.

The shoulder in both curves reflects the presence of manufacturers' warranties, typically 36 months or longer. The fact that there are some earnings in the first 36 months reflects "extras" covered by the extended warranty above the services provided by the factory warranty, and also the fact that some drivers "mile out" of the manufacturer's warranty before 36 months. There may be an issue of heterogeneity here also, with a few contracts on used cars with little or no manufacturers' warranty remaining.

The Experience curve lies above the Carried curve showing that the latter is slightly inadequate. To assign a measure to this inadequacy we use what we call a "steady state conservatism factor", which is the ratio of the carried to the indicated Test 2 UPR if policies had been written at a constant rate long enough (in this case six years) so that their maturity distribution would thereafter be stationary. In the absence of cancellations the steady-state UPR factor distribution would be precisely that shown by this graph, the steady-state carried and indicated UPR's would be proportional to the areas under the respective curves, and the ratio of carried to indicated would be the ratio of these areas. In the present example the conservatism factor calculated in this manner is 0.980, reflecting about 2% inadequacy. We can also calculate an alternate conservatism factor in which the steady state allows for cancellations; in this case it comes out 0.982, suggesting 1.8% inadequacy.

Bear in mind that this entire comparison is on a Test 2 basis. It is unaffected by loss ratios and says essentially nothing about Test 3, the relationship between the carried UPR and expected unincurred losses. In this example the loss ratio happened to be well below 100% and Test 3 was easily satisfied, but we have no way of knowing, from these Test 2 curves alone, anything about the loss ratio or the adequacy of the carried UPR relative to Test 3.

## 2.9 Effect of UPR Curves on Estimated Loss Ratios

It is very important to understand the effect on *estimated* loss ratios, especially for immature blocks of business, of carried UPR curves that do not match closely with experience. Using Figure 5 as a model, imagine that there had been a large spike in sales 24 months ago, dominating the contracts in the data set. Then this business would now be at lag 24 months and would have a carried UPR factor of about 0.9196 and an indicated factor of about 0.9298. But then the ratio of cumulative earned premiums would be 0.0804/0.0702, or about 1.145; if the apparent loss ratio were, say, 95% the true loss ratio would be about 108.8%.

If in this case we took the apparent inception-to-date loss ratio as a predictor of the ultimate loss ratio, we might imagine that we were extrapolating the immature business by a factor of 3, from 24 months to 72. In fact we would be extrapolating by a factor of more than 14, from 7% of ultimate losses to 100%. The estimated loss ratio is not only affected by any errors in the UPR curve, also but by random fluctuations in the early incurred losses, greatly leveraged.

In this case we assumed that, although the bulk of the business was clustered at maturity 24 months, there was enough other business to estimate the shape of the earnings curve reliably. The situation is exacerbated when all of the business is immature, for example, when writing started 24 months ago, for then we must rely on tail factors, and the average extrapolation might be from around 12 months (assuming 24 months of uniform writings), or about 1.8% emergence, to 100%, a factor of more than 56.

## 2.10 Trend and Seasonality

One kind of trend operates purely in the issue-month direction, as changes in product design affect claim frequency and severity in sometimes irregular ways; a second kind operates purely in the calendar-month direction that is along both the issue-month and the development-month axes. This is usually the result of inflation, affects both parts and labor, and drives severity generally upward while having little effect on frequency. There is actually a third kind of trend, purely in the development direction; this we usually regard as simply part of the earnings or UPR pattern, but if not controlled for it may affect our estimates of calendar trend.

Loss emergence patterns are affected by calendar trend, in that increasing severity in the development direction tends to defer the emergence of losses relative to the total. This has only a modest effect on SSAP 65 Test 2 but a much greater potential effect on Test 3. To illustrate, Figure 6 shows a 48-month pro rata UPR curve after a manufacturer's warranty of 12 months, with trends of 0%, 5%, 10%, and 15% per annum. The steady-state SSAP 65 Test 2 UPR's are greater than the no-trend case by about 2.2%, 4.2%, and 6.3%, respectively. Note that the fact that the trend extends through the manufacturer's warranty has no impact on these figures.

By contrast, the Test 3 expected unincurred losses, valued at inception of the contract, would increase by factors of about 15.5%, 32.8%, and 52.3%, respectively, and part of these increases would be the result of trend during the manufacturer's warranty. Moreover, using a UPR curve with incorrect trend built in would have the same amplified effect on loss ratios estimated from immature data as was described the indicated-versus-carried comparison above.

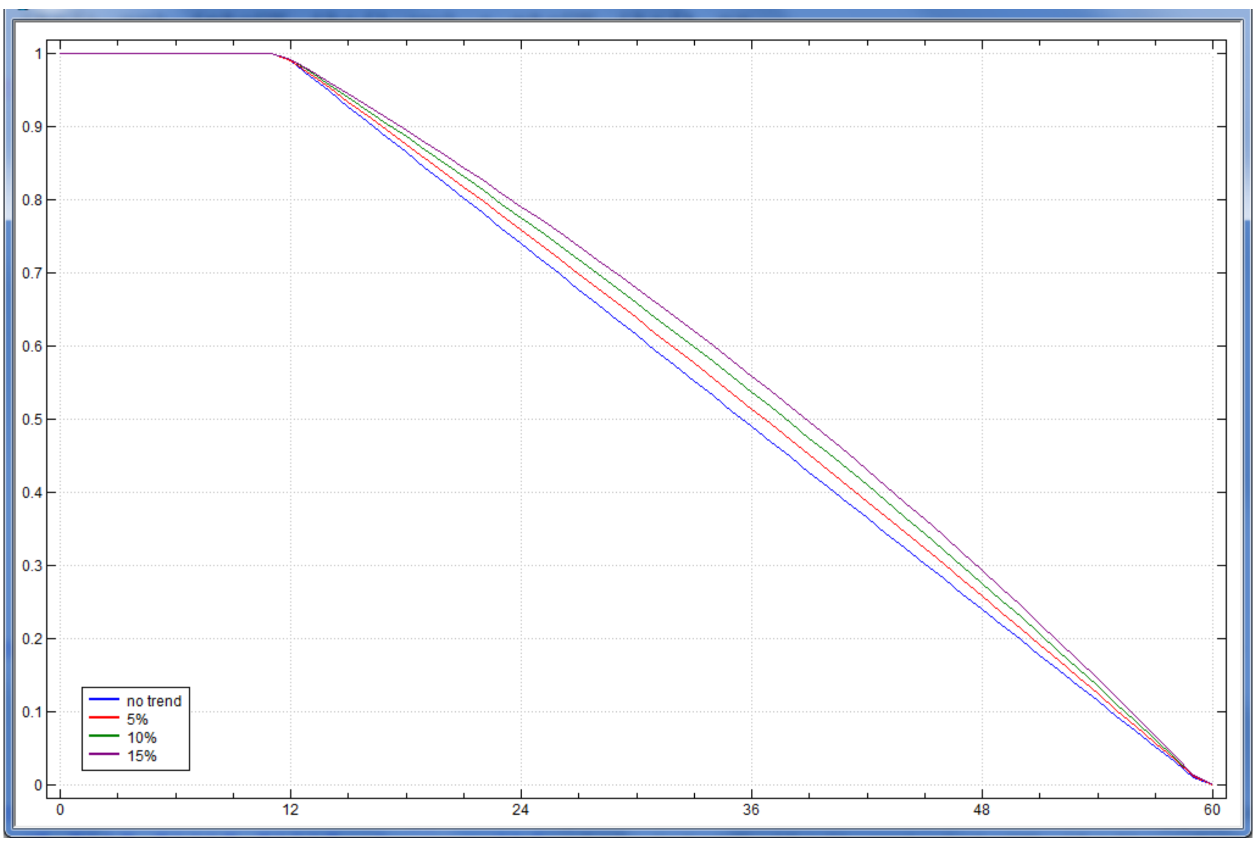



**Fig. 6.** Effect of 5%, 10% and 15% annual trend on pro rata UPR factors after MW

Carried UPR's by simple formulas like pro rata and Rule of 78 clearly do not anticipate trend, except possibly implicitly, as an offset to "breakage" (failure of eligible policyholders to present claims, which can be modeled as a negative trend). On the other hand, strings obtained empirically by analysis of experience are likely to have embedded in them whatever trend was present in that experience. If it is thought desirable to include a different trend in future strings, the experience can be deflated prior to estimating the earnings pattern, and the results re-inflated afterward. Or the strings from the usual analysis may be rendered more conservative by converting the UPR factors to earnings factors, inflating these by the difference between future and past trends, and converting once again to UPR factors.

Unlike trend, seasonality (which we take to have a period of 12 months) has rather a greater effect on Test 2 than on Test 3. We illustrate seasonality with an extreme, but nevertheless real, example: extended warranties on snowmobiles. In the Northern Hemisphere these may be expected to produce nearly all warranty claims during the season from November through, say, April, with the greatest concentration shortly after the equipment is first put in use for the season, say during December and January.

This pattern has little impact on expected ultimate losses, since most contracts are sold for a multiple of 12 months and thus include an equal number of Decembers and Januaries no matter in what month they are issued. On the other hand, if it is not provided for in the UPR factors, seasonality may lead to serious error in the estimation or projection of loss ratios. This problem may be avoided in two ways. One is to deseasonalize the loss data, estimate a constant loss ratio, then restore seasonality to the projections; the other is to build seasonality into the UPR strings and assign a different string to each contract depending on the month in which it is sold. Figure 7 illustrates the strings that might be assigned to two snowmobiles, one issued in November and one in March, under one plausible set of seasonality assumptions. Note that it is the slopes that are cyclical.

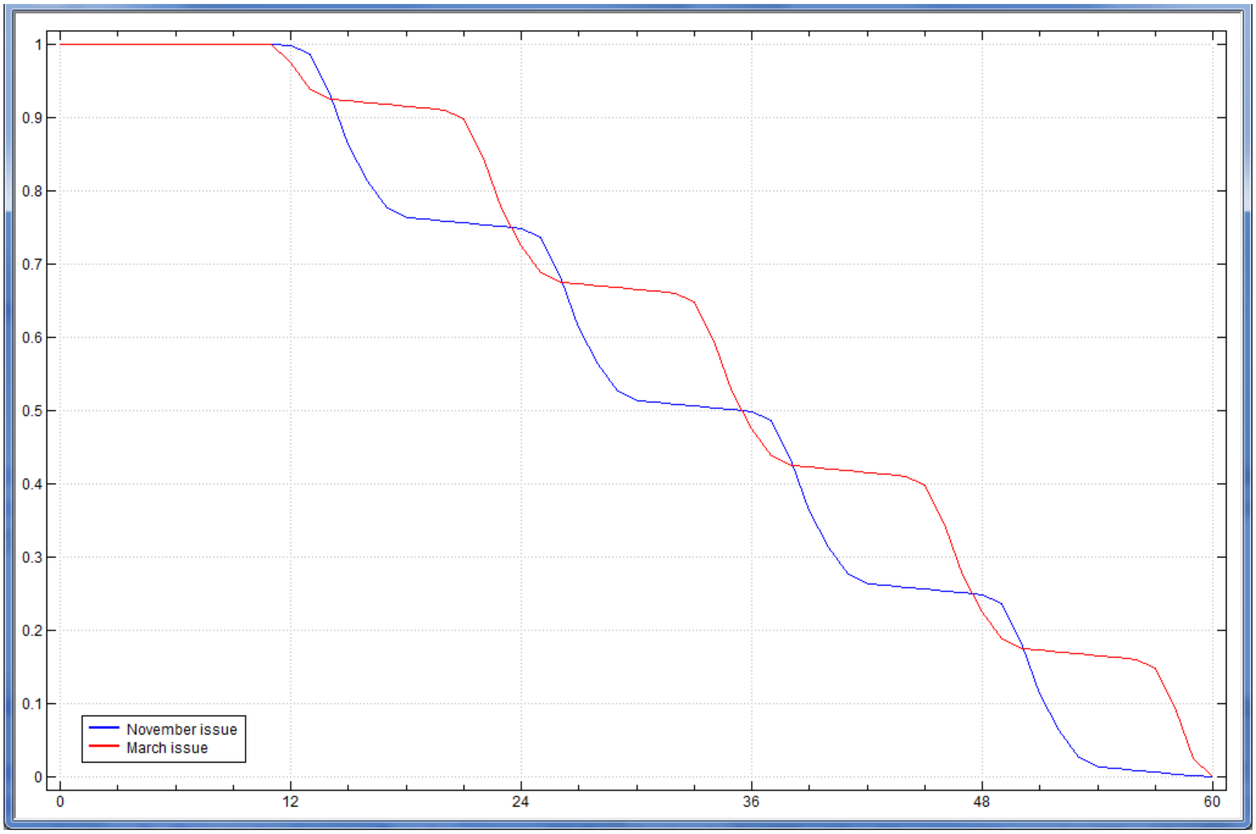



**Fig. 7.** Seasonalized UPR curves for snowmobiles, issued in November versus March

Implementing seasonalized UPR strings, even via a lookup-table formula, would appear to multiply storage requirements by 12; however, it may be done by factoring the strings into an underlying non-seasonal component and a single set of twelve seasonality factors. Estimating seasonality factors from the data may be done by first estimating the non-seasonal component with a 13-month moving weighted average and removing it from the raw data. If desired the residuals may be credibility-weighted against an a-priori set of seasonality factors and/or smoothed by application of a "circular" variation of Whittaker-Henderson graduation. One should to be cautious about smoothing, however, since some monthly spikes are real, such as those motivated by the timing of annual sales or service promotions.

## 2.11 Loss Development Factors Excluding the Effect of Cancellations.

An important difference between loss development for Warranty Insurance and for conventional lines is that for SSAP 65 Test 2 we need the lag factors that would be observed if there were no cancellations. In effect we need to decompose the total emergence pattern of losses into a component due to cancellations, which for UPR purposes we discard, and a component due to loss emergence proper, which becomes the basis for Test 2. This section describes one approach to this problem.

Assume we have a triangle of exposures $E_{ij}$ that reflects cancellations, so that the $E_{ij}$'s may decrease, but not increase, along each the row for each issue month *i*. By exposures we mean such things as in-force premium, loss fund (i.e., provision in the premium for losses), or contracts. $E_{i0}$ is the exposure at the moment of writing; $E_{i1}$ the exposure at the beginning of the next month, and so forth. Also assume that we have a triangle $L_{ij}$ of incremental losses. Here $L_{i0}$ are the losses in the month of writing, $L_{i1}$ in the next month, etc. Thus the exposures $E_{ij}$ give rise to the losses $L_{ij}$. To keep things simple we assume that a full month of exposure is earned by each contract in the month of cancellation. In practice we may refine the model slightly by adjusting the $E_{ij}$'s so as to recognize only partial exposure (usually one-half) in the month of cancellation.

We can convert $L_{ij}$ to a cumulative triangle with cells $C_{ij} = L_{i0} + L_{i1} + \ldots + L_{ij}$, and calculate development factors from cell $(i,j)$ to cell $(i,j+1)$ in the usual way as $F_{ij} = C_{i(j+1)} / C_{ij} = 1 + L_{i(j+1)} / C_{ij}$. These factors will reflect both ordinary loss development and cancellations. This is fine for some purposes, including obtaining a rough estimate of Test 3. But for other purposes, such as evaluating UPR curves, we need an efficient way to measure just the loss development for reasons *other* than cancellation, and these simple cumulative factors will not suffice.

A good start would be to adjust the $C_{i(j+1)}$'s and $C_{ij}$'s to the same exposure level. It is tempting to try to adjust $C_{ij}$ from its reduced exposure level $E_{ij}$ to the original exposure $E_{i0}$ before any cancellations. We can't do this directly, since $C_{ij}$ corresponds not just to $E_{ij}$ but to the whole history of exposures $E_{i0}, E_{i1}, \ldots, E_{ij}$, and these may vary independently. The only way to adjust $C_{ij}$ is to adjust each of its incremental pieces $L_{i0}, L_{i1}, \ldots, L_{ij}$ and recombine them. The adjusted $L_{ik}$ may be written as $L^*_{ik} = L_{ik}E_{i0}/E_{ik}$. Accumulating the $L^*_{ik}$'s to get $C^*_{i(j+1)}$ and $C^*_{ij}$ and taking their quotient we get a development factor $F^*_{ij}$ with the effect of cancellations removed.

But even this is not quite what we need. We are not usually interested in the triangle of cell by cell development factors, but in some sort of average down each column, to be used as an estimate of the underlying population development factors. For such an estimate to be efficient, we should give greater weight to those cells with greater losses or with greater exposure remaining after all earlier cancellations. Otherwise the larger variation in the smaller cells will contribute disproportionately to the variability of the average development factors.

The most common weighting for conventional loss development is by the losses in the denominator, which makes the average development factor in each column equal to the sum of the numerators divided by the sum of the denominators. (Other weightings involving powers of the losses in the denominator are appropriate for particular models of the loss process. Here we are concerned only with a preliminary weighting to correct for changes in exposure.) The conventional weighting works nicely when the denominators are the $C_{ij}$'s, but not when they are the $C^*_{ij}$'s, because the adjustment of $L_{ij}$ *to* $L^*_{ij}$ flattens out differences in volume of losses by issue month that result from different numbers of cancellations, and the weighting doesn't accomplish much.

A way around this problem is to put each $L_{ik}$ on the same exposure basis as $L_{i(j+1)}$ (that is $E_{i(j+1)}$) by letting $L'_{ik} = L_{ik}E_{i(j+1)} / E_{ik}$. The adjusted cumulative losses become $C'_{ik} = L'_{i0} + L'_{i1} + \ldots + L'_{ik}$, and the adjusted development factor becomes $F'_{ij} = C'_{i(j+1)} / C'_{ij}$. This is exactly the same as $F^*_{ij}$, but this time the weighted average works in the natural way:

$$\Sigma_i C'_{ij} F'_{ij} / \Sigma_i C'_{ij} = \Sigma_i C'_{i(j+1)} / \Sigma_i C'_{ij} = (\Sigma\text{numerators})/(\Sigma\text{denominators})$$

This looks as if it would be cumbersome to calculate. We have to loop through the $j$'s and for each one adjust all the earlier $L_{ik}$'s and sum them to get the $C'_{ij}$'s. But notice that

$$C'_{ij} = \Sigma_k(L_{ik}E_{i(j+1)}/E_{ik}) = E_{i(j+1)}\Sigma_k(L_{ik}/E_{ik}) = E_{i(j+1)}\Sigma_k P_{ik} = E_{i(j+1)}Q_{ij},$$

where $P_{ik}$ = partial loss ratio and $Q_{ik}$ = cumulative partial loss ratio. Similarly,

$$C'_{i(j+1)} = E_{i(j+1)}Q_{i(j+1)},$$

so that

$$F'_{ij} = C'_{i(j+1)} / C'_{ij} = Q_{i(j+1)} / Q_{ij}$$

No adjustment here, just developing the triangle of cumulative partial loss ratios, but we get the same factors!

We do not want to use an unweighted average of these factors since all differences in exposure, whether due to cancellations or otherwise, have been flattened out of them. Two approaches come to mind that make the weights reflect volume of experience.

If we weight the ratios $Q_{i(j+1)}/Q_{ij}$ by the $E_{i(j+1)}$'s, we get the following average:

$$(\text{Avg } F)_j = \Sigma_i(Q_{i(j+1)}/Q_{ij})E_{i(j+1)})/\Sigma_i E_{i(j+1)},$$

which equals (Σnumerators)/(Σdenominators) if we write each development factor as

$$F'_{ij} = (Q_{i(j+1)}/Q_{ij})E_{i(j+1)}) / E_{i(j+1)}$$

On the other hand, if we weight the ratios $Q_{i(j+1)}/Q_{ij}$ by the adjusted losses $C'_{ij}$, we get

$$(\text{Avg } F)_j = \Sigma_i(Q_{i(j+1)}/Q_{ij})(C'_{ij} / \Sigma_i C'_{ij}) = \Sigma_i Q_{i(j+1)}E_{i(j+1)}/\Sigma_i Q_{ij}E_{i(j+1)},$$

once again a "sum of numerators over sum of denominators" situation if we write each development factor as

$$F'_{ij} = Q_{i(j+1)}/Q_{ij} = Q_{i(j+1)}E_{i(j+1)}/Q_{ij}E_{i(j+1)}$$

Both of these formulas for (Avg $F$)$_j$ are easy to implement in software since we already have the exposure triangle. The first average is weighted by exposures, the second is weighted by losses. They are not identical, but they are consistent with each other because expected losses are proportional to exposures.

A simulation study involving extreme exposure patterns (some issue months nearly fully cancelled, others with few cancellations) shows that (a) the weighting by exposures appears to be more stable than the weighting by losses, (b) the weighting by losses at the exposure level of the numerator is practically indistinguishable from the weighting by losses at the initial written exposure level, (c) all of these variations of the loss development method are more stable than the Partial Loss Ratio (Stanard's "Additive") approach, (d) there is no evidence of bias in any of the methods, and (e) the differences among methods are small and would probably be immaterial with real-world data.

## 2.12 Calculations Involving Issue Date versus Breakdown Lag Triangles

In principle, the UPR covers the interval between issue and breakdown, the IBNR reserve covers the interval between breakdown and reporting, and the RBNP reserve covers the interval between reporting and payment.

We can estimate different components of the total reserve from different triangles, as follows:

| | **Cell contents** | **Time 1** | **Time 2** | **Resulting reserve** |
|---|---|---|---|---|
| A | Paid | Issue | Payment | UPR + IBNR + RBNP |
| B | Paid | Breakdown | Payment | IBNR + RBNP (i.e., IBNP) |
| C | Paid | Report | Payment | RBNP |
| D | Paid + Pending | Issue | Report | UPR + IBNR |
| E | Paid + Pending | Breakdown | Report | IBNR |
| F | Paid | Issue | Breakdown | UPR |
| G | Paid + Pending | Issue | Breakdown | UPR |

(A,D,F,G) UPR here means expected losses not yet incurred, or SSAP 65 Test 3, roughly equal to the usual Test 2 or gross UPR times an expected loss ratio.
(D,E) Relies on the pending reserve's being exactly equal to the eventual payment, in which case the paid-plus-pending data is equivalent to a case-basis incurred triangle.
(F) Unreliable without adjustment because some losses in the last few diagonals may not have been paid by the ending date of the data
(G) Unreliable without adjustment because some losses in the last few diagonals may not have been reported by the ending date of the data

We may use the technique described in 2.11 to handle triangles of types A and D, and ordinary loss development to handle B, C, and E. But triangles of types F and G require special treatment.

The problem is that the triangle of losses by issue month versus breakdown lag is not stable, in the sense that the next month's experience will not simply add a new diagonal and leave the existing triangle unchanged. Instead, the next month's experience may contain newly-reported losses that belong in interior cells of the existing triangle, usually somewhere in the last few diagonals. If we are not careful, developing an issue-versus-breakdown triangle may understate the true reserve because of these incompletely-paid or incompletely-reported losses in the recent diagonals. Note that case F (paid losses) and case G (reported losses) are structurally identical, except that F, in effect, treats each claim as if its report date were equal to its payment date.

One solution to this problem is to eliminate the latest diagonals from the triangle and obtain development and earnings factors from the curtailed triangle. This works, but it sacrifices information, especially for situations where report or payment lags stretch for more than three or four months.

A second approach is to use the breakdown-date versus payment-lag or report-lag triangle to obtain payment or report lag factors, cumulate these to obtain the expected fraction of losses paid or reported at the end of one, two, … months, apply these factors to the triangle of exposures (working in on each row from the latest diagonal) to obtain the "expected paid" or "expected reported" fraction of exposure, and calculate earnings factors from the original losses and these adjusted exposures. This is the most robust approach and makes full use of the latest information.

Still another approach is to use a single issue-date versus breakdown-lag triangle to estimate simultaneously the earnings factors (in the breakdown-date direction, a "column effect") and the fraction unpaid or unreported (in the calendar-month direction, a "diagonal effect"). From the latter we may derive lag factors for losses IBNP or IBNR, *even when we do not have payment or report dates of individual losses or any triangle by breakdown date versus payment or report lag.* Estimation proceeds iteratively, alternately using exposure-adjusted loss development as described in section 2.11 to estimate the earnings factors, then using the ratio of actual to fitted losses in the last few diagonals to estimate the cumulative fractions of losses reported at each lag, and then using these fractions to adjust the exposure in those diagonals for the next iteration.

This "one-triangle" approach is occasionally needed, produces good results with enough data, and should be part of the software tool kit of any Warranty insurer. But, whenever possible, it is better to use the breakdown-month versus report-lag triangle to estimate the report lag factors, since this makes use of all the claims in the data, not just those claims that happen to have been incurred within a few months of the valuation date. For example, in the case of 120 months' steady-state data with constant loss emergence, if all claims are settled within three months so there are just two independent payment lag factors to estimate, only the latest two diagonals of the issue-to-breakdown lag triangle will yield information relevant to estimating the "diagonal effect" of unreported losses, and these diagonals will contain only about 3.3% of the total claims that would be included in the full breakdown-to-report (or breakdown-to-payment) lag triangle.

## 2.13 Tail Factors

It might be thought that tails would be less important when estimating UPR factors for Warranty insurance than when projecting accident-year runoff for other lines, because the duration of Warranty runoff is limited by the term of the contracts. However, when we calculate tail factors we are not so much trying to answer "how much will be paid each month in the future" as "how much remains to be paid relative to what has already been paid". The results of an analysis of immature data are sensitive to this question no matter how long the tail may be in months. The only advantage of Warranty insurance over other lines in this respect is that the terms limit the number of distinct issue months for which tails need to be calculated.

The need for tail factors only arises when estimating indicated UPR factors, since the tail is already incorporated in strings or formulas used for carried UPR factors. The strings and formulas are usually known at the time of the analysis, so it is tempting to consider the problem solved: just append a tail with the same shape as the average tail of the strings or formulas. However, this does not answer the question of how large the entire tail should be relative to what came before.

To illustrate, suppose we have observed losses for 27 months in a new program with term 72 months and carried reserve pro rata. The losses have emerged irregularly but moderately increasing with lag from issue. If we assume that losses in the tail after 27 months will be constant, producing a pro rata UPR curve, we know that the shape of the tail will be a straight line, but we don't know how steep that line will be. Figure 8 shows three possibilities.

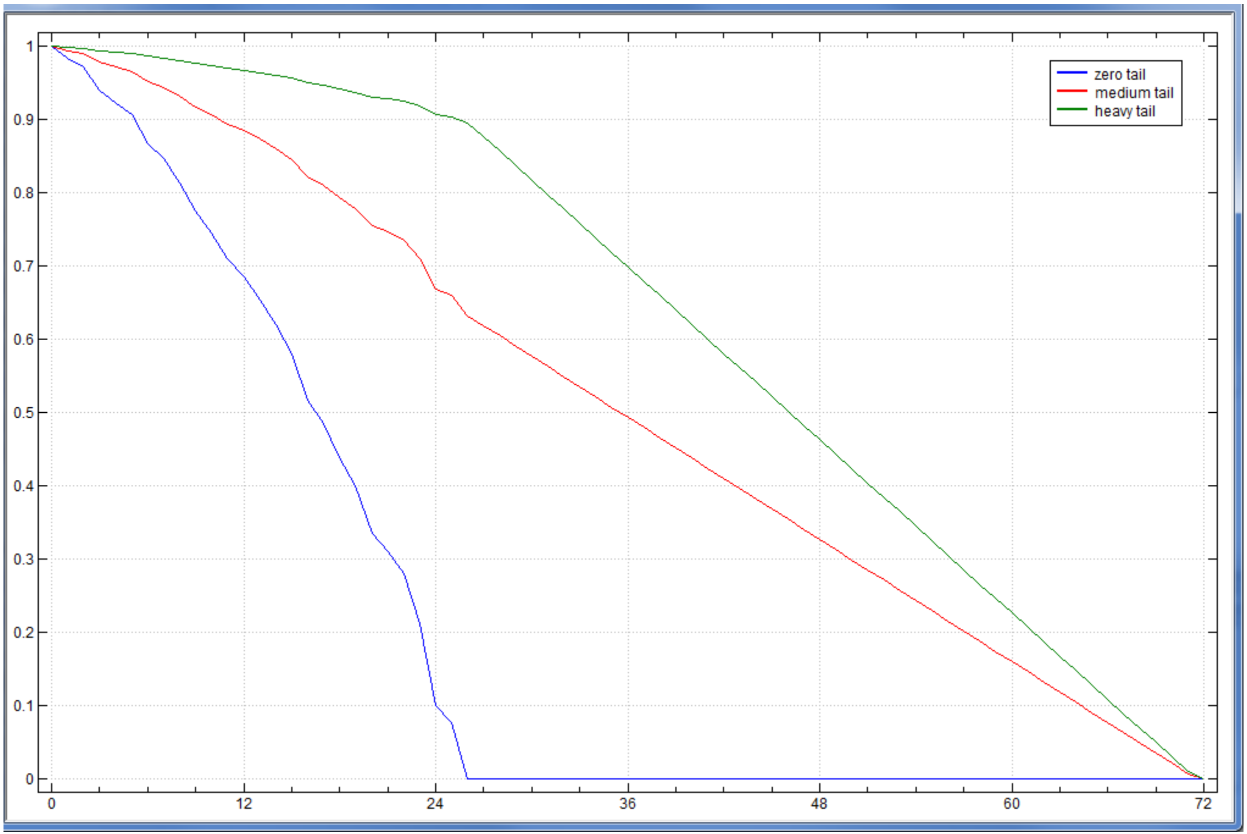



**Fig. 8.** Possible pro rata tails starting from the same observed losses through 27 months.

The "medium tail" in Figure 8 looks more reasonable than the other two. The key is that its earnings rate was derived from the known data; in fact it equals the mean rate of loss emergence over the entire 27 months of known data.

It might be objected that the slope of the tail thus selected is too gentle, since the known part of the curve exhibits a steepening in the last few months. One way to address this problem is to shorten the "lookback period" for averaging earnings to something shorter than all the known months. Another is to apply linear regression to the known earnings factors and either extrapolate them to 72 months or (as we prefer for stability) letting the earnings rate through the tail be constant at the fitted value for 27 months. A third way is to make the tail steeper using a factor selected by judgment. These possibilities are illustrated in Figure 9, using, respectively, the latest 6-month mean of the observed earnings, regression to month 27 of all 27 observed earnings, and a doubling of the earnings rate from the original 27-month mean. This does not double the slope of the tail, but doubles the ratio of the earnings in the tail to the earnings before the tail, since all of the UPR curves are based on earnings normalized to total 1.00.

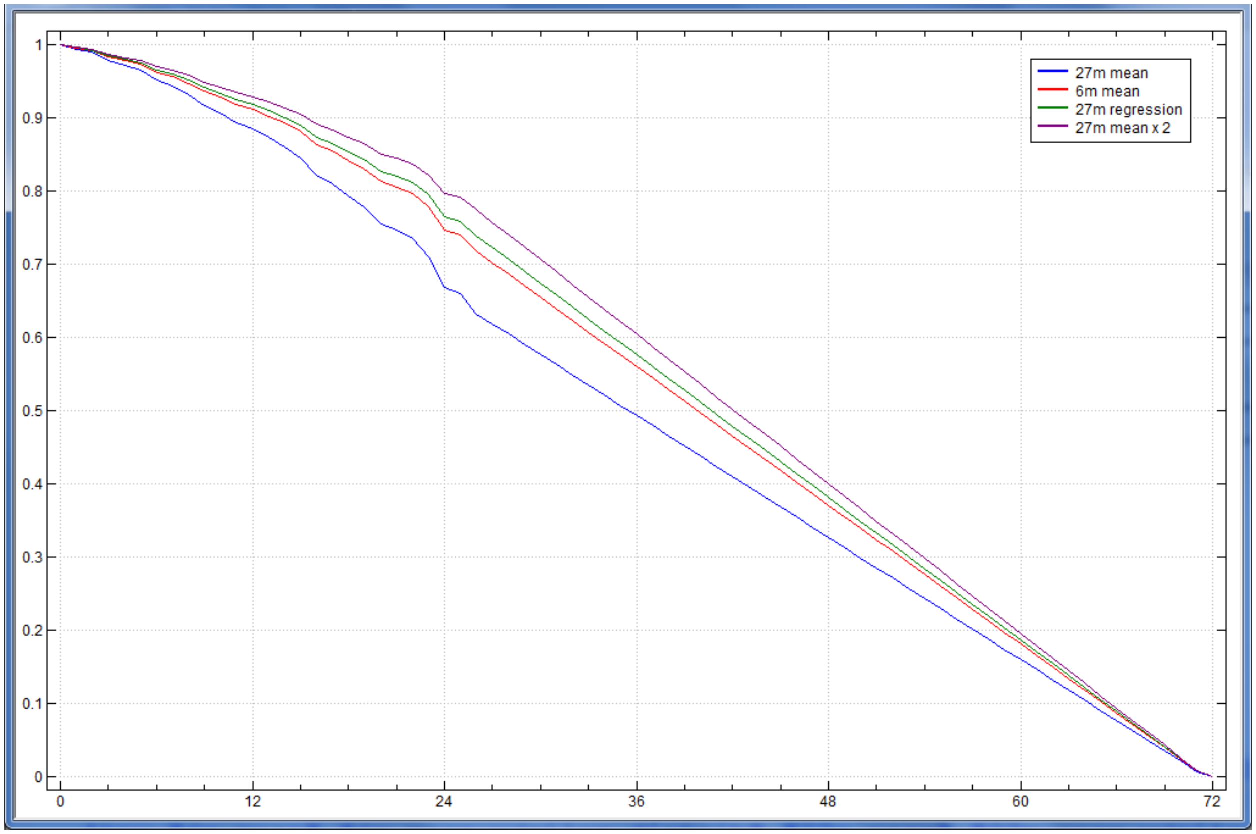



**Fig. 9.** Different determinations of initial tail slope

It is not necessary that the tail start with the first lag following the known data; in cases where the last few known lags are irregular, the tail may be made to start a few months earlier.

Nor is it necessary that the tail be linear. When dealing with products for which the natural UPR curve earns more rapidly at first than later, we may fit Rule of 78 tails; when breakage is apparent we may fit a curve with exponentially declining first differences; when "miling out" is a factor we may fit a curve that is a weighted average of tails for the various tail lengths corresponding to, say, a 50-point discrete distribution of miles driven per annum. The latter adjustment makes the UPR curves less conservative, since the miling-out of some contracts reduces the earnings in the later lags; it is illustrated in Figure 10. The two lower curves actually fit linear tails just like the two top curves, except using a weighted mixture of term lengths.

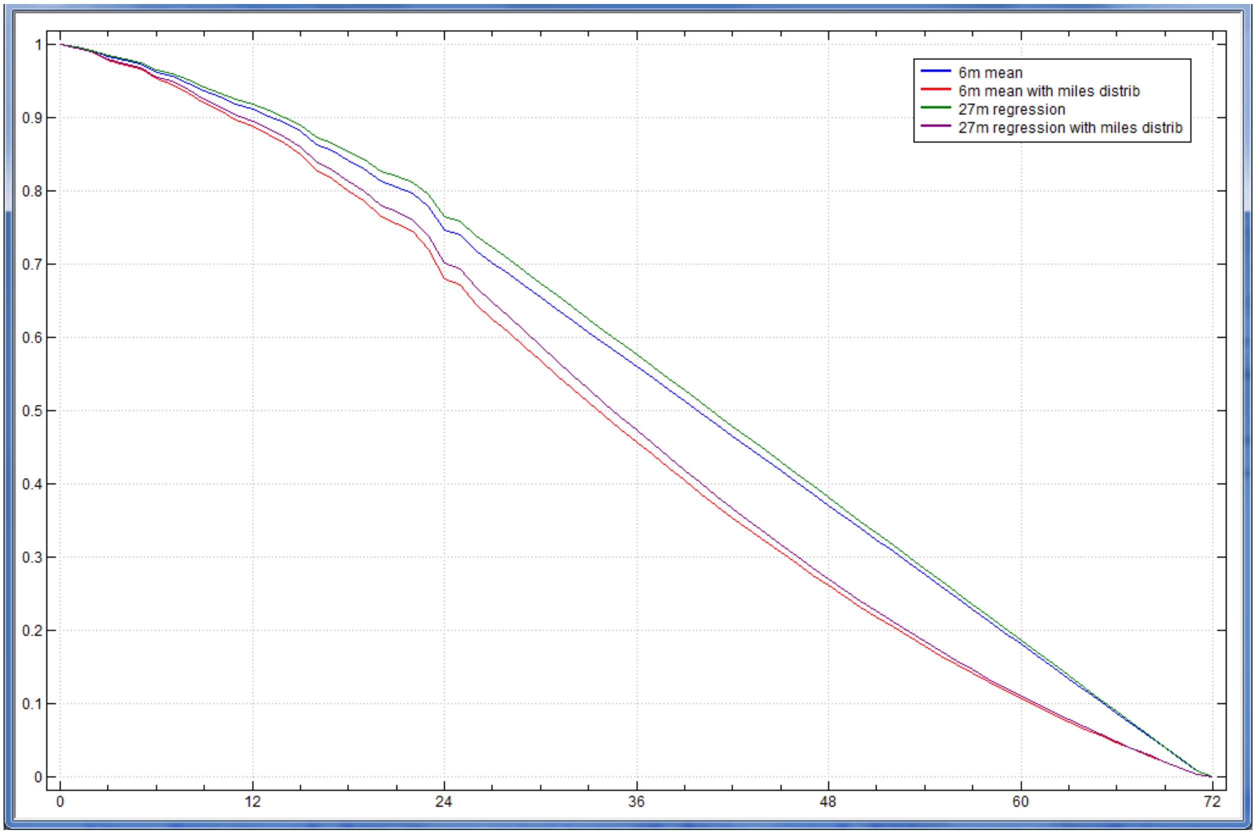



**Fig. 10.** Effect of using miles-per-annum distribution to create a mixture of tail terms.

In practical work the actuary may need rules for calculating tails, when necessary, for multiple programs and/or terms being analyzed in a batch, without halting execution for insertion of judgment parameters. Fitting tails for Warranty contracts is as much an art as a science (just as is the case for other lines) and cannot eliminate the uncertainty inherent in any extrapolation. But it can be systematized. Our preference is to divide the problem into three parts, determining for each tail (1) its starting point, (2) its shape, and (3) its weight.

When batch-processing multiple programs, we have generally found it satisfactory to start the tails a month or two earlier than the last available lag, to make the tail pro rata for each group of contracts expiring at the same time, using an estimated miles-per-annum distribution, and to average the earnings over a lookback period of about one quarter of the observations to determine the initial slope of the tail. When fine-tuning a single projection, we often prefer to use an initial earnings rate determined by regression, in order to join the tail smoothly to the known data. This requires inspection and possible judgment intervention, however, lest the fitted value at the last lag before the tail become negative or otherwise unreasonable.

Possible tail shapes are not limited to pro rata and pro rata summed across remaining terms of different lengths. For some lines of business, such as Tire & Wheel, we may let the earnings pattern for each remaining term decline linearly to zero, in a sum-of-digits fashion. And it is possible to let the *carried* UPR for each contract in a block of business, summed across contracts to the end of each term, determine the shape of the tail when otherwise estimating the UPR from experience and ignoring the carried UPR factors.

## 2.14 SSAP 65 Test 3 Estimates

For SSAP 65 Test 3 we need to convert the cancellation survivorship patterns and loss emergence patterns observed in estimating Test 2, together with incurred losses, into estimates of losses for existing contracts unincurred at the accounting date.

While the Chain-Ladder estimator, with the adjustments described in 2.11, is well adapted to estimating the Test 2 UPR pattern (and therefore the Test 2 UPR) of a block of Warranty business, it is not usually satisfactory for the Test 3 UPR, expected unincurred losses. The reason is that many extended Warranty contracts earn very little during their early months, causing projection of expected losses for the latest issue months to be impossible or erratic.

On the other hand, we can adapt the Bornhuetter-Ferguson method (BF) to produce stable projections of the Test 3 UPR, to accommodate known trends or seasonality in the calendar-month direction, and to be responsive to unknown or variable trends in the issue-month direction.

The adaptations involve (a) using Hans Bühlmann's Cape Cod estimator [2], adjusted for declining exposures, to obtain an expected loss ratio (ELR), (b) adjusting this procedure as necessary using Spencer Gluck's decay factors, to increase responsiveness to changes in the ELR by issue month, and (c) detrending and/or deseasonalizing the data entering the calculations and restoring trend and/or seasonality to the results.

### 2.14.1 Cape Cod

*In the traditional calendar-accident year context*, the Cape Cod estimate of the ELR equals losses reported (or paid) to date divided by the portion of premium (or other measure of exposure) expected to have emerged as losses over the same time period. The numerator may simply be summed over the entire incremental loss triangle, i.e., down the latest diagonal of the cumulative loss triangle. The denominator is usually summed across accident periods, with each accident period contributing its initial earned premium multiplied by the reciprocal of its cumulative development factor.

*In the Warranty context*, with issue months instead of accident periods, with in-force premium that declines across development months, and with recent diagonals deficient because of unreported losses, the denominator of the ELR must be adjusted. It is now summed across all cells of the triangle, with each cell contributing its initial in-force premium multiplied by the applicable incremental lag factor and, if we are working with an issue-month versus breakdown lag loss triangle, by the applicable factor for the expected fraction reported through the valuation date. No adjustment is necessary to the numerator; it remains just the sum of paid losses across all cells of the incremental loss triangle.

**2.14.2 Gluck factors**

In a 1997 PCAS article [3], Spencer Gluck proposed an enhancement of the Cape Cod estimator that puts BF/Cape Cod at one end of a spectrum with the pure Chain-Ladder at the other. Gluck's contribution may be thought of as the third in a series of simple, non-stochastic, but eminently practical approaches to loss reserving, the other two being the Bornhuetter-Ferguson method itself [1] and the Cape Cod estimator of the expected loss ratio (1983). We have found Gluck's procedure to be particularly useful for Warranty insurance, where product redesigns and technical changes result in irregular trends in the issue-month direction.

Gluck's approach produces a separate ELR for each issue month. Notice that the Cape Cod estimator, the sum of the losses divided by the sum of the adjusted exposures, is actually an average of the loss *ratios* in each cell, weighted by adjusted exposures. Gluck multiplies these weights by a "Gluck weight" dependent on the distance between each issue month and the target month via a geometric decay factor $g$. The scale is immaterial so we may assume that the target month receives Gluck weight 1. Then the adjacent months receive Gluck weight $g$, the next further months receive Gluck weight $g^2$, the next further months $g^3$, and so forth. For each target month, the ELR equals the sum of $g$ times losses divided by the sum of $g$ times adjusted exposure.

If the decay factor $g$ equals 1, this gives the pure BF/Cape Cod; if $g$ equals 0, it gives the pure Chain-Ladder; as $g$ moves from 1 toward 0, it gives an ELR that becomes increasingly responsive to local trends.

**2.14.3 Trending and seasonality**

The procedures to account for trend and seasonality apply to both Test 2 and Test 3, but these considerations have a much greater impact on Test 3 than on Test 2. Our preferred approach is to deflate and deseasonalize the losses in the source data, based on breakdown date, and, to the extent possible, recalculate premiums in the source data, usually putting both losses and premiums on the level of the valuation date. If we have correctly measured the trend and seasonality to remove, the data becomes stationary and yields UPR factors, ELR's, calendar-month projected losses, and Test 3 UPR's with no embedded trend or seasonality. We then invert the process, restoring the known losses to their historical values and trending and seasonalizing the future losses, producing a trended and seasonalized estimate of Test 3. The future trend and seasonality factors need not be identical to the historical ones.

This whole procedure depends on estimates of historical trend and seasonality. In principle there could be one cumulative trend factor – or even one cumulative seasonality and trend factor – for each historical month. In practice seasonality is usually expressed as a set of 12 factors, one for each calendar month, and trend is usually expressed via a single annual trend factor, presumably applicable to the entire time period from the earliest included month through the valuation date.

Estimating trend and seasonality is complicated by the fact that many quantities we would like to measure can only properly be compared with earned premium or earned contract counts, running the risk that errors in earnings pattern may be confused with trends. These quantities, for which measured trends are suspect, include frequency, loss costs, and loss ratios. On the other hand severity is not dependent on earnings and calendar-month severity trend may be estimated fairly easily and reliably and used as a proxy for total trend. We usually calculate average severity by breakdown month, fit both linear and exponential curves to it, and inspect the plots for evidence of discontinuities which might justify separate trends for different time periods. As for seasonality, we usually deseasonalize the paid severity data using a 13-month moving weighted average and estimate the seasonality by averaging the residuals over a multiple of whole years grouped in 12 columns by calendar month.

Because issue month, calendar month, and development lag are multicollinear we cannot obtain unique simultaneous estimates of pure trends all three of these directions, but with reasonable assumptions serving as constraints, we can use GLM's to estimate issue and development frequency trends and issue and calendar severity and loss-cost trends, and take the product of calendar trend for severity and the development trend for frequency as another estimate of total development trend to be used in deflating UPR curves.

### 2.14.4 Contractual Liability policies

In many states, service contracts issued by a retailer directly are not regulated as insurance and service contracts issued through an affiliated obligor are not regulated as insurance provided the obligor itself insures its obligations. For this reason Warranty insurers find themselves underwriting some Warranty programs from the ground up, others entirely conditional on the obligor's failure to perform (FTP), and still others so as to cover a share of the risk from the ground up and the remaining share on an FTP basis.

Coverage written from the ground up is assigned the Warranty line of business for annual statement purposes; coverage entirely on an FTP basis is considered Contractual Liability Insurance under the line of business Other Liability – Occurrence. The most reasonable analysis of coverage partially insured with the rest on an FTP basis is that part of the premium (often a small part) should be ascribed to Contractual Liability and the remainder to Warranty, all claims should be ascribed to Warranty unless and until the program is in FTP status, and thereafter the insured share of claims should be allocated to Warranty and the remaining claims to Contractual Liability.

Typically a Warranty insurer will impose financial requirements on any obligor insured on an FTP basis to protect its interest in that obligor's solvency, and the insurer may administer the entire program and its cash flow. With such protection in place the actual FTP premium may be only a small fraction of the amount that would be required on a first-dollar basis. For this reason, SSAP 65 Test 2 will never be more than a small fraction of SSAP 65 Test 2 applied to the obligor's underlying contracts – although the UPR factors should be somewhat greater than the factors for the underlying contracts because expected loss emergence is deferred by being conditional on future bankruptcy.

If the FTP is priced correctly (which on account of the catastrophe risk would justify a low expected loss ratio) then Test 3 should be even smaller than Test 2; but in the event the program enters FTP status, Test 3 rises immediately to the full Test 3 at the obligor level. In principle, the evaluation of Test 3 for a program not in FTP status requires estimation of the probability of first entering insolvency at each future month together with the expected unincurred losses at that time.

## 2.15 All-Terms Factors

Suppose we need strings for a set of target contracts differing from each other in term, manufacturer's warranty, odometer reading at issue, and so forth. Some or all of the target contracts may not be represented in the available data, or may be represented but in too small a volume to produce reliable indicated UPR factors. Or reasonably credible data may be available and each indicated string may appear reasonable in isolation but the strings for different contracts may not be related to each other in a logical way.

For example, Warranty contracts are usually marketed for a few distinct terms, such as 60 months or 60,000 miles (conventionally written 60060), 72 months or 90,000 miles (72090), and so forth. When these *nominal terms* are effective at issuance of the contract, the data may easily be broken into a few homogeneous term groupings. But when the nominal terms are effective on expiration of the manufacturer's warranty, there may be a great many *actual terms* from issue to expiration, depending on how much of the manufacturer's warranty remains at time of issue. In this case it is unlikely that the available data would contain enough contracts to estimate UPR curves for each separate actual term, not to mention also account for differences in remaining manufacturer's warranty.

Here we describe a technique that uses an entire body of data, with many terms, to derive a *family* of UPR curves, each curve reflecting the experience of *all* the terms in the data, weighted by proximity to the term being evaluated. We call the model that generates such a family of curves *All-Terms Factors,* or ATF. It is in fact a comprehensive model of the loss emergence process for contracts on automobiles, but it is also applicable in simplified form to other contracts.

The ATF model derives a theoretical loss emergence pattern based on our contract, the manufacturer's warranty (MW), trend, breakage (i.e., policyholders' failure to present eligible claims), and certain other factors, all conditional on the car's being driven a given number of miles per annum. It then integrates over an assumed distribution of miles per annum to obtain the expected loss emergence pattern for the usual case where the driver's average usage is not known in advance. This emergence pattern, when converted to a UPR curve, is by itself useful in reserving, and, as explained below, may be thought of as a broad generalization of the concept of earning pro rata from the end of the manufacturer's warranty. This part of our ATF model turns out to be a modest extension of Kerper and Bowron's exposure model [4], mentioned above. We explain it here in detail, using our own terminology and notation; it has certain enhancements but in its essentials is nearly identical to their model.

Our algorithm for applying the model derives the expected loss emergence pattern in units proportional to expected losses, before eventually normalizing it to total 1.00. Therefore in addition to the earnings pattern *per se*, it yields estimates of relativities between any pair of proposed contracts. These relativities, when taken to a common base contract, are very useful in ratemaking.

We may apply the model of loss emergence directly to a proposed contract, on the assumption that it accounts for all relevant factors. *Alternatively, we may fit the model to a body of data and obtain a family of curves, one for each term in months, measuring only the effects of any factors* not *included in the model.* We call these curves *residual ATFs*. We may then derive the UPR curve for any contract, whether part of the data base or not, by applying the contract's theoretical loss emergence pattern to the residual emergence pattern for the same term. The residual ATFs for any given term are derived not just from the experience of contracts with that term, but from *all* contracts in the data, weighted by proximity to the target term.

If the original model really does fully explain the variability of loss emergence patterns, the residual ATFs will all be straight lines and using them will make no difference. Otherwise, starting with the residual emergence pattern will improve the performance of the model on new contracts, provided the residuals were based on a large enough volume of relevant cxperience.

We can, if we wish, ignore some or all of the usual explanatory variables and let their effect (if any) appear in the residual ATFs. If we ignore all variables except term in months, the residual ATF's become a family of average UPR curves, one for each term. We call this the *simplified* ATF model. It is appropriate as an aid in reserving for whole term groupings in the same body of data from which the ATF's are derived, but it will not usually produce reliable UPR curves for individual contracts or for different bodies of data. For these purposes the general ATF model, with its explanatory variables, expected loss emergence pattern, and residual ATF's, is required.

The ATF model starts with the assumption that parts failures will occur whether or not a car is covered by a warranty, and that the number or cost of such failures may be subject to trend and may depend on the car's usage as well as on the passage of time. Then the model considers how parts failures translate into claims against our contract; this depends on the contract term in months and miles, the MW in months and miles, the relationship between services covered by our contract and by the MW, factors such as breakage and pre-expiration claims spikes, and the distribution of miles driven per annum. Finally the model provides procedures for estimating the residual UPR curves, and other parameters, from experience.

### 2.15.1 Emergence of failures

We assume that any car is subject to parts failures of two kinds:

a. Those that depend on use (e.g., power train)
b. Those that depend on time (e.g., paint and trim)

Here we consider all failures, whether ultimately paid for by the manufacturer, the insurer, or the customer. The cost and the earnings pattern of a contract to the insurer will eventually depend on how this responsibility is allocated, but it all starts from the initial failure of some part of the car. The failure emergence rate is therefore the starting point of the All-Terms Factors model.

Let $s$ be distance driven in miles and $t$ be time elapsed in months. Assume that failures of type (a) emerge at a rate $\lambda_a$ per mile and failures of type (b) emerge at a rate $\lambda_b$ per month. We may conceive of these rates as Poisson lambda parameters for failure frequency, or as means per unit distance or time of some distribution of costs.

For a car driven $m$ miles per annum, $ds/dt = m/12$ and, if $y$ = number or cost of failures in time $t$,

$$E(y) = ((m/12)\ \lambda_a + \lambda_b)t$$

or

$$E(y) = (\lambda_a + (12/m)\lambda_b)s$$

For our purposes we do not need the values of $\lambda_a$ and $\lambda_b$ in absolute terms so for simplicity we replace them with a single assumption: the fraction $f_{mi}$ of failures that emerge in proportion to miles driven, with the complement $1\text{-}f_{mi}$ emerging in proportion to time elapsed. If we let $m_{avg}$ = average miles driven per annum, we express this as

$$f_{mi} = (m_{avg}/12)\ \lambda_a\ /\ ((m_{avg}/12)\ \lambda_a + \lambda_b)$$

so that

$$\lambda_b = (m_{avg}/12)\ \lambda_a\ ((1\ /\ f_{mi})\ \text{-}1) = (m_{avg}/12)\ \lambda_a\ (1\ \text{-} f_{mi})\ /\ f_{mi}$$

and

$$E(y) = (\lambda_a\ (1 + (m_{avg}/m)\ (1\ \text{-} f_{mi})\ /\ f_{mi}))s$$

$$= (K\ (f_{mi} + (m_{avg}/m)\ (1\ \text{-} f_{mi})))s$$

which implies

$$E(y) = (K(m/12)\ (f_{mi} + (m_{avg}/m)\ (1 - f_{mi})))t$$
$$= (K((m/12)\ f_{mi} + (m_{avg}/12)\ (1 - f_{mi})))t$$

where $K = \lambda_a / f_{mi}$ is a constant of proportionality.

The 12's appear in this formula to convert mileage from annual to monthly, because the conventional measure of auto usage is miles per annum, but our time $t$ is measured in months. Notice that $K$ does not depend on miles driven, but that the remaining (parenthesized) factor does, so that the expected failures of cars driven at different rates are proportional to this factor. However, $K$ may be generalized to be a function of calendar month, months elapsed since the car was put in service, or attained odometer reading, so that the loss emergence pattern and/or the relative expected costs of two cars may reflect various trends. For convenience we implement these trends at the same time as we implement factors representing plan design, manufacturer's warranty, breakage, etc., but if you wish you may think of them as a distinct step modifying the emergence of failures.

### 2.15.2 Emergence of claims

The heart of our exposure model is how we adjust the failure emergence rate to convert it to a claims emergence rate, reflecting the conditions of our contract, the manufacturer's warranty, the miles-per-annum distribution, time trend, odometer trend, breakage, claims spikes, and "extras".

Because of the requirements of SSAP 65 Test 2, the emergence pattern of claims is a guide to the desired earnings pattern of premium, and we shall speak of earnings patterns and emergence patterns interchangeably. The ATF model generates these emergence patterns up to a constant of proportionality times total losses, comparable across plans, so we may obtain valid relativities. For the UPR pattern we normalize the emergence factors to total 1.00 and then sum stepwise backwards from the last elapsed month to the first.

In a sense our entire adjustment process is nothing more than an elaboration of the simple pro rata, or constant-earnings, model. It results in a step function for earnings, and a piecewise-linear function for UPR factors, but there may be so many steps that the UPR function is essentially indistinguishable from a smooth curve.

We illustrate the successive elaborations in turn.

*Straight pro rata.* Suppose the term of a contract is 60 months or 60,000 miles, with no remaining manufacturer's warranty, and that it is applied to a car driven 12,000 miles per annum. Then the contract will expire on account of both time and mileage at the end of its $60^{th}$ month. Assume for now that the average number of miles driven per annum (across all cars) is 15,000 and that claims emerge 80% in proportion to miles and 20% in proportion to time. Then all emerging failures are covered by the contract, the earnings pattern is constant for 60 months, equal to $K\,((m/12)\,f_{mi} + (m_{avg}/12)\,(1 - f_{mi})) = 1050K$, and is zero thereafter, and the UPR pattern is pro rata starting from 1 at 0 months and reaching 0 at 60 months. Its total earnings, proportional to its total expected loss cost, will be 60(1050*K*) or 63000*K*.

Now suppose we have another car driven 18,000 miles per annum but otherwise identical to the first. This car will mile out at 40 months, will earn $K\,((m/12)\,f_{mi} + (m_{avg}/12)\,(1 - f_{mi})) = 1450K$ per month for 40 months and zero thereafter, and its total earnings will be 40(1450*K*) = 58000*K*. The expected loss cost of the second car relative to the first is therefore 58000/63000 = 0.921. The UPR pattern of the second car is pro rata starting with 1 and reaching 0 at 40 months.

Similarly a third car driven 30,000 miles per annum but otherwise identical to the first will mile out at 24 months, will have total earnings of 54000*K*, and a relativity to the first car of 0.857.

The emergence pattern of the first car (reduced proportionally to start at 1) is shown in Figure 11a; the UPR patterns of all three cars are shown in Figure 11b.

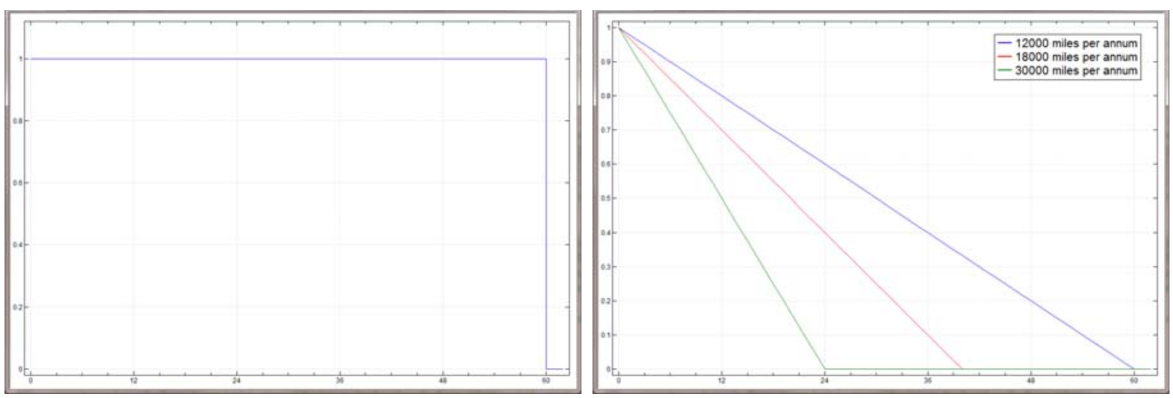



**Fig. 11a and 11b.** Constant earnings over term of contract, and straight pro rata UPR factors.

*Pro rata from end of MW*. Suppose each of these cars is subject to a manufacturer's warranty of 36 months or 36,000 miles. For the first car, this warranty will expire at the end of the 36th month. The second car will mile out of its MW after 24 months, and the third car after 14.4 months. We assume no earnings for our contract during the MW. The earnings pattern (Figure 12a) now has three steps, zero from issue to end of MW, non-zero constant from end of MW to expiration, and zero thereafter. The UPR pattern (Figure 12b) is flat until the MW expires, then pro rata to 60 months for the first contract, 40 for the second, and 24 for the third.

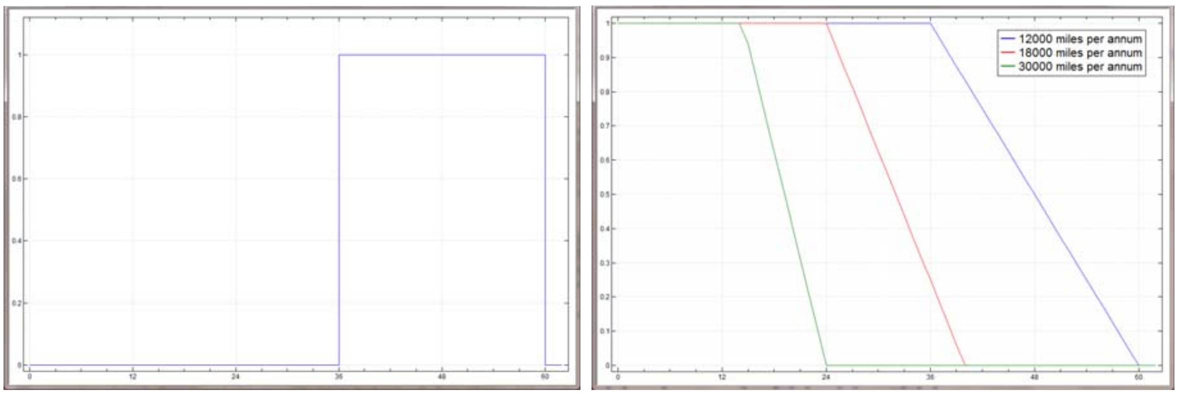



**Fig. 12a and 12b.** Constant earnings, and pro rata UPR factors, from end of MW.

*Power train and non-power train MW's*. Now suppose that each car has an extended power train (PT) warranty of 50 months or 50,000 miles, and that 35% of all failures relate to the power train. Now the earnings pattern (Figure 13a) has an additional step and the UPR pattern (Figure 13b) has an additional slope.

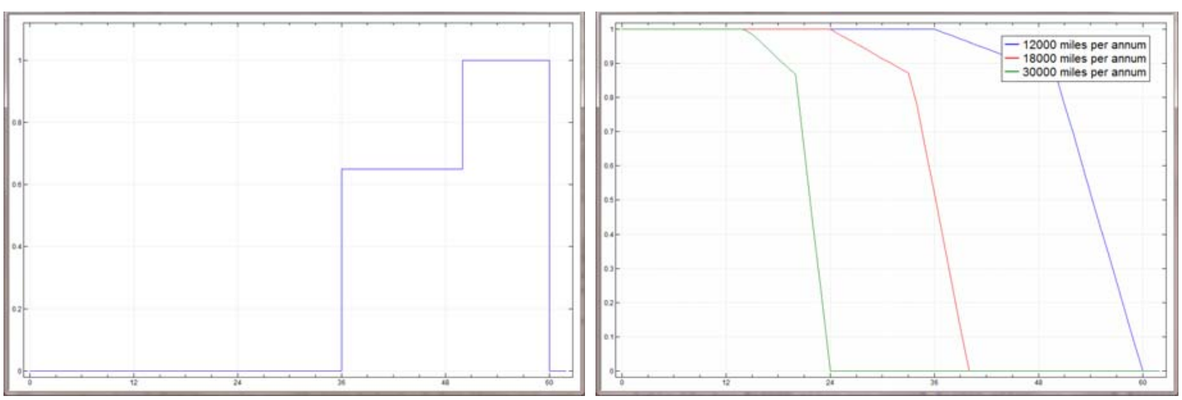



**Fig. 13a and 13b.** Earnings with PT and non-PT MW's and corresponding UPR factors.

*Extras and limitations*. Up to now we have assumed that the full cost of a failure, and nothing more, is borne either by the MW or by our contract. But our contract may in fact cover only part of what the MW covers. We adjust for these limitations via factors $F_{pt}$ and $F_{npt}$ applied to the PT and non-PT failures respectively. On the other hand, our contract may cover services, such as towing, that are not part of the MW. We relate these "extras" to $y$ via a factor $E$. For example, $F_{pt}$ might be

0.99, $F_{npt}$ 0.91, and $E$ 0.03. Figures 14a and 14b show the effect of these values on earnings pattern and UPR factors.

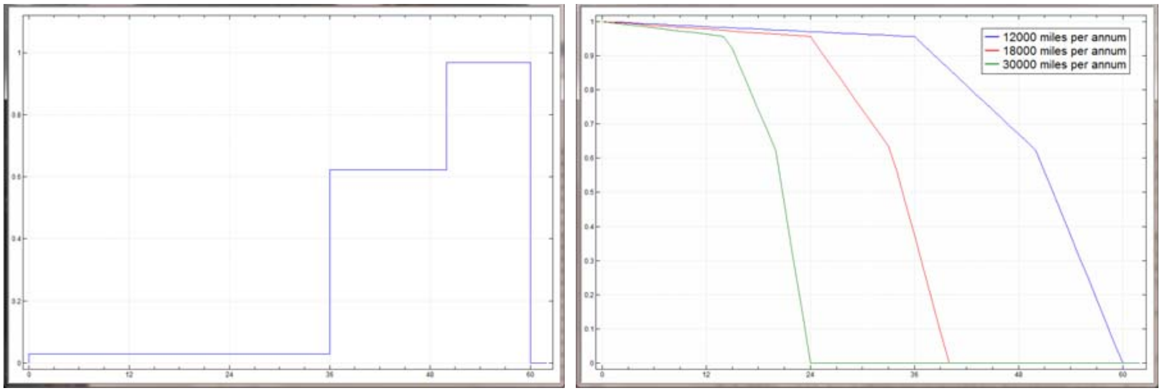



**Fig. 14a and 14b.** Earnings pattern with extras and limitations, and corresponding UPR curves.

Note that the first earnings step is greater than zero, and the initial slopes of the UPR curves are less than zero.

*Pre-expiration spike.* As a warranty contract approaches its expiration, policyholders tend to present claims for failures that may have accumulated over some months. If we allow that this increases by 10% the claims presented in the final month of a contract, and (for simplicity) make no adjustment to the claims presented in earlier months, then the resulting earnings and UPR patterns are as shown in Figures 15a and 15b.

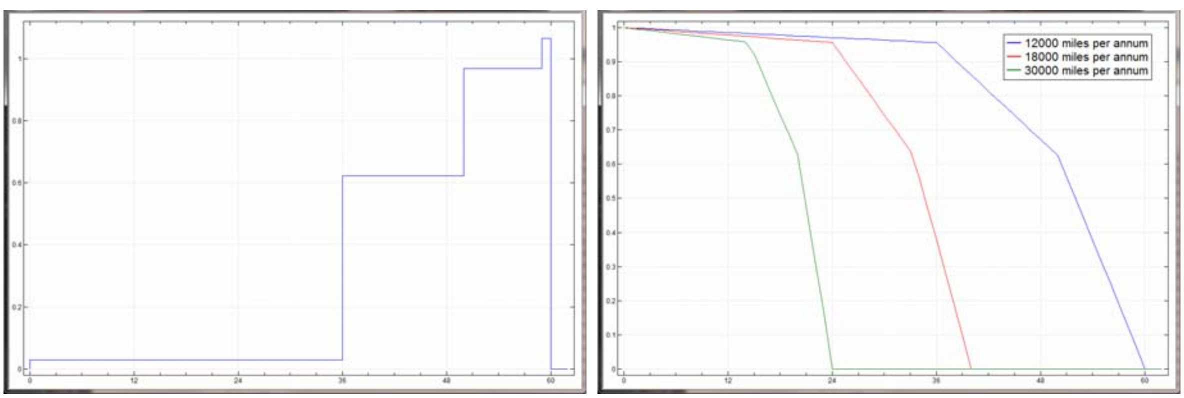



**Fig. 15a and 15b.** Earnings pattern and UPR factors with 10% pre-expiration spike.

*Trend and breakage.* Monthly trend and breakage may be represented by monthly growth or decay factors raised to the power of the number of elapsed months and netted against each other. This accounts for the effect of trend on the shape of the earnings and UPR curves. Odometer trend may be represented by a growth factor per (say) 10,000 miles, raised to an appropriate power based on the issue odometer and the number of miles driven per annum. Assuming a trend factor (more precisely a trend-net-of-breakage factor) of 1.05 per annum gives the results shown in Figures 16a and 16b (the vertical scale of Figure 16a is smaller than the previous earnings pattern tables, to accommodate the effect of five years' trend). The effect on the UPR curves is to make them more convex, but only slightly so; trend (when applied in the issue-month axis as well as the development axis) has a much greater effect on loss ratios.

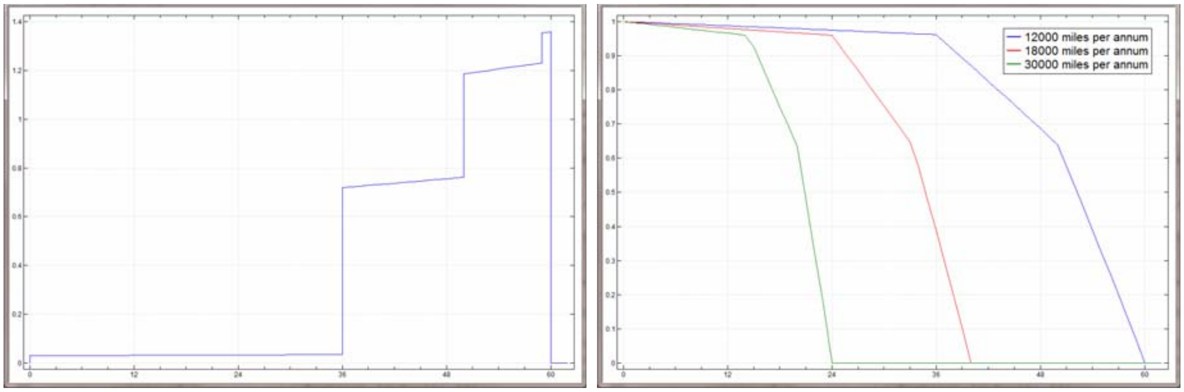



**Fig. 16a and 16b.** Earnings pattern with 5% annual trend.

*Combining miles per annum values.* Now suppose that there is a 60% chance that a car will be driven 12,000 miles per annum, a 35% chance of 18,000 and a 5% chance that it will be driven 30,000 miles. Weighting the curves from Figure 16b and averaging them gives us Figure 17.

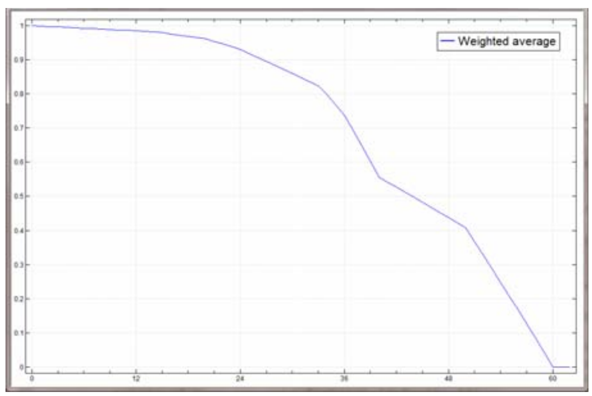



**Fig. 17.** Weighted-average combination of UPR curves for three miles-per-annum values

*Miles-per-annum distribution* . In practice we approximate the distribution of miles driven per annum as multinomial with some number N (say 50) of miles values each with an associated probability. Most commonly we select the miles values as midpoints of equiprobable intervals determined by analyzing a large collection of contracts presenting claims, using issue odometer, issue date, odometer at time of first claim, and breakdown date of first claim. This distribution may be adjusted by judgment if it is felt to be biased because of having excluded contracts with no claims. In the case N=50 the probabilities are each 0.02. The result using a representative 50-point miles-per-annum distribution is shown in Figure 18.

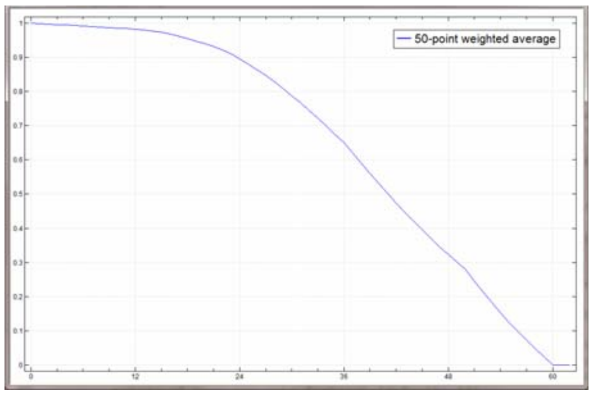



**Fig. 18.** UPR curve using 50-point miles-per-annum distribution.

The slight irregularity of the UPR curve in Figure 18 beyond lag 48 months reflects the fact that all cars driven less than 12,000 miles per annum will expire at 60 months; therefore there is a large probability mass weighting the 60-month curve, with its inflection point at 50 months when the manufacturer's power-train warranty expires. This is not entirely smoothed out by the contracts with other miles-per-annum values, since many of these mile out earlier than 50 months.

*One-half-month adjustment*. We normally assume that UPR curves are to be applied to all contracts written in a single month, and that writings are approximately uniform through the month. Figure 19 shows the curve from Figure 18 averaged with the same curve offset by one elapsed month, the so-called one-half-month adjustment.

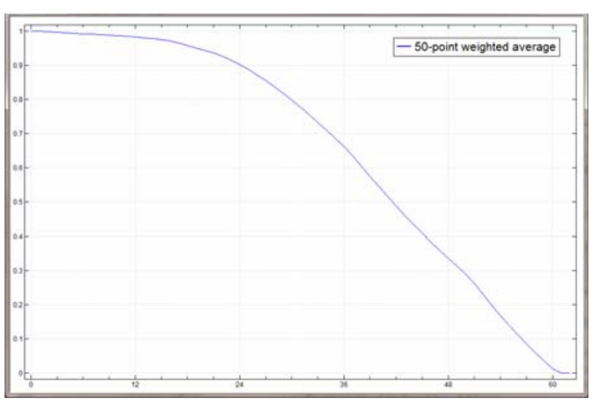



**Fig. 19.** UPR curve with one-half-month adjustment.

Note that this curve is slightly smoother than the preceding curve and also reaches zero at the end of the 61st month, rather than the 60th, but with slope from 60 to 61 about half that of the slope from 59 to 60. This is the characteristic pattern for one-half-month adjustments.

### 2.15.3 Residual emergence pattern

Up to now we have talked about how failures, and resulting claims, "ought" to emerge, based on known factors that seem important, such as term, MW, trends, and breakage. If claims really emerge in this way and if we divide actual losses by adjusted exposures lag by lag, for a large collection of contracts with the same term, the result should be constant incremental monthly loss ratios, or straight-line ("pro rata") residual ATF's.

But what if claims do not emerge exactly as expected? For example, what if we assumed no breakage but breakage really is important? Then our theoretical earnings pattern would not decline as much as it should with increasing lag, and our theoretical UPR curve would not be as concave as it should be. If in our body of data we "earn" the premiums according to our theoretical pattern, and then compare losses with this theoretical earned premium, the result will be a residual earnings pattern (proportional to the partial loss ratios to theoretical earned premium) that declines with increasing lag, or residual ATF's that are concave. Now suppose that, when we apply the exposure model to get a theoretical earnings pattern for some proposed contract, we multiply it by the residual earnings pattern for the term of the contract. Then we will be approximately where we would have been had we built the missing factors into our model to start with. Only approximately so, because the residual UPR curve is based on averages, while we are applying it to a particular contract, but in some cases it may be exact, if the missing factors picked up in the residual curve affect all contracts in the same way.

In this way we can use residual ATF's to "true up" our theoretical UPR curves by the ATF model – not just for factors that we might have inadvertently omitted or mis-estimated, but for factors not known to the model, or handled by the model in a simplified way. For example, if breakage were not described by simple exponential decay, but instead tended to increase sharply after, say, year 3, the residual ATF's will pick up this type of variation automatically and will transmit it back to UPR curves calculated for proposed contracts.

Using residual ATFs should improve the fit of the UPR curves for proposed contracts provided (a) the proposed contracts are similar to the contracts in the experience data and (b) the data is of sufficient volume that the residual ATFs represent signal and not noise.

### 2.15.4 Estimation

Our model provides for estimating residual ATFs for *any* term using *all* terms in the experience data. It does this by loss development using both loss and exposure triangles, assembled separately for each target term *t* from monthly lag triangles of all available terms, weighted by proximity to the target term, but (a) excluding losses that may have emerged after the "official" expiration of a contract (for which the model provides no natural exposure) and (b) with the early rows of the triangle for each contributing term *t** consolidated, if necessary, so that the number of rows equals the number of columns. The latter adjustment allows the triangle to be combined, in the weighted average, with triangles for terms greater than *t** without distorting development patterns. As usual, the exposure triangle reflects the possible decline along rows due to cancellations.

Because the emergence-of-failures patterns are independent of the MW and of our contracts, and the adjustments of the ATF model correct them to the expected emergence-of-claims patterns appropriate for our contracts using a constant of proportionality independent of term, both losses and adjusted exposures may be combined across terms in this way. In effect we first combine the data for all available terms, and then derive residual ATFs for *all* terms, including terms not found in the data. Combining the data and then deriving the curves is more consistent and systematic than deriving curves for each available term and then attempting to reconcile and interpolate them. Our use of a moving weighted average across terms almost always produces a smooth progression of residual ATFs from term to term; it is also tolerant of situations where there is a reasonable volume of data in aggregate but the data is sparse for individual terms (see Figure 20).

The only caveat is that development must essentially cease at the true end of each term. For business not satisfying this condition (e.g., prepaid maintenance, where some dealers informally extend the term to maintain goodwill), judgment adjustments may be necessary either to the data or to the results of the ATF model.

Figure 20 illustrates residual ATFs for a single large program. The residuals are close to straight lines for the shorter terms, but bow outward thereafter, especially for terms 84 months and greater. This suggests that the longer-term contracts may be qualitatively different from the rest in ways that are not explained by the model, and that, in the weighted average of data from different terms, these contracts are pulling those of, say, 60 months, outward. Also the data at the longer terms and later lags is sparse, as shown by the irregularity of the UPR curves.

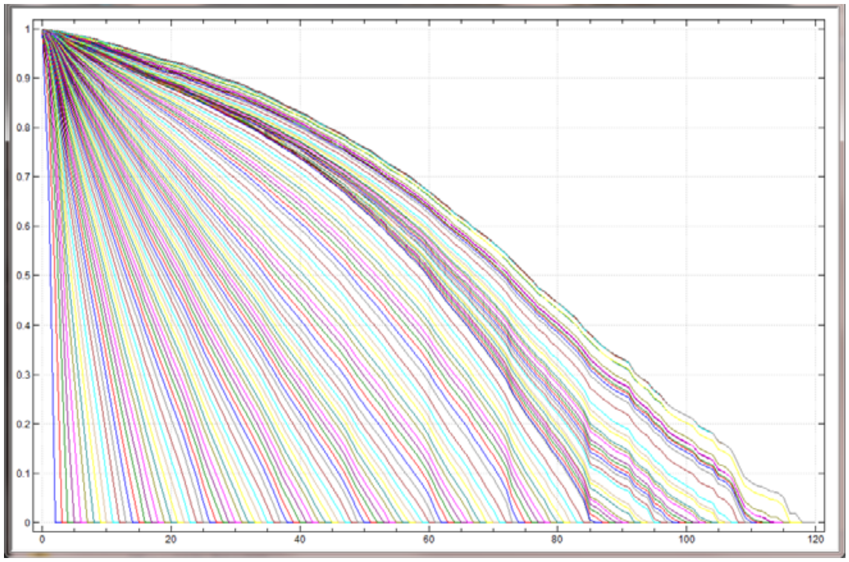


**Fig. 20.** A set of residual UPR curves based on all data.

Figure 21 shows the residual UPR curves for the same case but with the analysis subdivided so that the curves for terms 1-72 months depend only on data within that range, and likewise for terms 73-120 months. There are now two families of curves with space between.

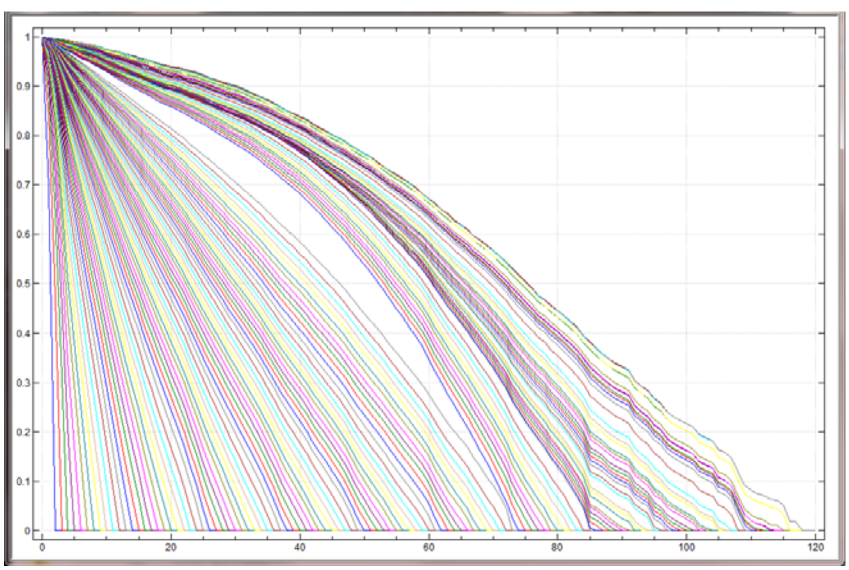


**Fig. 21.** Residual UPR curves based on data divided at term 72 months.

Figure 22 shows the same curves as in Figure 21, except limited to terms 1 through 72 months and rescaled for clarity. For the terms from about 48 months through 72 months there is still a modest amount of convexity. The ATF model as applied leaves a small amount of the variability of these UPR curves (from pro rata) unexplained.

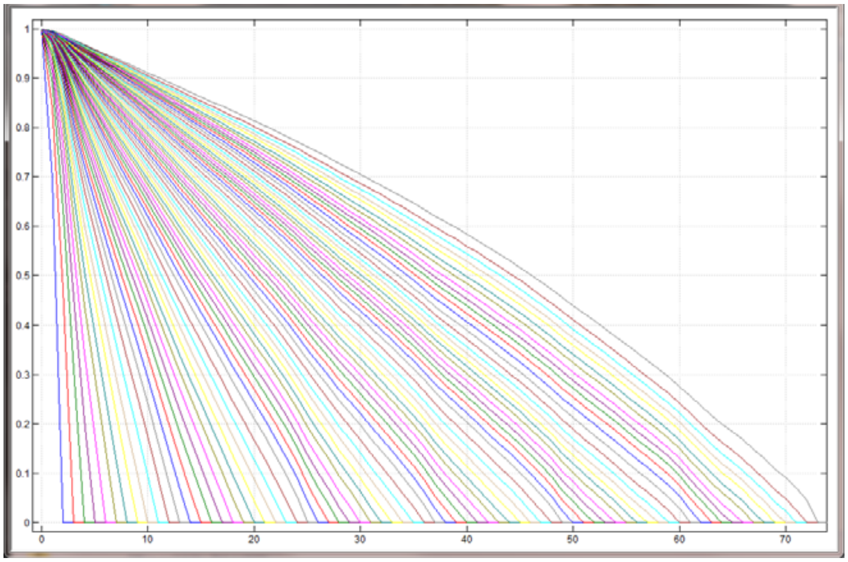



**Fig. 22.** Detailed look at terms 1-72 from Fig. 21.

Figure 23 shows the residual curves for terms 1 through 72 months after modifying the model to include an annual trend factor of 1.05, thereby removing that amount of trend from the residuals. This eliminates most of the apparent convexity.

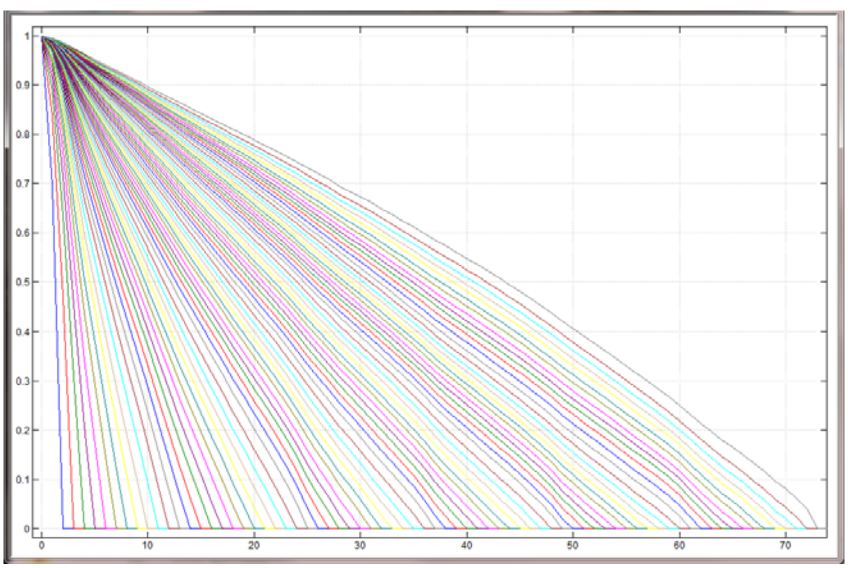



**Fig. 23.** Residuals after including annual trend of 1.05 in model.

Figure 24 compares the curves for terms 36, 48, 60, and 72 months, with and without the trend, against pro rata lines for the same terms. Clearly the model with the 5% trend parameter brings the residual curves closer to straight lines. However, in estimating UPR curves for proposed contracts, if trend is omitted from the model, it will emerge as part of the residuals, and thus find its way indirectly into the estimated curves, often making them close to what they would have been with trend included.

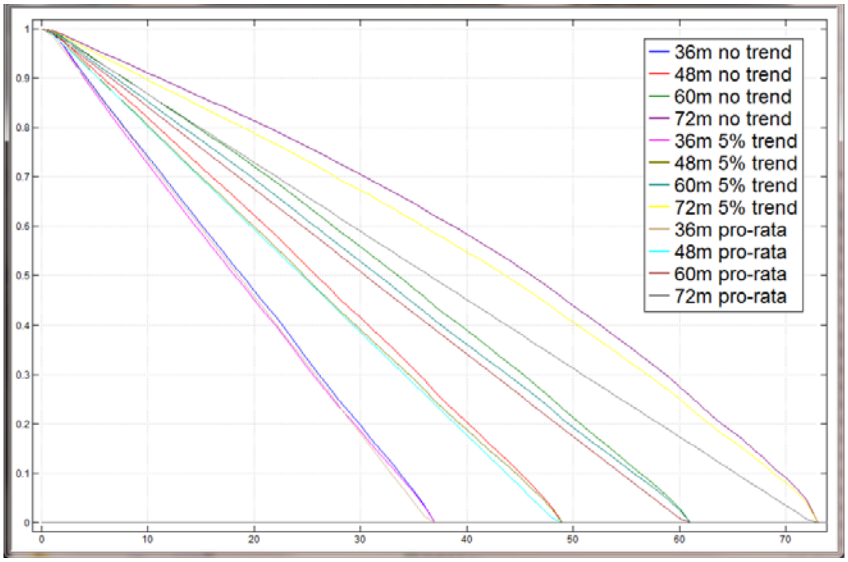



**Fig. 24.** Comparison of residuals after fitting models with no trend and 5% trend with pro rata.

In summary, when deriving UPR curves for proposed contracts, the ATF parameters involving our contract, the manufacturer's warranty, the power-train, non-power-train, and extras fractions, and the pre-expiration spike, should be chosen carefully, for they affect different contracts in different ways. If there exists a large enough body of relevant experience, it may be used to derive residual ATFs that capture the effect of all other factors, including trend and breakage, at least approximately. The UPR curves for the proposed contracts may be derived by applying the ATF model starting with the residuals. Otherwise, trend and breakage parameters should also be chosen carefully, and the ATF model applied directly to the proposed contracts without using residuals.

### 2.15.5 Relativities

Our ATF model generates its earnings factors, before normalizing them to total 1, as amounts which are proportional to expected losses across plans as well as across lags within a single plan. Therefore, if we sum these factors across all lags, we obtain sums proportional to the total expected losses for each plan. This facility works identically whether starting from a table of residual ATFs or not.

For example, a plan with term 48048 with underlying MW of 24024 power train and non power train is worth 45.2% more than a plan with term 60060 with MW of 50050 power train and 36036 non power train, using no residuals, or 42.6% more using the residuals from the model fitted to a particular large account with 5% trend.

We make use of relativities of this type when setting up manual rate tables for new programs. Typically such tables involve several different terms, and several different coverage levels, applied to vehicles with several different manufacturers' warranties. For used vehicles the rate table also includes different odometer "bands", for example, 0-999 miles, 1000-9999 miles, 10000-19999 miles, and so forth. Our rating procedure then applies the ATF procedure to proposed contracts with the given terms, MW's, coverage levels, and odometer-band midpoints or centers of gravity. For each such contract the model gives its relativity to a particular base contract. If we have separately estimated the expected loss cost for the base contract, these relativities lead to the expected loss cost for each of the proposed contracts. In this way the ATF model, primarily designed to derive families of UPR curves, may also derive families of expected loss costs for manual ratemaking.

Issues of homogeneity may arise when estimating ATFs. If the contracts for some terms differ from those for other terms in ways not being accounted for in the model, then the residual ATFs for terms close to the boundary may be distorted. We may address this issue by controlling the weighting of data for terms successively more remote from the term being estimated. Or, if we know a priori that the shape of the UPR curves should differ for certain groups of terms (for example because they do or do not have manufacturers' warranties), then we may request that our model calculate the family of curves within each of several term groupings using local data only.

It turns out that the final strings produced by **AllTermsFactors** are not sensitive to assumptions that affect all contracts in the same way, such as breakage and time-based trend. For example, omitting trend entirely, when it exists in the data, will simply transfer it from the model to the residual curves, from which it will be picked up again by the fitted curve for each particular contract.

On the next page is a schematic diagram of the All-Terms Factors process. The names inside blue ellipses are some functions in our system: **AllTermsFactors**, which uses experience to generate ATF curves for all contracts in the data, and a table of residual ATFs, **ATFFormula**, which applies ATFs to obtain UPR for each contract in the data, and **GetATF**, which uses residual ATFs to obtain UPR factors for arbitrary contracts. We do not need to be concerned here with how these functions operate, but need simply to recognize the roles they play in the model.

Note the distinction between contract attributes, global parameters, and estimator controls. The contract attributes are stored in our data. Some of the global parameters may be estimated from the data, or may be selected by judgment and checked for reasonableness by inspecting the residual string table. The estimator controls are largely a matter of judgment.

### 2.15.6 All-Terms Factors for non-auto contracts

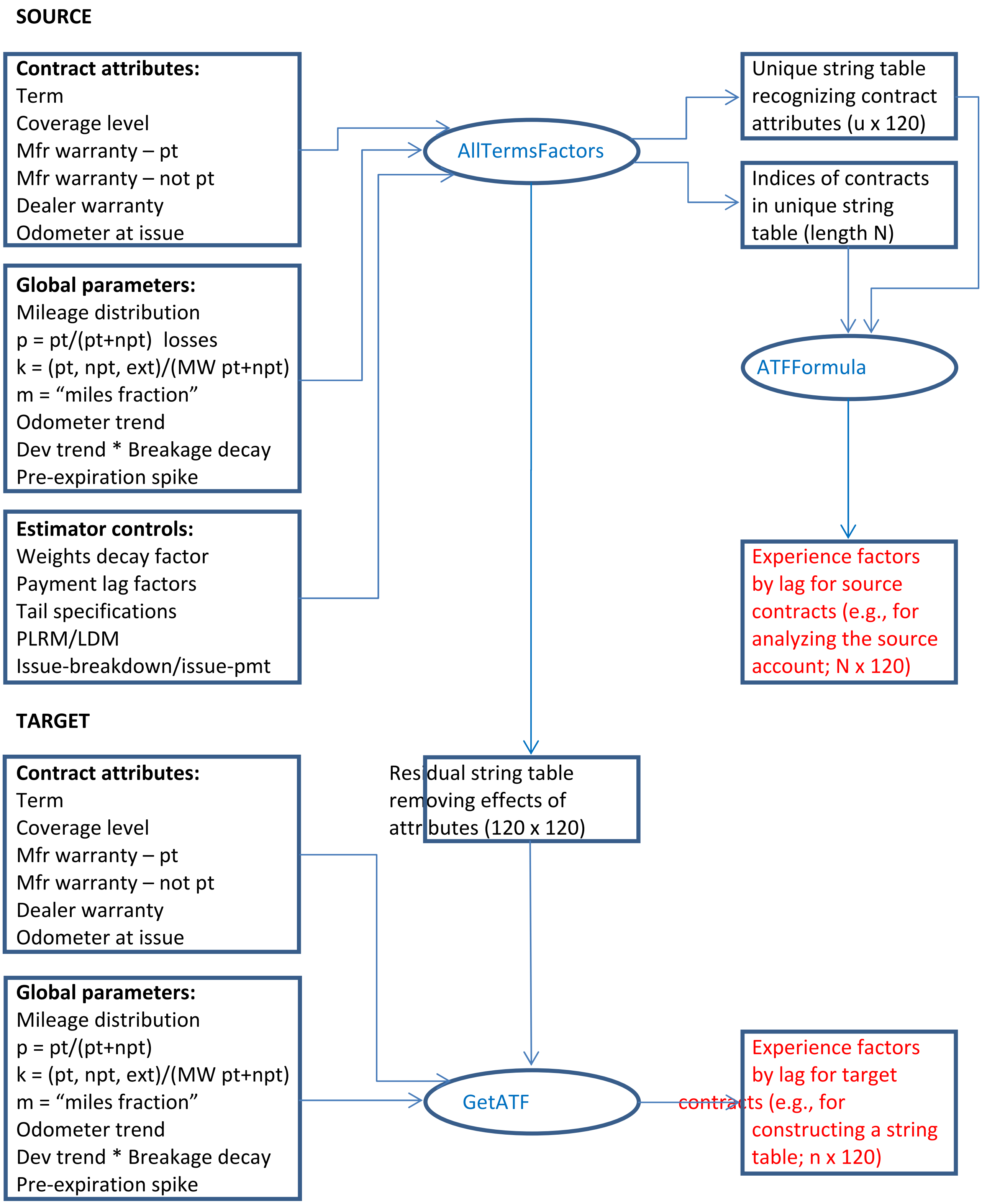
Many of the features described above, such as terms in months and miles, power-train versus non-power-train MW's, and extras, are designed for contracts on automobiles. However, the All-Terms Factors model also may be applied to non-auto business such as electronics or power sports.

One approach is to use the simplified ATF model, creating a table of ATFs by term only. This should be satisfactory when analyzing the UPR term by term for the same data set. If we wish to analyze individual contracts or groupings of contracts other than by term, but recognize differences among contracts in MW's, it may be preferable to use the regular model, interpreted slightly differently from auto, as follows:

- Distinguish parts versus labor instead of power train versus non-power train
- Code all terms and MW's in the data to have no miles limit
- Assume a one-point dummy miles distribution with a low miles value
- Use the "extras" input for Accidental Damage Handling coverage if applicable.

The last point requires explanation. Most non-auto Warranty contracts do not involve extra services, so the coverage factors for extras are normally zero. However, some contracts provide Accidental Damage Handling (ADH) coverage in addition to warranty coverage in case of defects. ADH is different from typical extras in that (a) its expected value may be much higher, often of the same magnitude as the regular repair or replacement coverage, and (b) it may be subject to a months limit shorter than the regular term of the contract. Our ATF model as implemented in software provides for (a) via the "extras" parameter and provides separately for (b), contract by contract.

**All-Terms Factors Process**



### 2.15.7 Miles-per-annum distributions

***Mileage distributions.*** The All-Terms Factors model requires that we take into account the number of miles driven per month by the holder of a contract; this is not known at the issue date but must be represented with a probability distribution. Such a distribution is naturally continuous but we have found it useful to approximate it with an n-point discrete distribution.

It is tempting to measure the annual mileage distribution using issue odometer and approximate age of vehicle, based on model year, both usually available in the data for contracts on used cars. Unfortunately this gives the average miles driven by the previous owners rather than by the owners purchasing the contracts in the data. Instead we consider contracts presenting claims, and use the odometer at issue, the odometer at breakdown, the issue date, and the breakdown date of the first claim for each such contract. Using only the first claim avoids counting the down time for service and thereby reducing the apparent mileage. However, using contracts with claims probably biases the result upward, as discussed below.

Figure 25 shows the distribution of miles per annum, to the nearest thousand, for contracts in a large representative program.

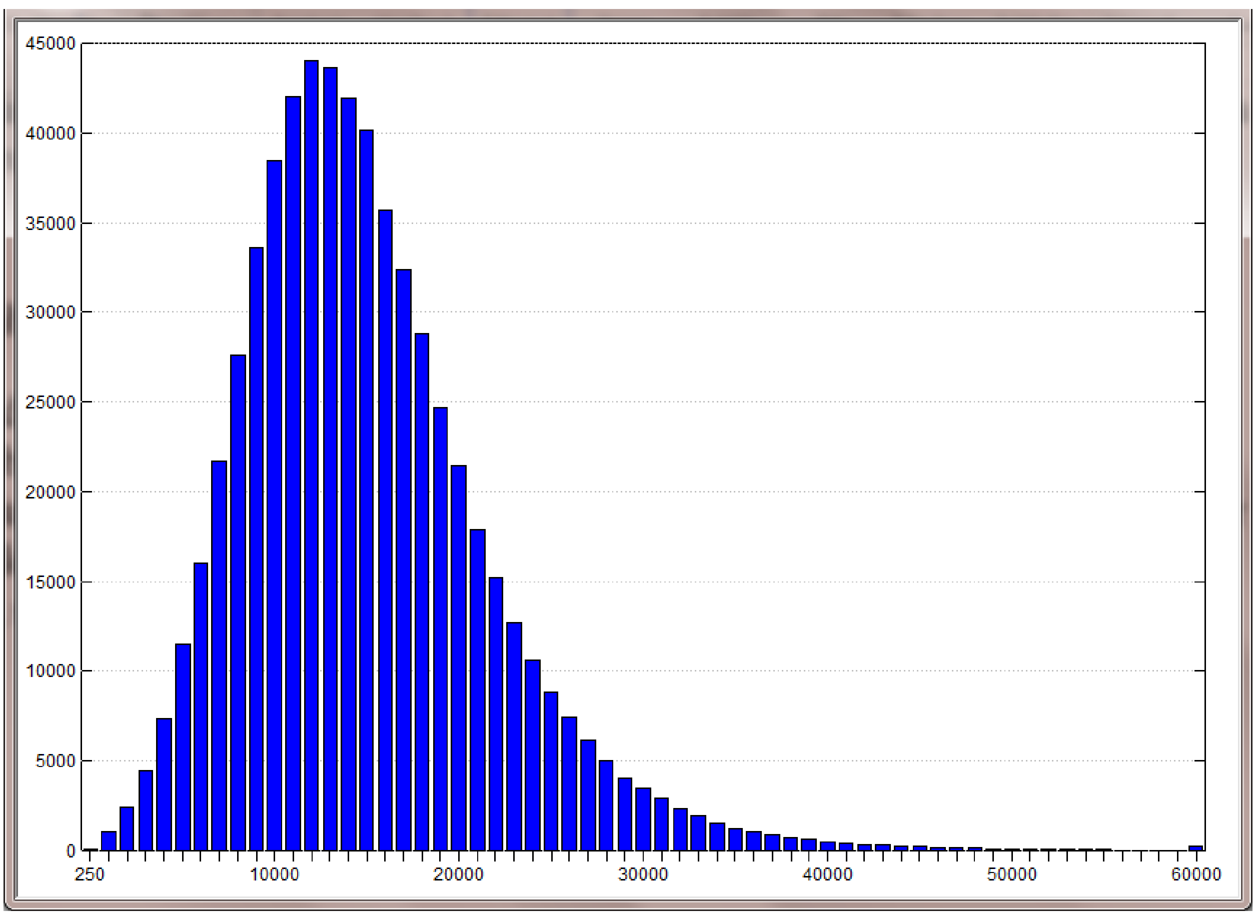


**Fig. 25.** Distribution of miles per annum, in thousands, based on first claims

The values represented by these bars may be normalized to total 1.00 by dividing by the total number of contracts observed, in this case 535,947, and these probabilities used with the midpoints of the x values in each group to define a discrete distribution. Here the first bar runs from 0 to 500 miles and the last bar covers 59,500 or more miles (so its midpoint should be greater than 60,000); the remaining bars are each of width 1,000, with boundaries 500, 1500, etc.

The function generating the above plot provided the following statistics (from a fragment of a J-language session screen):

```
   m=:20040701 20140630 20140630 1000 0 61 1 MilesDistribBands
''
  Mean (weighted by days) = 14930
  Mean (weighted by cars) = 16442
  Standard deviation (from variance weighted by days) = 6987
  Standard deviation (from variance weighted by cars) = 8173
  Number of observations = 535947
  Maximum = 179488
  Minimum = 365
  Probability weight for each band determined by total years
  Representative value in each band equals center of gravity
  Approximated mean: 14930.44683
  Approximated std dev: 6802.084723
```

Often, instead of using groups of equal width with different probabilities, we use N groups of variable width with equal probabilities. The following lines illustrate this with N=50:

```
   m50=.20040701 20140630 20140630 50 1 MilesDistribNpt ''
  Mean (weighted by days) = 14930
  Mean (weighted by cars) = 16442
  Standard deviation (from variance weighted by days) = 6987
  Standard deviation (from variance weighted by cars) = 8173
  Number of observations = 535947
  Maximum = 179488
  Minimum = 365
  Boundaries of bands determined by accumulated years
  Representative value in each band equals center of gravity
  Approximated mean: 14930.4052
  Approximated std dev: 6712.981379
```

The mean approximated by the discrete distributions is very close to the mean of the entire distribution while the standard deviation is a bit smaller. This is mainly the result of compressing all values in the right tail of the actual distribution – which range as high as 179,488 – into a single value of 38,463. The whole 50-point equally-probable distribution is:

```
   5 10$round >0{m50
   2984 4828 5804 6550 7151 7675 8138 8562 8954 9318
   9663 9992 10313 10628 10930 11224 11519 11808 12089 12373
  12658 12946 13234 13522 13815 14115 14418 14716 15025 15347
  15684 16032 16387 16755 17139 17543 17960 18412 18897 19416
  19973 20588 21279 22064 22971 24051 25402 27200 30004 38463
```

As expected, these values are much closer to each other toward the overall mean than in the extremes, especially in the long right tail.

The use of miles-to-first-claim as a starting point for estimating the miles-per-annum distribution may be the most practical available method but is not ideal. In particular, it ignores the miles driven by any cars that have not presented claims. Since claims emerge largely as the result of usage, cars presenting claims may be expected to be driven more, on average, than all cars, and produce a distribution that is biased upward. One way to offset this tendency is to determine the $N$ mileage intervals and their means using time as a weight, as shown above.

Miles per annum might well be represented by a continuous distribution and we use a discrete approximation mainly for simplicity. If we plot the distribution from Figure 25 as a curve and superimpose a lognormal distribution fitted by the method of moments, we obtain the comparison shown in Figure 32.

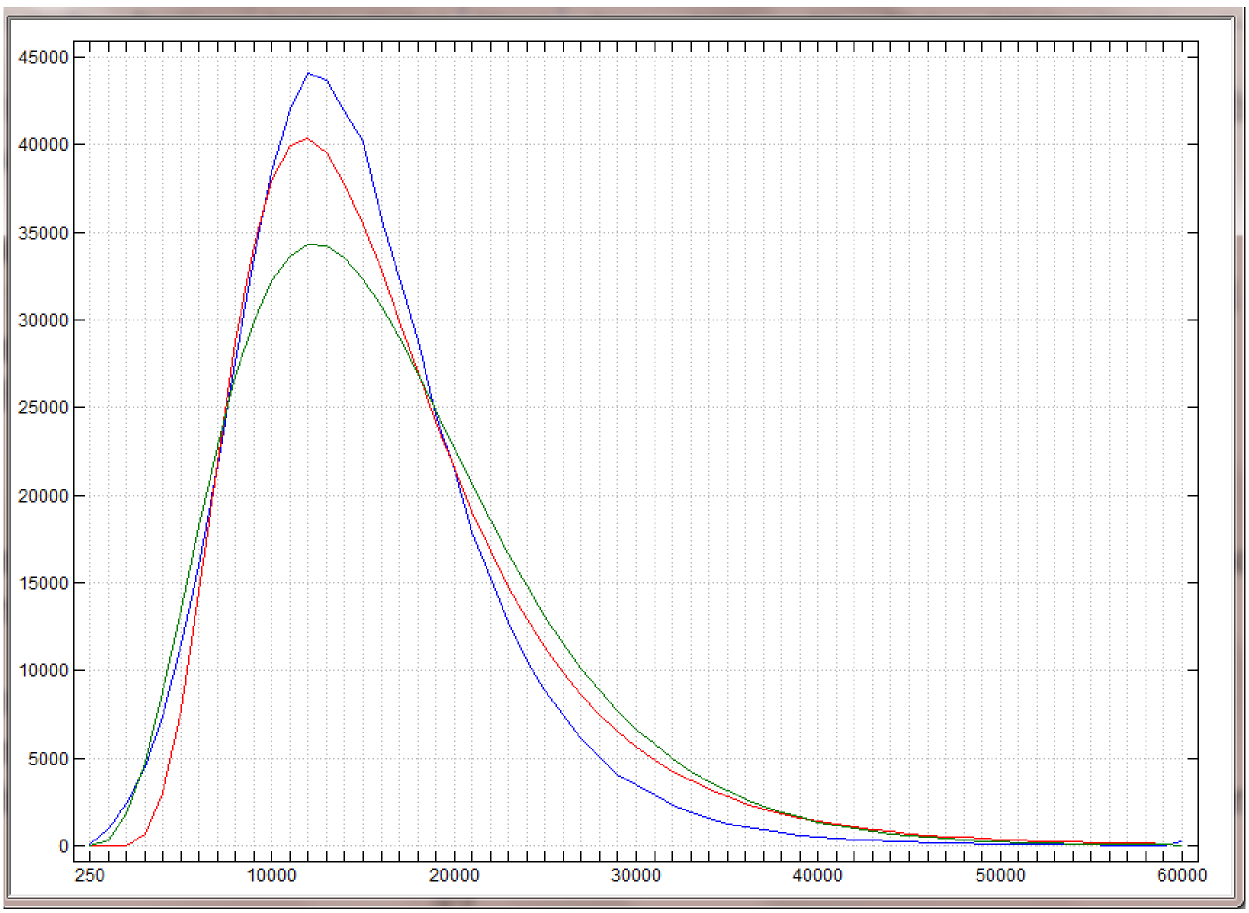



**Figure 32.** Observed miles-per-annum distribution (blue) compared with lognormal distribution (red) and Gamma distribution (green).

The lognormal does not fit this distribution well, especially in the tails; the Gamma is rather better in the left tail (in which we are not particularly interested) but worse in the central area and the right tail; moreover, there is some area under both of these fitted distributions arbitrarily far to the right, whereas miles driven have a practical upper bound based on speed limits and the finite number of hours in the day. For our purposes we find the discrete distributions more satisfactory.

## 3. CONCLUSIONS

In conclusion, the UPR is critically important for Warranty Insurance but there are reliable techniques for calculating it. These techniques are similar to those used for loss reserves in other lines of insurance, but require adaptations to address the special characteristics of Warranty contracts. Among these characteristics are exposure that declines over the life of a cohort of policies, loss development triangles that must be adjusted for incompletely reported losses in recent diagonals, tails that are of limited duration but contain significant probability mass, and the need to calculate internally consistent sets of UPR curves for entire families of contracts. When properly conceived and programmed, procedures for assigning and testing UPR factors may be efficiently applied to large numbers of Warranty programs each split into homogeneous subdivisions.

## 4. REFERENCES

[1] Bornhuetter, Ronald and Ronald Ferguson, The Actuary and IBNR, PCAS LIX, 1972: 181-195

[2] Bühlmann, Hans, Vereinigung Schweizerischer Versicherungsmathematiker / Association des Actuaires Suisses, Ecole d'éte 1983, Estimation of IBNR Reserves by the Methods Chain Ladder, Cape Cod, and Complementary Loss Ratio, unpublished

[3] Gluck, Spencer M., Balancing Development and Trend in Loss Reserve Analysis, PCAS LXXXIV, 1997: 482-532

[4] Kerper, John and Lee Bowron, An Exposure-Based Approach to Automobile Warranty Ratemaking and Reserving, CAS Forum 2007: 29-43

[5] Stanard, James N., A Simulation Test of Prediction Errors for Loss Reserve Estimation Techniques, PCAS LXXII: 124-153