Casualty Actuarial Society E-Forum, Fall 2012 Volume 1

The CAS *E-Forum,* Fall 2012-Volume 1

The Fall 2012, 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* includes four independently submitted papers.

E-Forum Committee

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

CAS *E-Forum,* Fall 2012-Volume 1

Table of Contents

Monitoring Industry Premium, Loss Ratios, and Loss Reserves

John Captain, FCAS, MAAA

Abstract: This paper presents a method to use US insurance industry information and economic data to monitor the relative adequacy of the earned premium volume and the calendar year loss ratios that are being booked. The economic data is updated monthly and the industry data being used is updated and available each quarter which allows for timely monitoring of the likely movement in the industry's loss reserve adequacy.

Keywords: reserving, APLR's, a priori loss ratios, expected loss ratios, calendar year, accident year.

1. INTRODUCTION AND BACKGROUND

This paper will suggest a technique for deriving estimates of what the industry loss ratios should be. These loss ratios are then compared to the actual loss ratios that the industry is booking in order to estimate the impact on the industry's loss reserve adequacy. This is done separately for total commercial lines and total personal lines as well as all lines in total.

Appendices A, B, and C display the premium and loss information for accident years 1995-2010 for the major commercial lines of Workers Compensation, Commercial Auto Liability, and Other Liability.[1] The original ultimate loss ratios booked as of 12 months as well as the most recently available booked ultimate loss ratios are also displayed. Examining those exhibits leads to the following 3 general observations for each of the major commercial lines:

1. The volume of ultimate losses increased during the 1996-1999 soft market years as we would expect with exposure and loss trend, but the volume of earned premium did not grow proportionally.

2. The initial ultimate loss ratios booked as of 12 months (accident year-end) did not show nearly as much variation across the years as do the most recently booked ultimate loss ratios. The initial booked loss ratios compared to what they were eventually booked as are too high in the hardest of the hard market years, and are much too low in the softest of the soft market years.

3. The volume of earned premium has been decreasing since the 2006 accident year rather than increasing as we might expect to keep up with exposure and loss trend.

Not surprisingly, the combination of the first 2 observations suggest that the aggregate premiums

[1] Data source is AM Best Aggregates and Averages

being charged during the soft market years became relatively and progressively less adequate as the market softened. However, the associated loss ratios that were booked did not increase enough to fully reflect the increasing degree of premium inadequacy. Although the core problem underlying the loss reserve inadequacy was the inadequacy of the premium, the loss reserve inadequacy build-up was due to booking loss ratios that were too low to fully reflect the inadequacy of the premium. The combination of the last 2 observations cited above raises the concern that a similar situation could be happening again in the later years of the data set.

2. METHODOLOGY

2.1 General Approach for Monitoring the Industry

The approach set forth in this paper uses US insurance industry information and economic data to monitor the relative adequacy of the earned premium volume and the calendar year loss ratios that are being booked. The economic data is updated monthly and the industry data being used is updated and available each quarter which allows for timely monitoring of the likely movement in the industry's loss reserve adequacy.

In simplistic terms, the approach is to choose a base year (2005) for which we think we know the actual accident year loss ratio with reasonable certainty and is a year we think is in a more neutral part of the cycle with the market being neither very hard nor very soft. In order to maintain the same loss ratio level in subsequent years, the change in earned premium would need to be sufficient to keep pace with the corresponding exposure trend and loss trend. To the extent the earned premium does not keep pace, there needs to be a change in the magnitude of the loss ratio being booked. If the loss ratio does not change as expected and an inaccurate calendar period loss ratio gets booked, then there is an impact on the relative loss reserve adequacy. A loss ratio inaccuracy in one direction for one calendar period can be offset by an inaccuracy in the other direction in another calendar year. Repeated inaccuracies in the same direction will have a cumulative impact on the industry reserve adequacy.

Note that this method is determining the accuracy of the normalized (i.e. adjusted to reflect normal rather than actual levels of catastrophe losses) calendar period loss ratio that was actually booked by comparing it to what the corresponding accident year ultimate loss ratio should be. The accuracy of the loss ratio being booked on the current accident year as well as changes in the accuracy of the loss ratios being booked for prior accident years would all impact the accuracy of the calendar period loss ratio. Therefore, if this approach indicates an inaccurate calendar period loss ratio was booked, we are not really identifying which accident years were booked inaccurately just that the aggregate impact produces an inaccurate calendar period loss ratio.

2.2 Details of the Approach

Each quarter, ISO releases a Chief Executive Circular on the Property/Casualty Insurance Industry Financial Results[2] that reports on the calendar period results for the US insurance industry. The ISO report provides the written premium, earned premium and incurred loss & LAE for the property/casualty industry for the latest 6 full calendar years and the year-to-date through the latest quarter partial years for the current and prior calendar years. The ISO report also shows the partial calendar year data separately for writers that predominate in Commercial Lines, writers that predominate in Personal Lines, and Balanced Writers. Splitting the Balanced Writers volume evenly between commercial and personal lines allows us to create an approximate industry commercial versus personal lines compilation.

The approach in this paper monitors how we believe the industry loss reserve adequacy has moved since the end of 2005. If the adequacy position of the industry were known for year-end 2005, then an estimate of the current adequacy position could be derived, but it is not necessary to know the precise adequacy level of the industry reserves at year-end 2005 to use this approach to say how the adequacy level has likely moved since then. The US industry loss reserves for both commercial lines and personal lines were probably in a strong position at year-end 2005.

We have deemed the calendar year 2005 industry personal lines reported loss ratio to be equal to the 2005 accident year ultimate loss ratio. That is, the 2005 calendar year loss ratio was not materially distorted by reserve changes on accident years 2004 and prior for personal lines. We are therefore deeming the 2005 calendar year loss ratio to be equal to the 2005 accident year ultimate loss ratio for personal lines.

The Insurance Information Institute reported that calendar year 2005 had $18.9B of prior year reserve strengthening.[3] Assuming the strengthening was all in commercial lines implies that the commercial lines calendar year 2005 loss ratio is 9.5 points above the corresponding 2005 ultimate accident year loss ratio. Therefore, we are deeming the 2005 ultimate accident year loss ratio for

[2] For example, ISO Chief Executive Circular CE-AA-2012-008 *Property/Casualty Insurance Industry Financial Results: First-Quarter 2012 Analysis*
[3] Insurance Information Institute November 7, 2007 presentation "P/C Insurance in an Era of Mega-Catastrophes" slide 19.

commercial lines to be equal to the 2005 calendar year loss ratio reduced by 9.5 points.

The loss ratios can vary by several points from year to year due to the fortuitous magnitude of catastrophe losses, such as from hurricanes. We want to normalize the industry loss ratios in order to state them as what they would be with a "normal" amount of catastrophe losses. The ISO Chief Executive Circular on the Property/Casualty Insurance Industry Financial Results[4] provides the magnitude of the catastrophe losses on a YTD basis each quarter. We remove the actual impact of catastrophes on the reported loss ratios and replace them with the "normal" impact on the YTD loss ratio each quarter. The result is the normalized loss ratio. The "normal" catastrophe loss ratio increases during the second half of the year when hurricanes, a significant contributor to catastrophe losses in the US, occur. For both personal and commercial lines, the "normal" loss ratio impacts are assumed to be 1.5 points through the first and second quarters, 1.9 points YTD through the third quarter, and 2.2 points for the full year. We made an exception in quarters 2-4 of 2011 because the unprecedented 2Q tornado losses impacted personal lines more than they did commercial lines- -for the full 2011 calendar year, the impact of cats increased the all lines loss ratio by 5.3 points more than normal, we assumed that the impact on the personal lines loss ratio was 6.3 points more than normal and the impact on the commercial lines loss ratio was 4.1 points more than normal.

To maintain a consistent degree of premium adequacy, we expect industry earned premium to change with exposure and loss trend. We assume that all business gets renewed somewhere within the US industry, it may not be with the same insurer but it gets renewed by some insurer within the industry. An embedded assumption is that the size of the insured US industry is not materially impacted by changes in use of captives, SIR's, or self-insurance, changes in coverage, or changes in terms and conditions.

Each month, Swiss Re's Economic Research & Consulting team publishes their US Economic Outlook[5]. This provides information on not only what recent changes in the real GDP (real gross domestic product) and CPI (consumer price index) have been, but also forecasts future periods. This gives us a way of using non-insurance industry data to project expected changes in the insurance industry, particularly exposure trend and loss trend.

We have assumed that industry exposure trend for overall Personal Lines is equal to zero. We have assumed that industry exposure trend for overall Commercial Lines is equal to the trend in real GDP.

---

[4] Ibid
[5] For example, Swiss Re Economic Research & Consulting, 6 July 2012 *US Economic Outlook*

The loss trend reflects both frequency and severity trend. This is the trend in the loss per unit of exposure that is sometimes called the pure premium trend. We have assumed that industry loss trends are equal to trends in the CPI plus a judgmental increment to reflect that insurance industry loss trends have traditionally been larger than the CPI trend. The judgmental adjustment also reflects the frequency trend which is not reflected in the CPI. For Personal Lines, the judgmental increment is 0.5 percentage points, except for 2008 where we reflected the favorable impact of the severe recession on personal lines loss trend by using a judgmental decrement of 1 percentage point. For Commercial Lines, the judgmental increment is 2 percentage points. One reason for the judgmental increment being smaller for personal lines is that decreasing frequency trends are more significant for personal lines in aggregate than for commercial lines in aggregate. The severe recession's impact on commercial lines is captured in the exposure trend (real GDP) rather than in the judgmental increment.

The previous assumptions and the information on exposure and loss trend combine to produce the figures in Chart 1 of the needed growth rates from the 2005 earned premium in order to maintain the 2005 normalized loss ratio level.

Total Commercial Lines **Total Personal Lines** **Chart 1**

Calendar Period	Exposure Trend	CPI % Change	Judgmental Increment to CPI for Ins. Trend	Needed[1] % Change in EP Since 2005		Exposure Trend	CPI % Change	Judgmental Increment to CPI for Ins. Trend	Needed[1] % Change in EP Since 2005
2005	3.2%	3.9%	2.0%	0.0%		0.0%	3.9%	0.5%	0.0%
2006	2.8%	3.2%	2.0%	8.1%		0.0%	3.2%	0.5%	3.7%
2007	2.0%	2.9%	2.0%	15.7%		0.0%	2.9%	0.5%	7.2%
2008	0.4%	3.8%	2.0%	22.9%		0.0%	3.8%	-1.0%	10.2%
2009	-2.6%	-0.3%	2.0%	21.8%		0.0%	-0.3%	0.5%	10.4%
2010	3.0%	1.6%	2.0%	29.9%		0.0%	1.6%	0.5%	12.8%
2011	1.7%	3.1%	2.0%	38.9%		0.0%	3.1%	0.5%	16.8%
1Q 2011	1.7%	3.1%	2.0%			0.0%	3.1%	0.5%	
1Q 2012	2.2%	2.0%	2.0%	41.0%		0.0%	2.0%	0.5%	17.6%

1 This is the % change needed to maintain the 2005 level of premium adequacy

Exhibit 1 shows how the actual change in earned premium levels since 2005 compares to the change needed to maintain the 2005 normalized loss ratios. The extent of the difference between actual and needed earned premium percentage changes for any given year tells us how much we could expect the loss ratio to be different from 2005. When the actual % change is less than the needed % change, then we should expect the normalized loss ratio to be higher than 2005's. We can compute the expected normalized loss ratio for each year by multiplying the 2005 loss ratio by (100 + % needed change in EP since 2005) / (100 + actual % change in EP since 2005).

To the extent the actual booked industry normalized loss ratios differ from the expected normalized loss ratios, we can infer that the loss ratios that the industry has booked are relatively less accurate than the 2005 loss ratios. The booking of inaccurate loss ratios has a direct impact on industry loss reserve strength. If the booked loss ratios are higher than the expected loss ratios then the industry reserves have become stronger. On the other hand, if the industry booked loss ratios are lower than expected then the industry loss reserves have become weaker and less adequate. The dollar impact on the industry's loss reserve adequacy from any calendar year can be computed by multiplying the earned premium for the year by the difference between the booked normalized loss ratio and the expected normalized loss ratio; positive impacts add to the industry's loss reserve strength and negative impacts weaken the industry's loss reserves. The impacts of individual calendar periods accumulate over time resulting in a cumulative impact on industry loss reserve adequacy. When the industry books calendar period loss ratios that are too low for an extended period of time, the industry's loss reserves can become quite inadequate. Likewise, when the industry books calendar period loss ratios that are too high for an extended period of time, the industry's loss reserves can become greater than necessary.

US P&C Insurance ($ Millions) **Exhibit 1**

Total Personal + Commercial Lines

Calendar Period	Earned Premium	Booked Normalized Loss Ratio	Needed % Change in EP Since 2005	Actual % Change in EP Since 2005	Expected Normalized Loss Ratio	Year's Impact on Reserve Adequacy	Cumulative Impact on Reserve Adequacy
2005	417,635	63.6%	0.0%	0.0%	63.6%	-	-
2006	435,484	64.1%	5.8%	4.3%	64.5%	(1,655)	(1,655)
2007	438,908	68.0%	11.3%	5.1%	67.3%	3,430	1,775
2008	438,316	74.3%	16.3%	5.0%	70.4%	17,796	19,572
2009	422,302	72.0%	15.8%	1.1%	72.8%	(2,646)	16,926
2010	422,200	72.4%	20.9%	1.1%	76.1%	(14,361)	2,565
2011	433,941	74.1%	27.3%	3.9%	77.9%	(15,006)	(12,441)
1Q 2011	105,232	69.9%					
1Q 2012	107,944	68.3%	28.7%	6.6%	76.8%	(8,804)	(21,245)

Total Commercial Lines

Calendar Period	Earned Premium	Booked Normalized Loss Ratio	Needed % Change in EP Since 2005	Actual % Change in EP Since 2005	Expected Normalized Loss Ratio	Year's Impact on Reserve Adequacy	Cumulative Impact on Reserve Adequacy
2005	198,757	60.1%	0.0%	0.0%	60.1%	-	-
2006	213,961	62.3%	8.1%	7.6%	60.3%	4,201	4,201
2007	218,956	66.0%	15.7%	10.2%	63.1%	6,434	10,635
2008	212,204	74.7%	22.9%	6.8%	69.1%	11,722	22,357
2009	200,905	70.2%	21.8%	1.1%	72.3%	(4,227)	18,129
2010	195,359	71.7%	29.9%	-1.7%	79.4%	(14,901)	3,229
2011	201,799	74.7%	38.9%	1.5%	82.1%	(14,957)	(11,729)
1Q 2011	49,190	72.9%					
1Q 2012	50,754	68.0%	41.0%	4.8%	80.8%	(6,501)	(18,230)

US P&C Insurance ($ Millions) **Exhibit 1 (cont.)**

Total Personal Lines

Calendar Period	Earned Premium	Booked Normalized Loss Ratio	Needed % Change in EP Since 2005	Actual % Change in EP Since 2005	Expected Normalized Loss Ratio	Year's Impact on Reserve Adequacy	Cumulative Impact on Reserve Adequacy
2005	218,878	66.8%	0.0%	0.0%	66.8%	-	-
2006	221,523	65.8%	3.7%	1.2%	68.4%	(5,855)	(5,855)
2007	219,952	69.9%	7.2%	0.5%	71.3%	(3,004)	(8,859)
2008	226,112	74.0%	10.2%	3.3%	71.3%	6,075	(2,785)
2009	221,397	73.7%	10.4%	1.2%	72.9%	1,581	(1,204)
2010	226,841	72.9%	12.8%	3.6%	72.7%	540	(664)
2011	232,142	73.6%	16.8%	6.1%	73.6%	(48)	(712)
1Q 2011	56,042	67.3%					
1Q 2012	57,190	68.5%	17.6%	8.2%	72.6%	(2,302)	(3,014)

3. OBSERVATIONS

3.1 Current Cycle

As seen in exhibit 1, the personal lines booked loss ratios have increased from the 2005 levels but they have done so by amounts close to what we should expect, meaning that there has not been a material impact on the overall industry loss reserve adequacy for personal lines.

The commercial lines in exhibit 1 show an increase in the booked normalized loss ratios from 2005 through 2008. These increases are slightly more than we might have expected. This indicates that the industry commercial lines loss reserves became stronger during the 2005-2008 calendar years. However, this situation abruptly reversed in 2009 with the booked normalized loss ratio being about 2 points less than expected. The situation deteriorated further in 2010, 2011, and so far in 2012 with the booked normalized loss ratios being about 8, 7, and 13 points less than expected, respectively. The industry booked a 68.0% normalized commercial lines loss ratio for the first 3 months of calendar year 2012 when we should have expected it to be 80.8%. This indicates that the industry commercial lines loss reserves have weakened by almost $41 billion during the 3.25 calendar years from 2009 through the first 3 months of 2012.

The indication that the calendar year 2011 commercial lines normalized loss ratio that was booked was 7 points lower than it should have been implies that the loss ratio booked for calendar year 2011 may have been 7 points lower than the ultimate loss ratio for AY (accident year) 2011. Note that this does not necessarily mean that AY 2011 itself was being booked 7 points low. The aggregate impact of AY 2011 and all the prior AY's on the 2011 calendar year resulted in the 2011 calendar year being booked 7 points lower than what the 2011 ultimate AY normalized loss ratio should be. However, since the calendar year normalized loss ratios that were booked until sometime in 2009 looked appropriate, there is a strong implication that AY's since 2009 are being booked too low.

3.2 Prior Cycle

We were curious to see how well this approach would have worked during the last soft market. We selected 1995 as the base year and used assumptions similar to those of the current market cycle except we used 2 percentage points for the judgmental increment to CPI for both personal and commercial lines. This is the adjustment to reflect that insurance industry loss trends have traditionally been larger than the CPI trend. The results for the prior soft market are displayed in exhibit 2. Rather than using the same excess catastrophe loss ratio impact on both total personal lines and total commercial lines, we assumed that the 2001 personal lines unadjusted loss ratio was at a normal level regarding cat losses and that the unusually large cat loss ratio was due to the 9/11 World Trade Center terrorist attack losses and the impact was assumed to be all in commercial lines.

US P&C Insurance ($ Millions) **Exhibit 2**

Total Personal + Commercial Lines

Calendar Period	Earned Premium	Booked Normalized Loss Ratio	Needed % Change in EP Since 1995	Actual % Change in EP Since 1995	Expected Normalized Loss Ratio	Year's Impact on Reserve Adequacy	Cumulative Impact on Reserve Adequacy
1995	254,172	74.8%	0.0%	0.0%	74.8%	-	-
1996	263,351	77.6%	6.1%	3.6%	76.6%	2,672	2,672
1997	271,502	73.7%	12.2%	6.8%	78.6%	(13,141)	(10,469)
1998	277,690	74.6%	17.9%	9.3%	80.7%	(16,977)	(27,446)
1999	282,791	77.7%	24.6%	11.3%	83.7%	(17,065)	(44,511)
2000	294,024	81.8%	32.6%	15.7%	85.7%	(11,523)	(56,034)
2001	311,529	85.0%	38.8%	22.6%	84.7%	987	(55,046)

Total Commercial Lines

Calendar Period	Earned Premium	Booked Normalized Loss Ratio	Needed % Change in EP Since 1995	Actual % Change in EP Since 1995	Expected Normalized Loss Ratio	Year's Impact on Reserve Adequacy	Cumulative Impact on Reserve Adequacy
1995	115,909	74.4%	0.0%	0.0%	74.4%	-	-
1996	118,489	78.6%	8.0%	2.2%	78.6%	(50)	(50)
1997	120,150	74.9%	16.7%	3.7%	83.8%	(10,712)	(10,762)
1998	123,357	75.9%	25.0%	6.4%	87.5%	(14,264)	(25,026)
1999	128,040	77.8%	35.0%	10.5%	91.0%	(16,871)	(41,897)
2000	135,088	79.9%	46.3%	16.5%	93.4%	(18,325)	(60,223)
2001	144,353	83.5%	53.7%	24.5%	91.9%	(12,136)	(72,358)

Total Personal Lines

Calendar Period	Earned Premium	Booked Normalized Loss Ratio	Needed % Change in EP Since 1995	Actual % Change in EP Since 1995	Expected Normalized Loss Ratio	Year's Impact on Reserve Adequacy	Cumulative Impact on Reserve Adequacy
1995	138,263	75.1%	0.0%	0.0%	75.1%	-	-
1996	144,862	76.8%	4.5%	4.8%	74.9%	2,722	2,722
1997	151,352	72.8%	8.5%	9.5%	74.4%	(2,429)	293
1998	154,333	73.5%	12.1%	11.6%	75.4%	(2,713)	(2,420)
1999	154,751	77.6%	16.3%	11.9%	78.0%	(193)	(2,613)
2000	158,936	83.5%	21.9%	15.0%	79.6%	6,802	4,189
2001	167,176	86.3%	27.2%	20.9%	79.0%	13,123	17,312

The personal lines industry loss ratios remained reasonably close to the expected loss ratios for the 1996-2000 calendar years of the last soft market, meaning that personal lines did not significantly contribute to the industry's loss reserve problem that developed during the last soft market.

The industry commercial lines booked normalized loss ratios started being significantly lower than the expected loss ratios as early as 1997 meaning that this monitoring approach would have signaled a problem fairly early in that soft market. The commercial lines booked loss ratio for calendar year 1997 was almost 9 points lower than the expected loss ratio, implying a weakening in the industry commercial lines loss reserves of $10.7 billion. The industry continued booking loss ratios through calendar year 2001 that did not fully reflect the inadequacy of the commercial lines premium and continued building up a commercial lines loss reserve inadequacy.

Exhibit 2 indicates that the industry reserves for all lines weakened by $55 billion between 1995 and 2001. An examination of the industry Schedule P Part 2 Summary[6] reveals that the industry strengthened the reserves held at year-end 2001 by $105.3 billion dollars between 2002 and 2009 with most of that strengthening occurring before 2006. The $105.3 billion included significant strengthening on asbestos and environmental reserves (A&E). We estimate that about $65 billion of the strengthening was for other than A&E. The $65B of strengthening taken on other than A&E would offset the indicated $55 billion of weakening that exhibit 6 shows built up between 1995 and 2001.

This approach for monitoring the industry would have worked well during the last soft market. It would have signaled a problem with the loss ratios being booked as early as 1997. It also would have computed a cumulative reserve weakening that agreed well with the strengthening subsequently taken.

4. CONCLUSIONS

The method for monitoring the industry explained in this paper suggests that the industry commercial lines booked loss ratios started being too low in calendar year 2009 with the gap growing in calendar year 2010 and continuing at least through the first three months of 2012. This implies a $41 billion weakening of the industry commercial lines loss reserves since year-end 2008. Back testing this method shows that it would have performed very well during the last soft market

[6] AM Best's Aggregates and Averages

by providing both an early signal that the industry was booking loss ratios that were too optimistic and a reasonably accurate estimate of the magnitude of reserve weakening that took place.

Appendix A

US P&C Industry
Workers Compensation ($000,000,000s)[7]

Accident Year	Net Earned Premium	Net Ultimate Loss & DCC[8]	Original Ult Loss Ratio (at 12 months)	Current[9] Ult Loss Ratio
1995	27.87	18.06	73%	65%
1996	28.74	20.45	73%	71%
1997	26.49	21.73	76%	82%
1998	25.57	23.58	80%	92%
1999	23.69	23.68	82%	100%
2000	26.68	26.52	80%	99%
2001	30.81	28.48	78%	92%
2002	36.10	28.33	72%	78%
2003	41.70	28.35	73%	68%
2004	46.25	26.80	71%	58%
2005	47.19	26.46	69%	56%
2006	47.74	29.09	68%	61%
2007	44.76	30.39	71%	68%
2008	41.51	30.54	73%	74%
2009	36.69	27.93	76%	76%
2010	34.64	27.54	80%	80%

7 Source: AM Best's Aggregates & Averages
8 As displayed in Schedule P Part 2
9 As of year-end 2010 or after 10 years development. Equals ratio of second column to net earned premium.

Appendix B

US P&C Industry
Commercial Auto Liability ($000,000,000s)[10]

Accident Year	Net Earned Premium	Net Ultimate Loss & DCC[11]	Original Ult Loss Ratio (at 12 months)	Current[12] Ult Loss Ratio
1995	11.42	8.92	78%	78%
1996	11.87	9.60	77%	81%
1997	12.04	10.08	78%	84%
1998	11.87	10.21	77%	86%
1999	11.83	10.93	79%	92%
2000	12.67	11.19	77%	88%
2001	13.88	10.76	73%	78%
2002	15.72	10.39	67%	66%
2003	17.47	10.45	64%	60%
2004	18.75	10.67	62%	57%
2005	19.17	11.03	61%	58%
2006	19.24	11.17	62%	58%
2007	19.07	11.67	62%	61%
2008	18.28	11.25	62%	62%
2009	17.01	10.32	63%	61%
2010	16.28	10.52	65%	65%

[10] Source: AM Best's Aggregates & Averages
[11] As displayed in Schedule P Part 2
[12] As of year-end 2010 or after 10 years development. Equals ratio of second column to net earned premium.

Appendix C

US P&C Industry
Other Liability Occurrence + Claims-Made ($000,000,000s)[13]

Accident Year	Net Earned Premium	Net Ultimate Loss & DCC[14]	Original Ult Loss Ratio (at 12 months)	Current[15] Ult Loss Ratio
1995	16.11	11.29	78%	70%
1996	17.10	12.37	78%	72%
1997	17.78	14.53	78%	82%
1998	18.72	17.88	78%	96%
1999	17.55	18.79	76%	107%
2000	18.72	19.54	76%	104%
2001	19.92	20.71	84%	104%
2002	26.78	22.82	71%	85%
2003	33.61	21.41	68%	64%
2004	39.66	20.21	67%	51%
2005	40.73	20.91	64%	51%
2006	43.52	23.80	64%	55%
2007	43.04	26.51	66%	62%
2008	41.31	28.03	67%	68%
2009	38.98	27.13	69%	70%
2010	37.60	25.76	69%	69%

13 Source: AM Best's Aggregates & Averages
14 As displayed in Schedule P Part 2
15 As of year-end 2010 or after 10 years development. Equals ratio of second column to net earned premium.

Biography of the Author

John Captain is Vice President and Reserving Actuary at Swiss Re in Schaumburg, IL. He is a Fellow of the CAS and a Member of the American Academy of Actuaries.

The Canadian Puzzle: Why Have the American and Canadian P/C Insurance Cost Structures Evolved Differently?

Marc-André Desrosiers, Ph.D. Candidate, FCAS, MBA

Abstract Kelly and Kleffner (2006) have documented that the structure in the Canadian P/C industry is materially different from that of the American P/C industry. As historical literature has rationalized the structure of the American P/C insurance industry, this represents a puzzle and a new explanation needs to be found. The attempt to solve the puzzle is relevant to actuarial practice as it directly impacts the business strategy of the insurer, to financial markets as it speaks to the efficient organization in the retail sector of the delivery of financial services, and to the theory of industrial organization as it speaks to the way the P/C insurance markets evolve. Using NAIC data from 1992-2010, the information on the distribution channel as documented in the A.M. Best's Aggregates and Averages for the matching period and quantile regression, the results on the structure of the American P/C insurance industry are reproduced, the rationalizations reviewed and the interpretations criticized. The P/C insurance industry is found to have been consolidating in the last two decades and this leads to the exploration of the role of economies of scale. The resolution of the puzzle finds its source in the role of economies of scale for various generic business strategies.

Keywords Financial service industry, economies of scale, differential evolution of markets

1. INTRODUCTION

For many actuaries, as they evolve throughout their careers, they will find themselves participating in the strategic decisions of the firms they work for, many of which will be insurers. In the role of strategic decision makers, actuaries will be presented with many theories about which course of actions are preferable to others. It is therefore invaluable for strategic decision makers to develop key concepts to anchor discussions, methodological understanding to establish support for arguments, and a repertoire of key results that can be readily called upon.

The present paper aims to assist decision makers in the P/C insurance industry on all three fronts. We will do so by re-examining the results of the historical literature on cost efficiency in the P/C insurance literature. While we explicitly focus on the P/C insurance industry, it is our belief that many results are equally applicable to many retail financial services industries, such as retail banking. In the past, it has been found that direct insurers tend to have lower underwriting expenses due to earlier implementation of cost saving technology and, in the USA, tend to dominate Personal Lines requiring less personalized service. If the nature of the distribution channel is the key determinant of cost efficiency and of the line of business choice, with Canada being quite similar to the USA, the same industry structure should be observed. However, it has been found that the Canadian P/C insurance industry structure is quite different from that of the USA Methodologically, we will present and make use of quantile[1] regression, as this will allow us to examine the impact of covariates on the whole distribution of the dependent variable of interest. Unfortunately, while we would prefer to generate causal models[2], quantile regression does not in and of itself always lend itself to causal interpretation and care will be taken in the interpretation of the results. Conceptually, we will anchor ourselves in the Porter generic strategies framework to attempt to formulate a reconciliation of the industrial organization puzzle regarding the differential evolution of the P/C insurance markets across the borders. Having established the critical importance of economies of scale to the P/C insurance industry and having thought through the potential sources of economies of scale, we will discuss why the choice of the generic business strategy should come before the choice of the marketing strategy, including the choice of the distribution channel.

For the ratemaking actuary, the following aspects of the present research should be of particular interest. One, it is the hope of the author that the actuary will be better equipped to understand the relationship between the growth/size of the insurer and the expense ratio, and how the rates could adjust (or not) as a function of the market structure in which the insurer evolves. In particular, the ratemaking actuary will be better equipped to think through whether economies of scale should be passed on to customers and to what extent. Second, the author wishes to demonstrate the usefulness of quantile regression when the actuary is attempting to understand the impact of covariates on the distribution of dependent variables[3].

1.1 Outline

The remainder of the paper proceeds as follows. Section 2 is fully dedicated to the setting up of the puzzle. We will start by describing the data used in section 2.1.1. Then we will describe the main econometric strategy in section 2.1.2. In section 2.1.3, we'll describe a proxy variable that we'll use *in lieu* of the distribution channel, when it will be convenient to do so. In sections 2.1.4 and 2.1.5, we'll validate and rationalize the results of the historical literature on cost efficiency in the P/C insurance market: the cost advantage of direct writers and the relative preference of broker writers for Commercial Lines. Section 2.1.6 will cover other relevant historical findings. In section 2.2, we'll review findings related to the structure of the Canadian P/C insurance industry. In section 2.3, we will discuss the econometric flaw in the interpretation of results in the historical literature. Section 3 will be fully dedicated to discussing economies of scale in the P/C insurance industry. We'll start by

[1] Quantiles are points taken at regular intervals from the cumulative distribution function of a random variable.
[2] Causal models are particularly preferable when the actuary is undertaking a budgeting exercise, as the actuary can then use the implied 'action-reaction' interpretation of causal models and use it for planning and forecasting purposes.
[3] As will be discussed below, quantile regression isn't ideal when the expected dollar value is of interest, but quantile regression could be quite useful for actuaries when they are, for example, attempting to understand how different insured's characteristics are affecting the distribution of the retention or conversion ratio.

providing the first hint that economies of scale are available in section 3.1. In section 3.2, we'll discuss potential sources of economies of scale. In section 3.3, we'll discuss some determinants of insurer size/growth. In section 3.4, we'll quickly discuss potential consequences of growth. In section 4, we'll provide a beginning of a reconciliation of the puzzle by appealing to Michael Porter's generic business strategies.

2. SETTING UP THE PUZZLE

In the section, we will set up the puzzle that we will attempt to resolve in the next sections. Significant portions of the text will be dedicated to discussing the available data and material hypotheses related to its treatment. We will also present the main econometric strategy of quantile regression. We will also review historical results and critic their interpretation.

2.1 The Cost Structure of the American P/C Insurance Industry

2.1.1 Data

The data that will be described here serves as the basis of most of the analysis found in the present and subsequent sections. The data comes from two main sources: (1) the National Association of Insurance Commissioners databases of regulatory Property/Casualty financial statements from years 1992 to 2010, and (2) the Best’s Aggregates & Averages: Property-Liability from year 1993 to 2011[4]. We will describe the material data gathering hypothesis starting with how the NAIC data was put together for the purposes of the current analysis.

First, insurers were considered on a group basis: that is, if a group code was present in the NAIC data, the data that was kept was the data coded at the group level; otherwise, if a group code was unavailable, the individual insurer was treated as a group. When multiple companies reported as "combined" for a given group code, we used the total for the group code[5].

Second, the following is a table that describes which expense exhibit lines were used to form different categories of expenses. Which code was used is year dependent, following the documentation of the NAIC databases.

Claims Adjustment Services, Direct	01A, 01.1
Commission, Direct	02A, 02.1
Contingent Commission, Direct	02D, 02.4
Advertising	04
Equipment	14, 15
Total Expenses Incurred	22, 25

[4] As the Aggregate & Averages book covers financial information for up to the preceding calendar year, the calendar year of the NAIC data match those of the Best's Aggregates & Averages.
[5] That was uncommon.

Third, in the following table, the rules of groupings of lines of business are documented.

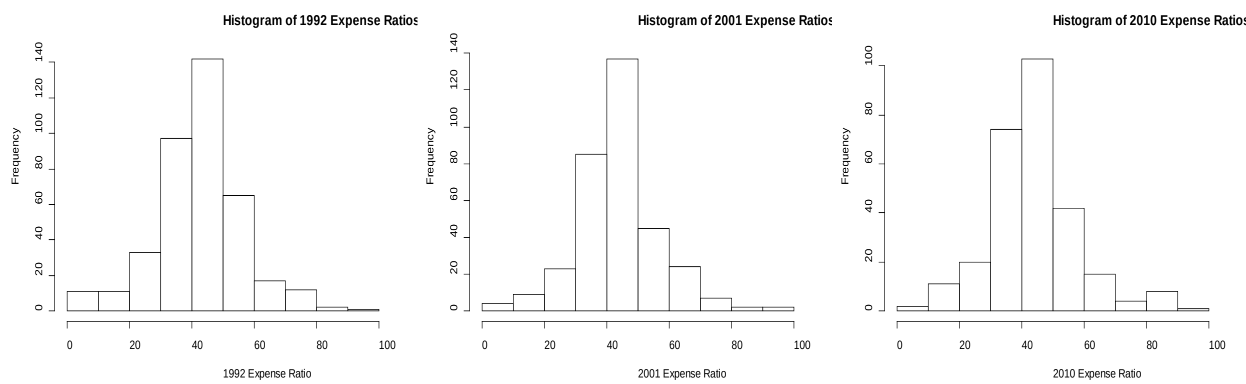
Automobile	Automobile liability, Automobile physical damage[6]
Commercial Lines Non-Auto	Commercial multiple peril, Ocean marine, Inland marine, Medical malpractice, Fidelity, Surety, Burglary and theft, Boiler and machinery, Other liability, Products liability, Farmowners multiple peril, Fire
Personal Lines Non-Auto	Homeowners multiple peril, Allied Lines, Earthquake

Fourth, ratios that relate to income are all computed using Earned Premium as the denominator. Two main ratios will serve to measure cost efficiency: (1) the underwriting expense ratio and (2) the underwriting income ratio. Using a measure of underwriting expense as a ratio to Earned Premium allows us to avoid needing to transform the measured expenses before being able to model them[7], as the resulting distribution of ratios is roughly symmetric and relatively light-tailed[8]. There are potentially some flaws with measuring cost efficiency using a ratio to premium. In a possible market structure[9], it could be the case that all gains in efficiency are entirely kept by firms in the way of profit such that the insureds never see any rate decrease associated with increased efficiency. In that case, the expense ratio would exactly reflect efficiency gains. In another possible market structure[10], insurance prices may shift without any related changes to the cost function such that the measured change in the expense ratio would not be reflective of (in-)decreased efficiency. The effective assumption made here is that neither pure scenario is reflective of reality: we acknowledge that the expense ratio is an imperfect measure of efficiency while maintaining its use, thus assuming that it is still a useful and practical measure of efficiency[11]. Another problem associated with using the underwriting expense ratio is that it ignores the fact that different insureds receive different levels of service. For example, it is sometimes assumed that insureds dealing with brokers receive

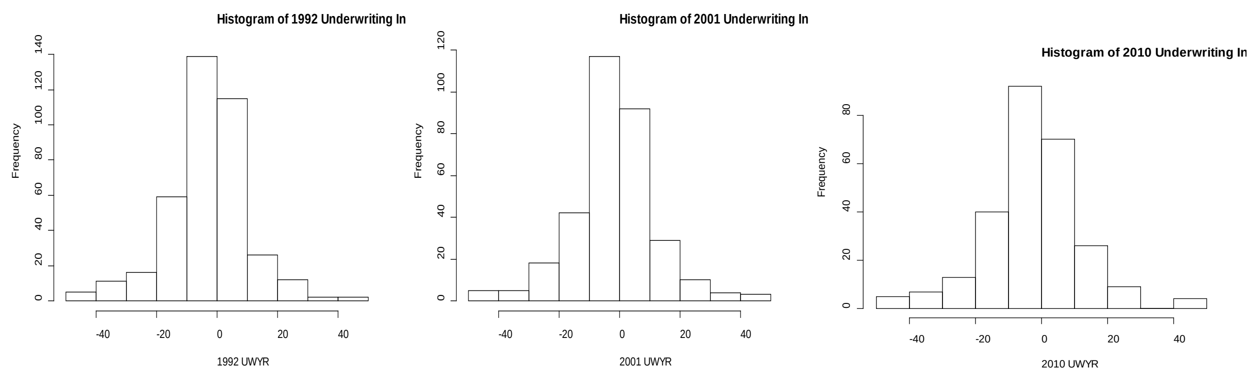
[6] Note that Commercial Automobile is included in Automobile because it cannot be separated for all considered years.
[7] In (Shi and Frees 2010, 307), the authors note that un-scaled and un-transformed expense have a long-tail distribution. The proposed approach used here is a re-scaling by Earned Premium. The authors note that re-scaling may not always be appropriate as, for predictive purposes, an estimate of the future values of the denominator first needs to be formed. However, in the case of Earned Premium, a significant portion of a one year ahead forecast is based on a realized value of Written Premium, such that the criticism loses some force.
[8] See tables below.
[9] See, for example, (Allen, Clark and Houde 2008) where the authors have assumed that any efficiency gains made by banks that are able to 'lead' customers to a more intensive use of less expensive electronic banking technology are not passed on to customers in decreased prices. In that model, banks are balancing the profit they are losing from customers driven away by decreased service with the profits gained on retained customers that switch to a lower cost technology. In that market structure, decreased competition and market dominance facilitate the adoption of internet banking.
[10] See, for example, (Brown and Goolsbee 2002), where prices may change without any change in the cost structure. In that model, the introduction of a technology that reduces search costs is assumed to force insurers to reduce their profits as markets become more competitive and insurers have to give up the rent they built up using price discrimination. The authors, unfortunately, fail to consider how an insurer that may generate significant growth using the internet channel may see its efficiency grow due to economies of scale.
[11] An ideal measure of cost efficiency would consider the different costs incurred by two insurers when they are servicing observationnally equivalent insureds. Unfortunately, the NAIC database does not contain number of insureds information and much less their characteristics.

supplementary assistance from the broker in the claims handling process[12]. To the extent that extra costs incurred are related to value-added activities, the increase in costs is not the result of an efficiency loss: thus, the underwriting expense ratio is also flawed in this way. A proposed remedy is to use the underwriting profit ratio instead, as this ratio would include an inflated denominator if all value-added activities were effectively paid for by the insured. Another way that the underwriting income ratio could serve to alleviate some of the flaws related to the underwriting expense ratio is that it can reflect different sources of economies of scale such as: (1) increased efficiency in loss adjustment, (2) lower loss ratio due to the impact of market power in the repair/replacement good market[13], (3) increased effectiveness in costing or modeling of insureds[14], *etc.*

Expense Ratio



Underwriting Income Ratio



Fifth, for quantile regression purposes, the biggest insurers that together compose 95% of the market share in any given year were kept for all years they are available. All together, these provide

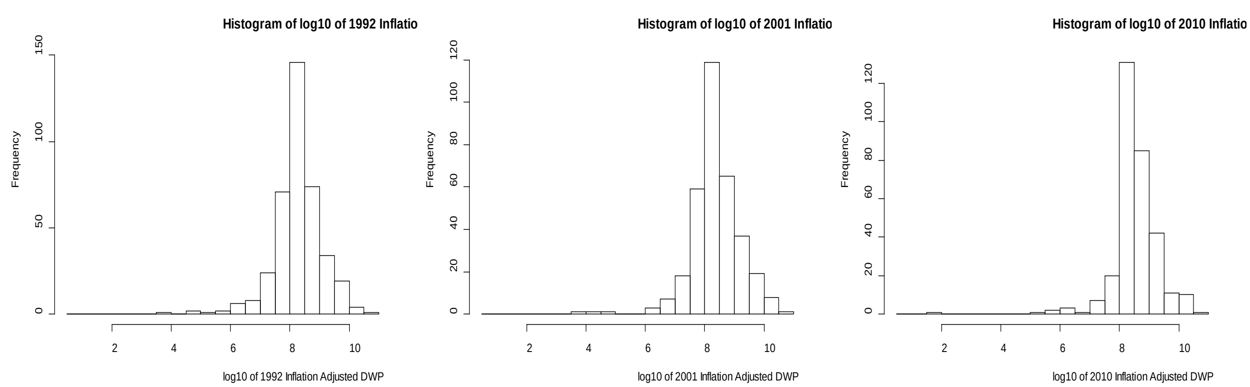
[12] See, for example, (Regan and Tennyson, Agent Discretion and the Choice of Insurance Marketing System 1996, 642).
[13] See, for example, (Nell, Richter and Schiller 2009).
[14] See, for example, (Intact Financial Corporation 2010, 6).

over sixty-five hundred insurer-year observations and over eight trillion inflation adjusted dollars of direct written premium.

Sixth, as there are times were some information is absent or composed ratios have a denominator of 0, the used quantile regression of R, rq, was set to omit missing information.

Seventh, inflation adjusted Direct Written Premium were put at 2010 level using the Consumer Price Index[15]. Inflation adjusted Direct Written Premium will serve as a measure of the size of a P/C insurer. Inflation adjustment is critical because it would be otherwise impossible to make inter-temporal comparisons. Inflation adjusted payroll or salaries could also have served as a measure of size. One advantage of payroll as a measure of size is its decreased sensitivity to the underwriting cycle and to rate levels[16]. Another advantage of using a non-claim related measure of size is that it avoids the introduction of a bias related to measurement of potential economies originating from the claims process[17] as, assuming economies of scale in the claiming are passed on to customers, premium growth will be a biased down measure of size. However, as can be visualized from the histograms presented below, inflation adjusted DWP is amply sufficient to allow us to discriminate between a very small insurer and a very large insurer, and everything in between.

$\log_{10} DWP_{inflation \quad adjusted}$



Eight, using (A.M. Best Company. 1993-2011), based on the "Total All Lines" sheet, where the distribution channel is documented for an insurer group, over 732 insurer years were assigned to the documented distribution channel. The following table documents how channels were consolidated in "Agency" or "Direct" when multiple codes were available.

---

15 As measured in (Bureau of Labor Statistics 2011).

16 Another measure of insurer size that does not necessarily suffer from those weaknesses is trend adjusted indemnity paid; however, as noted in (Skogh 1982, 219), the volatility of insurance losses causes a bias because of "the presence of a stochastic component in claims paid in various years."

17 This in part motivates why (Skogh 1982) uses payroll, or compensation paid, to measure insurer size.

Market Type	Simplified Market Type	COMMENT
A	AGENCY	
AB	AGENCY	(A FOR AGENCY; B FOR BROKER)
AD	MIXED	(A FOR AGENCY; D FOR DIRECT)
AK	MIXED	(A FOR AGENCY; K FOR OTHER DIRECT)
AR	REINSURER	(CODE FROM 1993 TO 2002)
B	AGENCY	B FOR BROKER
D	DIRECT	
DA	MIXED	(A FOR AGENCY; D FOR DIRECT)
DB	MIXED	(D FOR DIRECT; B FOR BROKER)
DL	MIXED	(D FOR DIRECT; L FOR GENERAL AGENT)
DR	REINSURER	(CODE FROM 1993 TO 2002)
E	DIRECT	
EA	MIXED	(E FOR EXCLUSIVE/CAPTIVE AGENTS; A FOR AGENCY)
ED	DIRECT	(E FOR EXCLUSIVE/CAPTIVE AGENTS; D FOR DIRECT)
GB	MIXED	

Ninth, the following table describes the variables used in quantile regressions. Note that each variable can be for the same year as the year considered and, in that case, the variable name is appended by $\_minus\_0$, it can be for the year prior to the year considered in which case the variable name is appended with $\_minus\_1$.

DWPt_onl	Direct Written Premium adjusted for inflation
log10_DWPt_onl	$\log_{10} DWP_{inflation\quad adjusted}$
CCCR	Commission and Contingent Commission Ratio (to Earned Premium)
AdvR	Advertising Ratio (to Earned Premium)
EquipR	Equipment Ratio (to Earned Premium)
ExpR	(Underwriting) Expense (to Earned Premium)
UWYR	Underwriting Income Ratio (to Earned Premium)
Auto_share	For the insurer group, the share of DWP coming from the Automobile Line of Business
CLNA_share	For the insurer group, the share of DWP coming from the Commercial Non-Auto Lines of Business
LLAER_diff	Differential of the group Loss and Loss Adjustment Ratio (to Earned Premium) compared to the industry, in a given year
growth_diff	Differential of the group DWP growth from the prior year compared to the industry, in a given year
simplified_channel	Distribution channel as identified using the (A.M. Best Company. 1993-2011) documentation

2.1.2 Econometric specification: the choice of quantile regression

Contrary to most of the existing literature examining cost efficiency in the American P/C insurance industry[18], we will not use either Ordinary Least Squares or Weighted Least Squares regression to study the effect of covariates on variables of interest. The main reason why we are

[18] (Shi and Frees 2010) being a notable exception.

choosing quantile[19] regression is that it will allow us to study how covariates change the distribution of the variable of interest. Ideally, we would prefer to provide a causal interpretation of the coefficients[20], but that may not always be possible. Provided that a direct or indirect causal link can be found, or at least imagined, a quantile regression treatment would allow us to identify which part of the distribution of the dependent variable is affected by the covariates. For example, as discussed in (Koenker and Machado, Goodness of Fit amd Related Inference Processes for Quantile Regression 1999, 1297), Chamberlain was able to find that union membership had significantly more effect for workers with lower wages compared to workers with higher wages. Other reasons why quantile regression may be preferred include (1) its robustness to outliers while maintaining high efficiency and (2) the ease with which transformed data can be used in estimation[21]. Also, fortunately, quantile regression also has a projection interpretation as a best linear predictor of the quantile of a conditional distribution.

There are three main routes to quantile regression. The first route[22] is quite convenient and practical when available and is based upon the Generalized Method of Moments. This method is only available when the covariates are discrete and data is abundant for each combination of covariates. The method basically consists of computing the quantile of interest of the dependent variables $y_\tau|\boldsymbol{x}$ for each combination of covariates and then running a Weighted Least Squares regression on the sample $\{(y_\tau|\boldsymbol{x}), \boldsymbol{x}\}_{\forall \boldsymbol{x}}$. The weights are computed as a function of the quantile τ, the proportion of observations that have combination of covariates $\boldsymbol{x}$, and the density of the residuals $\boldsymbol{\varepsilon}$. It is possible to stretch the application of the method when the data is continuous by discretizing the covariates and imputing a single value of $\boldsymbol{x}$ to the binned observations. Doing this requires that there is little variability in the covariates within a bin. Unfortunately, it is this condition that prevents us to use this simple yet powerful method for inference purposes. Nonetheless, as is exemplified in the tables below, to which we'll come back to later, this approach can be quite useful in data exploration, as it can allow us to quickly visualize how the distribution of a variable is affected by another variable. To facilitate this visual exploration, the author has used the conditional formatting function of Excel to make it more apparent that, in the first table, generally, the share of Commercial Lines Non-Auto line of business increases as the commission rate increases while, in the second table, the share of the Automobile lines of business generally increases when the commission rate decreases.

---

[19] See the appendix for a refresher on how to compute quantiles in the univariate case.
[20] In (Koenker and Machado, Goodness of Fit and Related Inference Processes for Quantile Regression 1999, 1296-1297), the authors discuss quantile treatment effects.
[21] See (Koenker and Bassett, Regression Quantiles 1978, 39).
[22] See (Buchinsky 1994, 409).

Binned Prop. Comm. Non-Auto Curr. Year (%) (→) *vs.* Binned Comm. & Cont. Comm. Curr. Year (%) (↓)

Year	(All)

	Columns										
	On-Level DWP Curr. Year %										Total On-Level DWP Curr. Year %
Row Labels	0	10	20	30	40	50	60	70	80	90	100	
[00.00-01.00)	84.3%	2.0%	1.8%	1.9%	1.7%	0.9%	0.5%	0.1%	0.0%	0.2%	6.5%	100.0%
[01.00-05.00)	32.2%	43.5%	1.1%	6.9%	2.9%	1.0%	4.6%	1.6%	0.7%	1.0%	4.6%	100.0%
[05.00-10.00)	36.5%	22.2%	12.5%	5.1%	3.9%	12.8%	2.1%	1.6%	0.4%	0.5%	2.3%	100.0%
[10.00-12.50)	17.7%	54.8%	4.6%	10.4%	2.6%	8.8%	0.7%	0.1%	0.0%	0.1%	0.2%	100.0%
[12.50-15.00)	12.0%	6.6%	40.2%	5.9%	10.3%	11.8%	11.3%	1.5%	0.0%	0.3%	0.0%	100.0%
[15.00-17.50)	3.6%	2.2%	20.0%	20.9%	18.5%	18.3%	12.6%	3.3%	0.4%	0.1%	0.0%	100.0%
[17.50-20.00)	4.2%	3.9%	5.3%	46.6%	18.2%	13.7%	4.8%	1.8%	0.9%	0.6%	0.0%	100.0%
[20.00-30.00)	6.8%	5.2%	6.3%	9.5%	20.5%	18.0%	7.2%	13.6%	6.4%	4.6%	1.9%	100.0%
[30.00-99.99)	17.8%	8.3%	7.8%	11.5%	8.5%	16.1%	11.2%	5.4%	6.6%	3.8%	3.1%	100.0%
Grand Total	**20.2%**	**21.0%**	**12.2%**	**14.2%**	**9.4%**	**11.8%**	**5.7%**	**2.4%**	**1.0%**	**0.8%**	**1.4%**	**100.0%**

Binned Prop. Auto Curr. Year (%) (→) *vs.* Binned Comm. & Cont. Comm. Curr. Year (%) (↓)

Year	(All)

	Columns										
	On-Level DWP Curr. Year %										Total On-Level DWP Curr. Year %
Row Labels	0	10	20	30	40	50	60	70	80	90	100	
[00.00-01.00)	32.4%	2.0%	2.4%	0.4%	1.1%	1.7%	12.9%	28.4%	8.0%	8.8%	1.8%	100.0%
[01.00-05.00)	24.9%	3.8%	5.8%	4.9%	8.8%	3.9%	2.4%	2.8%	25.8%	16.1%	0.7%	100.0%
[05.00-10.00)	9.4%	4.8%	16.1%	5.8%	7.6%	0.6%	4.5%	26.6%	11.0%	0.3%	13.3%	100.0%
[10.00-12.50)	2.5%	2.7%	7.6%	8.4%	3.7%	2.5%	28.2%	41.2%	1.3%	0.4%	1.4%	100.0%
[12.50-15.00)	4.8%	11.8%	13.8%	12.6%	4.2%	17.1%	23.3%	9.1%	1.0%	0.9%	1.4%	100.0%
[15.00-17.50)	3.3%	13.4%	19.4%	20.6%	9.7%	22.5%	6.6%	0.6%	0.6%	2.8%	0.7%	100.0%
[17.50-20.00)	3.8%	8.6%	9.6%	15.1%	43.5%	10.9%	2.7%	1.1%	0.8%	3.3%	0.5%	100.0%
[20.00-30.00)	16.9%	29.3%	18.3%	10.1%	7.9%	6.7%	3.2%	1.8%	1.8%	2.3%	1.7%	100.0%
[30.00-99.99)	30.2%	28.7%	15.8%	7.5%	2.7%	2.7%	1.7%	0.9%	3.1%	3.7%	3.1%	100.0%
Grand Total	**9.5%**	**9.3%**	**12.4%**	**10.6%**	**10.4%**	**8.4%**	**11.9%**	**16.3%**	**5.0%**	**3.0%**	**3.2%**	**100.0%**

A second route[23] to quantile regression uses the Generalized Method of Moments to solve the moment condition

$$\sum_{i=1}^{n}\left(\left(\left(I\{y_i \le \boldsymbol{\beta}^{\tau}\boldsymbol{x}_i\} - \tau\right)\boldsymbol{x}_i\right)'\left(\left(I\{y_i \le \boldsymbol{\beta}^{\tau}\boldsymbol{x}_i\} - \tau\right)\boldsymbol{x}_i\right)\right) = \mathbf{0} \qquad (2.1.2.1)$$

where τ is the quantile of interest, $\boldsymbol{\beta}^{\tau}$ is a vector of coefficients, $I\{\cdot\}$ is the indicator function, and n is the number of observations. Using this approach, the coefficients and standard errors can be computed using the standard Generalized Method of Moments machinery[24]. Finally, one can minimize the criterion function

$$S(\boldsymbol{\beta}^{\tau}) = \sum_{i=1}^{n}\left((1-\tau)|y_i - \boldsymbol{\beta}^{\tau}\boldsymbol{x}_i| I\{y_i \le \boldsymbol{\beta}^{\tau}\boldsymbol{x}_i\} + \tau|y_i - \boldsymbol{\beta}^{\tau}\boldsymbol{x}_i| I\{y_i \le \boldsymbol{\beta}^{\tau}\boldsymbol{x}_i\}\right) \quad (2.1.2.2)$$

by setting $\widehat{\boldsymbol{\beta}}^{\tau} = arg\min_{\boldsymbol{\beta}^{\tau}} S(\boldsymbol{\beta}^{\tau})$. This approach requires the implementation of a linear programming algorithm and is best done with a computer or vector algebra system. It is important to note that the $arg\min_{\boldsymbol{\beta}^{\tau=0.50}} S(\boldsymbol{\beta}^{\tau=0.50})$ is generally not the set of coefficients that return the

---

23 (Hansen 2011, 169-173) makes an introductory presentation of median and quantile regressions that is appropriate for mathematically inclined actuaries.

24 See, for example, (Hansen 2011, 108-114, 134-145).

conditional mean, or the Best Linear Prediction[25] of the conditional mean, but rather the conditional median, or the Best Linear Prediction of the conditional median. Actuarially speaking, that is generally an undesirable feature of quantile regression as the quantity of actuarial interest is very often the conditional mean itself. In this case, however, as we are not so much interested in the changes in conditional mean but in the changes of the distribution itself, this disadvantage of the quantile regression has no force.

For our purposes, we have chosen to use the rq implementation of quantile regression available in the R statistical software[26]. So doing, we have a choice of three possible ways to compute standard errors (SE) that don't assume that the error terms are independent and identically distributed or use a computation intensive bootstrap algorithm. One of the methods is based on (Koenker and Machado, Goodness of Fit amd Related Inference Processes for Quantile Regression 1999) but does not return p-values, but only a confidence interval. One of the methods is based on the more traditional "sandwich" form for standard errors but is computationally unstable on the considered data, as it regularly returns message errors. Finally, the here preferred method, "*ker*", is based on (Newey and Powell 1990, 302) and implements a non-parametric kernel estimation algorithm to compute the density of ε at the appropriate points, as required by theory.

In the quantile regression tables found below, the models were estimated for five quantiles: the $10th, 25th, 50th, 75th, 90th$ percentiles. In all cases, the models are separately estimated for the sake of convenience[27].

For the purposes of the current analysis, quantile regressions were computed using Direct Written Premium as weights[28]. Weights were introduced not for the sake of statistical efficiency, but for the purpose of better reflecting the impact of relative efficiency on the public and, most importantly, the insureds.

Also, again, contrary to most of the existing literature on the subject of cost efficiency in the P/C insurance market[29], we make use of the panel structure of the data. We use it only when it comes time to understand the drivers of the Loss and Loss Adjustment Expense Ratio and of Direct Written Premium growth. For these two dependent variables as opposed to the underwriting

---

[25] Depending on whether one wants to interpret the model as a regression or as a projection.
[26] See (R Documentation n.d.).
[27] Joint estimation of multiple quantiles is certainly feasible and relatively simple to implement using specifications based on the Generalized Method of Moment. One of the usefulness of joint estimation is to allow tests of equality of coefficients across quantiles. As this is not of interest to us here, it seems acceptable not to undertake joint estimation.
[28] Although quantile regression using insurer-years as weights were also computed. The results were generally similar and can be made available upon request to the author.
[29] (Shi and Frees 2010) and (Hecht 1999) being notable exceptions.

expense ratio, it is apparent that there are material year to year fluctuations and it is best to first neutralize the year effect before attempting a regression.

Year	Loss and LAE Ratio (%)	Expense Ratio (%)	DWP Growth (%)
1992	75.2	41.1	
1993	66.8	40.2	6.5
1994	68.6	41.5	1.0
1995	65.7	42.3	0.7
1996	65.4	40.4	5.6
1997	60.4	42.1	6.9
1998	63.1	41.8	5.2
1999	65.1	***46.3 ****	5.9
2000	67.8	40.8	1.5
2001	75.1	***44.3 *****	9.8
2002	68.3	40.4	17.8
2003	61.6	39.5	9.4
2004	59.9	39.1	3.8
2005	61.5	39.4	2.2
2006	53.2	40.0	4.1
2007	55.7	39.7	0.9
2008	65.4	39.7	-0.4
2009	59.3	40.7	-1.8
2010	61.1	41.7	1.5

When the insurer groups that make up 95% of DWP are used, the starred numbers are:
* 42.2
** 42.6

Note, however, that we do not otherwise really make use of the panel structure of the data for quantile regression and use all data as if it all came from one large cross-section because of the following rationale. Take the underwriting expense ratio as an example. In this case, for 80% of the Earned Premium available in the study, the year-to-year variability of the underwriting expense ratio is 5.7% or less, while the inter-group (all years combined) underwriting expense ratio has a standard deviation of 12.8%. This provides an indication that the expense ratio of the current year is largely determined by the expense ratio of the prior year for most insurer groups, especially under normal operations. This is *a priori* plausible because expenses, as opposed to losses, are largely in the control of the insurer and are subject to internal controls. Since we are interested in what features of the insurer drive the level of the underwriting expense ratio, since the level is approximately constant for most insurers under most circumstances, and since the features we'll be considering are also largely constant through time for most insurer groups, this justifies treating the entire dataset as being generated by one cross-section. This rationale applies also when we're considering the commission rate, and the Commercial Non-Auto and Automobile lines of business share of premium.

2.1.3 The Commission and Contingent Commission Ratio as a Proxy for the Distribution Channel

One of the key variables that have been examined in the literature concerning the cost efficiency of P/C insurers has been the distribution channel. In sub-sections 2.1.4 and 2.1.5, we will discuss and validate the historical findings.

As noted above, the exact distribution channel of insurers is only known for 732 insurer-years. Taking into account the size of the full database, it is apparent that it is desirable to identify a proxy for the distribution channel so as to enable us to use the full database when results not dependent on the exact knowledge of the distribution channel are required.

Target Variable:	CCCR_minus_0									
Percentile	10th		25th		50th		75th		90th	
DWP Weights; S.E.: "ker" Method	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value
intercept	2.420	0.4%	6.440	7.3%	10.830	0.0%	11.470	0.0%	14.290	0.0%
agency	7.320	0.0%	6.320	8.1%	4.800	0.0%	6.210	0.0%	5.460	0.0%
mixed	-0.310	80.7%	-4.030	28.3%	-3.530	0.1%	-2.790	0.2%	-5.290	0.0%
reinsurer	0.320	93.0%	-3.470	48.6%	2.840	0.0%	2.200	0.0%	-0.620	61.5%

As can be seen in the table above, agency writers, that do not distribute their insurance products directly to consumers, tend to have a higher Commission and Contingent Commission Ratio. This conclusion can be reached by examining the coefficients associated with the agency indicator variable for the quantile regression for the different quantiles: the coefficients are all positive and significant (as their p-value are all under 1%). This suggests that for the $10th, 25th, 50th, 75th,$ and $90th$ percentiles, the (DWP weighted) Cumulative Distribution Function of the $CCCR\_minus\_0$ variable for agency writers lies to right than the one for direct (non-agency, non-mixed, non-reinsurer) writers.

This is unsurprising because they are using external and independent parties to distribute their products. They will tend to have to compensate these parties in commissions more so than they would an employee. The difference arises because of the difference in the situation between an insurer and its brokers *versus* an insurer and its employees. In the case of a salaried workforce, while it is necessary to maintain incentive compatibility and offer a compensation package that rewards the employee for acting in the interest of the insurer, employees generally desire a significant portion of their compensation to be fixed and not subject to risk. This can be contrasted with the situation of an external contractor that is not salaried. Thus, it is not surprising to see that commissions are higher when insurers distribute through brokers. This leads us to formulate the following rule-of-thumb: as the CCCR of an insurer goes up, the likelihood that the insurer is a direct writer goes down.

2.1.4 Cost efficiency of direct writers

One of the key findings of the historical literature concerning itself with the cost efficiency in the P/C insurance industry is that, in the USA, direct writers tend to be more efficient than agency writers that distribute their products through independent brokers[30]. The table below illustrates that, for insurers that are in the upper half of the distribution of expenses conditional on their known distribution channel, insurers that distribute through independent brokers tend to have a higher underwriting expense ratio. A similar phenomenon can be found when we use the proxy variable CCCR as a predictor of the expense ratio.

Target Variable:	ExpR_minus_0									
Percentile	10th		25th		50th		75th		90th	
DWP Weights; S.E.: "ker" Method	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value
intercept	32.290	0.0%	36.040	0.0%	38.240	0.0%	41.690	0.0%	43.870	0.0%
agency	1.730	41.2%	0.600	60.5%	3.600	0.0%	4.130	0.0%	6.690	0.0%
mixed	-4.440	9.7%	-5.390	0.6%	-1.780	28.1%	-2.400	14.4%	-1.730	29.4%
reinsurer	2.460	70.8%	1.670	85.9%	2.570	0.3%	-0.880	38.4%	-3.060	0.1%

Target Variable:	ExpR_minus_0									
Percentile	10th		25th		50th		75th		90th	
DWP Weights; S.E.: "ker" Method	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value
intercept	2.709	9.0%	8.993	0.0%	15.497	0.0%	21.368	0.0%	28.122	0.0%
CCCR_minus_0	1.942	0.0%	1.942	0.0%	1.942	0.0%	1.941	0.0%	1.941	0.0%

When examining the impact of distribution channel on the overall Underwriting Income Ratio, known agency writers seem to do as well as direct writers, as can be seen in the table below. When we use the proxy variable CCCR, the values become significant, but the scale of the coefficients become economically neglectable.

Target Variable:	UWYR_minus_0									
Percentile	10th		25th		50th		75th		90th	
DWP Weights; S.E.: "ker" Method	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value
intercept	-14.670	0.0%	-9.360	0.0%	-3.270	0.0%	1.670	26.5%	8.080	0.2%
agency	-4.720	7.2%	1.680	46.9%	1.880	14.3%	3.180	11.6%	2.490	36.7%
mixed	12.040	0.0%	9.530	0.6%	8.940	0.2%	20.970	0.0%	21.910	0.0%
reinsurer	-22.000	48.2%	2.120	39.5%	-3.970	0.6%	-8.910	0.0%	-4.370	74.1%

Target Variable:	UWYR_minus_0									
Percentile	10th		25th		50th		75th		90th	
DWP Weights; S.E.: "ker" Method	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value
intercept	-17.435	0.0%	-8.230	0.0%	-2.264	0.0%	3.873	0.0%	10.265	0.0%
CCCR_minus_0	0.000	0.0%	-0.001	0.0%	-0.001	0.0%	-0.001	0.0%	-0.001	0.0%

This finding of greatest efficiency of the direct channel has lead some actuaries, like Sholom Feldblum, to criticize the agency way of distributing insurance[31]. Part of the expense advantage that

[30] Findings found, for instance, in (Cummins and VanDerhei 1979), (Barrese and Nelson 1992), (Berger, Cummins and Weiss 1997).

[31] "Independent agency companies pay level commissions, such as 15% or 20% of premium, in all years. The level commission structure is needed because the agent "owns the renewals" (National Fire Insurance case of 1904). (…) A

direct writers have built comes from early adoption of improved technology related to collection of premium[32]. However, part of the reason of the persistence of the broker distribution channel may well be due to the fact that brokers undertake value-added activities for the insureds[33].

2.1.5 Relative strength of agency writers in Commercial Lines

Another key finding of the historical literature concerning itself with the cost efficiency in the P/C insurance industry is that insurers that distribute through brokers tend to be more present in the Commercial Lines Non Auto lines of business in the USA. On the flip side, as is demonstrated in the second table below, direct writers tend to write more of the Automobile line of business. Unfortunately, as can be seen in the two tables of Appendix B, when the proxy variable CCCR is used, the findings are not conclusive; however, note that these relationships were visually explored in section 2.1.2 and the findings were supportive of historical findings. In this particular case, the results from the historical literature can be said to be confirmed by the current data.[34]

Target Variable:	CLNA_share_minus_0									
Percentile	10th		25th		50th		75th		90th	
DWP Weights; S.E.: "ker" Method	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value
intercept	3.688	0.0%	4.938	0.0%	5.690	0.0%	13.253	0.0%	23.501	0.0%
agency	7.283	20.8%	22.453	0.0%	36.212	0.0%	39.161	0.0%	36.852	0.0%
mixed	-1.406	74.3%	-2.434	58.8%	8.478	18.1%	45.581	0.0%	40.978	0.0%
reinsurer	5.417	72.7%	11.532	63.1%	18.386	0.0%	10.823	0.4%	0.575	82.2%

[35]

Target Variable:	Auto_share_minus_0									
Percentile	10th		25th		50th		75th		90th	
DWP Weights; S.E.: "ker" Method	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value
intercept	38.837	0.0%	60.339	0.0%	67.754	0.0%	71.158	0.0%	76.754	0.0%
agency	-30.431	0.0%	-43.543	0.0%	-42.274	0.0%	-27.844	0.0%	-10.664	16.6%
mixed	-38.837	0.0%	-60.339	0.0%	-7.530	30.4%	-1.756	78.0%	-3.453	60.8%
reinsurer	-33.935	46.0%	-46.661	41.2%	-17.384	0.0%	-20.788	0.0%	7.989	79.9%

lower commission in renewal years would induce the agent to move the policy to a competing insurer and obtain a "first year" commission.

The level commission structure does not reflect the actual incidence of acquisition expenses, since agents spend more effort writing new policies than renewing existing policies. **Because of this (and other reasons), many economists consider the independent agency system to be inefficient**. In the personal lines of business, direct writers are steadily gaining market share, and the level commission structure is becoming less important. As the asset share pricing model shows, a level commission structure works well for risks that terminate quickly. It works poorly for risks that endure with the carrier. " (Feldblum 1996, 205-206) [my emphasis]

32 See, for example, (Gron 1998, 410).

33 For example, see (Cummins and Doherty 2006, 361).

34 A potential explanation for the seemingly contradictory results could, in part, emanate from smaller insurers that choose to focus on a particular sub-market. To be competitive, they are more likely to distribute directly without having to pay commissions or they may use a broker that receives a lower commission rate because of the economies of scale that could accrue on the brokerage side.

35 Even though the "CLNA_share_minus_0" and "Auto_share_minus_0" variables are fundamentally variables that lie on a bounded [0,1] support, the author feels it is acceptable in this case to use quantile regression as it has been presented because the intent is only to show the existence of an association.

One way to rationalize this finding is to notice that some of the key roles of brokers are more valuable for insured businesses compared to insured individuals. Among these roles, one can think of the assistance the broker provides the insured in identifying the required coverages, of the matching of the insured with the insurer based on the insurer's appetite, of the risk 'branding' of the insured helping insurers to circumvent informational asymmetries in the insurance market, of the assistance that the broker provides the insured in the claiming process, *etc.*[36]

One of the ways to justify the continued coexistence of both distribution channels, direct and broker distributed, in both Personal and Commercial Lines of business is to note that different customers have different ways to shop for insurance. Some insureds that have high search costs prefer to take advantage of brokers to "avoid searching" by themselves.[37]

2.1.6 Ignored dimensions: geographic concentration, reinsurance usage, ownership form

Before moving to the exploration of the cost structure in the Canadian P/C insurance industry, the author wishes to complete the review of the historical literature. The items noted here are items that the author would be willing to stipulate without seeking further evidence and thus be willing to keep the items as part of the research blind spot. Fortunately, work has been done to gather evidence to support the findings.

First, some authors have considered the effect of geographic concentration on the cost structure of American P/C insurers. For the moment, suffice it to note that insurers more geographically diversified in the USA tend to be significantly bigger insurers, as can be seen from the table below where the Herfindahl Index has been computed as $\sum_{i=1}^{S} m_i^2$, where S is the number of states/territories found in total Direct Written Premium exhibits for the period from 2002 to 2010, and m_i is the proportion of the total Direct Written Premium that the insurer writes in state/territory i. The period 2002 to 2010 was selected by way of convenience. The results are highly similar for any given chosen year.

[36] See (Regan and Tennyson, Agent Discretion and the Choice of Insurance Marketing System 1996), (Kim, Mayers and Smith 1996), (Regan, An Empirical Analysis of Property-Liability Insurance Distribution Systems: Market Shares Across Lines of Business 1998), and (Cummins and Doherty 2006).

[37] See (Posey and Yavas, A Search Model of Marketing Systems in Property-Liability Insurance 1995, 669).

Second, some authors have inquired about the reinsurance cost portion of the underwriting expense. Mayers and Smith (1990) find that bigger insurers tend to purchase less reinsurance. In a follow up study, Cole and McCullough (2006) find that the demand for domestic reinsurance decreases as the size of the insurer increases, but the demand for foreign reinsurance increases as the size of the insurer increases.

Finally, Regan and Tzeng (1999) found and justify that the ownership structure of the insurer is related to its distribution channel. In particular, they found that the stock owned insurers tend to more commonly associate with the broker distribution channel. They find that:

> [c]ontrolling for ownership form as an exogenous variable, the authors find that independent agency insurers are likely to be associated with stock ownership, are characterized by higher liabilities relative to surplus, and are more likely to specialize in complex lines of business. (...) When ownership form is treated as an endogenous variable, however, no significant relation exists between ownership form and distribution system. This suggests that these elements are related, but only indirectly through the effect of risk and complexity. (Regan and Tzeng, Organizational Form in the Property-Liability Insurance Industry 1999, 253)

In short, there is a substantial body of work that demonstrates that some characteristics of the insurer are correlated with features of insurers that are of interest to us here.

2.2 The Cost Structure of the Canadian P/C Insurance Industry

Kelly and Kleffner (2006) conducted a study similar to the studies documented and reproduced in section 2.1 for the Canadian industry, but found quite surprising results. In effect, they found that, in the Canadian insurance P/C industry, direct writers do not enjoy a cost efficiency advantage like they do in the American P/C insurance industry, and direct writers do not dominate Personal Lines, although insurers that distribute through brokers have lost some market shares in Personal Lines.

Extracted from Table 2 of (Kelly and Kleffner 2006, 57)					
	Canadian writers 1995-2003				
Mean	Entire sample	Multiple-channel writers	Commodity writers	Exclusive writers	Agency writers
UWE / NPW	35.59%	34.30%	36.31%	35.58%	35.68%
(UWE + LAE) / NPW	44.91%	44.13%	43.14%	43.70%	45.54%

Extracted from Table 1 of (Kelly and Kleffner 2006, 56)				
	Personal Lines		Commercial Lines	
	1995	2003	1995	2003
Multiple-channel writers	9.55%	6.65%	13.54%	9.80%
Exclusive writers	15.33%	16.64%	4.92%	6.79%
Agency writers	67.30%	63.75%	77.44%	76.52%
Commodity writers	7.82%	12.95%	4.10%	6.89%

They also extract other statistics. First, contrary to the American P/C insurance industry direct and broker insurers have fairly similar commission rates. Second, like in the USA direct writers invest more in Electronic Data Processing expenses than broker insurers. Third, just like in the USA, direct writers tend to write less complex business than broker insurers.

Extracted from Table 4 of (Kelly and Kleffner 2006, 65)						
	Canadian writers 1995-2003				U.S. Insurers 1980-1998	
	Multiple-channel writers	Commodity writers	Exclusive writers	Agency writers	Exclusive writers	Agency writers
Commissions / DPW	12.67%	10.73%	15.08%	15.94%		
Advertising Ratio	0.44%	29.34%	3.40%	0.52%	0.32%	0.14%
EDP Ratio	0.84%	6.10%	1.42%	0.95%	1.19%	1.01%
Complexity Ratio	46.51%	38.12%	30.26%	48.79%	16.73%	41.39%

This leads us to reconsider the validity of the theory that supported the rationalizations of the market structure in the USA, as these theories are equally valid for the Canadian market. The authors believe that the smaller scale of the Canadian P/C insurance landscape is the key to understanding the different industry structures between the two markets. They point particularly towards the relative size of the Automobile market that is smaller in Canada due to increased governmental presence, to the decreased efficiency of mass advertising, and to the decreased efficiency of investment in information technology. We will further explore these in subsequent sections using the NAIC data.

2.3 Ignored Collinearity

While the historical literature has recognized that there are economies of scale in the P/C insurance industry, it was never recognized to be the leading driver of the magnitude of the expense ratio. The hope of the author is to establish that economies of scale are the principal force that leads to a decreased underwriting expense ratio. If that is established, the author has to explain why some insurers get to be significantly bigger than others. In the mean time, let us reconsider the findings from sub-section 2.1. First, examining the table below, it is highly probable that the distribution channel is materially correlated with the size of the insurer, as the upper half of the CCCR distribution decreases as insurer size increases. While, on the lower half of the distribution, insurer size seems to increase the CCCR, it does so with smaller values, such that the net effect is an increased likelihood to be a direct insurer conditional on being a large insurer and *vice versa*. So, if a regression was conducted using both the distribution channel and insurer size as covariates, due to the collinearity of insurer size with the distribution channel, it would be unclear what is the marginal contribution of insurer size for the coefficient relating to the direct writer indicator variable.

Target Variable:	CCCR_minus_0									
Percentile	10th		25th		50th		75th		90th	
DWP Weights; S.E.: "ker" Method	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value
intercept	-20.569	0.0%	-8.530	6.8%	40.578	0.0%	62.600	0.0%	127.141	0.0%
log10_DWPt_onl_minus_0	2.416	0.0%	1.751	0.0%	-2.748	0.0%	-4.661	0.0%	-10.658	0.0%

Before moving to section 3, where we will attempt to establish the importance of economies of scale in the P/C insurance industry, let us examine the large impact that insurer size has on the underwriting expense ratio of P/C insurers.

Target Variable:	ExpR_minus_0									
Percentile	10th		25th		50th		75th		90th	
DWP Weights; S.E.: "ker" Method	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value
intercept	-0.343	91.1%	35.109	1.2%	62.365	0.0%	82.600	0.0%	129.270	0.0%
log10_DWPt_onl_minus_0	3.247	0.0%	0.078	95.8%	-2.227	0.0%	-3.849	0.0%	-8.141	0.0%

3. ECONOMIES OF SCALE IN THE P/C INSURANCE INDUSTRY

In this section, we will set up what we believe to be the key of the resolution of the puzzle: economies of scale. First, we will demonstrate why we believe economies of scale play a critical role in the structure of the P/C insurance industry. Second, we will discuss the potential sources of economies of scale. Third, we will explore what are potential drivers of size and/or growth. Finally, we will explore the consequences of growth.

3.1 Signs of the Presence of Economies of Scale

As is noted in the finance and accounting literatures[38], the seeking of operational synergies can be a driving force behind Mergers and Acquisitions[39]. The argument can be extended to industry consolidations and, as a matter of fact, the American P/C insurance has been the subject of a major consolidation in the last 20 years[40].

In the tables below, "H.I." denotes the Herfindahl Index of the American P/C insurance industry. The columns "t" and "t^2" refer to a quadratic parametric model that is fitted to the values of interest. Ordinary Least Squares was used to fit the quadratic model. OLS is sufficient here because we are only looking for a Best Linear Predictor and we are not attempting to provide any causal or structural interpretation for the parameters.

Total: Total USA P/C Industry

Top 5	t	t^2	Year	Actual	Predicted
β(·)	9.20E-04	2.85E-05	1992	30.9%	30.9%
se(·)	1.33E-03	5.92E-05	2001	30.6%	32.1%
t-value	0.69	0.48	2010	33.0%	33.7%
Top 10	t	t^2	Year	Actual	Predicted
β(·)	2.64E-03	9.18E-05	1992	42.8%	42.1%
se(·)	1.40E-03	6.22E-05	2001	44.8%	45.5%
t-value	1.88	1.48	2010	49.5%	50.5%
Top 20	t	t^2	Year	Actual	Predicted
β(·)	5.66E-03	-7.17E-05	1992	57.6%	55.9%
se(·)	1.97E-03	8.72E-05	2001	61.0%	60.2%
t-value	2.88	-0.82	2010	62.8%	63.3%
H.I.	t	t^2	Year	Actual	Predicted
β(·)	-9.61E-05	9.92E-06	1992	3.0%	3.1%
se(·)	1.39E-04	6.16E-06	2001	3.1%	3.1%
t-value	-0.69	1.61	2010	3.3%	3.3%

Auto: Automobile

Top 5	t	t^2	Year	Actual	Predicted
β(·)	-6.78E-03	4.02E-04	1992	42.6%	43.1%
se(·)	1.42E-03	6.31E-05	2001	42.6%	41.7%
t-value	-4.77	6.37	2010	47.0%	46.8%
Top 10	t	t^2	Year	Actual	Predicted
β(·)	4.78E-03	1.39E-04	1992	50.9%	50.5%
se(·)	1.36E-03	6.05E-05	2001	57.7%	56.4%
t-value	3.51	2.29	2010	64.3%	64.6%
Top 20	t	t^2	Year	Actual	Predicted
β(·)	8.86E-03	-5.05E-05	1992	62.3%	61.8%
se(·)	1.07E-03	4.75E-05	2001	70.2%	69.2%
t-value	8.27	-1.06	2010	75.1%	75.8%
H.I.	t	t^2	Year	Actual	Predicted
β(·)	-1.93E-03	8.00E-05	1992	6.0%	6.2%
se(·)	4.26E-04	1.89E-05	2001	5.7%	5.4%
t-value	-4.54	4.23	2010	6.0%	5.9%

CLNA: Commercial Lines Non-Auto

Top 5	t	t^2	Year	Actual	Predicted
β(·)	1.57E-02	-5.07E-04	1992	25.2%	24.5%
se(·)	2.75E-03	1.22E-04	2001	30.0%	32.7%
t-value	5.69	-4.16	2010	31.5%	32.6%
Top 10	t	t^2	Year	Actual	Predicted
β(·)	1.63E-02	-5.50E-04	1992	39.8%	39.3%
se(·)	2.44E-03	1.08E-04	2001	45.4%	47.5%
t-value	6.67	-5.07	2010	46.1%	46.9%
Top 20	t	t^2	Year	Actual	Predicted
β(·)	1.09E-02	-3.89E-04	1992	59.3%	56.6%
se(·)	3.12E-03	1.38E-04	2001	62.0%	61.9%
t-value	3.50	-2.81	2010	59.9%	60.9%
H.I.	t	t^2	Year	Actual	Predicted
β(·)	1.97E-03	-6.90E-05	1992	2.6%	2.3%
se(·)	5.02E-04	2.23E-05	2001	2.8%	3.3%
t-value	3.93	-3.10	2010	2.9%	3.1%

PLNA: Personal Lines Non-Auto

Top 5	t	t^2	Year	Actual	Predicted
β(·)	8.26E-03	-4.76E-04	1992	42.3%	42.7%
se(·)	3.14E-03	1.39E-04	2001	48.4%	44.5%
t-value	2.63	-3.41	2010	40.2%	38.7%
Top 10	t	t^2	Year	Actual	Predicted
β(·)	1.01E-02	-4.24E-04	1992	51.9%	51.8%
se(·)	3.15E-03	1.40E-04	2001	60.0%	55.9%
t-value	3.19	-3.03	2010	54.5%	53.2%
Top 20	t	t^2	Year	Actual	Predicted
β(·)	1.13E-02	-4.56E-04	1992	64.5%	64.1%
se(·)	2.46E-03	1.09E-04	2001	72.4%	68.9%
t-value	4.58	-4.18	2010	67.1%	66.3%
H.I.	t	t^2	Year	Actual	Predicted
β(·)	1.39E-04	-5.53E-05	1992	6.5%	6.7%
se(·)	8.80E-04	3.90E-05	2001	7.1%	6.2%
t-value	0.16	-1.42	2010	5.1%	4.8%

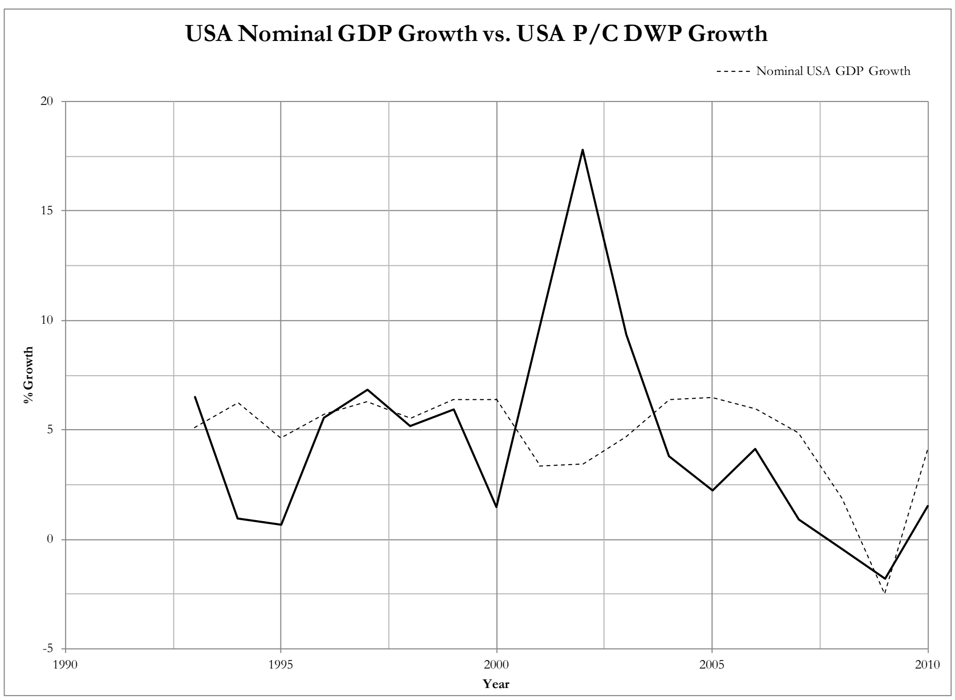
While it is apparent that the over American P/C industry has been consolidating in the last 20 years, it is also clear that the line of business that is the main source of the consolidation has been

---

[38] See, for example, (Grinblatt and Titman 2002, 699-701), (Palepu and Healy 2008, 11-1,11-2).
[39] See, for example, (Intact Financial Corporation 2011).
[40] See, for example, (Cummings n.d.).

the Automobile line of business. While the Commercial Lines Non-Auto has been the subject of some consolidation, it has been counter-balanced by the Personal Lines Non-Auto which has seen no increase in concentration and has also been increasing in importance in the last 20 years as a proportion of DWP[41].

The question then becomes one where we need to inquire about channels of growth for insurers. One possibility is that insurers are growing because the overall market is expanding. Based on the graph below, this is highly unlikely as the American P/C insurance industry seems saturated. Another source of growth could be from insurers forming a combine that forces prices up. While it is not easy to disprove that theory using archival data, like what is used here, the possibility will be rejected on the assumption that an insurance cartel would have likely lead to a 'major' class action against insurers and this class action has not been observed. We've already discussed the possibility of growth through Mergers and Acquisitions[42]. Unfortunately, it is not easy in the NAIC data to observe Mergers and Acquisitions activity. Finally, growth can occur organically. For example, this seems to be the current preferred growth channel of Progressive (The Progressive Corporation 2010).



Next, we will inquire about what are the potential sources of economies of scale.

[41] From about 10% of overall DWP in 1992 to about 20% in 2010.
[42] We are here mainly focusing on horizontal integration within the P/C insurance industry. Note however that vertical integration (*e.g.* insurers merging with/acquiring brokers, insurers forming strategic alliances with service providers, *etc.*) could also be considered.

3.2 Potential Sources of Economies of Scale

To better understand how large insurers can create a cost competitive advantage for themselves, we will explore some working hypotheses regarding economies of scale in the market for the manufacturing and distribution of financial products.

First, in the insurance industry, like in many financial sub-industries, the acquisition, processing, interpretation and usage of information is subject to economies of scale. Take the example of creating a report for a sub-portfolio and using the information discovered with the report to affect pricing strategy by implementing a rate change through the rating systems. In the considered example, the cost of labor required is quite possibly sub-proportional to the number of insureds in the sub-portfolio, while it is most probably an increasing function of the size of the portfolio.

Second, the viability of e-commerce investment in the financial sector is largely a function of the proportion of clients, or more generally affected stakeholders, that actually adopt the technology that saves costs to the financial firm. In the case of the P/C insurance industry, an investment in a Broker Management System or web quoting engine will likely only be a positive net present value project if brokers or clients adopt the technology. Like (Allen, Clark and Houde 2008) argue, less competitive markets and more dominant firms within the market tend to favor massive adoption of cost saving technology.

Third, a larger insurer can be in a much better position to influence prices in the market for repair goods, through the exercise of monopsony power. As (Nell, Richter and Schiller 2009, 350) note: "[t]aking the problems associated with incomplete insurance contracts into account, only institutional arrangements can increase welfare beyond a third-best situation. Especially the vertical integration of insurance and repair markets maybe an appropriate approach."

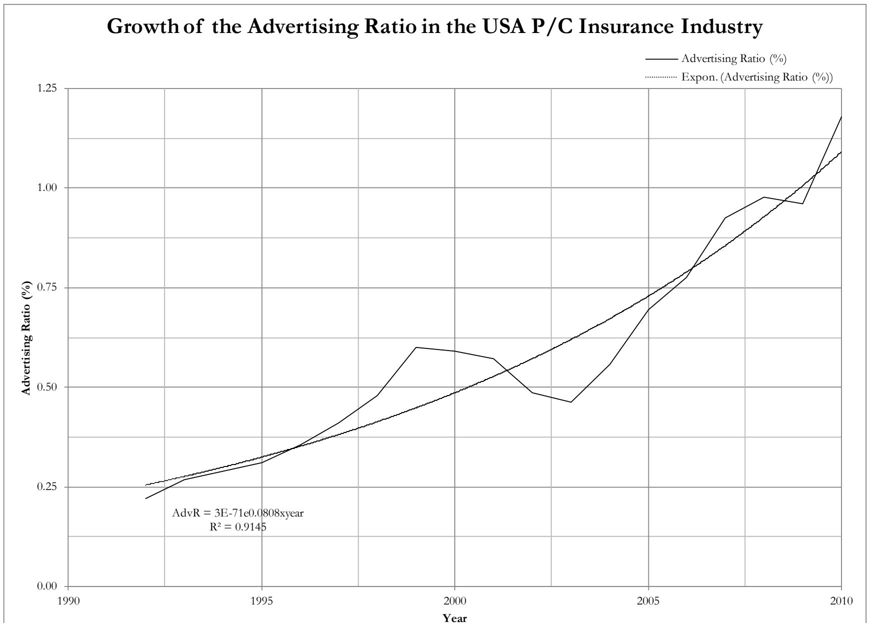
Fourth, more generally, insurer size may be associated with market power in the many markets insurers need to engage in, like the labor market.

Finally, as argued in (Intact Financial Corporation 2010, 6), larger insurers may be in a better position to form predictive models of consumer profitability.

Next, we will separately consider advertisement.

3.3 Determinations of Size or Growth

Before going further, let us examine the evolution of the advertising ratio in the American P/C industry in the last 20 years. Clearly, there has been a large positive trend of increased advertising expenditures. This large trend makes the comparison that (Kelly and Kleffner 2006) make in table 4 not as enlightening as they intended it to be, as the covered periods are long and do not overlap for Canada and USA



Mass advertising can be an effective tool for reaching a large number of persons at the same time, but it has the disadvantage that many people that see it may not have been the target audience. Also, mass advertising can become quite expensive[43]. Therefore, it is unclear, in an *a priori* way that advertising is subject to increased efficiency as insurer size increases even if it is likely that advertising effectiveness increases as the size of the advertising campaign increases, if it is executed appropriately within a marketing strategy coherent with the business strategy.

At this point, we empirically examine the quantile effect of advertising, controlling for the share of the insurer premium that is written in the Automobile lines of business[44]. As can be seen below, current period advertising seems to be positively correlated with DWP growth. It is somewhat surprising to find that prior year advertising is negatively correlated with current year DWP growth differential, but the advertising ratio should be correlated from one year to the next for most insurers. Note that the positive effect of advertising seems to stem from the upper half of the DWP growth differential distribution. Note also that it is unclear that a causal interpretation can be made of the result, because it could be that insurers that intend to pull out of a market decide to stop advertising in that market.

---

[43] See (Peter and Donnelly 2006, 111).

[44] Automobile insurance being the line of business most likely to see efficacious advertising, as products are standardized, purchased by a very large portion of the population, and competitive.

Target Variable:	growth_diff_minus_0									
Percentile	10th		25th		50th		75th		90th	
DWP Weights; S.E.: "ker" Method	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value
intercept	-26.126	67.5%	-31.787	89.5%	-156.036	0.0%	-69.572	0.0%	-27.986	0.0%
Auto_share_minus_0	0.279	78.4%	0.403	90.8%	2.141	0.0%	0.758	0.0%	0.340	0.2%
AdvR_minus_0	27.358	88.2%	58.240	91.7%	727.757	0.0%	727.126	0.0%	726.822	0.0%
Auto_share_minus_0 x AdvR_minus_0	-0.397	90.0%	-0.851	92.3%	-10.312	0.0%	-8.504	0.6%	-7.478	0.0%

Target Variable:	growth_diff_minus_0									
Percentile	10th		25th		50th		75th		90th	
DWP Weights; S.E.: "ker" Method	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value
intercept	-64.199	0.0%	-22.828	0.0%	-3.579	38.3%	6.137	4.5%	10.760	2.6%
Auto_share_minus_0	3.660	14.1%	2.606	31.0%	1.911	35.0%	-0.265	91.0%	-3.871	0.3%
AdvR_minus_0	733.364	0.0%	733.090	0.0%	732.497	0.0%	731.650	0.0%	730.093	0.0%
Auto_share_minus_0 x AdvR_minus_0	-10.057	0.0%	-10.531	0.0%	-10.533	0.0%	-9.001	0.0%	-7.685	0.0%
Auto_share_minus_1	-2.895	22.9%	-2.326	37.4%	-1.909	35.5%	0.125	95.8%	3.785	0.4%
AdvR_minus_1	-781.725	0.0%	-785.164	0.0%	-728.863	0.0%	-632.275	0.0%	-442.559	0.0%
Auto_share_minus_1 x AdvR_minus_1	10.189	0.0%	11.049	0.0%	10.496	0.0%	8.067	0.0%	4.777	0.0%

Before moving back towards the larger picture of business strategy, we will explore the potential effects of growth and investment in information technology of the Loss and Loss Adjustment Expense ratio side of underwriting profitability.

3.4 Consequences of Growth

As has been observed and justified in (D'Arcy and Doherty 1990), because new businesses tend to receive lowballed prices in a market where there are *ex ante* informational asymmetries, one could potentially expect that insurers that are growing rapidly will first experience a deteriorating loss ratio that would improve over time. But, as we mentioned earlier, a larger insurer may be able to generate economies of scale in loss adjustment expenses as well as in the loss ratio, by being able to negotiate better prices in the repair goods market. It is therefore an empirical matter of which force is strongest and the following tables attempt to answer that question.

Target Variable:	LLAER_diff_minus_0									
Percentile	10th		25th		50th		75th		90th	
DWP Weights; S.E.: "ker" Method	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value
intercept	-11.980	0.0%	-5.930	0.0%	-0.480	12.6%	4.670	0.0%	9.370	0.0%
growth_diff_minus_0	0.000	0.0%	0.000	0.0%	0.000	0.0%	0.000	0.0%	0.000	95.3%
growth_diff_minus_1	0.000	96.4%	0.000	20.5%	0.000	64.1%	0.000	38.1%	0.000	1.6%

As is observed in the preceding table, the DWP growth differential seems to have little net impact on the LLAER differential.

Target Variable:	LLAER_diff_minus_0									
Percentile	10th		25th		50th		75th		90th	
DWP Weights; S.E.: "ker" Method	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value
intercept	-11.919	0.0%	-5.891	0.0%	-0.351	60.0%	4.724	0.0%	9.443	0.0%
EquipR_minus_0	0.013	0.0%	-0.003	0.0%	-0.018	0.0%	-0.032	0.0%	-0.045	0.0%

The table found above was meant to serve the goal to explore the following working hypothesis: insurers that invest more heavily in information technology tend to be more sophisticated and disciplined than other insurers and are therefore experiencing a lower LLAER ratio than other insurers. Clearly, except for the 10th percentile, the working hypothesis is not infirmed by the data, although the economic significance of the coefficients associated with the $EquipR\_minus\_0$ term are not great.

4. BUSINESS STRATEGY

Before moving further, let's gather together the accumulated evidence. (1) The American P/C insurance industry has been consolidating in a material way in the last twenty years. (2) One way for the industry to consolidate is through Mergers and Acquisitions activities, and M&A activity has operational synergy as one of its key motivators. There is no *a priori* reason to believe that there is no similar motivation for insurers when they engage in organic growth; especially since we've identified potential sources of economies of scale. (3) When a one-way quantile regression of the underwriting expense ratio is run against P/C insurer size, larger insurers appear to have a lower expense ratio. (4) Insurer size is correlated with its distribution channel, as bigger American P/C insurer are more likely than average to distribute through the direct channel. (5) Provided that an insurer has chosen the broker distribution channel, Commercial Lines generally constitutes a larger portion of its book than it would otherwise be. Provided that an insurer has chosen the direct distribution channel, Automobile insurance generally constitutes a larger portion of its book than it would otherwise be. Both the Automobile and the Commercial lines of business have been consolidating; although the extent of consolidation has been much stronger in the Automobile lines of business. (6) Larger American P/C insurers tend to be much less geographically concentrated than average: their customer base is much larger and diversified geographically. (7) The Canadian P/C insurance market is not dominated by direct writers in Personal Lines insurance, that includes the Automobile line of business, and direct writers do not have an expense advantage. Let's add the following information. (8) In Canada, the P/C insurance has also been consolidating, but the consolidation has been lead by an insurer that mainly focuses on distributing its products through brokers[45].

The question can then be asked about what is the most likely dominant force leading to cost efficiency, given a business strategy. The historical answer from the P/C insurance efficiency literature, which was mainly written by Americans attempting to explain the structure of the American P/C insurance market, was that the distribution channel was the key driver of efficiency,

[45] "Proven acquisition strategy: We are an active acquirer in the industry, with 11 successful acquisitions since 1988. Our strategy focuses on fit, technological integration and increasing the profitability of the acquired book of business through our pricing, underwriting expertise and claims." (Intact Financial Corporation 2010, 6)

even while many noted that economies of scale were available. Assuming that the distribution channel was the driving force for efficiency, in Canada, direct writers should also be dominating in the sub-market in which they should naturally dominate: Automobile insurance. However, it is not the case. Plus, if the key force driving efficiency was the distribution channel, it would not provide a strong rationale for the material consolidation of the American P/C insurance industry.

If, instead, we suppose that economies of scale are the driving force behind cost efficiency in the P/C insurance industry, then (1) it is easy to rationalize the consolidation of the P/C insurance industry in Canada and in the USA and, (2) given that the distribution channel then becomes a secondary force, it not surprising to find that, in Canada, broker insurers do not have an expense disadvantage over direct writers, but that direct writers have nonetheless been gaining market shares.

Under this alternate rationalization, what has instead to be explained is why, in the USA, large P/C insurers are quite likely to choose a generic strategy of cost leadership while, in Canada, large P/C insurers are more likely to choose a generic strategy of differentiation?

Why do we say that large American insurers tend to prefer a cost leadership strategy? Cost leadership can be defined as "an integrated set of actions designed to produce or deliver goods or services with features that are acceptable to customers at the lowest cost, relative to that of competitors." (Hitt, et al. 2006, 147) The very motivation behind the direct distribution channel finds its roots in cost minimization. Historically, it has been expressed as direct insurers taking care of billing. More recently, it has expressed itself in large direct insurers pursuing initiatives related to usage of internet in the distribution of their products. Some of the cost savings technologies can have significant fixed costs associated with them and massive adoption of the technology can be a critical factor for success.

Why do we say that large Canadian insurers tend to prefer differentiation? Differentiation can be defined as a strategy designed "to produce or deliver goods or services (at an acceptable cost) that customers perceive as being different in ways that are important to them." (Hitt, et al. 2006, 153) Using Intact Financial Corporation as an example, we can see that the insurer intends to (1) be supportive of its broker sales force to provide clients with "customer choice, personalized service and trusted advice" (Intact Financial Corporation 2010, 6), (2) offer clients the choice of which distribution channel to use to approach the insurer, (3) offer superior claims service, and (4) use its scale advantage "to negotiate preferred terms with suppliers, priority repair service, quality guarantees and lower material costs." (Intact Financial Corporation 2010, 6) Similar examples could be found for other large Canadian P/C insurers.

Under both these generic business strategies[46], large insurer size is (1) possible and (2) useful. The way insurer size is used differs under the differentiation and the cost leadership strategies differ: under cost leadership, insurer size is used to channel economies of scale in reduced prices leading to further growth; under differentiation, insurer size is used to allow the insurer to offer more differentiating features (because the consumer pool increases) while not having prices explode (because of economies of scale).

It is worthy to note that, under both the differentiation and the cost leadership strategies, mass advertising and investment in information technology are sensible because, under both generic strategies, economies of scale help render the strategy more effective and efficient. Assuming that a properly strategized and executed advertising campaign actually favors growth, advertising helps insurers create economies of scale. Also, we saw that investment in information technology is likely to be associated with sophistication in the costing and pricing of insurance contracts, and pricing sophistication is necessary under both differentiation and cost leadership.

We can formulate two working hypotheses for why the American and the Canadian P/C insurance markets have evolved differently. As noted in (Kelly and Kleffner 2006, 66), in Canada, available premium in the Automobile line of business, historically favored by direct writers, is much smaller than in the USA because (1) the population is much smaller to start with, but also because (2) Automobile insurance is handled, at least in part, by government insurers in many provinces. As available economies of scale for direct writers are less important, it did not favor the growth of that distribution channel. Another working hypothesis would say that broker insurers in Canada found itself facing an insurance brokerage industry that was not as concentrated as in the USA and was therefore better able to embark brokers in the use of cost saving technology. The motivation for the second working hypothesis stems from noting that the American insurance brokerage is quite concentrated[47], and from noting that some large Canadian insurers work quite intensively with brokers to help them in their endeavors. Supporting evidence needs to be sought to support both working hypotheses.

5. CONCLUSION

We've identified two fatal flaws of the historical literature concerning itself with the cost efficiency P/C insurance market. One, we've identified that the historical literature has neglected the effect of collinearity when interpreting the results of regressions relating to the drivers of the underwriting expense ratio. Two, we've identified that the historical literature has neglected the

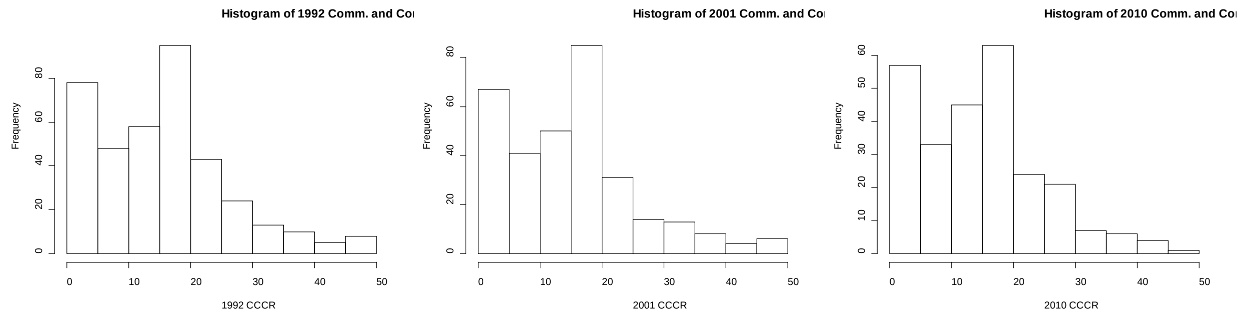
46 Generically speaking, both differentiation and cost leadership can be distinguished from a focus strategy where the firm focuses on "the needs of a particular competitive segment." (Hitt, et al. 2006, 159)
47 See (Cummins and Doherty 2006, 363-367).

possible effectiveness of the differentiation generic business strategy in the P/C insurance market. So doing, we've been lead to place economies of scale at the heart of a successful business strategy for insurers that do not choose the focus generic business strategy; thus, displacing the choice of distribution channel as subordinate to the choice of the generic business strategy.

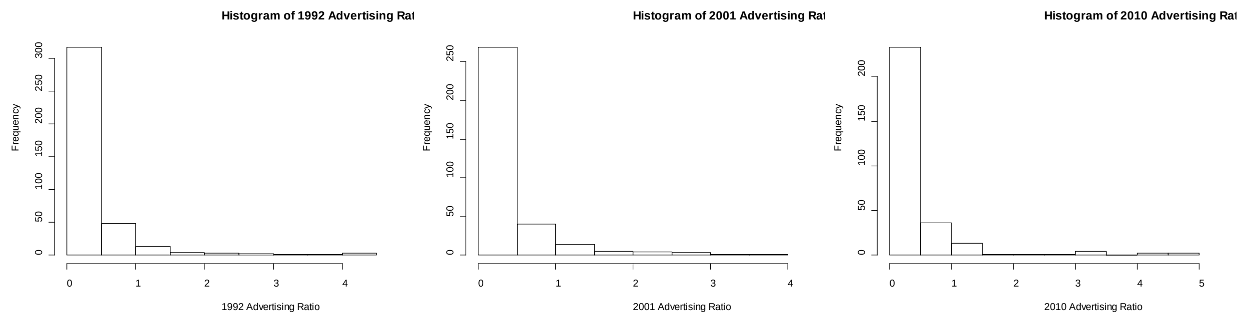
Acknowledgements The author wishes to thank Professors Mark Browne, Justin Sydnor, Anne Kleffner and Doctor Martin Halek for their support in the development of this paper.

Appendix A: Histograms of other key variables

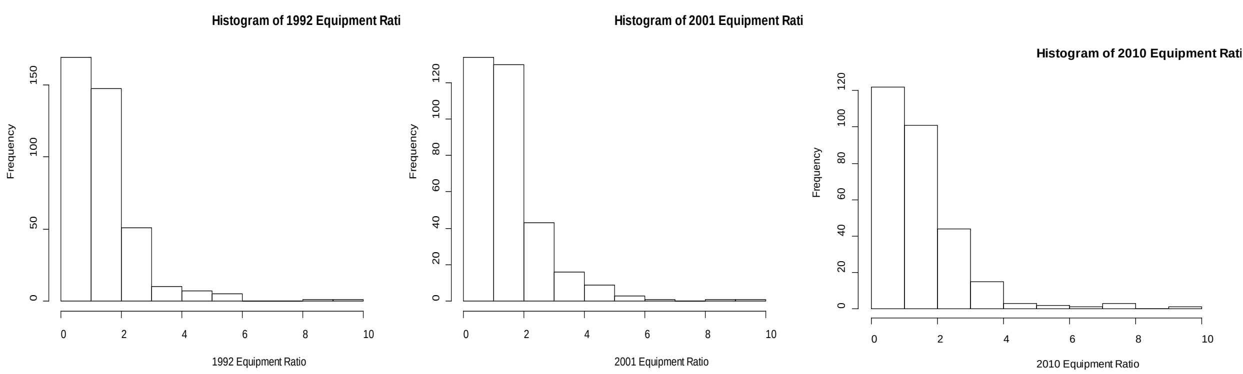
Commission and Contingent Commission Ratio



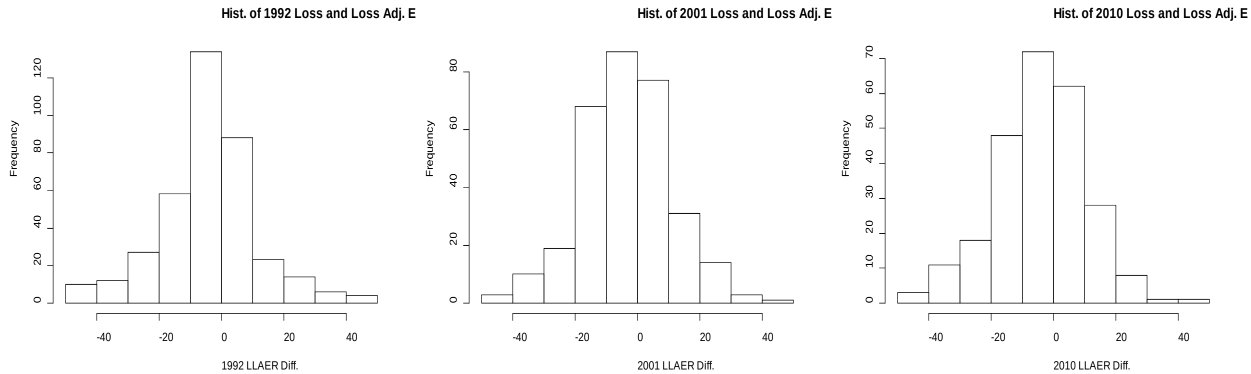
Advertising Ratio



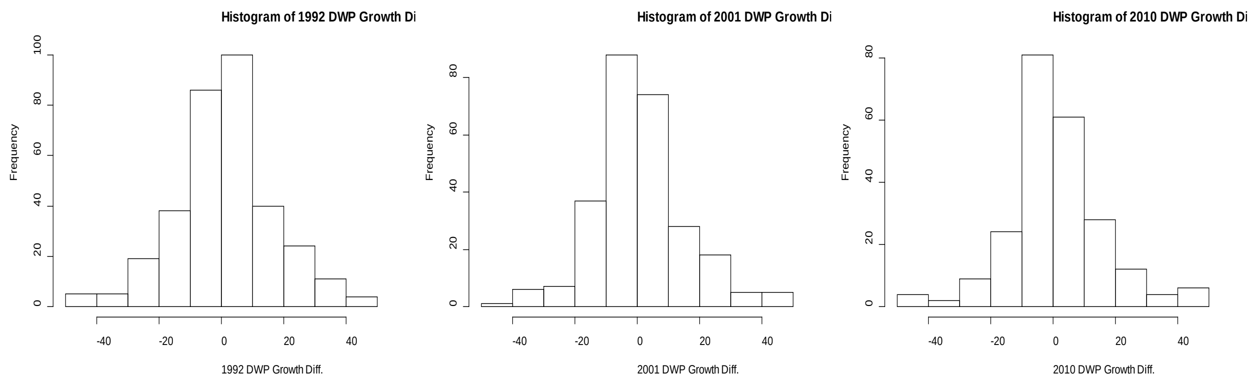
Equipment Ratio



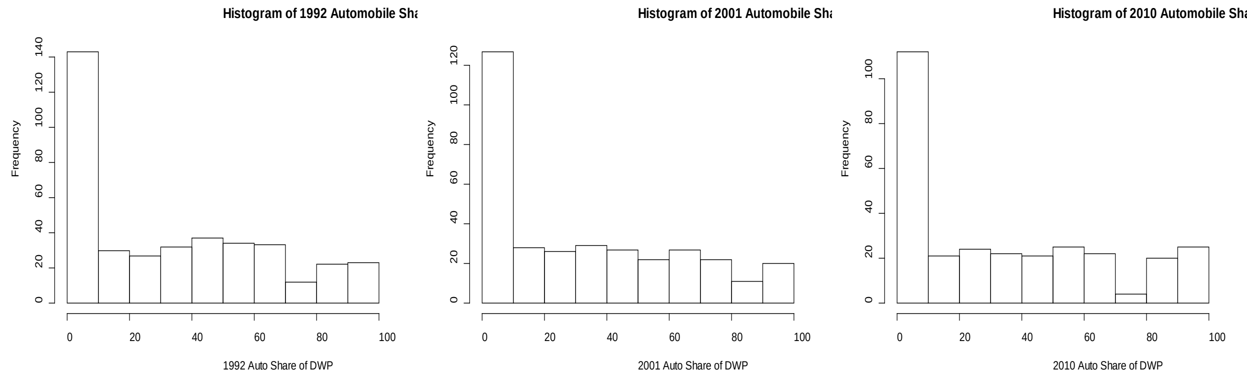
Loss and Loss Adjustment Ratio Differential to the Industry



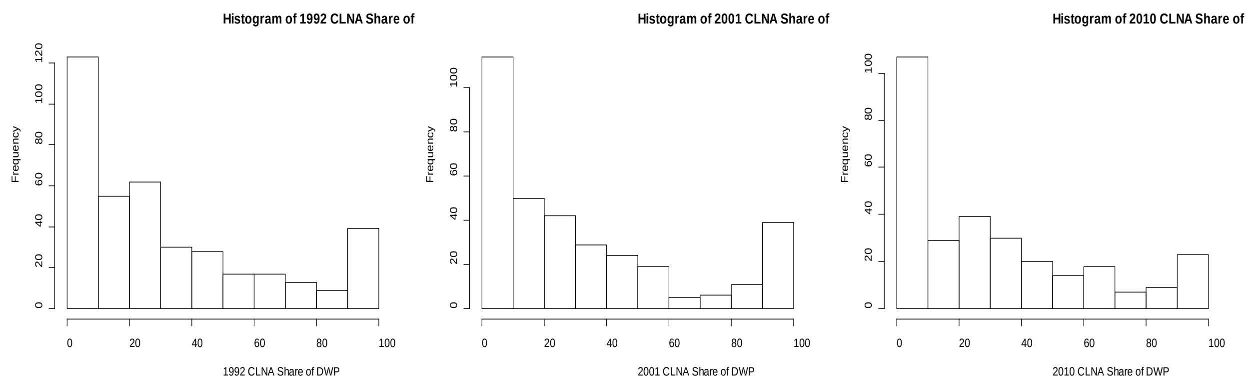
Direct Written Premium Growth Differential to the Industry



Automobile Share of Premium



Commercial Lines Non-Auto Share of Premium



Appendix B: Other quantile regression results

Target Variable:	CLNA_share_minus_0									
Percentile	10th		25th		50th		75th		90th	
DWP Weights; S.E.: "ker" Method	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value
intercept	2.283	0.0%	5.563	0.0%	22.937	0.0%	43.585	0.0%	55.779	0.0%
CCCR_minus_0	0.000	0.0%	0.000	0.0%	-0.001	0.0%	-0.001	0.0%	-0.002	0.0%

Target Variable:	Auto_share_minus_0									
Percentile	10th		25th		50th		75th		90th	
DWP Weights; S.E.: "ker" Method	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value	Coefficients	p-value
intercept	5.469	0.0%	18.988	0.0%	41.609	0.0%	67.454	0.0%	76.172	0.0%
CCCR_minus_0	0.000	0.0%	-0.001	0.0%	-0.001	0.0%	-0.002	0.0%	-0.002	0.0%

Appendix C: Quantiles in the univariate case

To better understand the second and third approach to quantile regression mentioned in section 2.1.2 "Econometric specification: the choice of quantile regression", we will recall how to compute quantile in the univariate case. Let us focus on the median. There are three ways to compute the median. One, one can plot the Cumulative Distribution Function of a random variable and find the point x for which $F_X(x) = 0.50$. Two, one could find the point m for which the quantity

$$\int_{-\infty}^{+\infty} |x - m| dF_X(x)$$

, or the absolute deviation from m, is minimized. Third, one could solve the following equation for m:

$$\int_{-\infty}^{+\infty} (I\{x \leq m\} - 0.5) dF_X(x) = 0$$

6. REFERENCES

A.M. Best Company. *Best's Aggregates & Averages: Property-Liability.* Oldwick, NJ: A. M. Best Co., 1993-2011.

Allen, Jason, Robert Clark, and Jean-François Houde. "Market Structure and the Diffusion of E-Commerce: Evidence from the Retail Banking Industry." *Bank of Canada Working Paper*, September 2008: 43.

Allen, jason, Robert Clark, and Jean-François Houde. "Price negotiation in differentiated product markets: An analysis of the market for insured mortgages in Canada." 02 2010: 19.

Barrese, James, and Jack M. Nelson. "Independent and Exclusive Agency Insurers: A Reexamination of the Cost Differential." *Journal of Risk and Insurance* 59, no. 3 (1992): 375-397.

Berger, Allen N., J. David Cummins, and Mary A. Weiss. "The Coexistence of Multiple Distribution Systems for Financial Services: The Case of Property-Liability Insurance." *Journal of Business* 70, no. 4 (October 1997): 515-546.

Brown, Jeffrey R., and Austan Goolsbee. "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry." *Journal of Political Economy* 110, no. 3 (2002): 481-507.

Buchinsky, M. "Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression." *Econometrica* 62 (1994): 405–458.

Bureau of Economic Analysis. *Current-Dollar and "Real" Gross Domestic Product.* Washington, DC, 11 22, 2011.

Bureau of Labor Statistics. *Consumer Price Index.* Washington, DC, 11 16, 2011.

Cole, Cassandra R., and Kathleen A. McCullough. "A Reexamination of the Corporate Demand for Reinsurance." *Journal of Risk and Insurance* 73, no. 1 (2006): 169-192.

Cummings, David. "Pricing Strategy and Risk Management."

Cummins, J. D., and N. A. Doherty. "The Economics of Insurance Intermediaries." *Journal of Risk and Insurance* 73, no. 3 (September 2006): 359-396.

Cummins, J. David, and Jack VanDerhei. "A Note on the Relative Efficiency of Property-Liability Insurance Distribution Systems." *Bell Journal of Economics* 10, no. 2 (Autumn 1979): 708-719.

D'Arcy, Stephen P., and Neil A. Doherty. "Adverse Selection, Private Information, and Lowballing in Insurance Markets." *Journal of Business* 63, no. 2 (April 1990): 145-164.

Feldblum, Sholom. "Personal Automobile Premiums: An Asset Share Pricing Approach for Property/Casualty Insurance." *Proceedings of the Casualty Actuarial Society* LXXXIII, no. 159 (November 1996): 190-296.

Green, Meg. "Accelerated Growth." *Best's Review*, May 2004: 45-52.

Grinblatt, Mark, and Sheridan Titman. *Financial Markets and Corporate Strategy.* 2nd. New York, NY: McGraw-Hill, 2002.

Gron, A. "Compensation and Industry Profitability: Evidence from the Property-Casualty Insurance Industry." *Journal of Business*, no. 71 (1998): 407-437.

Hansen, Bruce. *Econometrics.* Madison, Wisconsin, 2011.

Hecht, Jason. "Modeling Market Shares of the Leading Personal Automobile Insurance Companies." *Journal of Economics and Business* 51, no. 3 (May-June 1999): 279-296.

Hitt, Michael A., R. Duane Ireland, Robert E. Hoskisson, W. Glenn Rowe, and Jerry P. Sheppard. *Strategic Management: Competitiveness and Globalization.* Second Canadian. Toronto, ON: Nelson, 2006.

Intact Financial Corporation. "Annual Report 2010." 2010.

—. "Investor Presentation." August 2011.

Kelly, Mary, and Anne Kleffner. "The Distribution of Property/Liability Insurance in Canada: Costs and Market Structure." *Journal of Insurance Issues*, Spring 2006: 51-70.

Kim, Woo-Joong, David Mayers, and Clifford W., Jr. Smith. "On the Choice of Insurance Distribution Systems." *Journal of Risk and Insurance* 63, no. 2 (June 1996): 207-227.

Koenker, Roger, and Gilbert Jr. Bassett. "Regression Quantiles." *Econometrica* 46, no. 1 (January 1978): 33-50.

Koenker, Roger, and Jose A. F. Machado. "Goodness of Fit and Related Inference Processes for Quantile Regression." *Journal of the American Statistical Association* 94, no. 448 (December 1999): 1296-1310.

Mayers, David, and Cilfford W. Jr. Smith. "On the Corporate Demand for Insurance: Evidence from the Reinsurance Market." *Journal of Business* 63, no. 1 (January 1990): 19-40.

Nell, Martin, Andreas Richter, and Jorg Schiller. "When Prices Hardly Matter: Incomplete Insurance Contracts and Markets for Repair Goods." *European Economic Review* 53, no. 3 (April 2009): 343-354.

Newey, Whitney K., and James L. Powell. "Efficient Estimation of Linear and Type I Censored Regression under Conditional Quantile Restrictions." *Econometric Theory* 6, no. 3 (September 1990): 295-317.

Palepu, Krishna G., and Paul M. Healy. *Business Analysis & Valuation Using Financial Statements.* 4th. Mason, OH: Thomson South-Western, 2008.

Peter, J. Paul, and James H., Jr. Donnelly. *A Preface to Marketing Management.* 10th. New York, NY: McGraw-Hill/Irwin, 2006.

Posey, Lisa L., and Abdullah Yavas. "A Search Model of Marketing Systems in Property-Liability Insurance." *Journal of Risk and Insurance* 62, no. 4 (December 1995): 666-689.
Posey, Lisa L., and Sharon Tennyson. "The Coexistence of Distribution Systems under Price Search: Theory and Some Evidence from Insurance." *Journal of Economic Behavior and Organization* 35, no. 1 (1998): 95-115.
R Documentation.
Regan, Laureen. "An Empirical Analysis of Property-Liability Insurance Distribution Systems: Market Shares Across Lines of Business." *Journal of Insurance Issues* 21, no. 2 (1998): 151-171.
Regan, Laureen, and Larry Y. Tzeng. "Organizational Form in the Property-Liability Insurance Industry." *Journal of Risk and Insurance* 66, no. 2 (June 1999): 253-273.
Regan, Laureen, and Sharon Tennyson. "Agent Discretion and the Choice of Insurance Marketing System." *Journal of Law and Economics* 39, no. 2 (October 1996): 637-666.
Shi, Peng, and Edward W. Frees. "Long-tail Longitudinal Modeling of Insurance Company Expenses." *Insurance: Mathematics and Economics* 47 (2010): 303-314.
Skogh, Goran. "Returns to Scale in the Swedish Property-Liability Insurance Industry." *Journal of Risk and Insurance* 49, no. 2 (June 1982): 218-228.
The Progressive Corporation. "2010 Annual Report." 2010.

Biography of the Author

Marc-André Desrosiers is a Ph.D. candidate at UW-Madison in the Actuarial Science, Risk Management and Insurance program. He also completed his MBA at University of Calgary, after receiving his FCAS. Marc-André has studied Actuarial Mathematics and Philosophy at Concordia University, Montréal. The author also keeps contact with the industry as he is currently working as an external consultant for Intact Financial Corporation Actuarial Commercial Lines department. He is interested in pricing optimization, behavioral economics, customer behavior, and demand modeling. He can be joined at mdesrosiers@bus.wisc.edu.

An Experience Rating Approach to Insurer Projected Loss Ratios

Marc-André Desrosiers, Ph.D. Candidate, FCAS, MBA, B.A.

Abstract: The traditional approach to Property/Casualty rate indications starts with a methodology that uses internal data to forecast the Ultimate Loss Ratio, with losses making up about half of the expenses. For parties that are external to the insurer, this approach to forecasting a key component of future profitability is impractical as they generally do not have access to the necessary data. Using publicly available information, that is, the National Association of Insurance Commissioners Schedule P of the statutory financial statements from 1992 to 2010, we develop by line of business forecasts of the relativity to the industry Loss Ratio. To develop these forecasts, we use a weighted regression methodology that incorporates key ideas from fixed-effects regression, instrumental variables regression, credibility theory, as well as a flexible covariance structure for the residuals. Results indicate that the proposed approach of using lagged relativities from insurer own and other lines of business can provide adequate fits for many lines of business and for the combined results of the insurer as a whole.

Keywords. Experience Rating, Panel Data, Fixed-Effects Regression, Instrumental Variable Regression, Credibility Theory

1. INTRODUCTION

The traditional approach to Property/Casualty rate indications[1] (Werner and Modlin 2010) starts with a methodology that uses internal data to forecast the Ultimate Loss Ratio, with losses making up about half of the expenses. For parties that are external to the insurer, this approach to forecasting a key component of future profitability is impractical as they generally do not have access to the necessary data. External parties that are tasked with solvency surveillance, stock pricing, bond pricing, reinsurance underwriting, *etc*. need a Loss Ratio forecasting approach that relies on publicly available data. Even for the internal actuaries, using an alternate forecasting method can provide the actuary with a point of comparison that can supplement and complement forecasts supported by internal data.

Using publicly available information, that is, the National Association of Insurance Commissioners Schedule P of the statutory financial statements from 1992 to 2010, we develop by line of business forecasts of the relativity[2] to the industry Loss Ratio. To develop these forecasts, we

[1] Rate indications refers to approaches to the overall costing of a P/C insurance portfolio that rely mostly on the insurers own premium/exposure and loss data. Rate indications can be done using the Loss Ratio approach, where past LR are adjusted to be at the level of when the matching rates would be in-force, averaged out and compared with a Permissible Loss Ratio to attain a given level of profitability, or using the Loss Cost approach, where past insurance unit cost are adjusted to be at the level of when the matching rates would be in-force, average out and inflated for expected fixed and variable expenses.

[2] Relativity is a commonly used actuarial measure where a value of interest is compared to the same value of interest but for a larger set. For example, in ratemaking, it is common practice to breakdown manual rates into base rates and relativities. The said relativities can be calibrated by comparing the

use a weighted regression methodology that incorporates key ideas from fixed-effects regression, instrumental variables regression, credibility theory, as well as a flexible covariance structure for the residuals. From fixed-effects regression (Frees, Longitudinal and panel data: analysis and applications in the social sciences 2004, 51), we borrow the idea that the forecasts incorporate a (weighted) average of past results. From instrumental variables regression (Frees, Meyers and Cummings, Predictive Modeling of Multi-Peril Homeowners 2011, 3), we borrow the idea that other lines of business can share result-drivers in common, like similar strategies, similar clients or similar perils. From credibility theory, we borrow the idea that the experience rating values vary with the size of the individual. We also use a Toeplitz, or Moving Average, intra-insurer/line of business structure for the residuals over time (Frees, Longitudinal and panel data: analysis and applications in the social sciences 2004, 281).

Given that "[e]xperience rating recognizes the differences among individuals (…) by comparing the experience of individual (…) with the average (…) in the same classification" (National Council on Compensation Insurance 2007, R2), the proposed modeling approach can be thought of a form of experience rating. In line with more traditional experience rating methodologies, the forecasted relativities can be thought of as a modifier to a base rate, which is here the forecast of the by line industry Loss Ratio. These forecasts can reflect outlooks concerning the economy as a whole, the softness/hardness of the market, *etc*. We do not address the issue of how to forecast the state of the P/C industry market [by line of business] as a whole and instead presume that parties that may wish to follow our approach have developed an expertise in making these types of forecasts[3].

Contrary to the traditional use made of experience rating, our approach is not aimed at increasing incentive alignment between an insured and an insurer, decreasing the potential for adverse selection, or increasing fairness (Venter 1987, 1-2); instead, the main goal that our approach shares with traditional experience rating is predictive accuracy. These differences in goals make it such that, while we will have the chance to comment on modeling choices that also have to be made when calibrating an experience rating scheme, we will not comment on the potential micro-economic

actual Loss Ratio for a given value of a rating variable, in the numerator, to the overall actual Loss Ratio across all values of the variable, in the denominator.

[3] The author does not have specific expertise on that topic; nonetheless, the Loss Ratio projection methodology of the Loss Ratio approach to rate indications should be applicable to the industry as a whole, as long as the user can make assumptions about the future rate changes of the P/C insurance industry as whole, as well as future catastrophic loss activity.

importance of experience rating, like the rate at which insureds and insurers learn about the underlying riskiness of the insureds, the self-censoring of losses, and moral hazard avoidance.

Results indicate that the proposed approach of using lagged relativities from insurer own and other lines of business can provide adequate fits for many lines of business and for the combined results of the insurer as a whole. For solvency surveillance usage, we recommend that a regulator or a rating agency supplement the model with measured rate changes so as to better anticipate large changes in the Loss Ratio than are not due to smooth changes.

The rest of the paper will go as follows: section 2 will cover a short history of the actuarial development of experience rating, section 3 will cover a summarized version of elements that are normally included in an experience rating plan, section 4 will cover the modern statistical foundation of experience rating, section 5 will describe the data that was used for our current analysis, section 5.2 will cover the descriptive statistics, section 6 will cover the statistical analysis as such, including model selection and fit analysis, and section 7 will look back at practical choices that need to be made to calibrate an experience rating plan and we will be able to comment how our modeling choices can apply to such an exercise.

2. ACTUARIAL HISTORY OF EXPERIENCE RATING

Experience rating has been at the heart of Property/Casualty actuarial science ever since P/C actuarial science has developed has a separate sub-field of actuarial science. Early on, the foundation of what will come to be known as American credibility was developed by Mowbray (How Extensive a Payroll Exposure is Necessary to Give a Dependable Pure Premium 1914) who was attempting to answer the question of just how large an insured needed to be to generate, without using data related to other insureds, a forecast of future losses that had a given level of precision. To this day, P/C actuaries around the world know of the 1082 claims for full credibility rule-of-thumb (Hansen 1972) that can be derived using this approach.

As early as 1918, Whitney (The Theory of Experience Rating) used Bayesian and approximation arguments to derive the $Pn/(Pn+K)$ formula for credibility (Whitney 1918, 288), which is reminiscent of the traditional one-way random-effects analysis of variance models (Frees, Longitudinal and panel data: analysis and applications in the social sciences 2004, 126). This formula is still at the heart of many experience rating plans today (Gillam and Snader, Fundamentals of Individual Risk Rating, Part I 1992, 1-4).

In his 1934 Casualty Actuarial Society Presidential address, Dorweiler (A Survey of Risk Credibility in Experience Rating) presented the rating plan performance principle that was to become the foundation of what is known as the quintile test (Couret and Venter 2008, 82).

> A necessary condition for proper credibility is that the credit risks and debit risks equally reproduce the permissible loss ratio. Also, if the proper credibility has been attained, each sub-group of the credit and debit risks, provided it has adequate volume, should give the permissible loss ratio. While these conditions are necessary for a proper credibility of the experience rating plan, it does not follow that they are also sufficient. For a sufficient condition it would be required to establish that the risks within a group cannot be subdivided on any experience basis so as to give different loss ratios for the subdivisions, assuming the latter have adequate volume. (Dorweiler 1934, 100)

In 1959, Bailey and Simon (An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car) demonstrated[4] that experience rating was also pertinent for lines of business other than Workers' Compensation. Even when risks are fairly homogeneous to start with, the claiming history of an individual insured has predictive value that allows for rating that is more precise than that implied by classification rates.

The American history of credibility theory was complemented by what is sometimes called European credibility, as exemplified by the developments of Bühlmann, Bühlmann-Straub, Hachemeister, Jewell, (Frees 2004, 155) Dannenburg, and Goulet (Goulet 2001, 205-206). As is demonstrated by (Goulet 2001, 207), European credibility formulas can be interpreted as Best Linear Unbiased Predictors. As such, what is known as European credibility can be thought of as theoretical and practical developments that paralleled those made in North America by the econometrician Goldberger and associates (Frees 2004, 130).

3. CONCRETE EXAMPLES OF EXPERIENCE RATING PLANS

One of our aims with this proposal is to address practical modeling choices that would need to be made in the calibration of a more traditional experience rating algorithm; therefore, before going any further, we'll discuss elements that are traditionally included in an experience rating plan. For readers that are not already familiar with experience rating plans or with the material presented in the

[4] Using a non-parametric approach.

Advanced Ratemaking exam of the Casualty Actuarial Society, this section can safely be foregone at a first reading.

We will focus on the National Council on Compensation Insurance Workers' Compensation experience rating plan (2007) and on the Insurance Services Office's Commercial General Liability experience rating plan (2006). So doing, we won't be directly addressing other types of experience rating plans like driving records in Personal Automobile, fleet rating in Commercial Automobile, claims rating in Property insurance, *etc.*

A starting but key element of any experience rating plan is the definition of what counts as an 'individual' under the plan. Generally speaking, an 'individual' will be an insured, but there can be exceptions. For example, under the NCCI plan, an entity is defined with reference to ownership rules (R13) while, under the ISO plan, the definition of risk also refers to considerations relating to franchising (1).

Another key strategic rating consideration is the number of years of experience used. This can affect the way that the rating information is accumulated. Depending on the distribution channel used (*e.g.* direct or brokerage), the number of years of experience considered can also affect the burden put on parties involved in the distribution of insurance, especially if the used plan differs from industry standards. Under the NCCI plan, up to about four years of experience can be used (R10-R11) while, under the ISO plan, up to three years of experience are used (1). The use of the optimal quantity of experience implies that the plan must include rules about how to deal with the experience with other insurers: for example, under the NCCI plan, experience with other insurers can be included but is subject to verification (R11) and, under the ISO plan, special rules are formulated to deal with the fact that losses that occurred with another insurer are not revalued (10-11).

Properly actuarial elements also need to be grounded in rules. In particular, the losses and premium need to be put on-level[5] to ensure the comparability of the experience from multiple periods; therefore, commonly addressed elements include loss development and trends. Under the NCCI plan, the losses are extracted from the appropriate statistical plans (R6) while, under the ISO plan, factors are specifically provided to develop and detrend the losses (12-13). To properly address

[5] That is, in dollars of the forecasted-to period. The first rule that makes the experience on-level under a Loss Ratio based experience rating plan is the use of premium set at current rates in the denominator of the Loss Ratio.

the predictiveness of large losses, the plan can specify rules relating to the capping of losses and provide a way to compute an Expected Loss Ratio that covers only the lower layer of losses. For example, under the ISO plan, losses are capped at a rule-determined Maximum Single Loss (3) and the capped Loss Ratio is compared to an Expected Experience Ratio that reflects losses that are expected under the MSL. Under the NCCI plan, the actual experience of large losses is partially reflected in the rating modification m to manual rates: in this case,

$$m = \frac{\mathrm{A_p} + z_\mathrm{e}\mathrm{A_e} + (1 - z_\mathrm{e})\mathrm{E_e} + \mathrm{B}}{\mathrm{E_p} + z_\mathrm{e}\mathrm{E_e} + (1 - z_\mathrm{e})\mathrm{E_e} + \mathrm{B}}, \text{ where}$$

$\mathrm{A_p}$ refers to actual primary losses,
$\mathrm{A_e}$ refers to actual excess losses,
$\mathrm{E_p}$ refers to expected primary losses,
$\mathrm{E_e}$ refers to expected excess losses,
z_e refers to the credibility of actual excess losses, and
B is a ballast value. (R10)[6].

Other rules that can be included in an experience plan can include: rules relating to types of policies (*e.g.* rules to convert the experience of claims-made and occurrence-based[7] Commercial General Liability policies that have different development patterns (11-12)), schedule rating that relates to softer characteristics of the risk that may not be fully reflected in the experience as such (9), and rules relating to corrections of previously available information (R17).

Even though the context in which we want to apply the experience rating framework is different from a traditional experience rating application, it is our hope that we can comment on modeling choices that would be encountered in the calibration of a traditional experience rating plan. In particular, we hope to address how to handle the selection of the number of years of experience and the development of losses.

4. STATISTICAL FOUNDATIONS

The purpose of the section is to familiarize practicing actuaries with the statistical methods and hypothesis that will underlie our proposed models. Readers familiar with modern statistical

[6] The value that comes after the $(1 - z.)$ term is commonly called a complement of credibility. (Boor 1996) has documented commonly used complements of credibility. Moreover, at pp.36-37, he shows how to determine the optimal credibility weight as a function of the correlation between two unbiased estimators of the same parameter. This result can also be proven using a Generalized Method of Moments approach.

[7] More on that topic below.

techniques or readers that are mainly interested in the data and the results can safely skip this section at a first reading.

As was demonstrated in (Frees, Young and Luo, A Longitudinal Data Analysis Interpretation of Credibility Models 1999), many credibility models, like those that are used in experience rating, can be interpreted in terms of estimation in a longitudinal data context. It is not uncommon for credibility theory to be cast in terms of random-effects models. For example, one could write a model the following way:

$$y_{i,t} = \boldsymbol{x}_{i,t}\boldsymbol{\beta} + \boldsymbol{z}_{i,t}\boldsymbol{\alpha}_i + \varepsilon_{i,t} \text{ for } 1 \le i \le n, 1 \le t \le T \text{where } E\left[\varepsilon_{i,t}\right] = 0, V[\boldsymbol{\varepsilon}_i] = \boldsymbol{R}_i,$$

$$E[\boldsymbol{\alpha}_i] = \boldsymbol{0} \text{ and } V[\boldsymbol{\alpha}_i] = \boldsymbol{D} \text{ with } \boldsymbol{\alpha}_i \text{ independent and identically distributed.}$$

One possible way to interpret this model is to think of it as a mixture of fixed-effects $\boldsymbol{x}_{i,t}\boldsymbol{\beta}$ due to the observable variables $\boldsymbol{x}_{i,t}$ and random-effects due to unobserved individual heterogeneity $\boldsymbol{z}_{i,t}\boldsymbol{\alpha}_i$. Take Private Passenger Automobile insurance as an example. In this case, we can imagine that each driver is receiving a random draw that fixes the individual's 'driving abilities'. We then assume that this 'driving ability' is not directly observable but remains constant through time. Observing drivers that are consistently better/worse than average, we can infer that it is likely that these drivers were given better/worse driving ability draws. In effect, the unobserved 'driving ability' is inducing serial correlation between the observations made of the drivers: a better than average driver will tend to remain better than average and a worse than average driver will tend to remain worse than average. Going back to the mathematical formulation of the model, we can further interpret it as saying: (1) the expected observed average given the observable variables $\boldsymbol{x}_{i,t}$ is $\boldsymbol{x}_{i,t}\boldsymbol{\beta}$, (2) if one knew the values of the unobserved heterogeneity terms $\boldsymbol{\alpha}_i$ and the observable variables $\boldsymbol{z}_{i,t}$, then the unexplained portion the observations would form a potentially auto-correlated and heteroskedastic sample, and (3) there exists unobserved heterogeneity that drives serial intra-individual correlation and this unobserved heterogeneity $\boldsymbol{\alpha}_i$ forms a random sample.

For our purposes, however, we will instead anchor ourselves in a fixed-effects model. One way to think about fixed-effects models is as a classical regression that includes an indicator function for each of the included 'individual'[8]. As such, under traditional fixed-effects models, there is a unique

[8] Contrasting this with the random-effects models: under a fixed-effects model, other considered individuals do not provide information about the coefficients that need to be attached to time constant covariates and all time constant covariates become collinear with the individual-specific indicator covariate.

intercept terms for each of the considered 'individual'. Another way to think about fixed-effects models is as a regression that includes a straight average of the residuals (from the regression of $y_{i,t} - \bar{y}_i$ on the covariates $\boldsymbol{x}_{i,t} - \overline{\boldsymbol{x}}_i$) as a covariate that is affected with a slope of unity.

Leaping with this idea, we can think of more traditional Auto-Regressive time series models (Wikipedia n.d.)[9] as fixed-effects models that include only the individual specific intercepts, but that uses a weighted instead of a straight average of the residuals to estimate the individual specific intercepts. The connection with time series is particularly relevant for rating and forecasting purposes. It is critical in rating and forecasting applications that the used covariates constitute available information at the time of the forecast. In the probability literature, this has been captured by the filtration concept[10].

Pushing even further the connection with time series models, it is also possible to include a general structure for the correlation of the residuals. Of particular interest is the inter-temporal intra-individual covariance structure for the residuals. In our case, we will consider a flexible Moving Average model (Wikipedia n.d.) called a Toeplitz specification for $\boldsymbol{R}_i$ (Frees, Longitudinal and panel data: analysis and applications in the social sciences 2004, 281). One of the advantages of the Toeplitz specification is that it presumes homoskedasticity. For our purposes, this will greatly simplify the forecasting process, as the variance of the residuals will not first have to be forecasted for future periods before the forecasts can be computed. The hypothesis is far from perfect[11]. Take, for example, the case of Homeowners insurance that can be greatly affected by natural catastrophes. In a year where a great hurricane or earthquake hits, some insurer will have exposures in the affected region and have poor underwriting results, but insurers that do not have any exposures in the region will only be affected by 'normal' noise. Given that it is next to impossible to forecast these great catastrophes, the forecast of the future variance of the residuals is also very difficult. That is why we will focus on a covariance structure for the residuals that does not imply that we need to forecast the variance of the residuals before computing the forecasts as such.

---

[9] The author understands that many academic parties are uncomfortable with Wikipedia as a reference source. One traditional argument against Wikipedia is the non-certification of the source. As an actuary, the author is effectively endorsing any cited source as professional standards generally require that an actuary cannot cite references to other work for why the actuarial work product is not adequate. Another reason to support the resistance to the use of Wikipedia in academic work is it relative instability, in as much as this is a source that gets constantly updated. Here, the author is effectively making the practical assumption that ease of accessibility is more important than the stability of the source. Wikipedia, being a free web reference source, is imminently accessible to academic and professional populations.

[10] (Steele 2000, 50)

[11] Preliminary testing of the models indicates that models incorporating heteroskedasticity do better than models that imply homoskedasticity.

For our considered covariates, we will introduce the past results[12] in similar lines of business. An alternate modeling choice could have been to introduce a current period forecast of the results of similar lines of business as statistical instruments (Frees, Meyers and Cummings, Predictive Modeling of Multi-Peril Homeowners 2011)[13]. Given that we are working in Accident Year[14] and that about half of the results of the following Accident Year are driven by the same contracts as the current Accident Year, we believe that including the latest available Accident Year results serves substantially the same purpose. In effect, we are saying that if one wanted to make a 'back-of-the-envelop' forecast of the current Accident Year Ultimate Loss Ratio for a given line of business for a given insurer, one could use only the prior Accident Year Ultimate Loss Ratio as a covariate and come up with a good initial value for the forecast.

Because, under most experience rating formulas, the credibility factor z changes as the size of the account changes, we have included interaction terms that cross past insurer line of business past paid Ultimate Loss Ratio relativity with current insurer size, measured by a non-linear increasing concave down function of Earned Premium. This accomplishes the goal of varying the models for different insurer size. Further comments will be presented in section 6.

5. NAIC DATA

Before moving on to the selected models and the assessment of their predictiveness, let's first discuss the publicly available data that supports our methodology.

5.1 Data Preparation

At the heart of our analysis lies the National Association of Insurance Commissioners Schedule P of the statutory financial statements[15] from 1992 to 2010. As such, we are only focusing on

---

[12] Notice that past results are part of the information set of the person applying the rating or forecasting algorithm.

[13] The instrumental variable approach to dependency between lines of business can be thought of as an alternative to the copula approach (Frees, Meyers and Cummings, Dependent Multi-Peril Ratemaking 2009). The copula approach would be especially relevant for capital adequacy testing. For ratemaking purposes, because copula regressions preserve the conditional on the covariates models and we are only interested in the expected values, the only place where the copula could affect the results is in the joint estimation of the conditional on the covariates and copula models. A natural way to approach this is through Maximum Likelihood estimation that requires parametric modeling. Given model uncertainty that is inherent in the selection of the distribution of Ultimate Loss Ratio, this approach is not preferred here.

[14] More on what we mean by Accident Year below.

[15] Academic works that explored the relative efficiency of different P/C insurers also made reference directly or indirectly to the NAIC data. For example, (A Note on the Relative Efficiency of Property-Liability Insurance Distribution Systems 1979) (Independent and Exclusive Agency Insurers: A Reexamination of the Cost Differential 1992), (The Coexistence of Multiple Distribution Systems for Financial Services: The Case of Property-Liability Insurance 1997), and (Long-tail Longitudinal Modeling of Insurance Company Expenses 2010).

American Property/Casualty insurance exposures. For our purposes, one interesting feature of the Schedule P is its relative stability through time.

From the Schedule P Part 1, we extracted 'Premium Earned Direct and Assumed'[16], hereafter referred to as Earned Premium or EP. We extracted Part 3, which covers Paid Loss and ALAE[17]. This information is particularly useful because the key determinant of the loss development pattern is driven by the line of business and not by the insurer. As a consequence, it was possible for us to compute[18] Loss Development Factors (to Ultimate, or ULDFs), by maturity, by line of business, for the industry as a whole.

We also extracted Part 2, which refers to incurred loss and ALAE[19]; however, early tests demonstrated that, for different insurers, for a given line of business, the development patterns could be qualitatively and materially different: therefore, we chose ultimately to not use this information. Finally, we extracted Part 5 Section 3 'Cumulative Number of Claims Reported Direct and Assumed at Year End' and Part 6 Section 1 'Cumulative Premiums Earned Direct and Assumed at Year End'. Again, we ultimately chose not to use the information. For the claim counts, we chose not to use the information because it was not available for all the lines of business that were of interest to us[20]. As for the Earned Premium triangle, we are content in using the latest valuation of the Earned Premium.

We chose to work with insurer groups instead of the individual entities that report to the NAIC. One motivation for doing so was that internal strategic considerations can lead insurer groups to selectively assign risks to different insurers and this assignment can vary through time for endogenous reasons. Another motivation for this choice is that there should be fewer insurers entering and exiting when looking at the industry at the insurer group level. Note that no specific

---

16 Earned Premium refers to main revenue source of P/C insurers. Written Premium corresponds to the value of policy sold, while Earned Premium refers to the accrual of revenues relating to sold policies. 'Direct and assumed' refers that the said sold insurance policies can have been sold to the public directly (direct) or to another insurer (assumed).

17 'Cumulative Paid Net Losses and Defense and Cost Containment Expenses Reported at Year End'. ALAE refers to Allocated Loss Adjustment Expenses.

18 Using the Chain Ladder method (Werner and Modlin 2010, 105-109). One unfortunate aspect of the Chain Ladder method for Loss Development is the induced serial correlation of the residuals that results from the use of cumulative loss triangles. Generally, unaccounted for serial correlation of the residuals can lead to biased regression estimates. That being said, given that we are using the Chain Ladder method on the loss triangle generated by the industry as a whole and given that our covariates are more driven by the line of business than by the insurer, our Ultimate Loss estimates should not be materially inaccurate (taking into account the available information set). Also, whenever possible, for a given Accident Year, we use the latest available valuation, which is after 10 years for most lines of business (except Auto Physical Damage for which only 2 years of development is available.). In determining the latest valuation year, we have included a test that checks that the by insurer / line of business / accident year EP is not materially changing with new valuation. We have introduced this test because financial statements appear to be re-stated when the entities that form an insurer group change.

19 'Incurred Net Losses and Defense and Cost Containment Expense Reported at Year End'

20 More on that topic is to come.

treatment was made for entering or exiting insurers but, as will be seen below, we do indirectly account for some forms of entries.

We excluded insurer/line of business/Accident Year, on a per-observation basis, where the Earned Premium was less than 1M nominal USD. The net effect of that exclusion was measured to be in the order of the one tenth of a percent. We did so because these records generate missing, negative or highly volatile measured Loss Ratios.

We chose to focus on selected lines of business found in the Table 1 below. Another party could easily extend our results to include all available lines of business.

Line of Business		
Reference Letter	Description	both Occurrence and Claims-Made
A	Homeowners/Farmowners	
B	Private Passenger Auto Liability/Medical	
C	Commercial Auto/Truck Liability Medical	
D	Workers' Compensation	
E	Commercial Multi-Peril	
F	Medical Professional Liability	Y
G	Special Liability (Ocean Marine, Aircraft (All Perils), Boiler and Machinery)	
H	Other Liability	Y
J	Auto Physical Damage	
R	Products Liability	Y

Occurrence liability policies refer to liability policies where coverage is determined as a function of the **occurrence dates** of the alleged wrong-doing of the insured.

Claims-made liability policies refer to liability policies where coverage is determined as a function of the **reporting date** of the alleged wrong-doing of the insured.

Table 1: Line of Business Listing

For readers that are familiar with the Rate Indications methodology (Werner and Modlin 2010, 71-80), please note that no rate or exposure changes were available for extraction. The absence of rate changes creates a less than ideal environment for forecasting. Falling back on the internal Projected Loss Ratio methodology, the change in Ultimate Loss Ratio can generally be thought of to be the result of a loss trend[21], a premium trend, rate changes and mix changes[22]. Contrary to the other effects, rate changes are primarily the result of overt actions taken by the insurer and are not as much subject to momentum effects[23]. If an insurer decided to pass a rate increase of +25%, we would expect the Ultimate Loss Ratio to immediately begin to fall; *vice versa* for a rate decrease. Therefore, we expect that we will encounter instances where our predicted Loss Ratios would be off because they will not reflect rate change information.

Note that, in using the 'simple' industry-wide by line of business Chain Ladder methodology[24] for loss development, we are putting ourselves in a situation similar to an actuary that was calibrating a traditional Experience Rating algorithm: a traditional Experience Rating algorithm will generally not have rating factors that change insured by insured. Rating algorithms cannot generally be calibrated to the individual insured, unless the account is so large as to be able to be self-ratable. Even so, our practical assumption here is that no individual insurer group is so large that no information from the other insurers is necessary to forecast its Projected Loss Ratio.

5.2 Descriptive Statistics

Before justifying the exact nature of our modeling choices, let's explore the data. Note that 'Year' always refers to Accident Year. This can be contrasted with Policy Year, that refers to the inception year of the insurance contract, and with Accounting Year, that refers to the year in which the revenue and losses were recognized for accounting purposes. Accident Year is generally preferred for most P/C actuarial purposes because many factors that affect losses are best accounted for using the date of the accident: *e.g.* seasonality relating to natural catastrophes or driving conditions. If Policy Year was available in the NAIC data, it might be appropriate for our uses, but Loss Development requires extra care. Accounting Year is not suitable for our purposes as the year in

---

21 That can be decomposed into a frequency and a severity trend.

22 Mix changes sometimes refer to the effect of the change of the proportion of different types of insureds, instead here refers to 'other changes'.

23 Contrast with loss trends that are largely due to the direct inflation associated with insured 'objects' and the indirect inflation of changing insured 'objects'.

24 "The distinguishing characteristic of the development method is that ultimate claims for each accident year are produced from recorded values assuming that future claims' development is similar to prior years' development. In this method, the actuary uses the development triangles to track the development history of a specific group of claims. The underlying assumption in the development technique is that claims recorded to date will continue to develop in a similar manner in the future – that the past is indicative of the future. That is, the development technique assumes that the relative change in a given year's claims from one evaluation point to the next is similar to the relative change in prior years' claims at similar evaluation points." (Friedland 2010, 84)

which losses get recognized may have only to do with the timing of reserve changes and little with current policy wording, legal environment or general market conditions of the P/C insurance market.

First, as Figure 1 demonstrates, a key driver of an insurer Loss Ratio is the line of business mix[25], as different lines of business tend to have materially different Loss Ratios. With Figure 2, we can see that these differences are persistent through time.

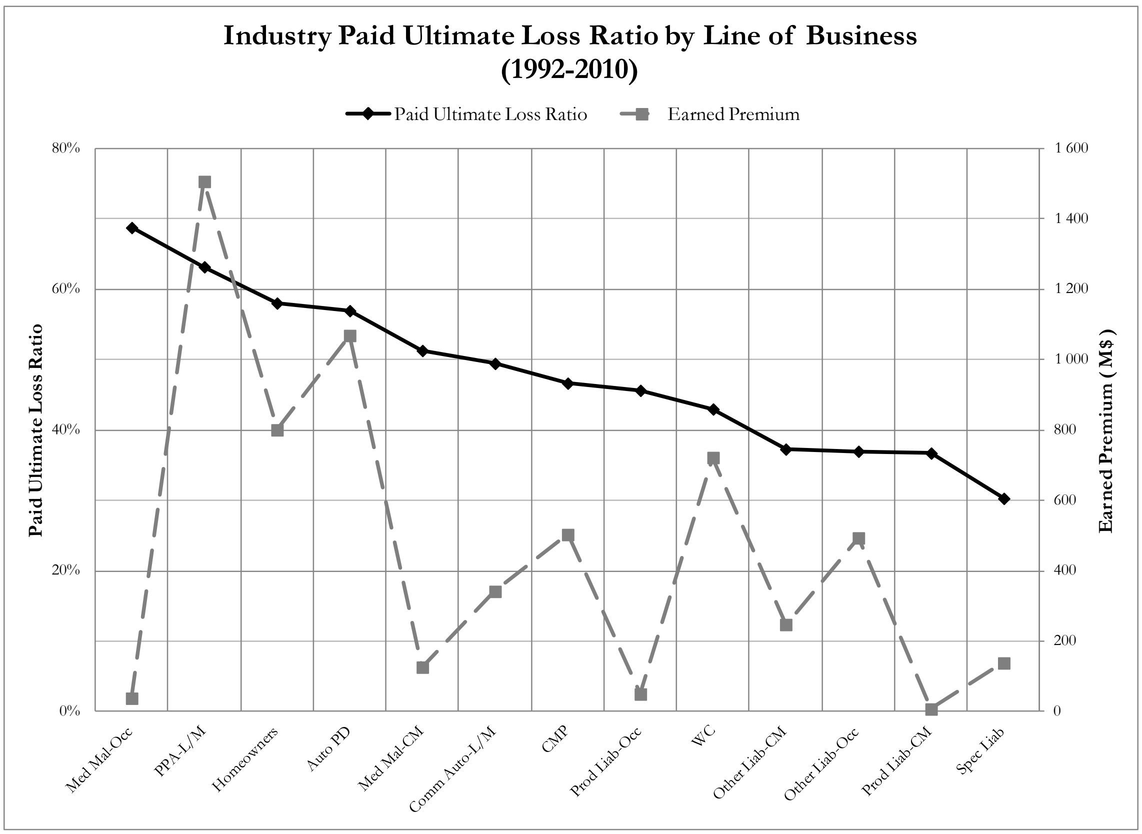


Figure 1: Industry Paid Ultimate Loss Ratio by Line of Business

[25] Our modeling presumption is that the line of business mix of an insurer is either stable or predictably changing.

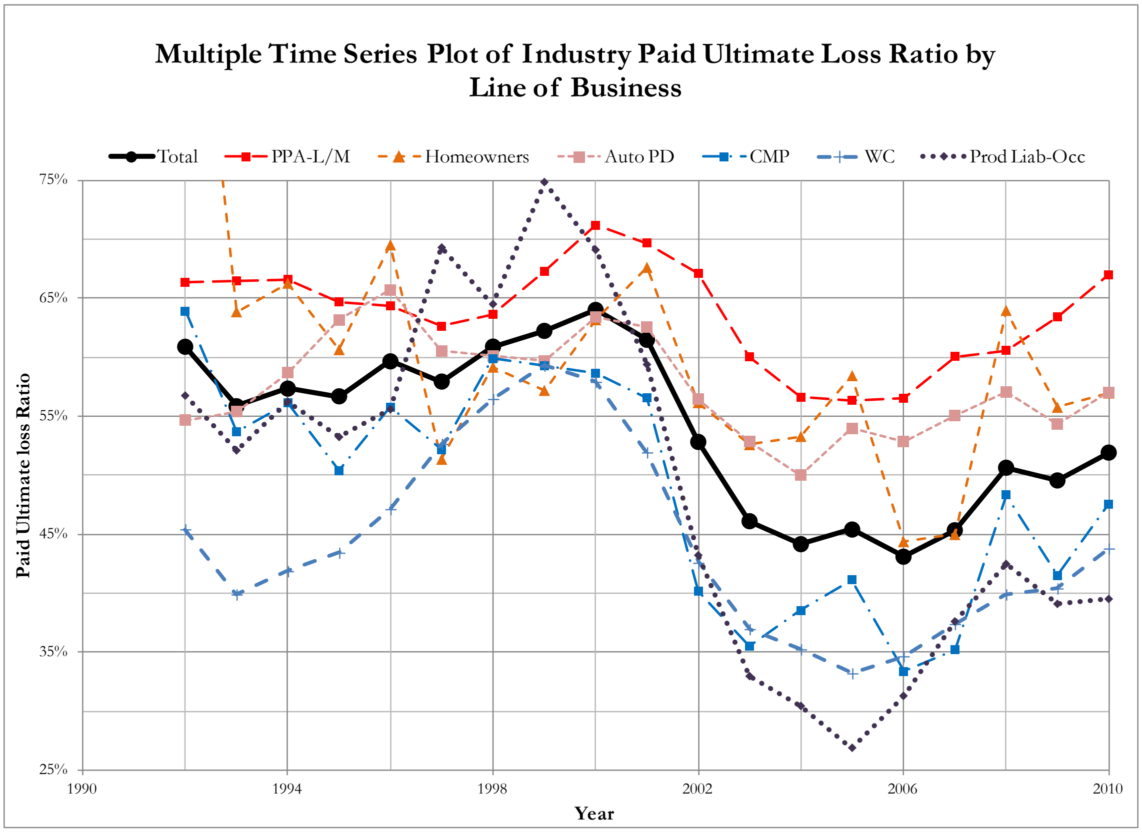


Figure 2: Multiple Time Series Plot of Industry Paid ULR by LOB

As the selected[26] lines of business (EP weighted) quantile and mean time series plots in Figures 3 and 4 demonstrate, the distribution of Ultimate Loss Ratio is fairly symmetric (if a little right-skewed) but heavier tailed than a Normal distribution in some years. For the Property lines of business, skewness and heaviness of the right tail can be affected by natural catastrophe like, for example, Hurricane Andrews.

[26] The plots for all lines of business were produced and can be presented upon request to the author. This applies for other presented plots also.

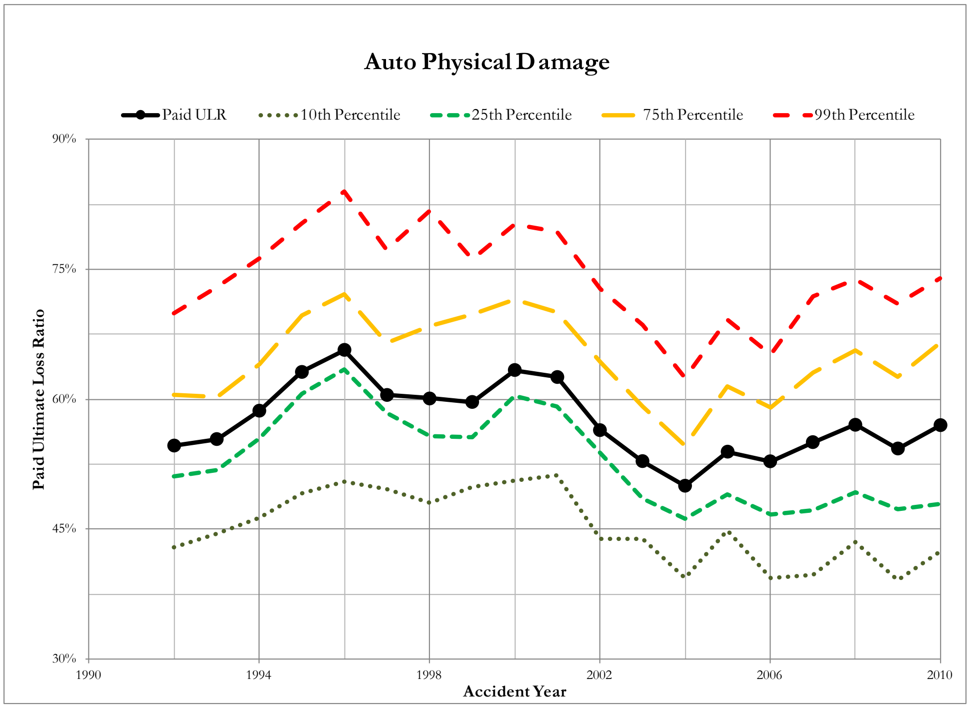


Figure 3: Multiple Time Series Plot of Features of the Auto PD Paid ULR Distribution

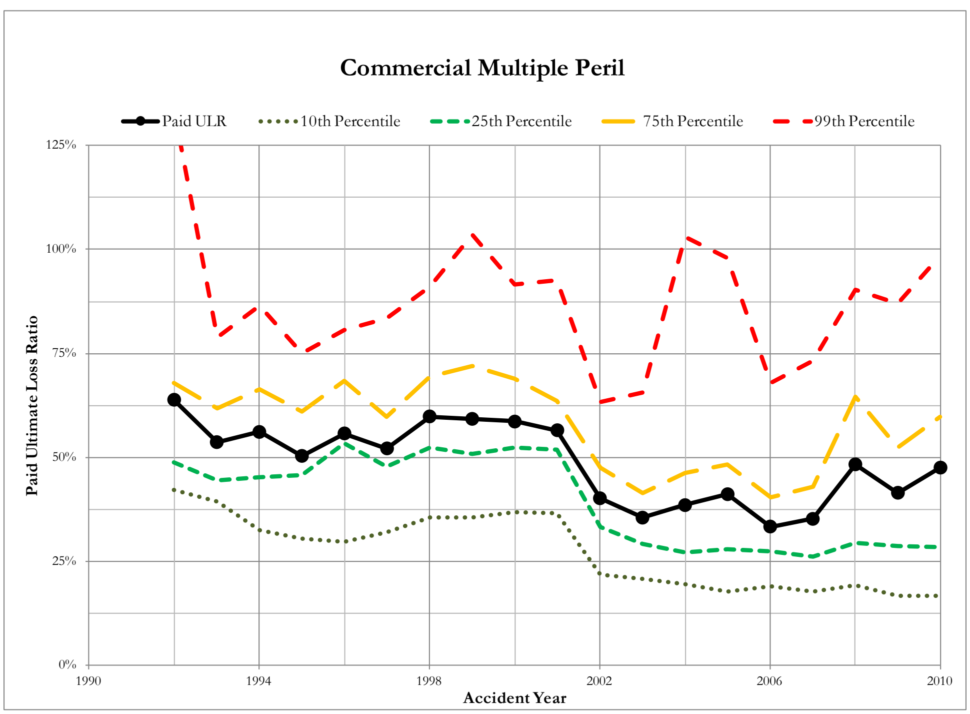


Figure 4: Multiple Time Series Plot of Features of the CMP Paid ULR Distribution

From the charts for the selected lines of business found in Figures 5 and 6, we can see that insurer Loss Ratio rankings are persistent through time, as the relative positions of the lines remain fairly stable.

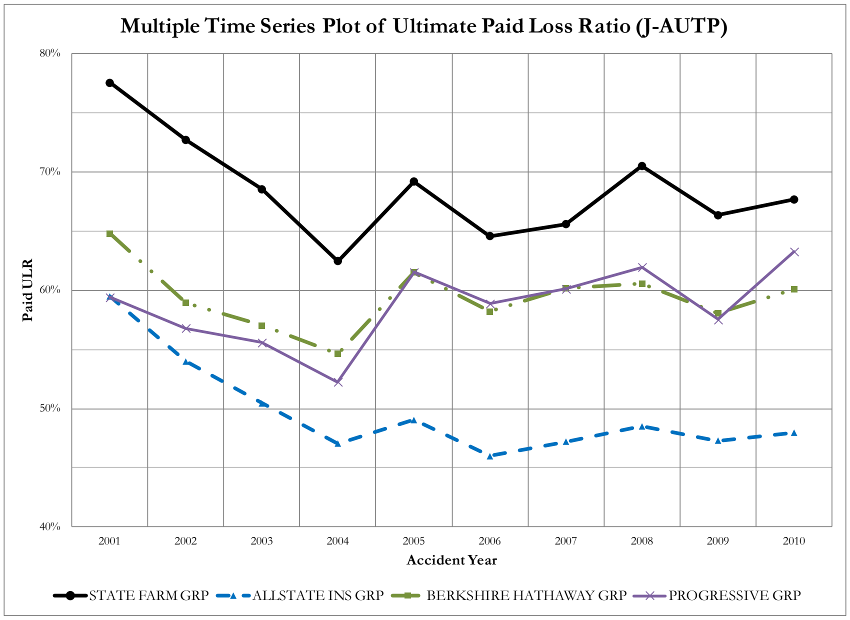


Figure 5: Multiple Time Series Plot of Paid ULR of Large Insurers: Auto PD

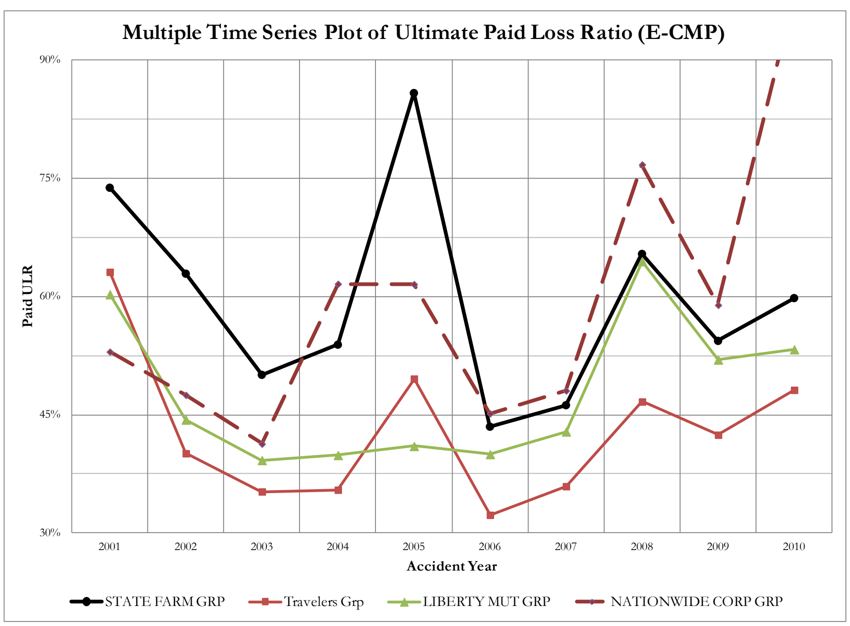


Figure 6: Multiple Time Series Plot of Paid ULR of Large Insurers: CMP

For the selected insurers found in Table 2, the tables below show summary statistics relating to the relativity of the ULR for the insurer group/line of business/Accident Year to the industry/line of business/Accident Year. We will model the Loss Ratio relativity instead of the Loss Ratio directly because we believe that a view on the future state of the P/C industry is generally easier to develop than a particular view for a given insurer group: this is analogous to why Experience Rating is generally calibrated with the practical assumption that Classification Rating has already appropriately reflected all factors other than claiming history. In that sense, Experience Rating can be thought of as the predictive modeling of the future profitability of an insurance account that uses the history of the 'individual' that has not already been accounted for by other known effects. In these cases, the Loss Ratio relativities of different lines of business appear quite linked: either because they have similar values or because the movements are correlated. The effect is quite general and applies to many other insurers, especially for the more important lines of business of larger insurers.

STATE FARM GRP

LOB	B_PPAL	J_AUTP	A_HMOW	TOTAL
EP (M$)	298 702	207 051	180 725	731 259
EP %	40.8%	28.3%	24.7%	100.0%
Avg Relativity	1.114	1.195	1.162	1.168

correlation with	J_AUTP	A_HMOW	EP
B_PPAL	16.0%	53.0%	298 702
J_AUTP		21.6%	207 051
EP	207 051	180 725	

PROGRESSIVE GRP

LOB	B_PPAL	J_AUTP	C_CA_L	TOTAL
EP (M$)	86 560	56 671	14 133	159 141
EP %	54.4%	35.6%	8.9%	100.0%
Avg Relativity	0.977	1.049	1.182	1.022

correlation with	J_AUTP	C_CA_L	EP
B_PPAL	64.4%	53.2%	86 560
J_AUTP		79.8%	56 671
EP	56 671	14 133	

Table 2: Typical Cross-Lines of Business Correlations
'B_PPAL' refers to line of business B 'Private Passenger Auto Liability/Medical', 'J_AUTP' refers to line of business J 'Auto Physical Damage', 'A_HMOW' refers to line of business A 'Homeowners/Farmowners', and 'C_CA_L' refers to lines of business C 'Commercial Auto/Truck Liability Medical'

In Table 3, we show (EP weighted) descriptive statistics by line of business for the Loss Ratio relativity for Accident Year 2006 as of 2010. Notice how the mean relativity is always 1.00: it is so by definition of a relativity. Notice also that comments made above about skewness and heavy-tails also apply here: the distributions are fairly symmetric but not quite Normal.

		Relativity					
			Percentile				
Line of Business Short-Hand	**Line of Business Long Description**	**Avg**	**10th**	**25th**	**50th**	**75th**	**90th**
B_PPAL	Private Passenger Auto Liability/Medical	1.00	0.83	0.98	1.03	1.10	1.12
J_AUTP	Auto Physical Damage	1.00	0.82	0.88	1.04	1.15	1.21
A_HMOW	Homeowners/Farmowners	1.00	0.77	0.88	1.01	1.09	1.32
D_WC__	Workers' Compensation	1.00	0.57	0.80	0.97	1.34	1.37
E_CMP_	Commercial Multi-Peril	1.00	0.56	0.90	1.08	1.23	1.32
H_OL_O	Other Liability - Occurrence	1.00	0.39	0.64	1.03	1.23	1.43
C_CA_L	Commercial Auto/Truck Liability Medical	1.00	0.57	0.93	1.03	1.17	1.31
H_OL_C	Other Liability - Claims-Made	1.00	0.28	0.59	1.31	1.34	1.59
G_SL__	Special Liability (Ocean Marine, Aircraft (All Perils), Boiler and Machinery)	1.00	0.31	0.70	0.87	1.32	1.77
R_PL_O	Products Liability - Occurrence	1.00	0.17	0.50	0.95	1.09	1.90
R_PL_C	Products Liability - Claims-Made	1.00	0.11	1.06	1.09	1.25	1.25
F_MM_C	Medical Professional Liability - Claims-Made	1.00	0.39	0.84	0.90	1.24	1.59
F_MM_O	Medical Professional Liability - Occurrence	1.00	0.43	0.67	1.11	1.29	1.66

Table 3: Summary Statistics, by Line of Business, of the Relativity Variables (AY 2006, as of 2010)

6. STATISTICAL ANALYSIS

We are finally at the point where we can discuss the fitted models. Our modeling approach was not entirely dissimilar to data mining, in as much as many models were fitted and compared. To fit our models, we used the SAS `proc mixed` procedure. The procedure was used on a line by line basis: that is, a model was fitted for each line of business. Only Accident Years 1997 to 2006 were used for the purposes of model selection. Even if the 1992 to 1996 Accident Years are known to us, given that we want to preserve inter-model comparability and that we will allow the use of up to 5 years of same line prior relativities, we need to start using our data starting in the 1997 Accident Year. We chose to stop using data past the 2006 Accident Year in an effort to balance responsiveness and stability (Werner and Modlin 2010, 80): using more recent data would increase responsiveness to current conditions but would be counter-balanced by the fact that more recent Ultimate Loss Ratios have a much greater portion that is estimated, as opposed to realized[27].

For model selection purposes, we used the empirical estimates of the standard errors (SAS n.d.), to allow for model misspecification. The model was used using the Residual Maximum Likelihood approach. The Bayesian Information Criterion was used for model selection as, among the common information criteria, it is the one that most penalizes for extra variables. For our purposes, it was particularly important to favor parsimony, as many hundreds of models were attempted for each line of business.

Again, to ensure comparability of models, we ensured that none of the covariates were missing by initializing them to a neutral value if they were otherwise missing and adding an indicator variable to indicate that the value was missing. From the fitted line of business, we included up to 5 prior realizations of the Loss Ratio relativity. We selected 3 lines of business that have similar strategies, clients or perils and included up to 2 years of prior realized relativities for those selected lines of business. As mentioned above, we have included EP-based interaction terms[28]. Interaction terms relating to older data were not included in the attempted models unless including the interaction term from more recent data improved the model. We allowed for serial correlation between intra-insurer residuals using a `repeated` statement (SAS n.d.). A Toeplitz specification was used because of the implied homoskedasticity and the flexible correlation structure. For forecasting, the implied homoskedasticity is particularly convenient, because we would otherwise first have to forecast the variability of the future relativities and that variability is largely determined by random factors, especially for Property lines of business. For the Toeplitz specification, we allowed ourselves a window of up to 5 years, consistent with our modeling choice for the Auto-Regressive component of the model. The regression was a weighted regression: with Earned Premium used as weights[29]. The choice of weights was not due to statistical efficiency considerations, but rather due to economic relevance considerations. As such, even if weights are used, the regression should not be

[27] Relating to footnote 10, we have chosen to always use the latest available maturity rather than demand that a minimum maturity. Doing so, we may create a bias because the ULDFs may not be as equally appropriate for all insurers, especially at early durations. We could have demanded a minimum maturity, but that would imply dropping a material quantity of data from the analysis. We could also use the always available earliest maturity, but that would imply that our modeled quantity is basically always an estimate and never materially realized.

[28] We used interaction terms based on $\log_{10} EP$, which is non-linear in EP.

[29] One may wonder at our modeling choice of including EP in both the interaction terms and the weights. Both uses have different rationales that are not mutually inconsistent. As mentioned above, the use of EP in the weights is aimed at reflecting the economic importance of the fit. The use of EP in the interaction terms is aimed at varying the models for different insurer size, just like in traditional credibility models. In effect, this implies that different models are fitted for different insurer sizes, but also that, among insurers of similar size, the bigger ones count more towards model fitting.

construed as a first approximation to a Feasible Generalized Least Squares estimator, but rather as an Ordinary Least Squares regression that puts equal weight on all dollars of Earned Premium[30]. A table summarizing the selected best fitting models is presented in Table 4. The fitted values for the parameters are also presented in Appendix 1.

Line of Business Short-Hand	B_PPAL	J_AUTP	A_HMOW	D_WC__	E_CMP_	H_OL_O	C_CA_L
Line of Business Long Description	Private Passenger Auto Liability/Medical	Auto Physical Damage	Homeowners/Farmowners	Workers' Compensation	Commercial Multi-Peril	Other Liability - Occurrence	Commercial Auto/Truck Liability Medical
Covariates	intercept	intercept	intercept	intercept	intercept	intercept	intercept
	same line - lag 1	same line - lag 1	interaction term - intercept	interaction term - intercept	same line - lag 1	interaction term - intercept	interaction term - intercept
	interaction term - same line - lag 1	interaction term - same line - lag 1	same line - lag 1	same line - lag 1	same line - lag 2	same line - lag 1	same line - lag 1
	same line - lag 2	same line - lag 2	interaction term - same line - lag 1	interaction term - same line - lag 1	same line - lag 3	interaction term - same line - lag 1	interaction term - same line - lag 1
	same line - lag 3	interaction term - same line - lag 2	same line - lag 2	line C_CA_L - lag 1	line C_CA_L - lag 1	same line - lag 2	same line - lag 2
	same line - lag 4	same line - lag 3	interaction term - same line - lag 2	interaction term - line C_CA_L - lag 1	interaction term - line C_CA_L - lag 1	interaction term - same line - lag 2	interaction term - same line - lag 2
	same line - lag 5	line B_PPAL - lag 1	same line - lag 3	line E_CMP_ - lag 1	line D_WC__ - lag 1	same line - lag 3	same line - lag 3
	line J_AUTP - lag 1	line C_CA_L - lag 1	same line - lag 4		line D_WC__ - lag 2	interaction term - same line - lag 3	same line - lag 4
	interaction term - line J_AUTP - lag 1	interaction term - line C_CA_L - lag 1	line B_PPAL - lag 1			same line - lag 4	same line - lag 5
	line J_AUTP - lag 2		interaction term - line B_PPAL - lag 1			interaction term - same line - lag 4	line D_WC__ - lag 1
			line J_AUTP - lag 1			same line - lag 5	interaction term - line D_WC__ - lag 1
			interaction term - line J_AUTP - lag 1			interaction term - same line - lag 5	line D_WC__ - lag 2
			line E_CMP_ - lag 1			line E_CMP_ - lag 1	line J_AUTP - lag 1
			interaction term - line E_CMP_ - lag 1			interaction term - line E_CMP_ - lag 1	line J_AUTP - lag 2
			line E_CMP_ - lag 2			line D_WC__ - lag 1	
			interaction term - line E_CMP_ - lag 2			interaction term - line D_WC__ - lag 1	
Intra-insurer Group Covariance Structure for the Residuals	TOEP(5)	TOEP(2)	TOEP(2)	TOEP(5)	TOEP(5)	TOEP(5)	TOEP(4)

Line of Business Short-Hand	H_OL_C	G_SL__	R_PL_O	R_PL_C	F_MM_C	F_MM_O
Line of Business Long Description	Other Liability - Claims-Made	Special Liability (Ocean Marine, Aircraft (All Perils), Boiler and Machinery)	Products Liability - Occurrence	Products Liability - Claims-Made	Medical Professional Liability - Claims-Made	Medical Professional Liability - Occurrence
Covariates	intercept	intercept	intercept	intercept	intercept	intercept
	same line - lag 1	interaction term - intercept	same line - lag 1	interaction term - intercept	same line - lag 1	interaction term - intercept
	same line - lag 2	same line - lag 1	interaction term - same line - lag 1	line H_OL_O - lag 1	interaction term - same line - lag 1	same line - lag 1
	line E_CMP_ - lag 1	interaction term - same line - lag 1	same line - lag 2	interaction term - line H_OL_O - lag 1	same line - lag 2	interaction term - same line - lag 1
	interaction term - line E_CMP_ - lag 1	same line - lag 2	same line - lag 3	line H_OL_O - lag 2	same line - lag 3	same line - lag 2
	line D_WC__ - lag 1	interaction term - same line - lag 2	same line - lag 4	line D_WC__ - lag 1	same line - lag 4	same line - lag 3
		same line - lag 3		interaction term - line D_WC__ - lag 1	same line - lag 5	same line - lag 4
		interaction term - same line - lag 3		line C_CA_L - lag 1	line F_MM_O - lag 1	line F_MM_C - lag 1
		same line - lag 4		interaction term - line C_CA_L - lag 1	line F_MM_O - lag 2	interaction term - line F_MM_C - lag 1
		interaction term - same line - lag 4		line C_CA_L - lag 2	line G_SL__ - lag 1	line F_MM_C - lag 2
		same line - lag 5			interaction term - line G_SL__ - lag 1	interaction term - line F_MM_C - lag 2
		interaction term - same line - lag 5			line G_SL__ - lag 2	line G_SL__ - lag 1
		line D_WC__ - lag 1			interaction term - line G_SL__ - lag 2	interaction term - line G_SL__ - lag 1
		interaction term - line D_WC__ - lag 1			line H_OL_C - lag 1	line G_SL__ - lag 2
		line C_CA_L - lag 1			line H_OL_C - lag 2	interaction term - line G_SL__ - lag 2
		interaction term - line C_CA_L - lag 1				line H_OL_O - lag 1
Intra-insurer Group Covariance Structure for the Residuals	TOEP(5)	TOEP(5)	TOEP(4)	TOEP(3)	TOEP(5)	TOEP(2)

Table 4: Table of Best Fitting Models

Add text: notice that many models include terms for other LOBs as well as interaction terms

At this point, it is worthwhile to mention that (1) the past (relative) results of other lines of business generally are significantly influential and (2) many coefficients statistically vary with insurer size. That other lines of business are predictive supports our expectations that lines with similar clients, perils or strategies should move together. That coefficients vary with insurer size is consistent with the expectations formed by a century of developments in credibility theory.

We are now in a position to comment on the quality of the best fitting models. As can be seen in the selected exhibits in Figures 7 and 8, the fitted Loss Ratio relativities generally preserve the relativity for the insurer in a given line of business. Given that the estimators are of the Auto-Regressive Moving Average family, they suffer from the same defect: the predicted values lag behind

[30] Which are hopefully roughly proportional to the underlying insurance exposure. Another alternative might have been to put equal weight on all insureds. This measure would be even less perfect than our chosen weights as, even though the Loss Ratio varies by line of business, it does so materially less than the loss cost does across lines of business.

if a trend is present. Again, if rate changes were known, the hope is that this particular shortcoming could be dampened.

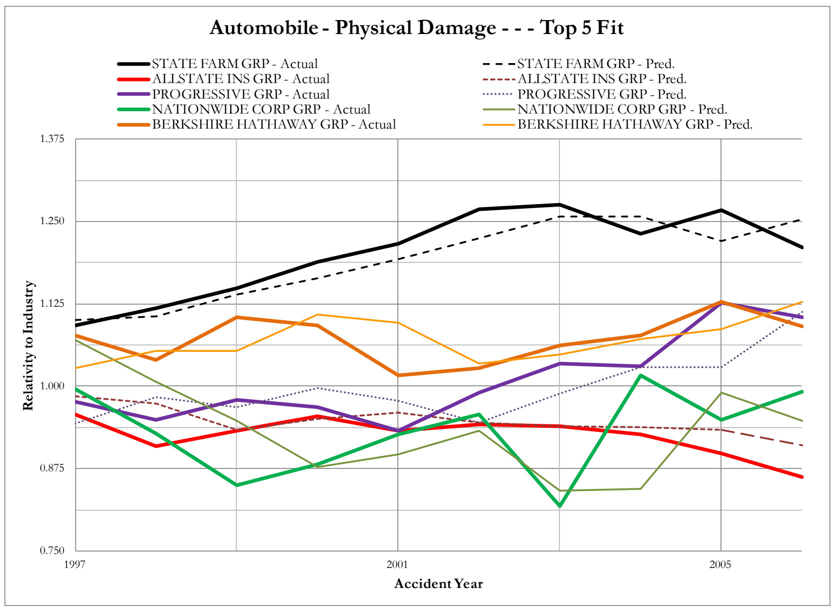


Figure 7: Multiple Comparative Time Series Plot of Large Insurer Actual *vs*. Predicted Relativities (AY 1997-2006, Auto PD)

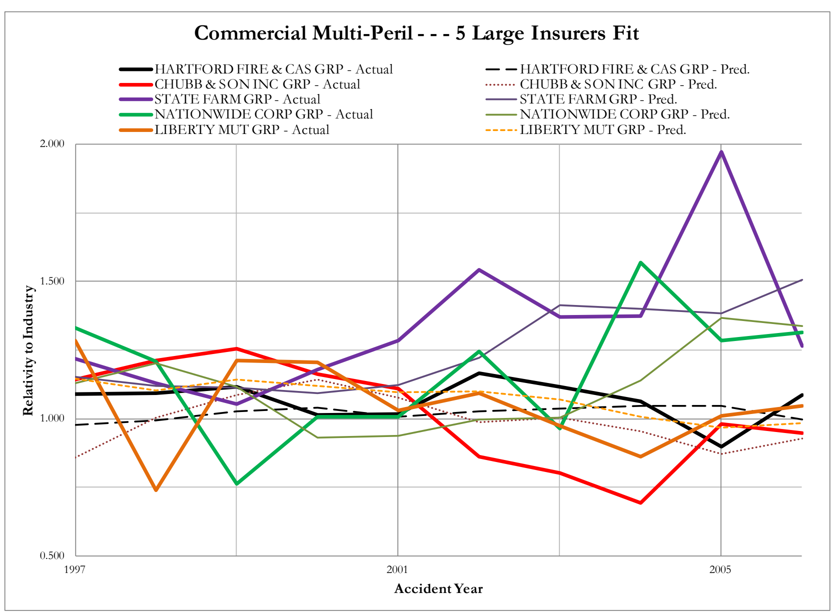
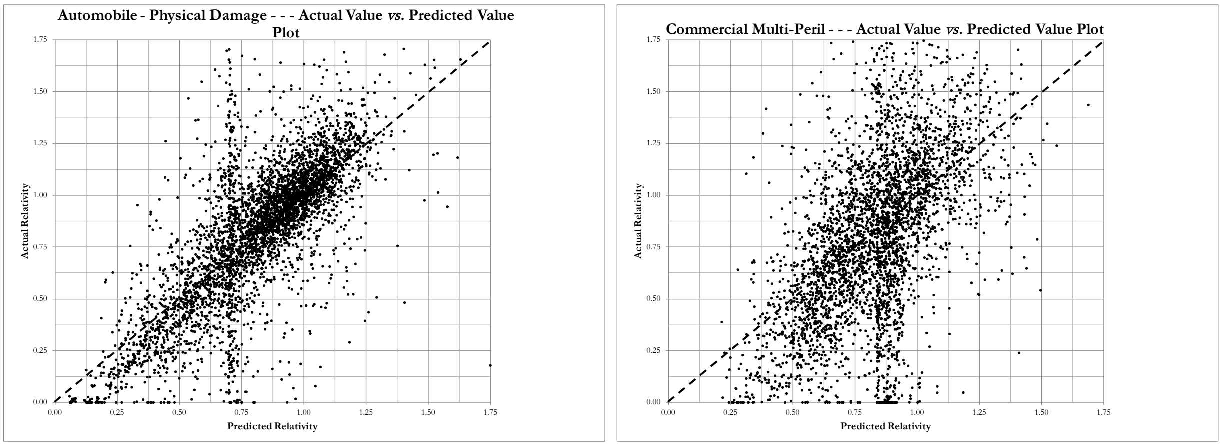
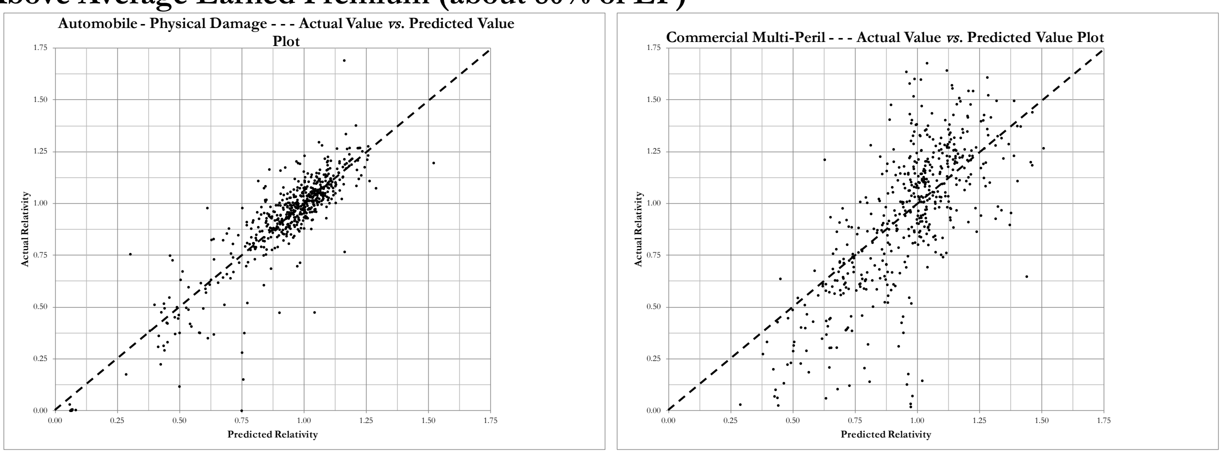


Figure 8: Multiple Comparative Time Series Plot of Large Insurer Actual *vs*. Predicted Relativities (AY 1997-2006, CMP)



Above Average Earned Premium (about 80% of EP)



Below Average Earned Premium

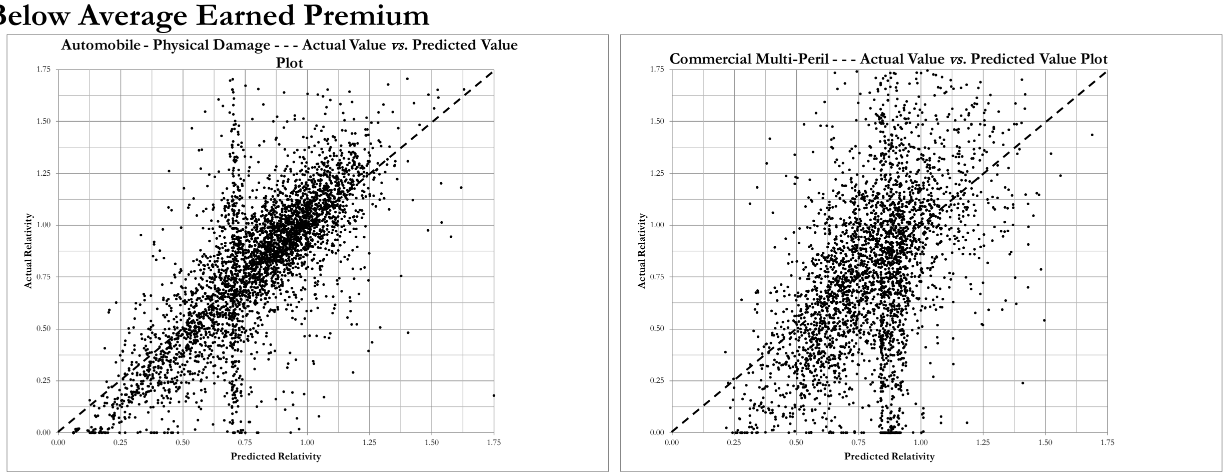
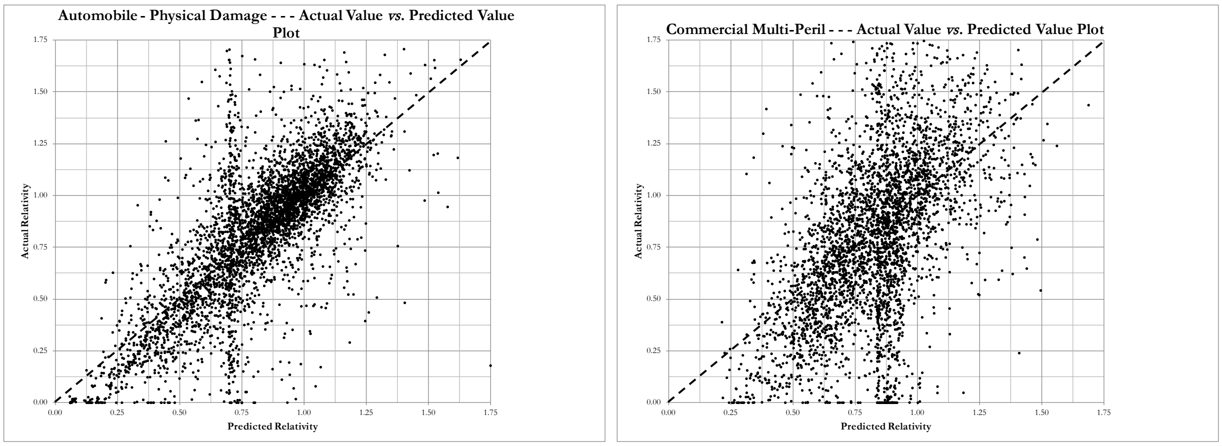
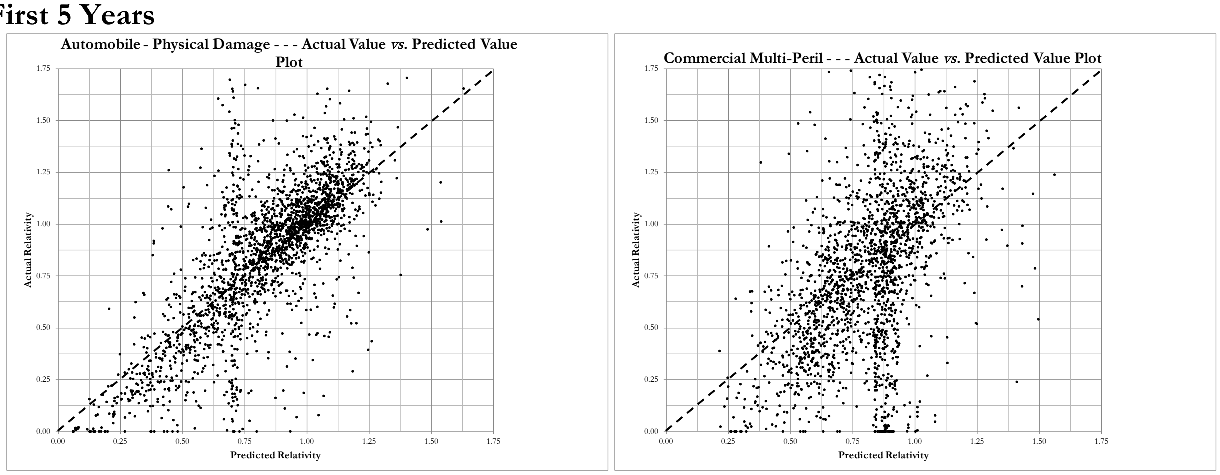


Figure 9: Actual *vs.* Predicted Relativity Plots (AY 1997-2006, Auto PD and CMP)



First 5 Years



Second 5 Years

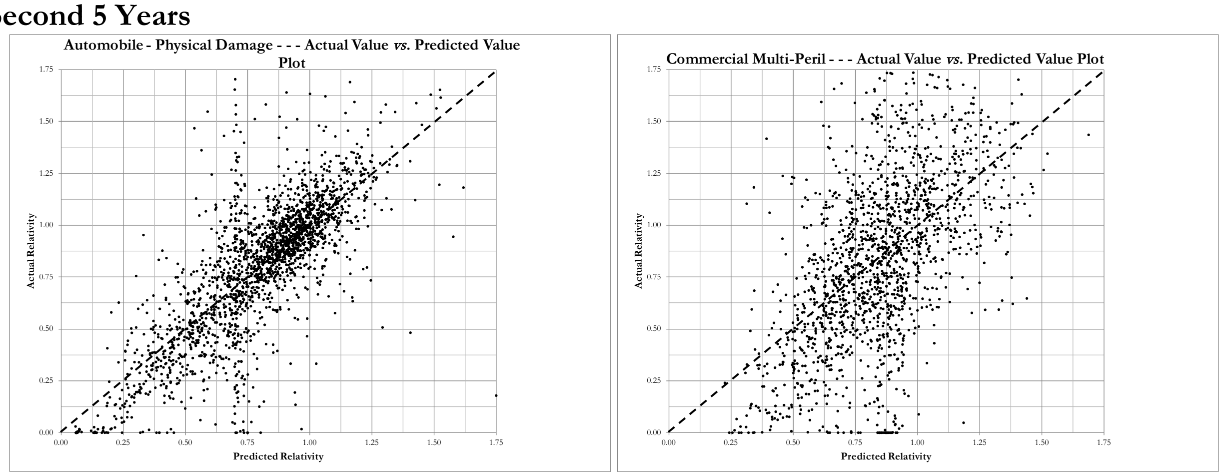
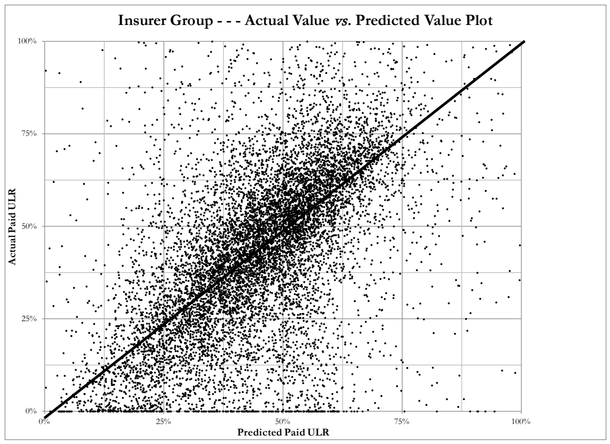


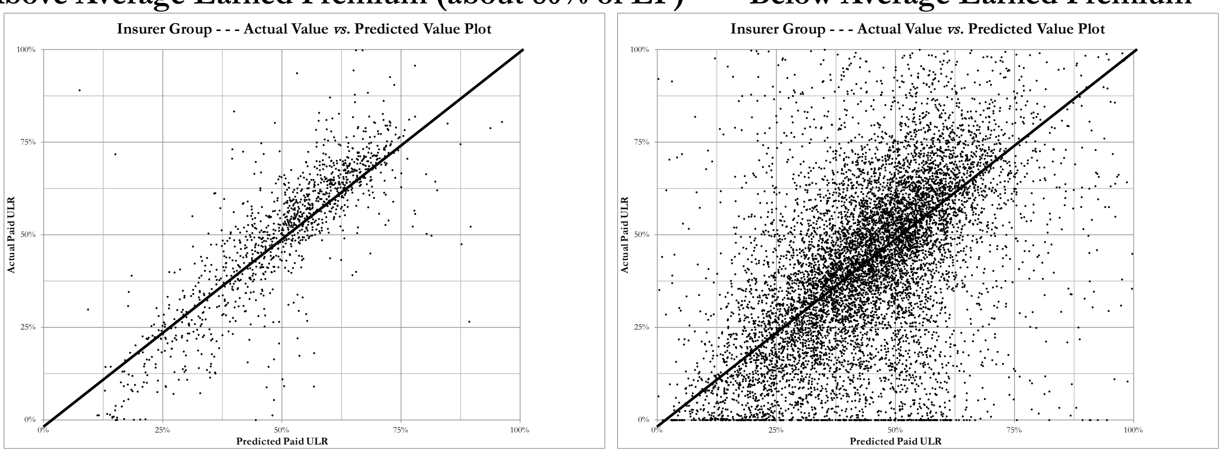
Figure 10: Actual *vs.* Predicted Relativity Plots (AY 1997-2006, Auto PD and CMP)

Insurer Group	2004	2005	2006	Grand Total	Last 5 Years	Last 3 Years	Last 2 Years
STATE FARM GRP							
Earned Premium	45 940 231	46 249 303	47 103 431	392 096 073	223 009 968	139 292 965	93 352 734
Actual ULR	62.6%	71.2%	60.2%	69.3%	66.6%	64.7%	65.7%
Pred. ULR	62.9%	66.4%	62.5%	68.5%	66.0%	63.9%	64.4%
ULR Diff.	**-0.2%**	4.8%	-2.3%	**0.7%**	**0.6%**	**0.7%**	1.2%
ALLSTATE INS GRP							
Earned Premium	25 066 983	26 263 811	26 899 116	223 012 544	124 437 058	78 229 910	53 162 927
Actual ULR	57.1%	64.9%	50.9%	61.9%	58.1%	57.6%	57.8%
Pred. ULR	55.6%	60.5%	52.3%	61.2%	57.6%	56.1%	56.4%
ULR Diff.	1.5%	4.3%	-1.4%	**0.7%**	**0.4%**	1.4%	1.4%
AMERICAN INTL GRP							
Earned Premium	31 729 086	31 795 842	32 810 846	193 972 203	135 646 692	96 335 774	64 606 688
Actual ULR	39.0%	42.4%	46.4%	47.1%	43.0%	42.6%	44.4%
Pred. ULR	39.3%	42.0%	42.2%	45.3%	40.8%	41.2%	42.1%
ULR Diff.	**-0.4%**	**0.4%**	4.1%	1.9%	2.3%	1.4%	2.3%
NATIONWIDE CORP GRP							
Earned Premium	13 443 304	14 516 434	15 245 735	124 270 757	72 470 460	43 205 473	29 762 169
Actual ULR	56.4%	52.9%	51.3%	54.8%	51.5%	53.4%	52.1%
Pred. ULR	48.1%	54.4%	50.3%	55.5%	51.5%	51.0%	52.3%
ULR Diff.	**8.3%**	-1.6%	1.0%	**-0.7%**	**-0.1%**	2.4%	**-0.2%**
LIBERTY MUT GRP							
Earned Premium	15 387 164	15 804 196	16 817 465	114 148 362	74 317 891	48 008 825	32 621 661
Actual ULR	41.0%	42.4%	40.3%	51.1%	43.0%	41.2%	41.3%
Pred. ULR	43.3%	43.2%	41.6%	53.0%	45.3%	42.7%	42.4%
ULR Diff.	-2.3%	**-0.8%**	-1.3%	-1.9%	-2.3%	-1.4%	-1.0%
Travelers Grp							
Earned Premium	19 350 200	18 734 052	18 869 201	110 119 067	73 768 428	56 953 453	37 603 253
Actual ULR	36.8%	41.2%	36.0%	48.7%	40.0%	38.0%	38.6%
Pred. ULR	39.1%	41.2%	39.6%	47.9%	41.0%	39.9%	40.4%
ULR Diff.	-2.3%	**0.1%**	-3.6%	**0.8%**	-1.0%	-1.9%	-1.8%
BERKSHIRE HATHAWAY GRP							
Earned Premium	12 243 116	14 124 855	15 327 065	91 477 201	62 223 013	41 695 036	29 451 920
Actual ULR	52.2%	51.2%	52.7%	58.7%	53.6%	52.0%	52.0%
Pred. ULR	51.9%	55.3%	52.7%	59.7%	54.9%	53.4%	54.0%
ULR Diff.	**0.3%**	-4.2%	**0.0%**	-1.0%	-1.3%	-1.3%	-2.0%
PROGRESSIVE GRP							
Earned Premium	13 386 629	13 959 011	14 233 252	91 080 785	62 134 914	41 578 892	28 192 263
Actual ULR	52.0%	57.3%	57.4%	57.9%	55.8%	55.6%	57.4%
Pred. ULR	55.0%	56.4%	58.3%	59.7%	56.9%	56.6%	57.4%
ULR Diff.	-3.0%	**0.9%**	**-0.9%**	-1.8%	-1.1%	**-1.0%**	**0.0%**
HARTFORD FIRE & CAS GRP							
Earned Premium	9 541 892	10 317 618	10 714 211	74 954 966	46 764 468	30 573 721	21 031 829
Actual ULR	43.7%	42.4%	43.0%	52.1%	45.0%	43.0%	42.7%
Pred. ULR	43.7%	45.2%	40.3%	50.7%	44.6%	43.0%	42.7%
ULR Diff.	**0.1%**	-2.8%	2.7%	1.3%	**0.4%**	**0.0%**	**0.0%**
Ace Ltd Grp							
Earned Premium	5 123 521	6 210 710	6 340 060	34 039 948	25 736 712	17 674 291	12 550 770
Actual ULR	22.3%	23.2%	20.7%	28.4%	23.3%	22.0%	22.0%
Pred. ULR	22.0%	25.0%	26.0%	31.7%	25.9%	24.5%	25.5%
ULR Diff.	**0.3%**	-1.8%	**-5.2%**	-3.3%	-2.6%	-2.4%	-3.5%

Table 5: Overall Insurer Back Testing of Actual *vs.* Predicted Relativity



Above Average Earned Premium (about 80% of EP) - - - Below Average Earned Premium



First 5 Years - - - Second 5 Years

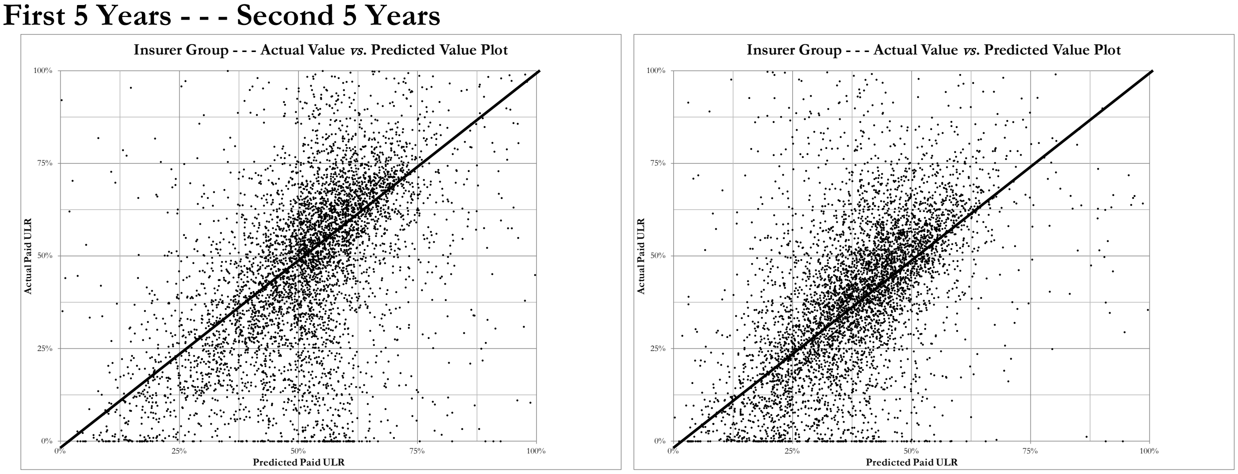


Figure 11: Actual *vs.* Predicted Paid ULR Plots AY 1997-2006, all LOB combined)

As can be seen in Figure 9, we show the plot the actual values against the predicted values. An ideal model would have all of its points lining up on the 45° line. We also show the residual plot for larger and smaller insurer separately. As can be seen from the graphs, the models seem equally valuable for both larger and smaller insurers. For smaller insurers, there is a cluster for predicted value that is apparent in both graphs. These clusters are due to the way missing covariates were treated. Clearly, the variance of the residuals is affected by the size of the insurers but, other than for the cluster of missing covariates, the conditional variance seems unaffected by the predicted value. Given that the predicted values are a form of an average, the increased variability of the residuals for smaller insurers is not unexpected[31]. Note also the quality of the fit seems equally good for the first 5 and second 5 years of the 10 year horizon that is considered.

As can be seen in Figure 10, we show the actual *versus* predicted relativity graphs for two sub-periods (the first 5 years and the second 5 years) and we find that the models appear equally valid for each sub-periods.

In Table 5, we also present the actual *versus* predicted Loss Ratios at the insurer-group level. Contrary to a future forecast, this back-testing exercise starts with the realized industry/line of business/Accident Year Ultimate Loss Ratio as its basis[32]. As can be seen, the fits are generally good. The author is unaware of a statistical study that would allow for a comparison with the performance of Loss Ratio projection methods that rely only on internal data. The author conjectures that internal budgets can be missed by several Loss Ratio points and so not only because of undue aggressiveness or conservatism. Interestingly, the estimator seems to perform even better when several years are compared together. Therefore, for most of the purposes mentioned in the introduction, the proposed forecasting methodology seems particularly relevant.

Figure 11 presents the actual *versus* predicted paid ULR by insurer group and Accident Year. We present it overall, for smaller and larger insurers separately, and for the first and last 5 years. The findings are similar to those found by line of business.

---

[31] In a follow-up to this project, this will be pursued.

[32] This leads us to schedule more appropriate out-of-sample performance testing for a further phase of the project. We intend to use (Frees, Meyers and Cummings, Summarizing Insurance Scores Using a Gini Index 2011) as a basis for that research.

7. ANALOGY TO CALIBRATION OF EXPERIENCE RATING PLANS

With these promising results in hand, we can now come back and comment on some guidance that can be given for the calibration of a more traditional Experience Rating plan.

Regarding the definition of an individual under the rating plan, we have proposed to use a definition of an individual as an entity that exercises control over activities that influence the loss potential: the insurer-group in our example.

Regarding the selection of the used number of years of experience, we have proposed to use semi-parametric predictive modeling[33] to make that selection but, just like would be done in practice, we have chosen not to use information past a certain age for practical reasons. Also, like in practice, we developed some rules to allow us to deal with missing information.

Regarding the issue of how to best reflect Loss Development, we proposed to always use the latest available valuation and use a definition of claim that makes the Loss Development pattern most similar across individuals.

Regarding the most adequate formula for the Experience Rating modification, we have departed from tradition to the extent that we have proposed a formula that did not incorporate explicit credibility considerations. We feel that, while credibility-type formulas can have the advantage of parsimony of the rating factors that need to vary by size of account, our coefficient-based approach can be quite parsimonious and has the benefit that it can be explained in terms of the more widely known regression framework. Nonetheless, even though our proposed models were inspired in part by fixed-effects regression, a traditional credibility interpretation is possible because an intercept term was always included. In this case, the complement of credibility is effectively always 1.0: that is, the overall average relativity. All the models can then always be re-written as $z \cdot \overline{rel_{i,t}} + (1-z) \cdot 1$, where $\overline{rel_{i,t}}$ refers to a weighted[34] average of past own line and other lines past Loss Ratio relativities and z is the credibility.

We have forgone commenting on the issue of trending and on-leveling[35], as well as on the issue of loss capping. Regarding the issue of trending and on-leveling, we recognize the value of creating an estimate of the losses and premium as if they were experienced in the current period.

---

[33] In that sense, our approach is not entirely unlike the one proposed by (Bailey and Simon 1959).
[34] Where the weights can vary with the size of the insurer.
[35] Because we used a relativity approach and because we allowed for a flexible structure for the residuals.

Regarding loss capping, we believe that the proposed predictive framework would be as valuable in the selection of the appropriate loss capping as it was to us in the selection of the loss experience horizon.

We have also proposed that the experience from other lines of business could potentially carry information. To our knowledge, that has not been commonly been incorporated in Experience Rating algorithms.

There remains the issue of whether an Experience Rating algorithm needs to (approximately) balance to a 1.0 relativity. The author mentions the issue because he is aware of many plans where the average debit/credit is not 0.0%. From a logical point of view, an Experience Rating scheme that does not balance to a 1.0 relativity implies that the classification rates are not adequate. Although this is not inconsistent as such, it implies incoherence in the rating algorithm. A non-balanced Experience Rating plan is more likely to occur if the when the credibility/size of account is correlated with the bias in the classification rates.

8. CONCLUSION

For this research project, we have chosen to present a close simile to the calibration of an experience rating scheme that could be used for Loss Ratio projection purposes for a party external to an insurer. Doing so allowed us to comment on practical modeling choices that would need to be made by a practicing actuary calibrating an experience rating scheme. We have departed from the traditional credibility-type approach to experience rating to instead anchor ourselves in a predictive modeling approach. We modeled the relativity to the industry Loss Ratio by Accident Year and Line of Business and found that, generally, (1) the own line of business past results were relevant predictors with factors varying by size of insurer, and (2) that past results in lines of business with similar clients, perils or strategies were also relevant predictors, again with factors potentially varying with the size of the insurer.

Acknowledgments

The author wishes to acknowledge Prof. Jed Frees and Claude Nadeau, ACAS, for their assistance on this project.

9. REFERENCES

Bailey, Robert A., and LeRoy J. Simon. "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car." *Proceedings of the Casualty Actuarial Society* XLVI (1959): 159-164.

Barrese, James, and Jack M. Nelson. "Independent and Exclusive Agency Insurers: A Reexamination of the Cost Differential." *Journal of Risk and Insurance* 59, no. 3 (1992): 375-397.

Berger, Allen N., J. David Cummins, and Mary A. Weiss. "The Coexistence of Multiple Distribution Systems for Financial Services: The Case of Property-Liability Insurance." *Journal of Business* 70, no. 4 (October 1997): 515-546.

Boor, Joseph A. "The Complement of Credibility." *Proceedings of the Casualty Actuarial Society* LXXXIII, no. 158 (May 1996): 1-32.

Couret, Jose, and Gary Venter. "Using Multi-Dimensional Credibilityto Estimate Class Frequency Vectors in Workers Compensation." *ASTIN Bulletin* 38, no. 1 (May 2008): 72-85.

Cummins, J. David, and Jack VanDerhei. "A Note on the Relative Efficiency of Property-Liability Insurance Distribution Systems." *Bell Journal of Economics* 10, no. 2 (Autumn 1979): 708-719.

Dorweiler, Paul. "A Survey of Risk Credibility in Experience Rating." *Proceedings of the Casualty Actuarial Society* XXI, no. I (November 1934): 90-114.

Frees, Edward W. *Longitudinal and panel data: analysis and applications in the social sciences.* New York, New York: Cambridge Univ Press, 2004.

Frees, Edward W., and Ping Wang. "Credibility Using Copulas." *North American Actuarial Journal* 9, no. 2 (2005): 31-48.

Frees, Edward W., Glenn Meyers, and A. David Cummings. "Dependent Multi-Peril Ratemaking." 2009.

Frees, Edward W., Glenn Meyers, and A. David Cummings. "Predictive Modeling of Multi-Peril Homeowners." 2011: 33.

Frees, Edward W., Glenn Meyers, and A. David Cummings. "Summarizing Insurance Scores Using a Gini Index." *Journal of the American Statistical Association* 106 (2011): 1085-1098.

Frees, Edward W., Virginia R. Young, and Yu Luo. "A Longitudinal Data Analysis Interpretation of Credibility Models." *Insurance: Mathematics and Economics* 24 (1999): 229-247.

Friedland, J. F. *Estimating Unpaid Claims Using Basic Techniques.* 3rd. Casualty Actuarial Society, 2010.

Gillam, William R. "Workers' Compensation Experience Rating: What Every Actuary Should Know." *Proceedings of the Casualty Actuarial Society* LXXIX (1992): 215-239.

Gillam, William R., and R. H. Snader. "Fundamentals of Individual Risk Rating, Part I." Study Note, National Council on Compensation Insurance, 1992.

Goulet, Vincent. "A Generalized Crossed Classification Credibility Model." *Insurance. Mathematics & Economics* 28, no. 2 (2001): 205-216.

Hansen, J. Ernest. "A Note on Full Credibility for Estimating Claim Frequency." *Proceedings of the Casualty Actuarial Society*, 1972: 51-56.

Insurance Services Office, Inc. "Commercial General Liability Experience and Schedule Rating Plan." 2006.

Klugman, Stuart A., Harry H. Panjer, and Gordon E. Willmot. *Loss Models: From Data to Decisions.* New York, New York: John Wiley and Sons, 1998.

Mahler, Howard C., and Curtis Gary Dean. "Credibility." In *Foundations of Casualty Actuarial Science*. Casualty Actuarial Society.

Mowbray, Albert H. "How Extensive a Payroll Exposure is Necessary to Give a Dependable Pure Premium." *Proceedings of the Casualty Actuarial Society* I, no. 1 (1914): 24-30.

National Council on Compensation Insurance. "Experience Rating Plan Manual for Workers Compensation and Employers Liability Insurance." 2007.

SAS. *PROC MIXED Statement.* http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_mixed_sect008.htm (accessed 05 21, 2012).

—. *REPEATED Statement.* http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_mixed_sect019.htm (accessed 05 21, 2012).

Shi, Peng, and Edward W. Frees. "Long-tail Longitudinal Modeling of Insurance Company Expenses." *Insurance: Mathematics and Economics* 47 (2010): 303-314.

Steele, J. Michael. *Stochastic Calculus and Financial Applications.* New York, NY: Springer, 2000.

Venter, Gary. "Experience Rating—Equity and Predictive Accuracy." *NCCI Digest* II, no. I (April 1987): 27-35.

Werner, Geoff, and Claudine Modlin. *Basic Ratemaking.* Casualty Actuarial Society, 2010.

Whitney, Albert W. "The Theory of Experience Rating." *Proceedings of the Casualty Actuarial Society* IV, no. 10 (May 1918): 274-292.

Wikipedia. *Autoregressive model.* http://en.wikipedia.org/wiki/Auto-regressive_process (accessed 05 20, 2012).

—. *Moving-average model.* http://en.wikipedia.org/wiki/Moving_average_model (accessed 05 20, 2012).

Biography of the Author

Marc-André Desrosiers is a Ph.D. candidate at UW-Madison in the Actuarial Science, Risk Management and Insurance program. He also completed his MBA at University of Calgary, after receiving his FCAS. Marc-André has studied Actuarial Mathematics and Philosophy at Concordia University, Montréal. The author also keeps contact with the industry as he is currently working as an external consultant for Intact Financial Corporation Actuarial Commercial Lines department. He is interested in pricing optimization, behavioral economics, customer behavior, and demand modeling. He can be joined at mdesrosiers@bus.wisc.edu.

APPENDIX 1. COEFFICIENTS OF THE PREDICTIVE MODELS

B_PPAL

Predictors	Estimate	Standard Error	t Value	Pr > |t|
intercept	0.35	0.06	6.05	< 0.01%
same line - lag 1	-0.26	0.16	-1.57	11.59%
M.V. Ind. - same line - lag 1	-0.05	0.09	-0.52	60.44%
interaction term - same line - lag 1	0.11	0.03	3.78	0.02%
same line - lag 2	-0.03	0.03	-0.98	32.65%
M.V. Ind. - same line - lag 2	0.05	0.06	0.89	37.45%
same line - lag 3	-0.01	0.01	-0.56	57.42%
M.V. Ind. - same line - lag 3	0.07	0.04	1.70	8.92%
same line - lag 4	0.03	0.02	1.11	26.62%
M.V. Ind. - same line - lag 4	-0.04	0.05	-0.83	40.90%
same line - lag 5	0.10	0.04	2.78	0.54%
M.V. Ind. - same line - lag 5	-0.08	0.04	-2.01	4.47%
line J_AUTP - lag 1	0.50	0.14	3.50	0.05%
M.V. Ind. - line J_AUTP - lag 1	-0.03	0.06	-0.52	60.05%
interaction term - line J_AUTP - lag 1	-0.08	0.02	-3.03	0.25%
line J_AUTP - lag 2	0.06	0.04	1.75	8.00%
M.V. Ind. - line J_AUTP - lag 2	0.00	0.03	-0.17	86.71%
Toeplitz(2)	2 239.29	111.22	20.13	< 0.01%
Toeplitz(3)	1 497.42	94.71	15.81	< 0.01%
Toeplitz(4)	1 108.23	83.62	13.25	< 0.01%
Toeplitz(5)	565.27	74.92	7.54	< 0.01%
Residual	3 808.87	115.84	32.88	< 0.01%

J_AUTP

Predictors	Estimate	Standard Error	t Value	Pr > |t|
intercept	0.06	0.02	3.44	0.06%
same line - lag 1	0.01	0.13	0.05	96.05%
M.V. Ind. - same line - lag 1	-0.15	0.05	-2.78	0.54%
interaction term - same line - lag 1	0.13	0.03	4.95	< 0.01%
same line - lag 2	0.52	0.13	3.92	< 0.01%
M.V. Ind. - same line - lag 2	-0.07	0.03	-2.21	2.69%
interaction term - same line - lag 2	-0.08	0.02	-3.21	0.14%
same line - lag 3	0.05	0.02	2.26	2.40%
M.V. Ind. - same line - lag 3	0.00	0.02	-0.09	92.53%
line B_PPAL - lag 1	0.05	0.02	2.33	1.97%
M.V. Ind. - line B_PPAL - lag 1	-0.02	0.02	-0.92	35.69%
line C_CA_L - lag 1	0.16	0.06	2.51	1.21%
M.V. Ind. - line C_CA_L - lag 1	0.01	0.01	2.15	3.20%
interaction term - line C_CA_L - lag 1	-0.03	0.01	-2.51	1.20%
Toeplitz(2)	148.40	30.99	4.79	< 0.01%
Residual	1 253.44	27.72	45.21	< 0.01%

A_HMOW

Predictors	Estimate	Standard Error	t Value	Pr > |t|
intercept	-0.21	0.18	-1.16	24.46%
interaction term - intercept	0.07	0.03	2.00	4.51%
same line - lag 1	0.28	0.32	0.87	38.70%
M.V. Ind. - same line - lag 1	0.27	0.07	4.18	< 0.01%
interaction term - same line - lag 1	-0.03	0.06	-0.45	65.07%
same line - lag 2	0.55	0.25	2.14	3.22%
M.V. Ind. - same line - lag 2	-0.04	0.06	-0.75	45.63%
interaction term - same line - lag 2	-0.08	0.05	-1.57	11.76%
same line - lag 3	0.19	0.05	3.47	0.05%
M.V. Ind. - same line - lag 3	-0.26	0.07	-3.94	< 0.01%
same line - lag 4	0.25	0.07	3.67	0.02%
M.V. Ind. - same line - lag 4	-0.15	0.05	-2.78	0.55%
line B_PPAL - lag 1	-0.92	0.35	-2.60	0.93%
M.V. Ind. - line B_PPAL - lag 1	-0.11	0.08	-1.52	12.97%
interaction term - line B_PPAL - lag 1	0.20	0.07	2.62	0.88%
line J_AUTP - lag 1	0.51	0.20	2.51	1.22%
M.V. Ind. - line J_AUTP - lag 1	0.11	0.08	1.38	16.92%
interaction term - line J_AUTP - lag 1	-0.09	0.04	-2.36	1.81%
line E_CMP_ - lag 1	-0.29	0.34	-0.85	39.51%
M.V. Ind. - line E_CMP_ - lag 1	0.08	0.05	1.48	14.03%
interaction term - line E_CMP_ - lag 1	0.06	0.07	0.90	36.73%
line E_CMP_ - lag 2	0.46	0.30	1.51	13.14%
M.V. Ind. - line E_CMP_ - lag 2	-0.10	0.05	-1.85	6.39%
interaction term - line E_CMP_ - lag 2	-0.09	0.06	-1.53	12.63%
Toeplitz(2)	728.22	111.06	6.56	< 0.01%
Residual	4 452.56	98.64	45.14	< 0.01%

D_WC__				
Predictors	Estimate	Standard Error	t Value	Pr > |t|
intercept	1.70	0.39	4.40	< 0.01%
interaction term - intercept	-0.26	0.07	-3.44	0.06%
same line - lag 1	-0.44	0.26	-1.69	9.15%
M.V. Ind. - same line - lag 1	-0.02	0.06	-0.34	73.04%
interaction term - same line - lag 1	0.20	0.06	3.43	0.06%
line C_CA_L - lag 1	-0.25	0.31	-0.80	42.22%
M.V. Ind. - line C_CA_L - lag 1	-0.11	0.04	-2.92	0.35%
interaction term - line C_CA_L - lag 1	0.04	0.06	0.65	51.89%
line E_CMP_ - lag 1	0.09	0.04	2.17	2.98%
M.V. Ind. - line E_CMP_ - lag 1	0.07	0.05	1.61	10.84%
Toeplitz(2)	2 968.58	219.98	13.49	< 0.01%
Toeplitz(3)	2 195.32	201.70	10.88	< 0.01%
Toeplitz(4)	1 841.19	180.14	10.22	< 0.01%
Toeplitz(5)	854.54	172.21	4.96	< 0.01%
Residual	8 365.16	235.52	35.52	< 0.01%

E_CMP_				
Predictors	Estimate	Standard Error	t Value	Pr > |t|
intercept	0.27	0.05	5.63	< 0.01%
same line - lag 1	0.24	0.06	3.95	< 0.01%
M.V. Ind. - same line - lag 1	0.07	0.10	0.68	49.38%
same line - lag 2	0.15	0.07	2.06	3.93%
M.V. Ind. - same line - lag 2	-0.09	0.10	-0.87	38.41%
same line - lag 3	0.16	0.07	2.34	1.94%
M.V. Ind. - same line - lag 3	0.06	0.09	0.68	49.75%
line C_CA_L - lag 1	-0.29	0.06	-4.61	< 0.01%
M.V. Ind. - line C_CA_L - lag 1	-0.05	0.04	-1.20	22.83%
interaction term - line C_CA_L - lag 1	0.06	0.01	5.30	< 0.01%
line D_WC__ - lag 1	0.12	0.04	3.42	0.06%
M.V. Ind. - line D_WC__ - lag 1	-0.09	0.03	-2.87	0.41%
line D_WC__ - lag 2	0.01	0.03	0.27	78.91%
M.V. Ind. - line D_WC__ - lag 2	0.11	0.06	1.86	6.36%
Toeplitz(2)	1 364.39	113.47	12.02	< 0.01%
Toeplitz(3)	246.03	115.11	2.14	3.26%
Toeplitz(4)	472.77	121.62	3.89	0.01%
Toeplitz(5)	502.16	100.27	5.01	< 0.01%
Residual	4 678.28	120.95	38.68	< 0.01%

H_OL_O				
Predictors	Estimate	Standard Error	t Value	Pr > |t|
intercept	-0.99	0.59	-1.66	9.74%
interaction term - intercept	0.24	0.11	2.12	3.44%
same line - lag 1	0.47	0.24	1.99	4.68%
M.V. Ind. - same line - lag 1	0.12	0.08	1.41	15.95%
interaction term - same line - lag 1	-0.05	0.04	-1.15	25.18%
same line - lag 2	0.27	0.26	1.03	30.43%
M.V. Ind. - same line - lag 2	0.01	0.07	0.20	84.13%
interaction term - same line - lag 2	-0.02	0.06	-0.41	68.47%
same line - lag 3	0.22	0.54	0.41	68.33%
M.V. Ind. - same line - lag 3	0.06	0.12	0.48	63.23%
interaction term - same line - lag 3	-0.03	0.11	-0.30	76.68%
same line - lag 4	0.29	0.41	0.72	47.30%
M.V. Ind. - same line - lag 4	-0.05	0.08	-0.59	55.41%
interaction term - same line - lag 4	-0.05	0.08	-0.61	53.90%
same line - lag 5	1.45	0.42	3.48	0.05%
M.V. Ind. - same line - lag 5	-0.37	0.26	-1.40	16.10%
interaction term - same line - lag 5	-0.28	0.08	-3.66	0.03%
line E_CMP_ - lag 1	0.29	0.50	0.58	56.24%
M.V. Ind. - line E_CMP_ - lag 1	0.21	0.13	1.58	11.37%
interaction term - line E_CMP_ - lag 1	-0.03	0.10	-0.35	72.43%
line D_WC__ - lag 1	-0.60	0.25	-2.43	1.51%
M.V. Ind. - line D_WC__ - lag 1	0.00	0.08	-0.02	98.29%
interaction term - line D_WC__ - lag 1	0.14	0.05	2.97	0.30%
Toeplitz(2)	5 631.71	473.06	11.90	< 0.01%
Toeplitz(3)	4 511.05	432.00	10.44	< 0.01%
Toeplitz(4)	2 531.73	406.25	6.23	< 0.01%
Toeplitz(5)	1 783.30	459.55	3.88	0.01%
Residual	20 914.00	506.42	41.30	< 0.01%

C_CA_L				
Predictors	Estimate	Standard Error	t Value	Pr > |t|
intercept	1.44	0.42	3.44	0.06%
interaction term - intercept	-0.21	0.08	-2.77	0.56%
same line - lag 1	-0.60	0.19	-3.22	0.13%
M.V. Ind. - same line - lag 1	0.08	0.11	0.71	47.66%
interaction term - same line - lag 1	0.12	0.04	3.07	0.22%
same line - lag 2	-0.61	0.20	-3.05	0.23%
M.V. Ind. - same line - lag 2	-0.12	0.06	-1.88	5.96%
interaction term - same line - lag 2	0.16	0.04	3.63	0.03%
same line - lag 3	0.07	0.03	2.27	2.35%
M.V. Ind. - same line - lag 3	-0.15	0.06	-2.42	1.54%
same line - lag 4	-0.02	0.02	-1.39	16.36%
M.V. Ind. - same line - lag 4	0.00	0.07	-0.02	98.18%
same line - lag 5	0.02	0.04	0.45	65.34%
M.V. Ind. - same line - lag 5	0.14	0.05	2.71	0.69%
line D_WC__ - lag 1	0.49	0.28	1.74	8.22%
M.V. Ind. - line D_WC__ - lag 1	-0.06	0.04	-1.28	19.94%
interaction term - line D_WC__ - lag 1	-0.09	0.06	-1.61	10.71%
line D_WC__ - lag 2	0.03	0.03	0.88	37.64%
M.V. Ind. - line D_WC__ - lag 2	0.02	0.04	0.63	52.76%
line J_AUTP - lag 1	0.22	0.05	4.78	< 0.01%
M.V. Ind. - line J_AUTP - lag 1	-0.03	0.06	-0.54	58.65%
line J_AUTP - lag 2	0.10	0.03	3.20	0.14%
M.V. Ind. - line J_AUTP - lag 2	-0.06	0.04	-1.66	9.63%
Toeplitz(2)	2 878.02	143.25	20.09	< 0.01%
Toeplitz(3)	1 611.96	110.61	14.57	< 0.01%
Toeplitz(4)	724.10	72.07	10.05	< 0.01%
Residual	4 738.42	155.58	30.46	< 0.01%

H_OL_C				
Predictors	Estimate	Standard Error	t Value	Pr > |t|
intercept	0.40	0.13	3.01	0.29%
same line - lag 1	0.11	0.07	1.59	11.10%
M.V. Ind. - same line - lag 1	0.43	0.14	3.00	0.28%
same line - lag 2	0.21	0.06	3.65	0.03%
M.V. Ind. - same line - lag 2	-0.23	0.14	-1.68	9.39%
line E_CMP_ - lag 1	-0.46	0.28	-1.69	9.19%
M.V. Ind. - line E_CMP_ - lag 1	-0.32	0.10	-3.08	0.21%
interaction term - line E_CMP_ - lag 1	0.11	0.05	2.15	3.20%
line D_WC__ - lag 1	0.13	0.08	1.51	13.16%
M.V. Ind. - line D_WC__ - lag 1	0.02	0.08	0.22	82.98%
Toeplitz(2)	7 970.58	631.92	12.61	< 0.01%
Toeplitz(3)	4 073.46	597.59	6.82	< 0.01%
Toeplitz(4)	3 048.24	499.47	6.10	< 0.01%
Toeplitz(5)	1 679.18	413.00	4.07	< 0.01%
Residual	16 577.00	689.27	24.05	< 0.01%

G_SL__				
Predictors	Estimate	Standard Error	t Value	Pr > |t|
intercept	-0.52	0.99	-0.53	59.94%
interaction term - intercept	0.21	0.22	0.98	32.77%
same line - lag 1	1.26	0.85	1.49	13.75%
M.V. Ind. - same line - lag 1	0.22	0.15	1.47	14.14%
interaction term - same line - lag 1	-0.22	0.19	-1.14	25.50%
same line - lag 2	-1.93	1.00	-1.93	5.41%
M.V. Ind. - same line - lag 2	-0.18	0.14	-1.29	19.86%
interaction term - same line - lag 2	0.44	0.22	1.95	5.12%
same line - lag 3	1.37	0.50	2.72	0.67%
M.V. Ind. - same line - lag 3	0.06	0.09	0.72	47.27%
interaction term - same line - lag 3	-0.26	0.11	-2.47	1.38%
same line - lag 4	0.97	0.36	2.68	0.75%
M.V. Ind. - same line - lag 4	-0.10	0.10	-1.05	29.22%
interaction term - same line - lag 4	-0.19	0.07	-2.64	0.84%
same line - lag 5	-0.65	0.52	-1.23	21.79%
M.V. Ind. - same line - lag 5	-0.03	0.12	-0.26	79.73%
interaction term - same line - lag 5	0.12	0.10	1.16	24.51%
line D_WC__ - lag 1	0.58	0.52	1.12	26.26%
M.V. Ind. - line D_WC__ - lag 1	-0.01	0.08	-0.14	88.87%
interaction term - line D_WC__ - lag 1	-0.10	0.11	-0.93	35.11%
line C_CA_L - lag 1	0.19	0.36	0.52	60.50%
M.V. Ind. - line C_CA_L - lag 1	0.11	0.07	1.49	13.65%
interaction term - line C_CA_L - lag 1	-0.04	0.07	-0.63	53.02%
Toeplitz(2)	5 108.87	638.02	8.01	< 0.01%
Toeplitz(3)	3 912.92	574.73	6.81	< 0.01%
Toeplitz(4)	2 129.37	529.33	4.02	< 0.01%
Toeplitz(5)	1 478.48	486.34	3.04	0.24%
Residual	14 068.00	659.67	21.33	< 0.01%

R_PL_O				
Predictors	Estimate	Standard Error	t Value	Pr > |t|
intercept	0.38	0.10	3.87	0.02%
same line - lag 1	0.49	0.39	1.26	20.78%
M.V. Ind. - same line - lag 1	0.21	0.17	1.25	21.25%
interaction term - same line - lag 1	-0.09	0.08	-1.13	26.06%
same line - lag 2	0.13	0.05	2.32	2.08%
M.V. Ind. - same line - lag 2	-0.30	0.17	-1.75	8.03%
same line - lag 3	0.23	0.06	3.58	0.04%
M.V. Ind. - same line - lag 3	-0.18	0.23	-0.80	42.66%
same line - lag 4	0.16	0.04	4.13	< 0.01%
M.V. Ind. - same line - lag 4	0.34	0.37	0.94	34.74%
Toeplitz(2)	9 623.39	1 043.54	9.22	< 0.01%
Toeplitz(3)	5 925.52	875.16	6.77	< 0.01%
Toeplitz(4)	1 632.53	684.78	2.38	1.71%
Residual	19 328.00	1 091.68	17.70	< 0.01%

R_PL_C				
Predictors	Estimate	Standard Error	t Value	Pr > |t|
intercept	3.59	2.05	1.75	8.69%
interaction term - intercept	-0.76	0.47	-1.62	10.80%
line H_OL_O - lag 1	1.81	0.85	2.13	3.47%
M.V. Ind. - line H_OL_O - lag 1	0.71	0.17	4.07	< 0.01%
interaction term - line H_OL_O - lag 1	-0.37	0.18	-2.06	4.08%
line H_OL_O - lag 2	0.21	0.13	1.57	11.81%
M.V. Ind. - line H_OL_O - lag 2	-0.54	0.13	-4.31	< 0.01%
line D_WC__ - lag 1	2.07	1.31	1.58	11.64%
M.V. Ind. - line D_WC__ - lag 1	-0.15	0.16	-0.93	35.13%
interaction term - line D_WC__ - lag 1	-0.41	0.29	-1.40	16.34%
line C_CA_L - lag 1	-6.59	2.02	-3.26	0.13%
M.V. Ind. - line C_CA_L - lag 1	0.62	1.21	0.51	60.73%
interaction term - line C_CA_L - lag 1	1.62	0.50	3.22	0.15%
line C_CA_L - lag 2	-0.07	0.03	-2.15	3.32%
M.V. Ind. - line C_CA_L - lag 2	-0.92	1.15	-0.79	42.79%
Toeplitz(2)	3 267.34	802.27	4.07	< 0.01%
Toeplitz(3)	2 923.03	608.49	4.80	< 0.01%
Residual	9 664.57	975.72	9.91	< 0.01%

F_MM_C				
Predictors	Estimate	Standard Error	t Value	Pr > |t|
intercept	0.17	0.20	0.84	40.07%
same line - lag 1	-1.35	0.42	-3.19	0.15%
M.V. Ind. - same line - lag 1	-0.01	0.06	-0.11	91.41%
interaction term - same line - lag 1	0.33	0.10	3.38	0.08%
same line - lag 2	0.07	0.06	1.20	23.22%
M.V. Ind. - same line - lag 2	-0.14	0.08	-1.75	8.13%
same line - lag 3	0.14	0.05	2.94	0.33%
M.V. Ind. - same line - lag 3	0.08	0.08	0.95	34.01%
same line - lag 4	0.16	0.07	2.43	1.51%
M.V. Ind. - same line - lag 4	-0.06	0.08	-0.76	44.78%
same line - lag 5	0.11	0.08	1.36	17.55%
M.V. Ind. - same line - lag 5	0.04	0.09	0.50	61.97%
line F_MM_O - lag 1	0.00	0.01	0.49	62.59%
M.V. Ind. - line F_MM_O - lag 1	-0.09	0.09	-0.99	32.39%
line F_MM_O - lag 2	-0.04	0.01	-2.78	0.55%
M.V. Ind. - line F_MM_O - lag 2	0.00	0.07	0.00	99.62%
line G_SL__ - lag 1	3.11	1.50	2.07	3.90%
M.V. Ind. - line G_SL__ - lag 1	0.18	0.15	1.19	23.39%
interaction term - line G_SL__ - lag 1	-0.72	0.35	-2.07	3.90%
line G_SL__ - lag 2	-2.31	1.42	-1.62	10.48%
M.V. Ind. - line G_SL__ - lag 2	-0.28	0.22	-1.28	20.24%
interaction term - line G_SL__ - lag 2	0.56	0.34	1.66	9.80%
line H_OL_C - lag 1	0.03	0.04	0.87	38.19%
M.V. Ind. - line H_OL_C - lag 1	0.09	0.05	1.77	7.70%
line H_OL_C - lag 2	0.06	0.03	2.09	3.68%
M.V. Ind. - line H_OL_C - lag 2	0.11	0.06	1.63	10.27%
Toeplitz(2)	7 865.46	671.63	11.71	< 0.01%
Toeplitz(3)	4 580.28	597.73	7.66	< 0.01%
Toeplitz(4)	2 637.96	485.50	5.43	< 0.01%
Toeplitz(5)	1 310.39	331.46	3.95	< 0.01%
Residual	12 901.00	697.13	18.51	< 0.01%

F_MM_O				
Predictors	Estimate	Standard Error	t Value	Pr > |t|
intercept	3.57	3.39	1.05	29.37%
interaction term - intercept	-1.05	0.74	-1.43	15.26%
same line - lag 1	7.34	1.79	4.09	< 0.01%
M.V. Ind. - same line - lag 1	0.93	0.50	1.85	6.45%
interaction term - same line - lag 1	-1.99	0.50	-4.00	< 0.01%
same line - lag 2	0.50	0.17	2.97	0.31%
M.V. Ind. - same line - lag 2	-1.28	0.50	-2.57	1.04%
same line - lag 3	0.45	0.25	1.81	7.14%
M.V. Ind. - same line - lag 3	0.32	0.38	0.84	40.31%
same line - lag 4	0.68	0.31	2.19	2.91%
M.V. Ind. - same line - lag 4	-0.54	0.35	-1.56	12.04%
line F_MM_C - lag 1	-7.76	2.23	-3.49	0.05%
M.V. Ind. - line F_MM_C - lag 1	0.90	0.53	1.70	9.03%
interaction term - line F_MM_C - lag 1	1.96	0.54	3.66	0.03%
line F_MM_C - lag 2	4.85	1.90	2.55	1.10%
M.V. Ind. - line F_MM_C - lag 2	0.14	0.37	0.39	69.89%
interaction term - line F_MM_C - lag 2	-1.17	0.42	-2.79	0.54%
line G_SL__ - lag 1	0.92	3.61	0.25	79.89%
M.V. Ind. - line G_SL__ - lag 1	0.72	0.38	1.90	5.82%
interaction term - line G_SL__ - lag 1	-0.12	0.90	-0.13	89.34%
line G_SL__ - lag 2	-11.22	3.62	-3.10	0.20%
M.V. Ind. - line G_SL__ - lag 2	-1.57	0.64	-2.46	1.42%
interaction term - line G_SL__ - lag 2	2.65	0.89	2.98	0.30%
line H_OL_O - lag 1	0.16	0.10	1.67	9.60%
M.V. Ind. - line H_OL_O - lag 1	1.13	0.40	2.84	0.47%
Toeplitz(2)	11 553.00	5 332.77	2.17	3.03%
Residual	115 396.00	5 959.67	19.36	< 0.01%

A Note On Mixed Distributions

Rasa Varanka McKean

October 23, 2012

1 Introduction

The author intends to outline and clarify a basic application of mixed distributions. The equations are based on a life insurance publication written more than fifty years ago. By a change in perspective, the same model can be applied to workers compensation insurance for the fitting of probability density curves to a mixture of injury types.[1]

The original life insurance research paper did not consider workers compensation as an application; we can in the following way. Our example will be cash flows and their stopping times for different workers compensation and employers liability injury types. In this model, the cash flow stops or fails when the claim is closed. The WC and EL application is not necessarily based on mortality tables. The claim can close when the employee is healed and returns to work.

It should be noted that although mixed distributions are in use countrywide for workers compensation business, the application described in this paper may or may not be the same as the countrywide model.

The basic equations for the life insurance model are taken from statistical methods in the testing of failure rates. The failures can be due to a variety of causes. As one example, think of a group of cohorts in health insurance, each group of claimants having a certain illness. As another example, think of a population of automobiles, each failing due to a mechanical failure, electrical failure or normal deterioration.

First, we address basic notation.

Consider a mixture of failure sub populations. Denote the number of sub populations by the variable s. There will be s=5 different types of claims in our model and Employers Liability claims also, in a separate 6th sub population. Let r_i be the number of units belonging to the i^{th} sub population. For example, the first sub population contains r_1 units, the second sub population contains r_2 units, the i^{th} sub population contains r_i units, and the last sub population contains r_s units. Given a random sample of n units, the failure of r_1 units is due to cause (1), r_2 units fail due to cause (2), and so on up to r_s.

[1]"Estimation of Parameters of Mixed Exponentially Distributed Failure Time Distributions from Censored Life Test Data" by William Mendenhall and R.J. Hader Source: Biometrika Vol. 45 No. 3/4 (Dec., 1958) pp. 504-520

A random sample of n units is tested up to time $t = T$. Then $\Sigma_{i=1}^{s} r_i = r$ is the total number of units failing before time T and $(n - r)$ units which can't be identified as to sub population survive the test. The data is similar to censored life data.

In visual terms, think of the size of loss distribution as a matrix. Column (1) shows the fatalities, column (2) shows permanent total claims, column (3) all permanent partial, column (4) temporary total, column (5) medical only claims; here an additional column (6) will be included for Employers Liability. The rows of the matrix are the loss limits. The loss limits can start as low as $5,000$ and end as high as $10,000,000$. Note that the subscript i refers to the columns, the sub populations.

The matrix is populated with the number of claims by injury type whose ultimate payout is the size of the loss limit. The cash flow stops or fails when the claim is closed and the loss has reached a limit. We make one assumption to adapt property and casualty insurance to this model, that the claim incurred amounts increase with time. We'll ignore subrogation or other types of reimbursement. In other words, the claims are at their ultimate value. Recall from page 2 that the survival function $G(x)$ accounts for IBNR claims.

Back to notation.

Denote the failure times for the i^{th} sub population by t_{ij}. Then the r_i claims which close in sub population (1) can be ordered as t_{11}, $t_{12}, \ldots,$ $t_{1j}, \ldots,$ t_{1r_1}. In other words, there are r_1 claims in sub population (1) and they close in a certain order in time. The r_2 claims in the second sub population can be ordered as t_{21}, $t_{22}, \ldots,$ $t_{2j}, \ldots,$ t_{2r_2}. The i^{th} sub population contains r_i claims ordered as t_{i1}, $t_{i2}, \ldots,$ $t_{ij}, \ldots,$ t_{ir_i}.

Let the sub populations be mixed in proportions p_1, $p_2, \ldots,$ p_s. The p_i are constants.

Note that the number of different ways the claims can be ordered is: $\dfrac{n!}{r_1!r_2!\ldots r_i!\ldots r_s!(n-r)!}$

To simplify the computation, define a new variable $x_{ij} = t_{ij}/T$. Recall that T is the total allotted time for the experiment and that the t_{ij} are times to failure for each individual cash flow. Necessarily, each t_{ij} is less than the total time T. Then each x_{ij} is less than unity.

Now consider an arbitrary cumulative distribution function $F_i(x)$, the associated density function $f_i(x)$, and the survival function $G_i(x) = 1 - F_i(x)$. This general CDF can be either the exponential, the Weibull or the log normal distributions. Let $F(x) = \Sigma_{i=1}^{s} p_i F_i(x)$ and $G(x) = 1 - F(x)$. In the model below, the survival function $G(x)$ will account for the IBNR claims.

It should be noted that we are accustomed to thinking of $F(1) = 1$ and $G(1) = 0$ for $F(x)$ and $G(x)$ valued at $x = 1$. Here it isn't true because the x_{ij} have a maximum value of unity

As an example, consider the exponential distribution: $F_i(x) = 1 - \exp[-x/\beta_i]$

$$G(1) = 1 - \Sigma_{i=1}^{s} p_i F_i(1) = 1 - \Sigma_{i=1}^{s} p_i + \Sigma_{i=1}^{s} \exp[(-1/\beta_i))] = \Sigma_{i=1}^{s} \exp[(-1/\beta_i)]$$

since $\Sigma_{i=1}^{s} p_i = 1$. Thus $G(1)$ is not equal to zero.

2 The Basic Theory

We consider an arbitrary cumulative distribution function in this section. The calculation of the likelihood function will be clearer without specific detail. Some of the terms in the numerator and denominator of the likelihood function will cancel. The cancellations will be seen more clearly if detail is left out.

In Sections 3, 4, and 5, we consider examples of the mixed exponential, the mixed Weibull, and the mixed log normal distributions. The basic theory holds for an arbitrary CDF.

We need:

1. The probability of the ordered sequences of failure times,
2. The joint probability density functions,
3. The conditional probability of the ordered observations,
4. The likelihood function, and
5. The maximum likelihood estimates of the parameters.

The formula for the probability of the ordered sequences includes the number of possible ordered sequences, the probability of failure for claims in each of the sub populations, and the survival probability at time T for claims still open at the end of the experiment. The probability is evaluated at time $x = t/T$ for $t = T$. The value of x is then $x = 1$.

Given a random sample of n units comprised of i sub populations and total number of claims $r = r_1 + r_2 + \ldots + r_i + \ldots + r_s$, the probability that r_1 units will fail due to cause (1), that r_2 units will fail due to cause (2), that r_i units will fail due to cause (i), and that $(n - r)$ units will survive the test is given by a multinomial distribution.

Denote the above probability by $P(r_1, r_2, \ldots, r_s|n)$ then for $x = 1$ at time T:

$$P(r_1, r_2, \ldots, r_s|n) = \frac{n!}{r_1!r_2! \ldots r_s!(n-r)!} \prod_{i=1}^{s} [p_i F_i(1)]^{r_i} [G(1)]^{(n-r)} \tag{1}$$

At this point, the reader may want to review the joint density functions of order statistics. Good references may be found in the CAS Exam 1 syllabus. Recall that the joint probability density function is equal to the product of the density functions if and only if the random variables are independent.

Now we select the i^{th} sub population conditional on the event that there are r_i claimants in that sub population in order to derive a likelihood equation.

Denote the conditional probability distribution by $P(x_{i1}, x_{i2}, ..., x_{ir_i}|r_i)$ and the conditional probability density function by $p(x_{i1}, x_{i2}, ..., x_{ir_i}|r_i)$.

For the i^{th} sub population the joint density function for the ordered statistics conditional on the probability of r_i claimants in the i^{th} sub population at the end of the experiment is:

$$p(x_{i1}, x_{i2}, ..., x_{ir_i}|r_i) = r_i! \prod_{j=1}^{r_i} f_i(x_{ij})/[F_i(1)]^{r_i}$$

The joint conditional density for all of the s sub populations is the product of the s sub populations:

$$\prod_{i=1}^{s} p(x_{i1}, x_{i2}, ..., x_{ir_i}|r_i) = \prod_{i=1}^{s} r_i! \prod_{j=1}^{r_i} f_i(x_{ij})/[F_i(1)]^{r_i} \tag{2}$$

The likelihood function is the product of equations (1) and (2) above:

$$p(r_1, r_2, \ldots, r_s|n) \prod_{i=1}^{s} p(x_{i1}, x_{i2}, ..., x_{ir_i}|r_i) =$$

$$\frac{n!}{r_1!r_2!\ldots r_s!(n-r)!} \prod_{i=1}^{s} [p_i F_1(1)]^{r_i} [[G(1)]^{(n-r)} \prod_{i=1}^{s} r_i! \prod_{j=1}^{r_i} f_i(x_{ij})/[F_i(1)]^{r_i}$$

Notice that the terms $[F_i(1)]^{r_i}$ in both numerator and denominator cancel. The same is true for the product of the $r_i!$.

We are left with the likelihood and the log likelihood functions:

$$L = \frac{n!}{(n-r)!} \prod_{i=1}^{s} p_i^{r_i} \prod_{j=1}^{r_i} f_i(x_{ij}) [G(1)]^{(n-r)} \tag{3}$$

$$\ln L = \ln \frac{n!}{(n-r)!} + \Sigma_{i=1}^{s} p_i^{r_i} + \Sigma_{j=1}^{r_i} \ln f_i(x_{ij}) + (n-r) \ln[G(1)] \tag{4}$$

The p_i will be redefined here to clarify their relationship in the curve fitting process. It's important to note that there is no effect on the model or the calculations. As we will see in the maximum likelihood examples below, redefining the p_i clarifies that the most weight in the tail is given to the serious injury types. For instance, at some point in the actuarial data, the proportion p_2 of permanent total claims dominates the other injury type weights. Stay tuned...

Thus, we define a new functional form for the proportions p_i. Define $p_i(x) = p_i G_i(x)/G(x)$ for $G(x)$ the survival function as defined in the Introduction above. The proportions of the injury type curves now depend on the survival function. Define $k_i = p_i(1) = p_i G_i(1)/G(1)$.

At this point, maximum likelihood estimates will be computed for specific examples. The MLEs depend on the number of parameters in the distribution and we continue with specific forms of the equations.

3 The Mixed Exponential Distribution

Take the case of the mixed exponential distribution with $s = 2$ sub populations.

The number of distinct sequences of claims is $\dfrac{n!}{r_1!r_2!(n-r)!}$

The density functions are $f_1(x_{1j}) = (1/\beta_1)exp[-(x_{1j}/\beta_1)]$ and $f_2(x_{2j}) = (1/\beta_2)exp[-(x_{2j}/\beta_2)]$

The CDFs are $F_1(x_{1j}) = 1 - exp[-(x_{1j}/\beta_1)]$ and $F_2(x_{2j}) = 1 - exp[-(x_{2j}/\beta_2)]$

The survival functions are $G_1(x_{1j}) = exp[-(x_{1j}/\beta_1)]$ and $G_2(x_{2j}) = exp[-(x_{2j}/\beta_2)]$

Given a random sample of n units comprised of two sub populations and total number of claims $r = r_1 + r_2$, the probability that r_1 units will fail due to cause (1), that r_2 units will fail due to cause (2), and that $(n - r)$ units will survive the test is given by a multinomial distribution.

Denote the above conditional probability by $P(r_1, r_2|n)$ then for $x = 1$ at time T:

$$P(r_1, r_2|n) = \frac{n!}{r_1!r_2!(n-r)!}[p_1F_1(1)]^{r_1}[p_2F_2(1)]^{r_2}[G(1)]^{(n-r)} \tag{5}$$

The joint distribution in this example before conditioning is given by the product of the $f_1(x_{1j})$ and the $f_2(x_{2j})$ for the two sub populations $i = 1, 2$ and for all j.

For the 1^{st} and 2^{nd} sub populations the respective joint conditional density functions are:

$$p(x_{11}, x_{12}, ..., x_{1r_1}|r_1) = r_1!\prod_{j=1}^{r_1} f_1(x_{1j})/[F_1(1)]^{r_1} \tag{6}$$

$$p(x_{21}, x_{22}, ..., x_{2r_2}|r_2) = r_2!\prod_{j=1}^{r_2} f_2(x_{2j})/[F_2(1)]^{r_2} \tag{7}$$

Taking the product of the above three equations (5), (6) and (7) yields the likelihood function and the log likelihood function:

$$L = p(r_1, r_2|n)\prod_{i=1}^{2} p(x_{i1}, x_{i2}, ..., x_{ir_i}|r_i) = \frac{n!}{(n-r)!}\prod_{i=1}^{s} p_i^{r_i}\prod_{j=1}^{r_1} f_1(x_{1j})\prod_{j=1}^{r_2} f_2(x_{2j})[G(1)]^{(n-r)} \tag{8}$$

$$\ln L = \ln\frac{n!}{(n-r)!} + \Sigma_{i=1}^{s} r_i \ln p_i + \Sigma_{j=1}^{r_1}\ln f_1(x_{1j}) + \Sigma_{j=1}^{r_2}\ln f_2(x_{2j}) + (n-r)\ln[G(1)] \tag{9}$$

In order to derive maximum likelihood parameters, start by taking the partial derivative of the log likelihood function with respect to the first parameter.

$$\frac{\partial \ln L}{\partial \beta_1} = \frac{\partial \ln \frac{n!}{(n-r)!}}{\partial \beta_1} + \Sigma_{i=1}^{s} \frac{\partial r_i \ln p_i}{\partial \beta_1} + \Sigma_{j=1}^{r_1} \frac{\partial \ln f_1(x_{1j})}{\partial \beta_1} + \Sigma_{j=1}^{r_1} \frac{\partial \ln f_2(x_{2j})}{\partial \beta_1} + \frac{\partial (n-r) \ln[G(1)]}{\partial \beta_1} \tag{10}$$

Note that the first term in equation (9), the factorial, is a constant. The derivative of a constant is zero and the first term in equation (10) will disappear. The same is true for the second term since the p_i are constants. The derivative of the function $f_2(x_{2j})$ with respect to β_1 will also disappear since it is a function of β_2 but not β_1.

The following terms in the derivative of the log likelihood function with respect to β_1 remain:

$$\frac{\partial \ln L}{\partial \beta_1} = \Sigma_{j=1}^{r_1} \frac{\partial \ln f_1(x_{1j})}{\partial \beta_1} + (n-r) \frac{\partial \ln[G(1)]}{\partial \beta_1} \tag{11}$$

Consider the first term in equation (11):

$$f_1(x_{1j}) = \frac{1}{\beta_1} \exp[-(\frac{x_{ij}}{\beta_1})]$$

$$\ln f_1(x_{1j}) = -\ln \beta_1 - \frac{x_{1j}}{\beta_1}$$

$$\frac{\partial \ln f_1(x_{1j})}{\partial \beta_1} = -\frac{1}{\beta_1} + \frac{x_{1j}}{\beta_1^2}$$

$$\Sigma_{j=1}^{r_1} \frac{\partial \ln f_1(x_{1j})}{\partial \beta_1} = \Sigma_{j=1}^{r_1} - (\frac{1}{\beta_1}) + \Sigma_{j=1}^{r_1} (\frac{x_{1j}}{\beta_1^2}) = -\frac{r_1}{\beta_1} + (-\frac{r_1}{\beta_1^2}) \Sigma_{j=1}^{r_1} (\frac{x_{1j}}{r_1})$$

$$\Sigma_{j=1}^{r_1} \frac{\partial \ln f_1(x_{1j})}{\partial \beta_1} = -\frac{r_1}{\beta_1} + (\frac{r_1}{\beta_1^2}) \bar{x_1} \tag{12}$$

where $\bar{x_1}$ is the average of the r_1 values x_{1j}.

Consider the second term in the derivative of the log likelihood equation:

$$(n-r) \frac{\partial \ln[G(1)]}{\partial \beta_1} = (n-r) \frac{\partial \ln[1 - \Sigma_{i=1}^{2} p_i F_i(1)]}{\partial \beta_1} = \frac{(n-r)(p_1/\beta_1^2) \exp[-\frac{1}{\beta_1}]}{(p_1 \exp[-\frac{1}{\beta_1}] + p_2 \exp[-\frac{1}{\beta_2}])} \tag{13}$$

Substituting these results, equations (12) and (13) back into equation (11), we have so far:

$$\frac{\partial \ln L}{\partial \beta_1} = -\frac{r_1}{\beta_1} + (\frac{r_1}{\beta_1^2})\bar{x_1} + \frac{(n-r)(p_1/\beta_1^2)\exp[-\frac{1}{\beta_1}]}{(p_1\exp[-\frac{1}{\beta_1}]+p_2\exp[-\frac{1}{\beta_2}])}$$

To see the results more clearly, define the variable:

$$k_1 = \frac{(p_1)\exp[-\frac{1}{\beta_1}]}{(p_1\exp[-\frac{1}{\beta_1}] + p_2\exp[-\frac{1}{\beta_2}])} \tag{14}$$

Then we have the result:

$$\frac{\partial \ln L}{\partial \beta_1} = \frac{k(n-r)}{\beta_1^2} - \frac{r_1}{\beta_1} + \frac{r_1\bar{x_1}}{\beta_1^2} \tag{15}$$

We can compute the derivative with respect to β_2 in a similar way:

$$\frac{\partial \ln L}{\partial \beta_2} = -\frac{r_2}{\beta_2} + (\frac{r_2}{\beta_2^2})\bar{x_2} + \frac{(n-r)(p_2/\beta_2^2)\exp[-\frac{1}{\beta_2}]}{(p_1\exp[-\frac{1}{\beta_1}] + p_2\exp[-\frac{1}{\beta_2}])}$$

And since

$$k_2 = (1-k_1) = 1 - \frac{p_1\exp[-1/\beta_1]}{(p_1\exp[-1/\beta_1] + p_2\exp[-1/\beta_2])} = \frac{p_2\exp[-1/\beta_2]}{(p_1\exp[-1/\beta_1] + p_2\exp[-1/\beta_2])} \tag{16}$$

we can then write:

$$\frac{\partial \ln L}{\partial \beta_2} = \frac{(1-k)(n-r)}{\beta_2^2} - \frac{r_2}{\beta_2} + \frac{r_2\bar{x_2}}{\beta_2^2} \tag{17}$$

There remains one more derivative to take before setting the above derivatives in equations (15) and (17) equal to zero and solving for the optimal parameters. Note that there is a constraint in the system that the proportional amounts p_i add to unity when summed.

Recall from the introduction that the p_i will be redefined:

$p_i(x) = p_iG_i(x)/G(x)$, that $p_i(1) = p_iG_i(1)/G(1)$, and that $1 = \Sigma_{i=1}^s p_i$.

Recall also that for the exponential distribution: $F_i(x) = 1 - \exp[-x/\beta_i]$

Then, for the case of two sub populations where $s = 2$:

$$G(1) = 1-\Sigma_{i=1}^2 p_iF_i(1) = 1-\Sigma_{i=1}^2 p_i[1-\exp[-x/\beta_i]] = 1-\Sigma_{i=1}^2 p_i+\Sigma_{i=1}^2 p_i\exp[-\frac{1}{\beta_i}] = \Sigma_{i=1}^2 p_i\exp[-\frac{1}{\beta_i}] \tag{18}$$

$$k_i = p_i(1) = \frac{p_i G_i(1)}{G(1)} = \frac{p_i[1 - F_i(1)]}{G(1)} = \frac{p_i \exp -[\frac{1}{\beta_i}]}{\Sigma_{i=1}^{2} p_i \exp[-\frac{1}{\beta_i}]} \tag{19}$$

Before calculating the maximum likelihood equation in its entirety, firstly consider the term in the log likelihood equation (9) that involves $G(1)$. Utilizing equations (18) and (19):

$$\frac{\partial \ln G(1)}{\partial p_i} = \frac{1}{1 - F(1)} \frac{\partial[1 - p_1 F_1(1) - p_2 F_2(1)]}{\partial p_1} = \frac{1}{1 - F(1)} \frac{\partial[1 - p_1 F_1(1) - (1 - p_1) F_2(1)]}{\partial p_1}$$

$$\frac{\partial \ln G(1)}{\partial p_1} = \frac{[-F_1(1) + F_2(1)]}{1 - F(1)} = \frac{\exp(-\frac{1}{\beta_1}) - \exp(-\frac{1}{\beta_2})}{p_1 \exp(-\frac{1}{\beta_1}) + p_2 \exp(-\frac{1}{\beta_2})} = \frac{k_1}{p_1} - \frac{k_2}{p_2} \tag{20}$$

Now we'll compute the entire equation for $\frac{\partial \ln L}{\partial p_1}$ to reflect the constraint in the system of two sub populations.

$$\frac{\partial \ln L}{\partial p_1} = (n - r)\frac{\partial \ln G(1)}{\partial p_1} + \frac{\partial r_1 \ln p_1}{\partial p_1} + \frac{\partial r_2 \ln(1 - p_1)}{\partial p_1} = (n - r)[\frac{k_1}{p_1} - \frac{k_2}{p_2}] + \frac{r_1}{p_1} - \frac{r_2}{p_2} \tag{21}$$

Gathering the terms together from equations (15), (17), and (21), we can state that the system of maximum likelihood equations for two sub populations is the following. Each of these equations will be set to zero to derive the optimal set of parameters with a constraint:

$$\frac{\partial \ln L}{\partial \beta_1} = \frac{k_1(n-r)}{\beta_1^2} - \frac{r_1}{\beta_1} + \frac{r_1 \bar{x_1}}{\beta_1^2} = 0$$

$$\frac{\partial \ln L}{\partial \beta_2} = \frac{(1-k_1)(n-r)}{\beta_2^2} - \frac{r_2}{\beta_2} + \frac{r_2 \bar{x_2}}{\beta_2^2} = 0$$

$$\frac{\partial \ln L}{\partial p_1} = (n - r)[\frac{k_1}{p_1} - \frac{k_2}{p_2}] + \frac{r_1}{p_1} - \frac{r_2}{p_2} = 0$$

At this point, let's review which of the quantities are known and which are unknown.

The quantity n is the total number of cohorts in the population. The quantity r is the total number of known claims at the end of the experiment $t = T$. The quantities r_1 and r_2 are the number of known claims in the first and second sub populations respectively. The quantities $\bar{x}_i$ are the average values of the claims in each sub population. These are the known quantities.

The unknown quantities are the optimal β_i and the optimal p_i. At this point, we have three equations in three unknowns.

For the case of an arbitrary number of sub populations below, the seriously interested reader can work out similar equations, following the same steps and techniques as above.

Here we state the equations for an arbitrary number of sub populations:

$$L = \frac{n!}{(n-r)!}[G(1)]^{(n-r)} \prod_{i=1}^{s} p_i^{r_i} \prod_{i=1}^{s} \prod_{j=1}^{r_i} f_i(x_{ij}) \tag{22}$$

$$\ln L = \ln \frac{n!}{(n-r)!} + (n-r)\ln[G(1)] + \Sigma_{i=1}^{s} r_i \ln p_i + \Sigma_{i=1}^{s} \Sigma_{j=1}^{r_i} \ln f_i(x_{ij}) \tag{23}$$

$$\frac{\partial \ln L}{\partial \beta_i} = (n-r)\frac{\partial \ln[G(1)]}{\partial \beta_i} + \Sigma_{j=1}^{r_i} \frac{\partial \ln f_i(x_{ij})}{\partial \beta_i} = \frac{k_i(n-r)}{\beta_i^2} - \frac{r_i}{\beta_i} + \frac{r_i \bar{x}_i}{\beta_i^2} = 0 \tag{24}$$

$$\frac{\partial \ln L}{\partial p_i} = (n-r)\frac{\partial \ln G(1)}{\partial p_i} + \frac{\partial r_i \ln p_i}{\partial p_i} + \frac{\partial r_s \ln(1 - \Sigma_{i=1}^{s-1} p_i)}{\partial p_i} = (n-r)[\frac{k_i}{p_i} - \frac{k_s}{p_s}] + \frac{r_i}{p_i} - \frac{r_s}{p_s} = 0 \tag{25}$$

Note that the equation for $\frac{\partial \ln L}{\partial p_i}$ holds for $i = 1, 2, \ldots, (s-1)$. The partial derivative with respect to p_s has been eliminated by the constraint $\Sigma_{i=1}^{s} p_i = 1$.

4 The Mixed Weibull Distribution

The probability density function, cumulative distribution function, and survival function for the Weibull distribution differs slightly in the exponential. Each of the functions is shown below.

The Weibull density functions are $f_i(x_{ij}) = (c_i/\beta_i)(x_{ij}/\beta_i)^{(c_i-1)} exp[-(x_{1j}/\beta_1)^{c_i}]$

The Weibull CDFs are $F_i(x_{ij}) = 1 - \exp[-(x_{ij}/\beta_i)^{c_i}]$

The Weibull survival functions are $G_i(x_{ij}) = \exp[-(\frac{x_{ij}}{\beta_i})^{c_i}]$ and $G(1) = \Sigma_{i=1}^{s} p_i \exp[-(1/\beta_i)^{c_i}]$

The likelihood and log likelihood functions are the same as before but the exact form of the density and survival functions will differ:

$$L = p(r_1, r_2, \ldots, r_i|n) \prod_{i=1}^{s} p(x_{i1}, x_{i2}, ..., x_{ir_i}|r_i) = \frac{n!}{(n-r)!} \prod_{i=1}^{s} p_i^{r_i} \prod_{i=1}^{s} \prod_{j=1}^{r_i} f_i(x_{ij})[G(1)]^{(n-r)}$$

$$\ln L = \ln \frac{n!}{(n-r)!} + \Sigma_{i=1}^{s} r_i \ln p_i + \Sigma_{i=1}^{s} \Sigma_{j=1}^{r_i} \ln f_i(x_{ij}) + (n-r)\ln[G(1)]$$

The derivatives with respect to β_i and p_i for both the Weibull and the exponential are similar and will seem familiar to the reader. However, as we will see, the derivative with respect to c_i is very different for the Weibull than for the exponential. In practical terms, the implementation will be more difficult.

Firstly, we'll take the derivative with respect to β_i:

$$\frac{\partial \ln L}{\partial \beta_i} = (n-r)\frac{\partial \ln[G(1)]}{\partial \beta_i} + \Sigma_{j=1}^{r_i}\frac{\partial \ln f_i(x_{ij})}{\partial \beta_i}$$

$$= \frac{(n-r)}{1-F(1)}\frac{\partial \Sigma_{i=1}^{s} p_i \exp[-(\frac{1}{\beta_i})^{c_i}]}{\partial \beta_i} + \Sigma_{j=1}^{r_i}\frac{\partial}{\partial \beta_i}[\ln c_i - \ln \beta_i + (c_i-1)(\ln x_{ij} - \ln \beta_i) - (\frac{x_{ij}}{\beta_i})^{c_i}]$$

$$= (n-r)\frac{p_i(\frac{c_i}{\beta_i})(\frac{1}{\beta_i})^{c_i}\exp[-(\frac{1}{\beta_i})^{c_i}]}{\Sigma_{i=1}^{s} p_i \exp[-(\frac{1}{\beta_i}^{c_i})]} + \Sigma_{j=1}^{r_i}[-\frac{1}{\beta_i} - (\frac{c_i-1}{\beta_i}) + (\frac{c_i}{\beta_i})(\frac{x_{ij}}{\beta_i})^{c_i}]$$

As before, define

$$k_i = \frac{p_i \exp[-(\frac{1}{\beta_i})^{c_i}]}{\Sigma_{i=1}^{s} p_i \exp[-(\frac{1}{\beta_i}^{c_i})]}$$

Then:

$$\frac{\partial \ln L}{\partial \beta_i} = \frac{c_i k_i (n-r)}{\beta_i^{c_i+1}} - \frac{c_i r_i}{\beta_i} + (\frac{c_i}{\beta_i^{c_i+1}})\Sigma_{j=1}^{r_i} x_{ij}^{c_i} \tag{26}$$

Secondly, we'll take the derivative with respect to c_i:

$$\frac{\partial \ln L}{\partial c_i} = (n-r)\frac{\partial \ln[G(1)]}{\partial c_i} + \Sigma_{j=1}^{r_i}\frac{\partial \ln f_i(x_{ij})}{\partial c_i}$$

$$= \frac{(n-r)}{1-F(1)}\frac{\partial \ln \Sigma_{i=1}^{s} p_i \exp[-(\frac{1}{\beta_i})^{c_i}]}{\partial c_i} + \Sigma_{j=1}^{r_i}\frac{\partial}{\partial c_i}[\ln c_i - \ln \beta_i + (c_i-1)(\ln x_{ij} - \ln \beta_i) - (\frac{x_{ij}}{\beta_i})^{c_i}]$$

$$= (n-r)(\frac{1}{\beta_i})^{c_i}(\ln \beta_i)\frac{p_i \exp[-(\frac{1}{\beta_i})^{c_i}]}{\Sigma_{i=1}^{s} p_i \exp[-(\frac{1}{\beta_i}^{c_i})]} + \Sigma_{j=1}^{r_i}[\frac{1}{c_i} + (\ln x_{ij} - \ln \beta_i) - (\frac{x_{ij}}{\beta_i})^{c_i}\ln(\frac{x_{ij}}{\beta_i})]$$

$$= (n-r)(k_i)(\frac{1}{\beta_i})^{c_i}(\ln \beta_i) + \frac{r_i}{c_i} - r_i \ln \beta_i + \Sigma_{j=1}^{r_i}[\ln x_{ij} - (\frac{x_{ij}}{\beta_i})^{c_i}\ln(\frac{x_{ij}}{\beta_i})] \tag{27}$$

The derivative of the Weibull with respect to p_i is similar to the exponential since the density function is not a function of p_i.

$$\frac{\partial \ln L}{\partial p_i} = (n-r)\frac{\partial \ln G(1)}{\partial p_i} + \frac{\partial r_i \ln p_i}{\partial p_i} + \frac{\partial r_s \ln(1-\Sigma_{i=1}^{s-1} p_i)}{\partial p_i}$$
$$= (n-r)[\frac{k_i}{p_i} - \frac{k_s}{p_s}] + \frac{r_i}{p_i} - \frac{r_s}{p_s} \tag{28}$$

Here we summarize the maximum likelihood equations for the mixed Weibull distribution, referring back to equations (26), (27), and (28):

$$\frac{\partial \ln L}{\partial \beta_i} = \frac{c_i k_i (n-r)}{p_i \beta_i^{c_i+1}} - \frac{c_i r_i}{\beta_i} + (\frac{c_i}{\beta_i^{c_i+1}})\Sigma_{j=1}^{r_i} x_{ij}^{c_i}$$

$$\frac{\partial \ln L}{\partial c_i} = (n-r)(k_i)(\frac{1}{\beta_i})^{c_i}(\ln \beta_i) + \frac{r_i}{c_i} - r_i \ln \beta_i + \Sigma_{j=1}^{r_i}[\ln x_{ij} + (\frac{x_{ij}}{\beta_i})^{c_i} \ln(\frac{x_{ij}}{\beta_i})]$$

$$\frac{\partial \ln L}{\partial p_i} = (n-r)[\frac{k_i}{p_i} - \frac{k_s}{p_s}] + \frac{r_i}{p_i} - \frac{r_s}{p_s}$$

Again, note that the equation for $\frac{\partial \ln L}{\partial p_i}$ holds for $i = 1, 2, \ldots, (s-1)$.

5 The Mixed Log Normal Distribution

Each of the probability density functions, cumulative distribution functions, and survival functions for both the normal and the log normal distributions will be shown below for easy reference.

The normal density function is given in the standard notation:

$$\phi(t) = \frac{1}{\sqrt{2\pi}} \exp[-t^2/2]$$

By a change of variable, replacing t with $w = (t-\mu)/\sigma$ and replacing dt with $dw = \frac{1}{\sigma}dt$, the normal density function is:

$$\phi(\frac{t-\mu}{\sigma}) = \frac{1}{\sqrt{2\pi}} \exp[-\frac{(t-\mu)^2}{2\sigma^2}](\frac{1}{dt})d(\frac{t-\mu}{\sigma}) = \frac{1}{\sigma}\phi(t)$$

By a different change of variable, replacing t with $w = (\ln t - \mu)/\sigma$ and dt with $dw = \frac{1}{t\sigma}dt$, the log normal density function is:

$$\phi(\frac{\ln t - \mu}{\sigma}) = \frac{1}{\sqrt{2\pi}} \exp[-\frac{(\ln t - \mu)^2}{2\sigma^2}](\frac{1}{t\sigma dt})d(\frac{(\ln t - \mu)}{\sigma}) = \frac{1}{t\sigma}\phi(t) \tag{29}$$

Now consider the normal distribution function, which is the integral of the density function:

$$\Phi(x) = \int_{-\infty}^{x} \phi(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp^{[-t^2/2]}\, dt$$

Continuing in our notation with the change of variables above:

$$\Phi_i(\frac{x_{ij} - \mu_i}{\sigma_i}) = \int_{-\infty}^{(x_{ij}-\mu_i)/\sigma_i} \phi(t)dt = \frac{1}{\sigma_i} \int_{-\infty}^{w_{ij}} \phi(t)dt$$

$$\Phi_i(\frac{\ln x_{ij} - \mu_i}{\sigma_i}) = \int_{-\infty}^{(\ln x_{ij}-\mu_i)/\sigma_i} \phi(t)dt = \frac{1}{\sigma_i} \int_{-\infty}^{w_{ij}} \frac{1}{t}\phi(t)dt \tag{30}$$

Notice in the first equality of each of the two equations immediately above, that the integrand is the same as that for the normal distribution. What has changed is the upper limit of integration. In practical terms, to calculate the value of the log normal distribution function for a given value of x_{ij}, compute the value of $w_{ij} = (\ln x_{ij} - \mu_i)/\sigma_i$ and then look up w_{ij} in a table for the normal distribution. No additional tables are necessary for the log normal distribution function.

The survival function for the normal distribution is known as the Q-function in engineering textbooks. The survival function for the log normal distribution is expressed similarly:

$$Q(x) = \int_{x}^{\infty} \phi(t)dt = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} \exp^{[-t^2/2]}\, dt = 1 - \Phi(x)$$

$$Q(\frac{x_{ij} - \mu_i}{\sigma_i}) = \int_{(x_{ij}-\mu_i)/\sigma_i}^{\infty} \phi(t)dt = \frac{1}{\sqrt{2\pi}} \int_{(x_{ij}-\mu_i)/\sigma_i}^{\infty} \exp^{[-t^2/2]}\, dt = 1 - \Phi_i(\frac{x_{ij} - \mu_i}{\sigma_i})$$

$$Q(\frac{\ln x_{ij} - \mu_i}{\sigma_i}) = \int_{(\ln x_{ij}-\mu_i)/\sigma_i}^{\infty} \phi(t)dt = \frac{1}{\sqrt{2\pi}} \int_{(\ln x_{ij}-\mu_i)/\sigma_i}^{\infty} \exp^{[-t^2/2]}\, dt = 1 - \Phi_i(\frac{\ln x_{ij} - \mu_i}{\sigma_i}) \tag{31}$$

We'll continue the derivation focused only on the log normal distribution. The likelihood and log likelihood functions are the same as before but the exact form of the density and survival functions will differ for the log normal.

Recall that equations (3) and (4) give us the likelihood and log likelihood functions for the general mixed distribution. In the standard notation for the log normal density and distribution functions, the analogous equations are now:

$$L = \frac{n!}{(n-r)!}\prod_{i=1}^{s} p_i^{r_i} \prod_{i=1}^{s}\prod_{j=1}^{r_i} \phi(\frac{\ln x_{ij}-\mu_i}{\sigma_i})[Q(\frac{-\mu_i}{\sigma_i})]^{(n-r)} \tag{32}$$

$$\ln L = \ln\frac{n!}{(n-r)!} + \Sigma_{i=1}^{s} r_i \ln p_i + \Sigma_{i=1}^{s}\Sigma_{j=1}^{r_i} \ln\phi(\frac{\ln x_{ij}-\mu_i}{\sigma_i}) + (n-r)\ln[Q(\frac{-\mu_i}{\sigma_i})] \tag{33}$$

Where, as before: $\Phi(x) = \Sigma_{i=1}^{s} p_i \Phi_i(x)$ and $Q(x) = 1 - \Phi(x)$.

The partial derivatives are taken with respect to the variables μ_i and σ_i. As before, the derivatives of the first two terms in $\ln L$ vanish when the partials are taken. The first two terms in $\ln L$ contain factorials and the variables p_i but not the variables μ_i and σ_i.

$$\frac{\partial \ln L}{\partial \mu_i} = \Sigma_{j=1}^{r_i} \frac{\partial \ln \phi_i(\frac{\ln x_{ij}-\mu_i}{\sigma_i})}{\partial \mu_i} + (n-r)\frac{\partial \ln[Q(\frac{-\mu_i}{\sigma_i})]}{\partial \mu_i} \tag{34}$$

Consider the partial with respect to μ_i of the density function in the first summation. By the chain rule:

$$\frac{\partial \ln \phi(\frac{\ln x_{ij}-\mu_i}{\sigma_i})}{\partial \mu_i} = \frac{1}{\phi(\frac{\ln x_{ij}-\mu_i}{\sigma_i})}\frac{\partial \phi(\frac{\ln x_{ij}-\mu_i}{\sigma_i})}{\partial \mu_i} = \frac{1}{\phi(\frac{\ln x_{ij}-\mu_i}{\sigma_i})}[\frac{1}{\sqrt{2\pi}}][\frac{2(\ln x_{ij}-\mu_i)}{2\sigma_i^2}]\exp[\frac{-(\ln x_{ij}-\mu_i)^2}{2\sigma_i^2}]$$

$$\frac{\partial \ln \phi(\frac{\ln x_{ij}-\mu_i}{\sigma_i})}{\partial \mu_i} = \frac{(\ln x_{ij}-\mu_i)}{\sigma_i^2} \tag{35}$$

since the term $\phi(\frac{(\ln x_{ij}-\mu_i)}{\sigma_i})$ cancels from both the numerator and the denominator.

Similarly, for the partial with respect to σ_i:

$$\frac{\partial \ln \phi(\frac{\ln x_{ij}-\mu_i}{\sigma_i})}{\partial \sigma_i} = \frac{1}{\phi(\frac{\ln x_{ij}-\mu_i}{\sigma_i})}\frac{\partial \phi(\frac{\ln x_{ij}-\mu_i}{\sigma_i})}{\partial \sigma_i} = \frac{1}{\phi(\frac{\ln x_{ij}-\mu_i}{\sigma_i})}[\frac{1}{\sqrt{2\pi}}][\frac{(\ln x_{ij}-\mu_i)^2}{\sigma_i^3}]\exp[\frac{-(\ln x_{ij}-\mu_i)^2}{2\sigma_i^2}]$$

$$\frac{\partial \ln \phi(\dfrac{\ln x_{ij} - \mu_i}{\sigma_i})}{\partial \sigma_i} = \frac{(\ln x_{ij} - \mu_i)^2}{\sigma_i^3} \tag{36}$$

Before proceeding, recall the operation of differentiation under the integral sign. Don't feel bad about looking it up in Wikipedia if you don't remember the formula.

For the function $F(x)$, with the proper conditions of continuity and differentiability allowing us to interchange a derivative and an integral, we have from the fundamental theorem of calculus:

$$F(x) = \int_{a(x)}^{b(x)} f(x,t)dt$$

$$\frac{\partial F(x)}{\partial dx} = f(x, b(x))\frac{\partial b(x)}{\partial x} - f(x, a(x))\frac{\partial a(x)}{\partial x} + \int_{a(x)}^{b(x)} \frac{\partial f(x,t)}{\partial x} dt \tag{37}$$

Now consider the last term in equation (34), the term with the survival function. We'll see that keeping the integrand as a function of only the variable t is a definite advantage here. If the integrand is not a function of μ_i or σ_i then differentiation under the integral sign will be particularly easy since the partial with respect to the integrand will vanish.

$$Q(\frac{-\mu_i}{\sigma_i}) = 1 - \Phi(\frac{-\mu_i}{\sigma_i}) = 1 - \Sigma_{i=1}^s p_i \int_{-\infty}^{(-\mu_i/\sigma_i)} \phi(t)dt$$

For clarity, let's first compute:

$$\frac{\partial Q(\dfrac{-\mu_i}{\sigma_i})}{\partial \mu_i} = -\Sigma_{i=1}^s p_i \frac{\partial}{\partial \mu_i} \int_{-\infty}^{(-\mu_i/\sigma_i)} \phi(t)dt \tag{38}$$

Then, insert (38) into:

$$\frac{\partial \ln Q(\dfrac{-\mu_i}{\sigma_i})}{\partial \mu_i} = \frac{1}{Q(\dfrac{-\mu_i}{\sigma_i})} \frac{\partial Q(\dfrac{-\mu_i}{\sigma_i})}{\partial \mu_i} = \Sigma_{i=1}^s \frac{-p_i}{Q(\dfrac{-\mu_i}{\sigma_i})} \frac{\partial}{\partial \mu_i} \int_{-\infty}^{(-\mu_i/\sigma_i)} \phi(t)dt$$

The integrand and the lower limit of integration are not functions of μ_i. By equation (37), the differentiation reduces to that of the upper limit of integration:

$$\frac{\partial \ln Q(\frac{-\mu_i}{\sigma_i})}{\partial \mu_i} = \Sigma_{i=1}^{s} \frac{-p_i}{Q(\frac{-\mu_i}{\sigma_i})} \frac{\partial(-\mu_i/\sigma_i)}{\partial \mu_i} = \Sigma_{i=1}^{s} \frac{(p_i/\sigma_i)}{Q(\frac{-\mu_i}{\sigma_i})} \tag{39}$$

Similarly, taking the partial with respect to σ_i, yields:

$$\frac{\partial \ln Q(\frac{-\mu_i}{\sigma_i})}{\partial \sigma_i} = \Sigma_{i=1}^{s} \frac{-p_i}{Q(\frac{-\mu_i}{\sigma_i})} \frac{\partial(-\mu_i/\sigma_i)}{\partial \sigma_i} = \Sigma_{i=1}^{s} \frac{(-p_i/\sigma_i^2)}{Q(\frac{-\mu_i}{\sigma_i})} \tag{40}$$

Gathering the terms in equations (35) and (39), we have from equation (34), the derivative of the log likelihood function with respect to μ_i:

$$\begin{aligned} \frac{\partial \ln L}{\partial \mu_i} &= \Sigma_{j=1}^{r_i} \frac{\partial \ln \phi_i(\frac{\ln x_{ij} - \mu_i}{\sigma_i})}{\partial \mu_i} + (n-r)\frac{\partial \ln[Q(\frac{-\mu_i}{\sigma_i})]}{\partial \mu_i} \\ &= \Sigma_{j=1}^{r_i} \frac{1}{\sigma_i}(\frac{\ln x_{ij} - \mu_i}{\sigma_i}) + (n-r)\Sigma_{i=1}^{s} \frac{(p_i/q_i)}{Q(\frac{-\mu_i}{\sigma_i})} \end{aligned} \tag{41}$$

Gathering the terms in equations (36) and (40), we have the derivative of the log likelihood function with respect to σ_i:

$$\begin{aligned} \frac{\partial \ln L}{\partial \sigma_i} &= \Sigma_{j=1}^{r_i} \frac{\partial \ln \phi_i(\frac{\ln x_{ij} - \mu_i}{\sigma_i})}{\partial \sigma_i} + (n-r)\frac{\partial \ln[Q(\frac{-\mu_i}{\sigma_i})]}{\partial \sigma_i} \\ &= \Sigma_{j=1}^{r_i} \frac{1}{\sigma_i}(\frac{\ln x_{ij} - \mu_i}{\sigma_i})^2 + (n-r)\Sigma_{i=1}^{s} \frac{(-p_i/\sigma_i^2)}{Q(\frac{-\mu_i}{\sigma_i})} \end{aligned} \tag{42}$$

The derivative of the Log Normal with respect to p_i is similar to the exponential and Weibull functions since the density function is not a function of p_i.

$$\frac{\partial \ln L}{\partial p_i} = (n-r)\frac{\partial \ln Q(\frac{-\mu_i}{\sigma_i})}{\partial p_i} + \frac{\partial r_i \ln p_i}{\partial p_i} + \frac{\partial r_s \ln(1 - \Sigma_{i=1}^{s-1} p_i)}{\partial p_i}$$

$$= -(n-r)\Sigma_{i=1}^{s}\frac{1}{Q(\frac{-\mu_i}{\sigma_i})}[\frac{\partial}{\partial p_i}p_i\int_{-\infty}^{(-\mu_i/\sigma_i)}\phi_i(t)dt] + \frac{r_i}{p_i} - \frac{r_s}{p_s}$$

$$= -(n-r)\Sigma_{i=1}^{s}\frac{1}{Q(\frac{-\mu_i}{\sigma_i})}[\int_{-\infty}^{(-\mu_i/\sigma_i)}\phi_i(t)dt] + \frac{r_i}{p_i} - \frac{r_s}{p_s}$$

$$= -(n-r)[\Sigma_{i=1}^{s}\frac{\Phi_i(\frac{-\mu_i}{\sigma_i})}{Q(\frac{-\mu_i}{\sigma_i})}] + \frac{r_i}{p_i} - \frac{r_s}{p_s}$$

$$= -(n-r)[\Sigma_{i=1}^{s}\frac{\Phi_i(\frac{-\mu_i}{\sigma_i})}{1-\Sigma_{i=1}^{s}p_i\Phi_i(\frac{-\mu_i}{\sigma_i})}] + \frac{r_i}{p_i} - \frac{r_s}{p_s} \qquad (43)$$

The maximum likelihood equations (41), (42), and (43) for the mixed log normal distribution are a challenge. The integral in the formula of the distribution has no closed form solution. This integral appears in both the numerator and denominator of the summation in equation (43). The values $\Phi(x)$ can be approximated very accurately for asymptotic (large) values of x. However, equation (43) could involve several approximations at each step of the MLE iteration. Thus, the algorithm could be lengthy. Professional optimization software is highly advisable.

6 Summary

A model of mixed distributions pertinent to workers compensation insurance is adapted from life insurance. Maximum likelihood equations for the mixed exponential, mixed Weibull, and mixed log normal distributions are derived for the fitting of a mixture of probability density curves by injury type. Implementation of the model requires optimization software.