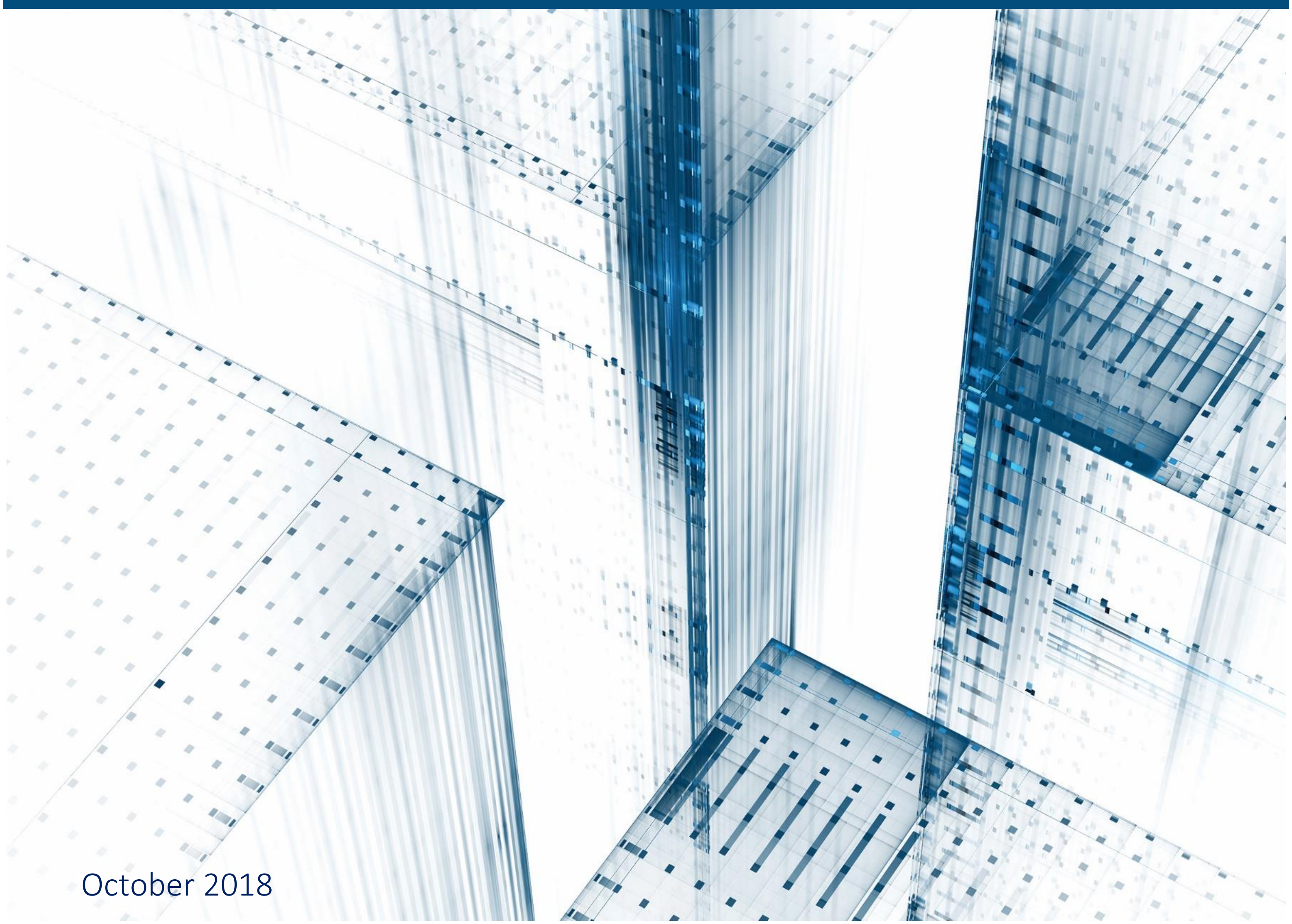
2018 Group Long-Term Disability Experience Study - New Variables and Predictive Analytics Applications



October 2018

2018 Group Long-Term Disability Experience Study

AUTHORS David Wall and Thomas Corcoran, Group Long-Term Disability Experience Committee
Mervyn Kopinsky, Society of Actuaries

TABLE OF CONTENTS

Section 1: Executive Summary

In November 2016 the SOA released the 2016 Group Long-Term Disability Experience Study Preliminary Report. Associated with this were five pivot tables that provided experience results in various combinations of the prior and new variables. The 2016 LTD Experience study contained many new variables that had not been included in prior LTD termination studies. Documentation of the specifications for all the variables was included.

In April 2018, the LTD Experience Study Committee released a report based on the analysis of the data contained in the pivot tables. This report utilized traditional analysis techniques and concentrated on three aspects of the experience:

- Trends since the 2008 study
- Insights into areas of new findings
- Experience for new variables; i.e., not in the 2008 Study.

Following this, the LTD experience study committee researched the practicality of developing a single Consolidated Data Base (CDB), that would include all the 2016 Study variables in one database. In addition, some new data fields were calculated using the submitted variables. The CDB was intended to allow companies to do their own analysis of the full complexities of the new data. It would also facilitate companies' use of Predictive Analytics (PA) techniques. The committee devised a design for the CDB that maintained individual contributor confidentiality. (The CDB, along with appropriate documentation, is in the process of being released by the SOA at the time of the writing of this report.)

A Work Group reporting to the Study Committee was formed to perform some analyses that the new CDB made feasible. In particular, the Work Group considered:

- Experience for variables not included in the previously released pivot tables
- Insights into potential drivers of termination experience trends observed in the 2016 Study
- Demonstrating how the CDB could be utilized in Predictive Analysis applications

The Work Group consisted of Mervyn Kopinsky of the Society of Actuaries and David Wall and Thomas Corcoran of Willis Towers Watson. Our report has 6 main sections:

1. Executive Summary - i.e., this section
2. Description of the CDB - the SOA created a CDB that contains all the data elements in the pivot tables, plus several new variables that were calculated from those. Section 2 describes the CDB and, in particular, the variables not included or described in the previously released pivot tables.
3. Analysis of New Variables – The CDB contains four variables not included in the pivot tables. Analysis of those shows some interesting results, summarized in section 3. In particular, COLA was noted to have a material impact on termination experience.
4. Trend Analysis – Traditional Approach – Section 4 describes efforts by the Work Group to analyze trend using traditional regression analysis techniques. Of note was that the SOA could analyze the impact of change in mix of business by contributor on trend while maintaining contributor confidentiality.
5. Predictive Analytics Methodology – The CDB is particularly well suited to PA techniques, whose main strength is the ability to analyze experience across large numbers of variables. One of the main objectives of the Work Group was to apply PA techniques to the CDB. However, since PA techniques have not been commonly used to analyze LTD terminations, we felt a detailed description of the process our Work Group used would be valuable. That is provided in Section 5; several valuable insights are also documented as part of this process.
6. Predictive Analytics Trend Attribution Analysis – Significant improvement in recovery A/E trend was evident when using the 2008 Study as the expected. PA techniques were used to evaluate whether changes in the mix of business were contributing to the apparent trend. The process is described in Section 6; it shows that approximately one half of the apparent trend in recoveries could be attributable to business mix changes. A similar process was used to analyze death trends; that showed that business mix changed appeared to have little impact.

Section 2: Description of the CDB

The CDB contains approximately 39 million unique data elements representing the combination of 24 different variables. The CDB employs the same general approach as the previously released pivot tables; i.e., it contains every combination of the 24 variables that had data. As with the pivot tables, no specific contributor's data can be identified or inferred.

The CDB includes data fields for 24 variables; 20 of these were included in the previously released pivot tables; four are newly available. A summary of the variables is provided in Table 1 below. More detailed information may be found on the SOA website.

Table 1: List of Consolidated Data Base Variables

	2016 GLTD Experience Study - Description of Consolidated Data Base			Inclusion in 2016 GLTD
	data fields	source	how treated	Pivots A-E
1	Company size (S, M, L)	submitted&grouped	new in 2016 study	A,C,D,E
2	[Elimination_Period] int	submitted	in 2008 LTD study	A,B,D,E
3	[Calendar_Year] [varchar](10)	submitted	in 2008 LTD study	A,B,C,D,E
4	[Duration] int	submitted	in 2008 LTD study	A,B,D
5	[Age_at_Disability] int	submitted	in 2008 LTD study	A,B,C,D,E
6	[Disability_Category] [varchar](50)	submitted&grouped	in 2008 LTD study	A,B,D,E
7	[Limited_Own_Occupation_Period] [varchar](50)	submitted	in 2008 LTD study	A,C
8	[OwnOccToAnyTransition] [varchar](10)	submitted	in 2008 LTD study	A,B,C
9	[Gender] [varchar](10)	submitted	in 2008 LTD study	B
10	[Attained_Age] int	calculated	new in 2016 study	B
11	[Annual_Duration] [varchar](3)	submitted&grouped	new in 2016 study	B,C,E
12	[Mental_and_Nervous_Period] [varchar](50)	submitted	in 2008 LTD study	B
13	[M_N_Limit_Transition] [varchar](7)	calculated&grouped	in 2008 LTD study	B
14	[Gross_Indexed_Benefit_Amount] [varchar](50)	submitted&grouped	in 2008 LTD study	C
15	[Industry] [varchar](70)	submitted&grouped	new in 2016 study	C,D.E
16	[Indexed_Monthly_Salary] [varchar](50)	submitted&grouped	new in 2016 study	C,D
17	[Taxability_Benefits] [varchar](50)	submitted	new in 2016 study	D
18	[Integration_with_STD] [varchar](50)	submitted	new in 2016 study	D
19	[CaseSize] [varchar](96)	submitted&grouped	new in 2016 study	E
20	[StateofResidence] [varchar](50)	submitted&grouped	new in 2016 study	E
21	[COLA_Indicator] [varchar](50)	submitted	new in 2016 study	none
22	[Benefit_Max_Limit_Proxy] [varchar](34)	calculated	new in 2016 study	none
23	[Benefit_Duration] [varchar](8) =Limited, Life, To Age	calculated&grouped	new in 2016 study	none
24	[ratiogroup] [varchar](25) = replacement ratio	calculated&grouped	new in 2016 study	none

2.1 Description of New Variables

- COLA Indicator – COLA indicators were submitted by contributors in the GLTD study data request, but they were not included as variables in the pivot tables previously released. They are included in the CDB, and an analysis is shown in section 3. Valid fields include “Y,” “N” and “U” (unknown).

- Benefit Max Limit Proxy – This indicator was defined and calculated as a proxy for identifying claims with benefit amounts that had been capped by plan maximums; i.e., where an insured’s salary and the plan replacement ratio would have produced a benefit amount greater than the plan maximum. (This information was not part of the study data request.) Valid fields are benefit amounts that end in “000,” benefit amounts that end in “00,” and “Not ending in 000 or 00.” The underlying assumption was that a benefit amount that was a % of salary would usually not end in “00” or “000” unless it had been capped. (It was recognized that this assumption would not be 100% accurate.)

- Benefit Duration Proxy – This indicator was defined and calculated as a proxy for identifying each claim’s plan maximum benefit duration; e.g., whether the plan was To Age (65) vs. short term vs. lifetime. (This information was not part of the study data request; the maximum claim end date was requested instead.) Best efforts were made to sort claims into the above three categories based on their Ages at Disability and their Benefit Maximum End Dates. “To Age” claims are intended to include to ages 60, 65, 70, 75 and Social Security Normal Retirement Age (SSNRA). All claims were sorted (there were no “Unknowns).

- Replacement Ratio Proxy – This indicator was defined and calculated as a proxy for identifying each claim’s replacement ratio; i.e., the ratio of the claim gross benefit amount to the claimant’s salary. (This information was not part of the study data request.) Best efforts were made to calculate the ratios and then sort the results into ranges that would provide the most useful basis for analysis. There are 11 valid ranges plus “NA,” as follows:

    - 01: Less than 40%
    - 02: 40%
    - 03: GT 40% and LT 50%
    - 04: 50%
    - 05: GT 50% and LT 60%
    - 06: 60%
    - 07: GT 60% and LT 66 2/3%
    - 08: 66 2/3%
    - 09: GT 66 2/3% and LT 70%
    - 10: 70%
    - 11: GT 70%
    - 12: N/A

Section 3: Analyses of New Variables

We have provided graphs and associated brief commentary for each of the new variables discussed in Section 2. There are three graphs for each new variable:

- By observation year
- By disability category, and
- By duration

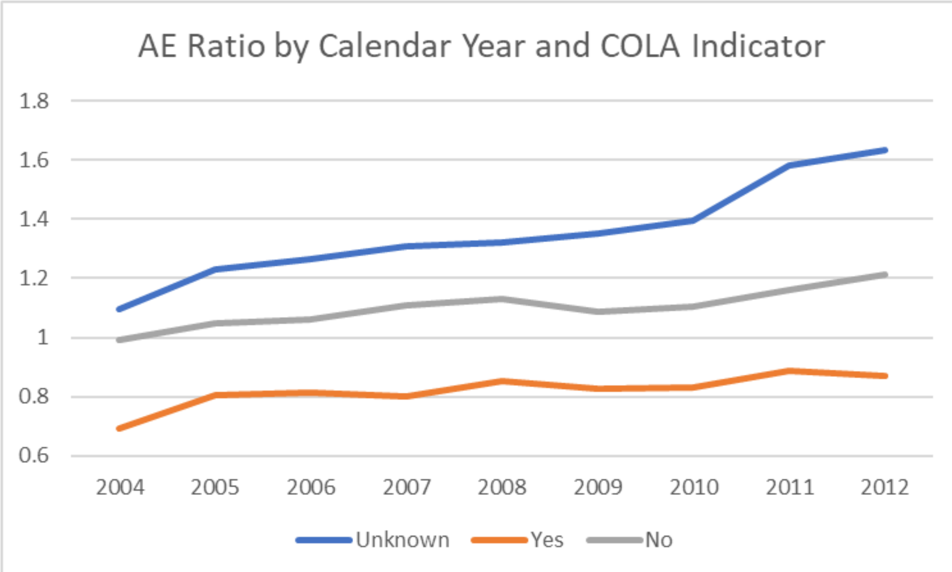
We have also included the table below to give a sense of the lives exposed for each of the categories of these new variables. Clearly some results will be more relevant than others, given the skewed exposure distributions.

Table 2: Exposure Data for Variables not included in Pivot Tables

New Variable	Exposures	New Variable	Exposures
COLA		**Ratio Group**	
Unknown	7,360,000	01: Less than 40%	1,380,000
Yes	1,490,000	02: 40%	420,000
No	29,140,000	03: GT 40% and LT 50%	520,000
Total	*37,990,000*	04: 50%	3,270,000
Benefit Duration		05: GT 50% and LT 60%	750,000
Lifetime	210,000	06: 60%	24,890,000
Limited	3,910,000	07: GT 60% and LT 66%	820,000
To Age	33,870,000	08: GE 66% and LT 68%	4,260,000
Total	*37,990,000*	09: GE 68% and LT 70%	30,000
Benefit Max Limit Proxy		10: 70%	1,220,000
Benefit Multiple 100	1,740,000	11: GT 70%	360,000
Benefit Multiple 1000	1,300,000	N/A	70,000
Benefit Not a Multiple of 100/1000	34,950,000		
Total	*37,990,000*	*Total*	*37,990,000*

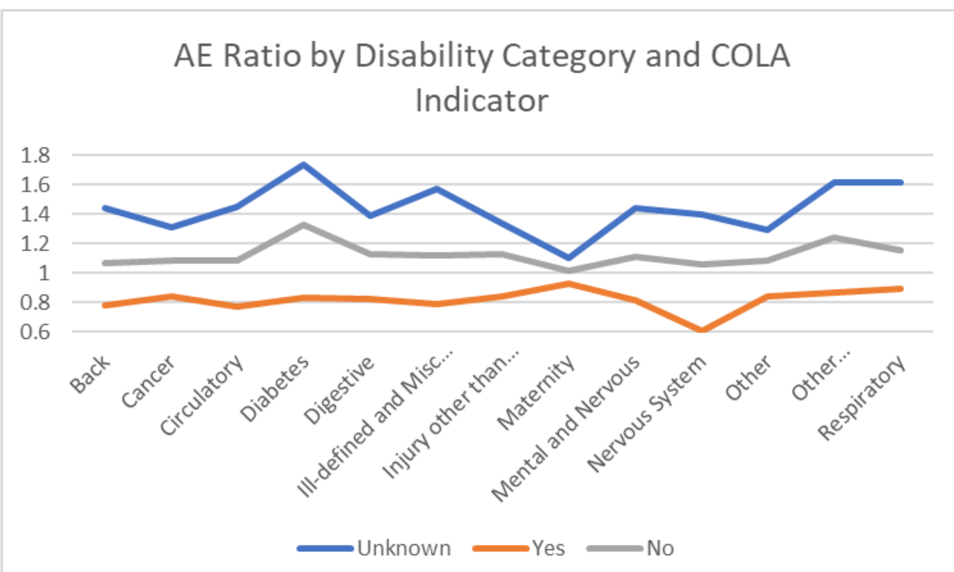
Charts 3, 4 and 5 show the impact that a COLA policy provision has on recovery rates.

Chart 3: COLA by Calendar Year



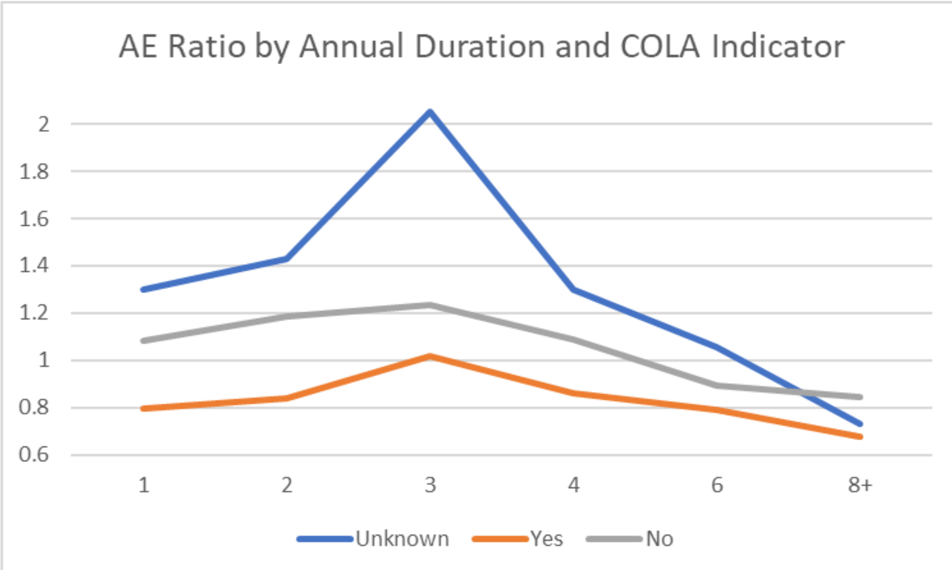
The results above line up with expectation – the presence of a COLA results in lower recovery rates. Note that the exposures are very skewed as shown by the previous table, with only 1.5 million having a COLA of "Yes." Based on some additional examination by company (data not publicly available), we note that the higher AE for the "Unknown" category appears to be more reflective of individual company experience vs the COLA status itself.

Chart 4: COLA by Disability Category



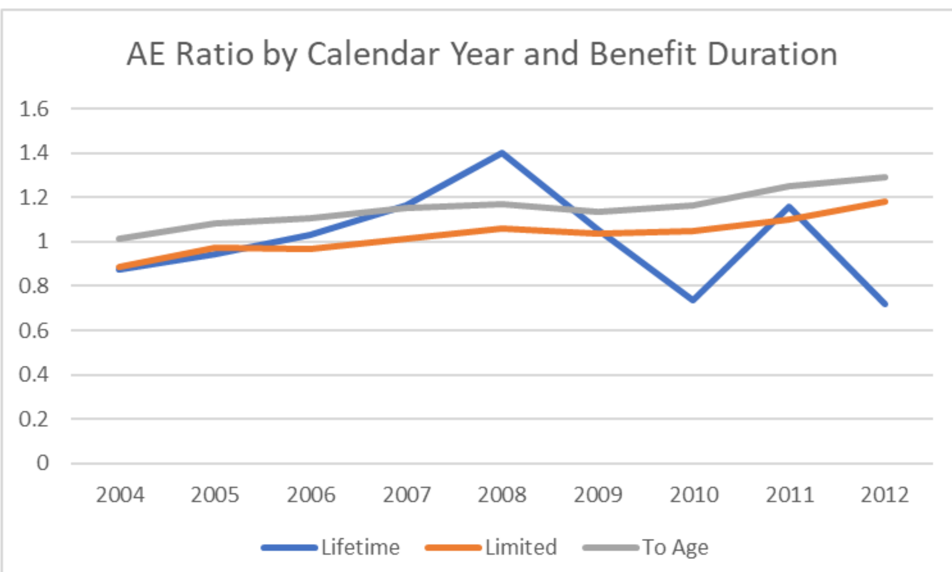
Similar to Chart 3, there is a clear separation of the AE ratios by COLA indicator.

Chart 5: COLA by Annual Claims Duration (X Axis in months)
(Durations are grouped by year as 1, 2, 3, 4, [5 & 6 & 7] shown as 6, and 8+)



As noted in the commentary for Chart 3, the "Unknown" COLA category reflects specific experience of a small number of companies, which may not translate to the industry in general. It does appear that the presence of a COLA reduces recovery rates at all durations.

Chart 6: Benefit Duration by Calendar Year



The "Lifetime" category has only 210,000 exposures, so any conclusions would not be meaningful. Note that almost all policies are of the "To Age" type with 33.9 of the 38 million exposures being in that category. It does appear that the "To Age" policies display a higher recovery rate across all observation years, and across most disability categories (below).

Chart 7: Benefit Duration by Disability Category

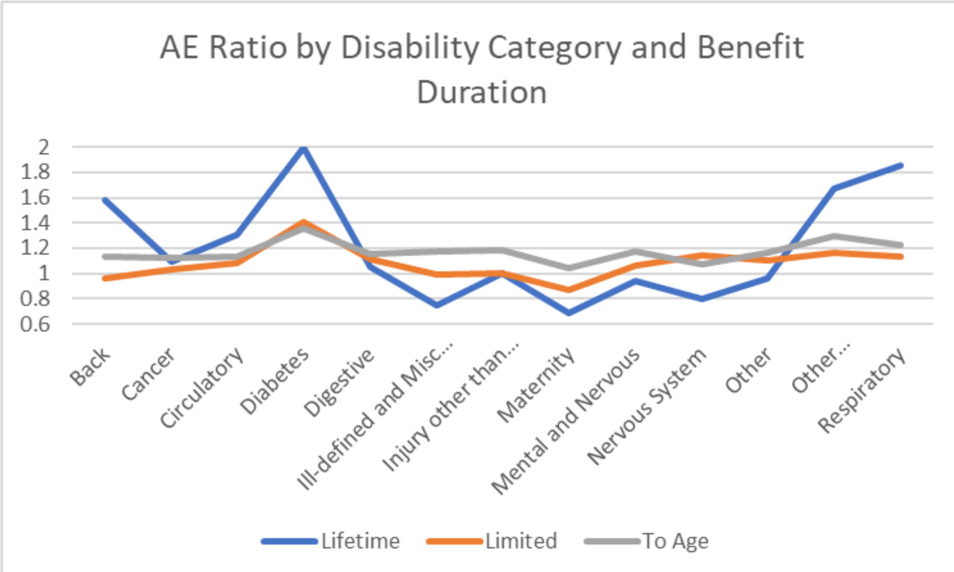
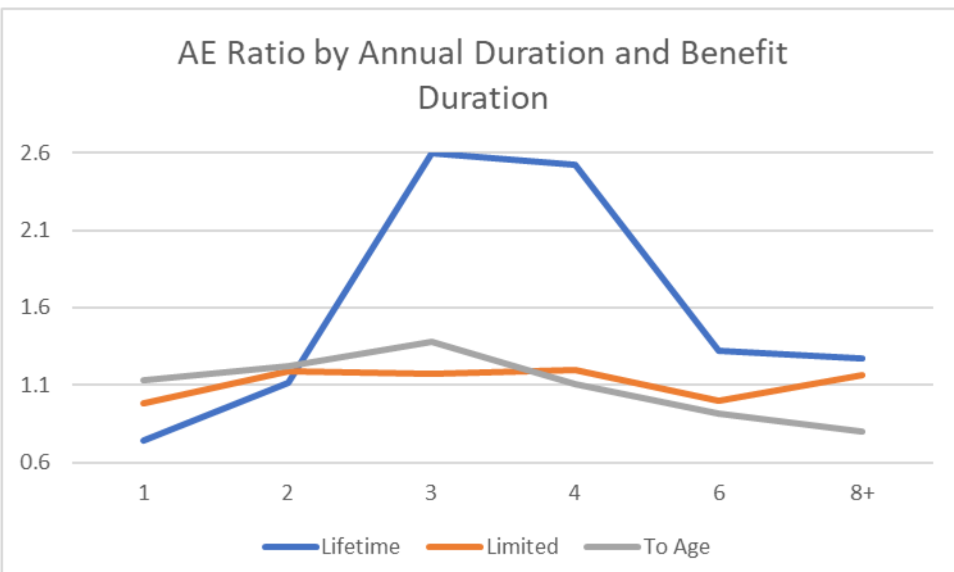
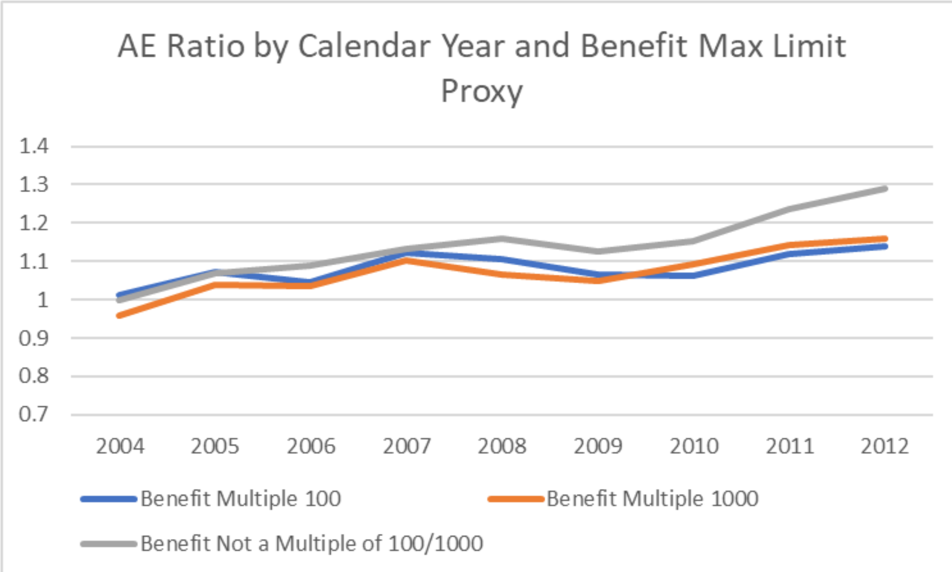


Chart 8: Benefit Duration by Annual Claims Duration
(Durations are grouped by year as 1, 2, 3, 4, [5 & 6 & 7] shown as 6, and 8+)



The "To Age" category is the only one with significant exposures when the data is split by Duration. Not surprisingly, the "To Age" follows the shape of the overall AE ratio when reviewed by Duration. The total exposure for "Lifetime" benefit durations at 3- and 4-year durations is only 8,000 and 6,000 respectively, so the spike in the AE ratio at these durations could be spurious.

Chart 9: Benefit Max Limit Proxy by Calendar Year



It appears that overall those policies that are not multiples of 100 or 1000 do show a slightly higher AE ratio than the other 2 categories, but keep in mind that those policies where the benefit is not a multiple of 100/1,000 have by far the largest share of the exposures.

When the data is analyzed by disability category or by claims duration, no clear picture emerges.

Chart 10: Benefit Max Limit Proxy by Disability Category

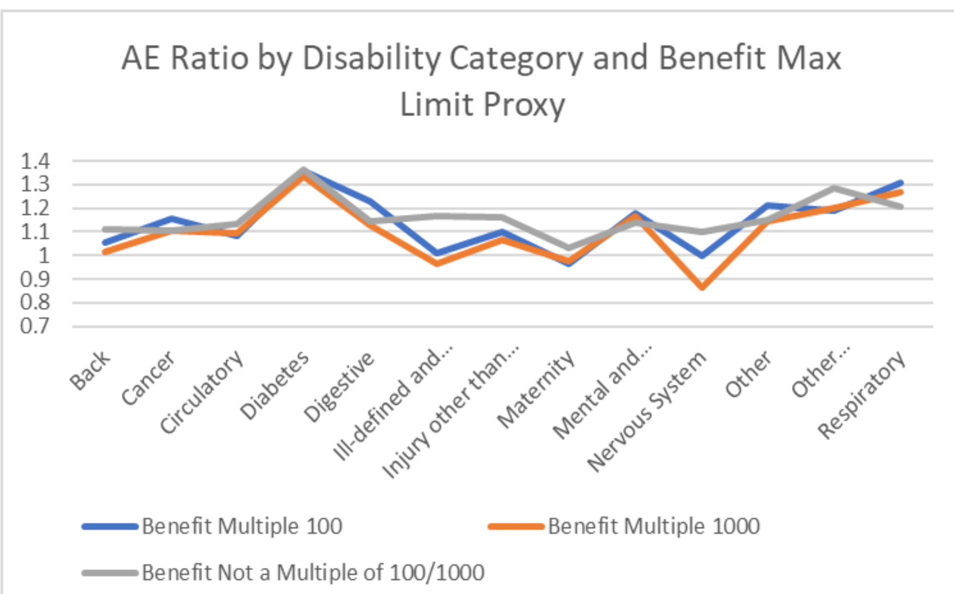


Chart 11: Benefit Max Limit Proxy by Annual Claims Duration
(Durations are grouped by year as 1, 2, 3, 4, [5 & 6 & 7] shown as 6, and 8+)

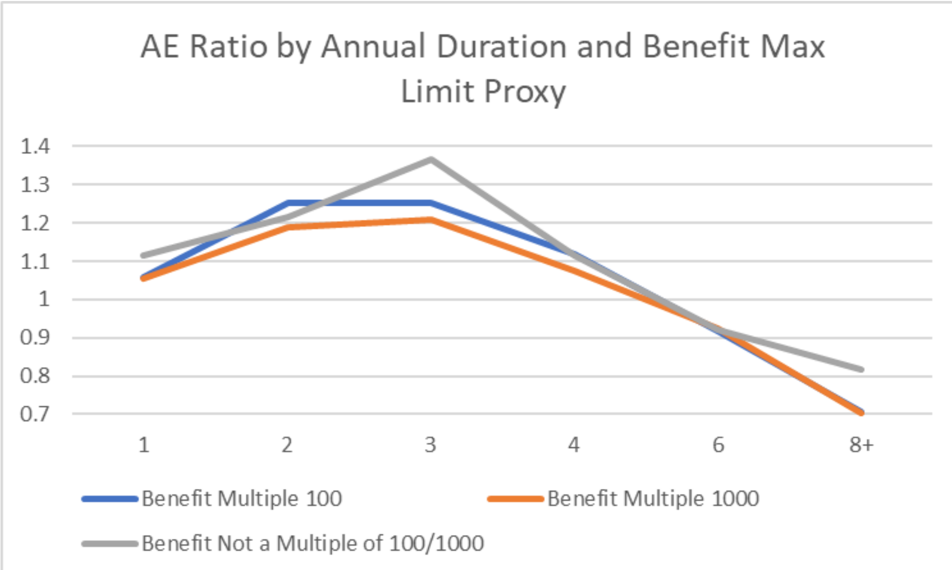
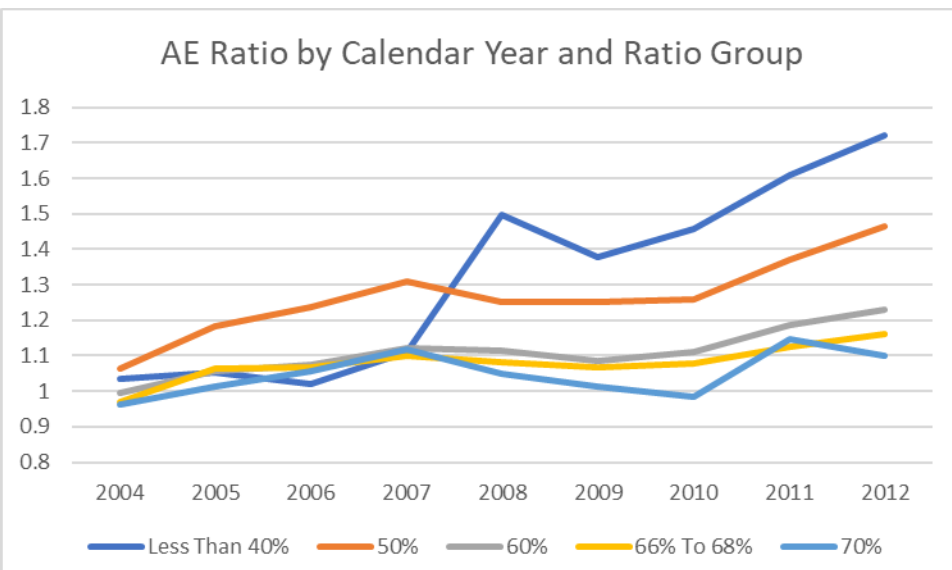


Chart 12: Ratio Group by Calendar Year
(Showing significant exposure groups only)



Since two thirds of the exposures are in the 60% ratio group, it's helpful to think of the gray line in Charts 11, 12 and 13 as being the anchor, with the other lines potentially showing significant noise. There does appear to be a noticeable difference in experience between the 50% and 60% groups. Some of this is likely due to individual company differences, but we note that almost all companies did provide data that included both the 50% and 60% ratio group, so this is more than just an intercompany variation. It appears that the highest recovery rates are associated with the lowest ratio group, which is expected given that these policies result in steeper compensation decreases when on disability.

Chart 13: Ratio Group by Disability Category
(Showing significant exposure groups only)

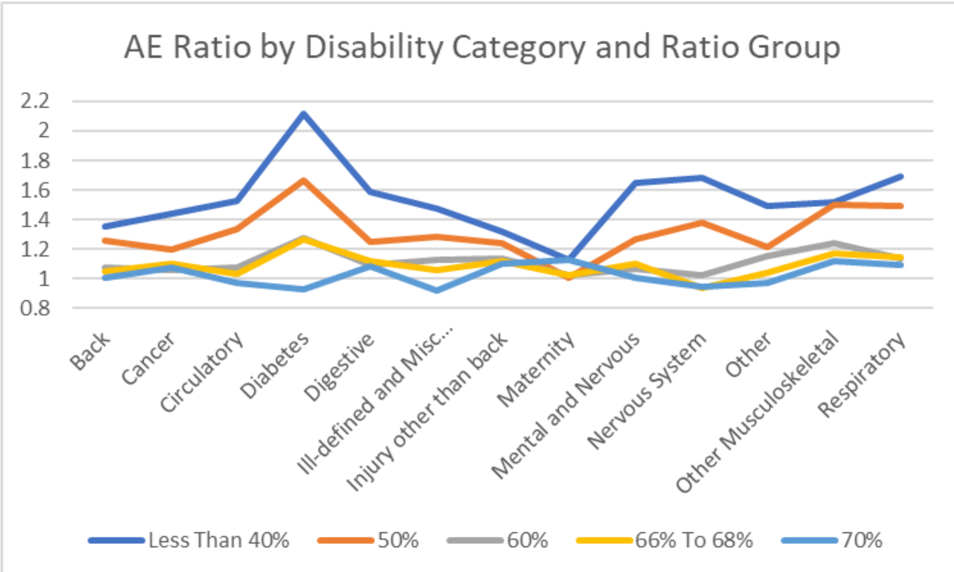
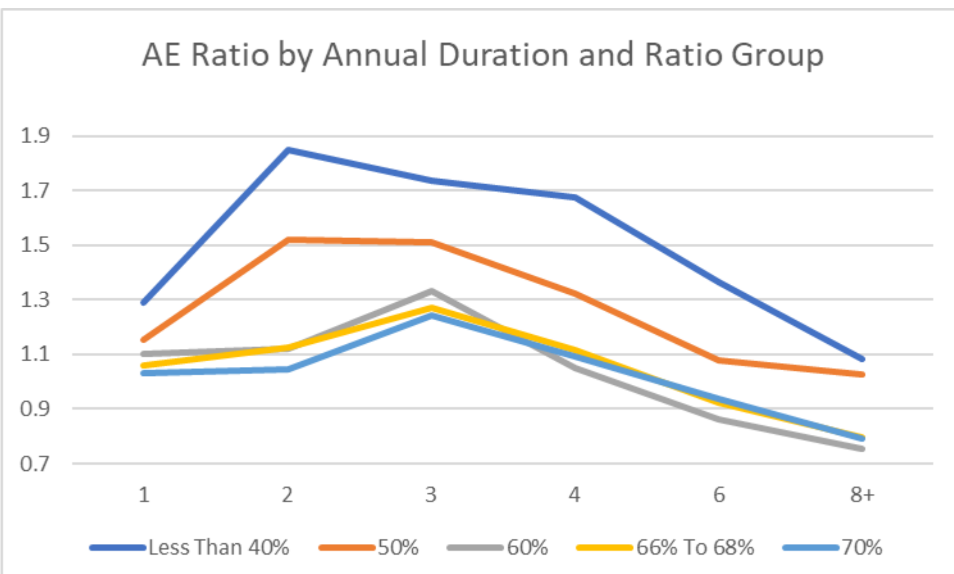


Chart 14: Ratio Group by Annual Claims Duration
(Durations are grouped by year as 1, 2, 3, 4, [5 & 6 & 7] shown as 6, and 8+)



We note that even at the later durations, the differing experience between the ratio groups persists.

Section 4: Trend Analysis – Traditional Approach

The 2018 study showed a significant trend of improvements in recoveries, and we worked with the SOA to perform some analyses to try to identify potential drivers of this. We looked for situations where the experience varied materially for the different values of a specific variable, and the exposure mix by that variable changed materially over the study period. (Note: there are theoretically some other situations that could drive material changes in trend, but we did not see evidence of those.)

For most variables that showed changes in experience, it did not appear that the mix of business changed materially over the study period. However, we did identify that both the required conditions existed for the STD Integration variable. Charts 15 and 16 show how STD Integration status affected experience over the study period. Chart 15 shows experience for the first 48 months, and chart 16 shows the corresponding exposure distribution.

Chart 15

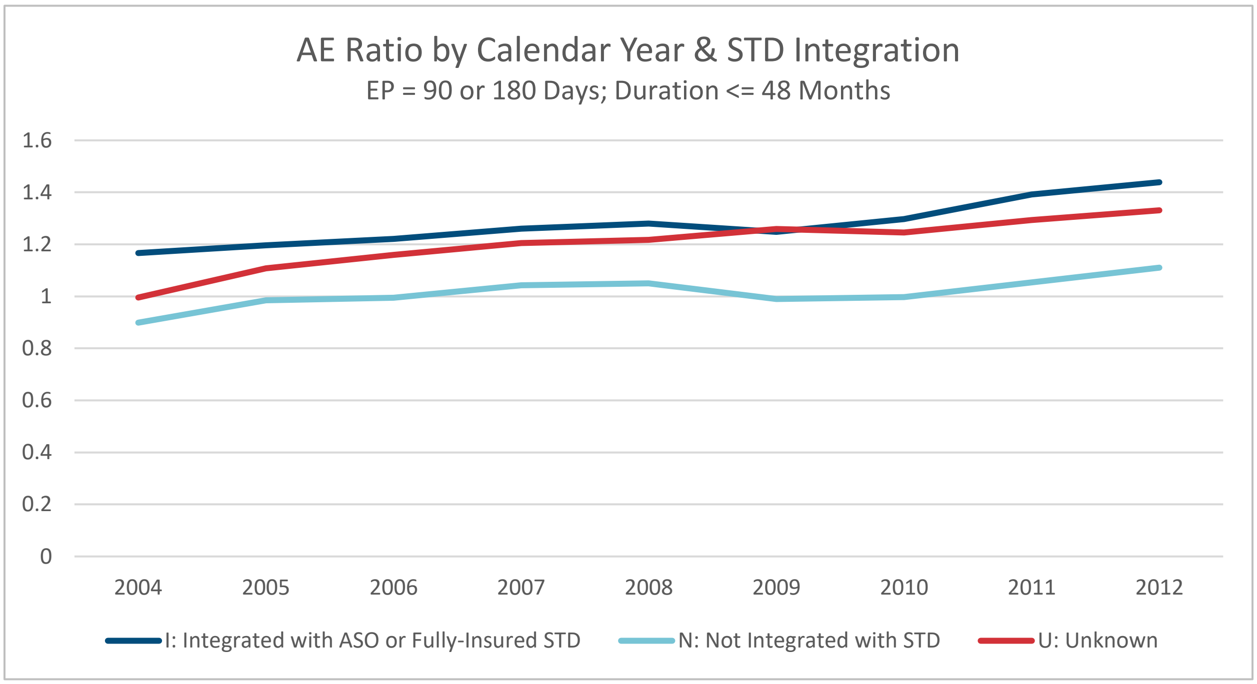
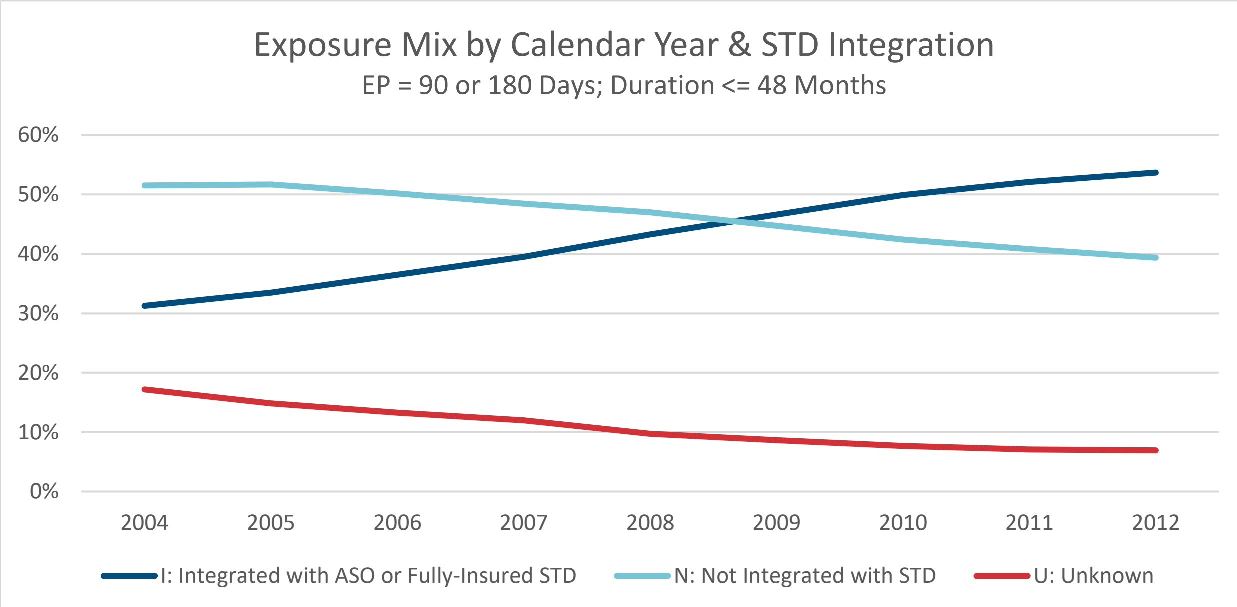


Chart 16



Note: We also looked at experience after 48 months and it continued to show integrated A/Es higher than non-integrated.

We also wondered whether the change in the mix of companies over the study period had a material impact on trend. In particular, we asked the SOA to look at experience by contributor to see if both the experience and exposure mix varied materially. (The SOA could do so while maintaining confidentiality of the specific contributors.) The SOA verified that the mix of business changed materially over the study period and that the experience also varied materially by company. This implied that further analysis would be appropriate.

To review the possibility STD Integration was masking results that were more driven by changing company mix, we adjusted the expected recoveries to reflect actual company experience during the study period and recalculated the AE ratios. (This utilized data not available to the POG or the general public.) If company experience was impacting the results, we would anticipate that adjusting the expected recoveries for company experience would significantly change the AE Ratios when viewed by calendar year. However, Chart 19 below indicates that allowing for Company experience has negligible impact on the relationship between STD Integration and duration. Hence, we concluded that the change in company mix was not being masked by the STD Integration variable.

Chart 17: Ratios Adjusted for STD Integration, Duration and Company

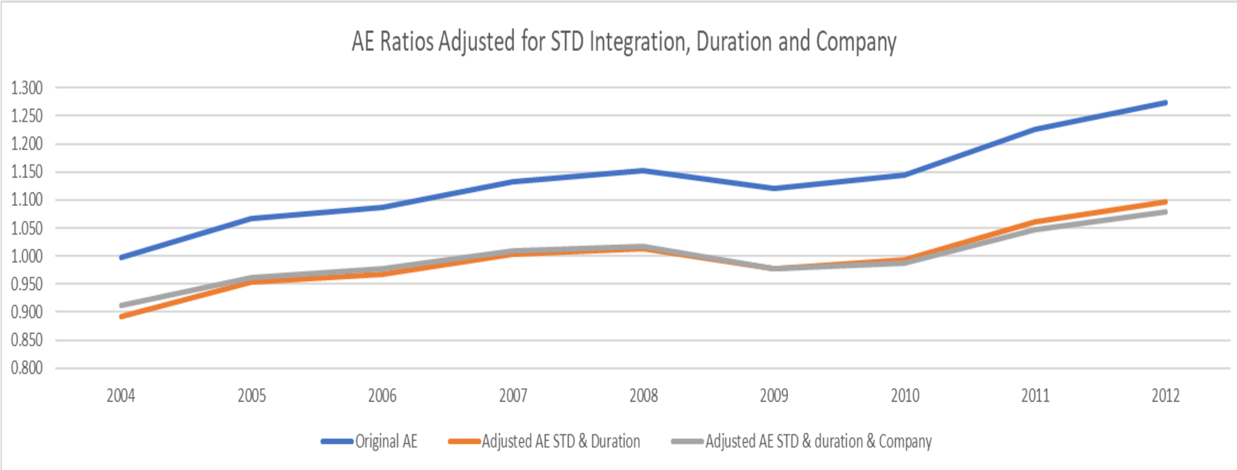


Chart 17 shows that, once the experience trends are normalized for STD Integration, the effect of company mix on trend appears to be minor.

4.1 Conclusions

It was evident that STD Integration had had a material impact on overall trend, while the impact of company mix was minor. However, we also had many variables, any number of which could be making minor contributions to overall observed trend, possibly resulting in a material aggregate contribution. We decided that predictive analytics was the way to address this possibility.

Section 5: Predictive Analytics Methodology

PA is an approach that identifies a specific variable's contribution to overall experience. It is particularly useful when there are many variables available. We noted that claim recovery rates exhibited a distinct trend of increases over the study period. To help us to understand the drivers, we developed a Generalized Linear Model (GLM)of the claim recovery rates. (A similar approach was later developed for death rates.) We note that GLMs are one of many types of predictive analytics models. GLMs are commonly used for many risks in the insurance industry, as they provide a framework that:

- Is relatively easy to understand
- Allows for a wide range of statistical diagnostics
- Is extremely flexible, accommodating to a wide range of distributions, input variables and applications.

Detailed explanations of GLM theory that address the underlying theory of the approach are widely available for the interested reader[1]. We provide a basic description below to help the reader understand some of the implementation considerations and results.

The GLM we used contains a Log link function and a Poisson error structure. The predicted recovery rate for a given claim exposure is assumed to have the following form:

$$E[\mathrm{Re}\operatorname{cov}ery] \equiv \mu = \mu_0 * \exp(\sum_{j=1}^{k} \beta_j * x_j)$$

Where,

μ_0 equals the mean loss ratio at base level[2], *k* equals the number of parameters, β_j is the regression coefficient corresponding to the *jth* independent variable*, and* x_j is the *jth* independent variable.

The regression coefficients and the base level value were estimated to maximize the log-likelihood for the observed recoveries in the data (using the assumed distribution) using iterative numerical techniques. More detail on the model form and assumed distribution follows.

5.1 Choice of Model Form

When modeling claim recovery behavior, the most rigorous mathematical choice is a logistic model which pairs a logit link function with a binomial error structure. This form is naturally suited to modeling binary responses such as claim recovery rates, where there are many trials (months of claims exposure) and a binary result (either the claim recovers or does not recover). The result of the model will be the probability (between 0 and 1) that a claim recovers during a period of exposure. Unfortunately, results from the logistic model are typically hard to visualize and interpret.

An alternative model form is the one we used, which has a Log Link function and a Poisson error structure. With a Log Link function and Poisson error structure, the log of the expected recovery rate is predicted by a linear combination of the independent variables. This makes the model of expected recovery rates multiplicative (the different parameters in the model are multiplied together), and therefore the results are very easy to interpret and to implement, as in the example of a simple recovery rate model shown below in Figure 18. This model predicts the number of recoveries but is not capped at 100% like the logistic model, so it can theoretically predict recovery rates greater than 100%. We have not found this to be a problem in practice, and we have confirmed that our recovery rate model resulted in a good fit to the actual data.

Figure 18: Example of a multiplicative recovery rate model

1. Base Rate
0.015

2. Gender	
Male	0.79
Female	1.00

3. Annual Duration	
1	3.48
2	1.00
3	1.00
4	0.28
5	0.15
6	0.10
7	0.07
8	0.06
9	0.05
10+	0.04

4. Age at Disability	
<30	2.79
30-39	2.09
40-49	1.48
50-59	1.00
60+	0.80

5. Age at Disability * Annual Duration	1	2 - 3	4 - 5	6 - 7	8 - 9	10+
<30	1.30	1.00	1.28	1.30	1.17	0.69
30-39	1.16	1.00	1.38	1.54	1.35	0.93
40-49	0.90	1.00	1.26	1.36	1.21	0.93
50-59	1.00	1.00	1.00	1.00	1.00	1.00
60+	1.06	1.00	0.83	1.25	1.59	1.43

Predicted Recovery Rate for a Sample Claim:

Gender:	Male
Annual duration:	5 Years
Age at disability:	40
Monthly recovery rate:	**0.015** * **0.79** * **0.15** * **1.48** * **1.36** = 0.0042

5.2 Variable Selection and Model Development

We used a randomly selected three quarters of the experience study data to fit the model, with the remaining one quarter held back to validate the model. As illustrated in the example shown above in Figure 18, the model predicted recovery rates are equal to the product of a base rate and factors for each of the claim characteristics that were determined to be significant drivers. The model we developed includes a base rate plus fifteen single variable factors and ten interactions. Interactions are situations where the influence of one variable varies according to the value of another. Figure 19 on the next page shows the variables included in the final model.

Figure 19: Variables Included in Final Recovery Rate Model

Single Variable Factors

- Claim Duration
- Gender
- Age at Disability
- Disability Category
- Elimination Period
- Limited Own Occupation Period
- Ownocc To Any Occ Transition
- Integration with STD Indicator
- COLA Indicator
- Contractual Benefit Duration
- Gross Indexed Monthly Benefit Amount (GIB)
- Ratio Group (Benefit Amount / Salary)
- Case Size
- Industry
- Region

Interaction Factors

- Elimination Period * Claim Duration
- Age at Disability * Claim Duration
- Disability Category * Claim Duration
- Integration with STD * Claim Duration
- Gross Indexed Monthly Benefit Amount * Claim Duration
- Case Size * Claim Duration
- Disability Category * Age at Disability
- Ownocc To Any Occ Transition * Limited Own Occupation Period
- Ownocc To Any Occ Transition * Age at Disability
- Ownocc To Any Occ Transition * Disability Category

We developed the model in three steps:

- Review of Raw Data - Reviewing the raw data for reasonableness and making preliminary observations about the influence of the variables.
- Model Fitting – Deciding which variables to include in the model.
- Model Validation – Reviewing the fit of the model to the actual data.

The following three sections describe these steps including selected analyses and results.

5.3 Review of Raw Data

We reviewed the raw data using tables and graphs of the actual claim exposures and actual recovery rates for each of the variables included in the study. We verified that the underlying data appeared to be reasonable; we also made preliminary observations about which variables appeared to have significant impacts on claim recovery rates.

Chart 20 shows the graph of claim exposures and actual recovery rates by Claim Duration. Based on this chart, we believe that both the distribution of claim exposures by duration, as well as the actual level of recovery rates, appear to be reasonable. We observed that, in general, both the level of exposures and recovery rates decline as claim duration increases. We note that the spike in the recovery rates between months 26 and 30 coincides with the end of the own occupation period for most claims.

Chart 20: Raw Recovery Rate Data – Claim Duration

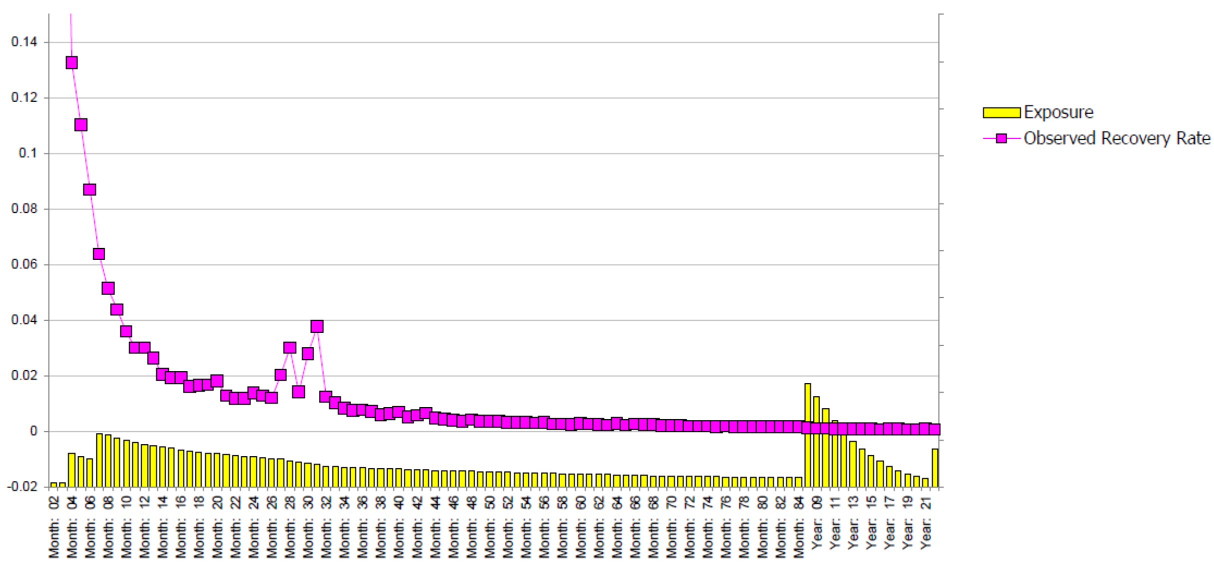
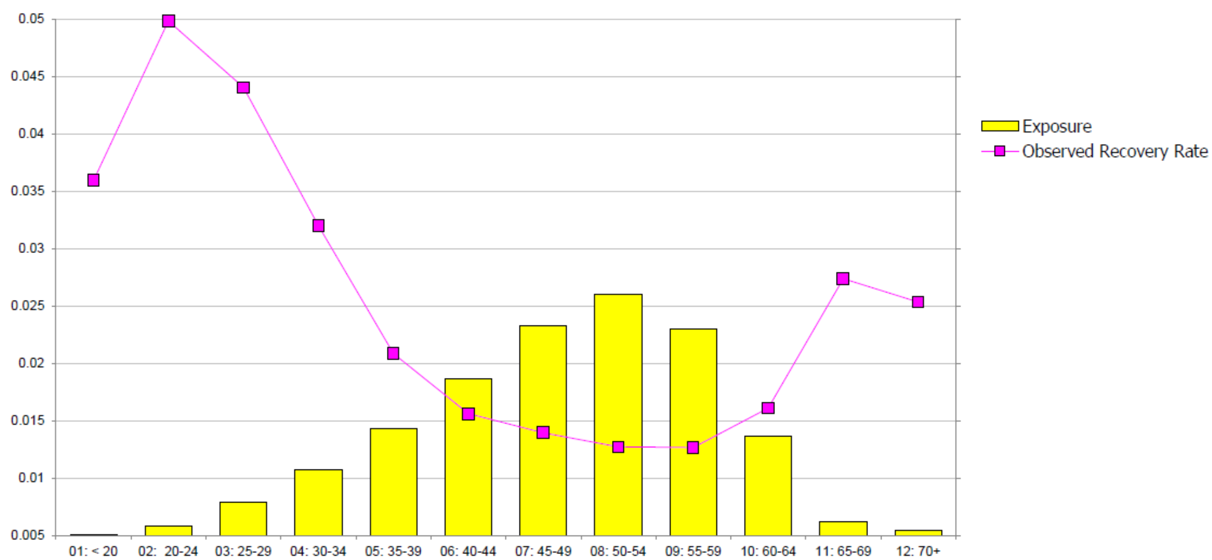


Chart 21 shows the graph of claim exposures and actual recovery rates by Age at Disability. Based on this chart, we believe that both the distribution of claim exposures by Age at Disability, as well as the actual level of recovery rates, appear to be reasonable. We observed that, in general, raw recovery rates decline as Age at Disability increases. We believe the higher raw rates for ages 60+ are due to those ages having more short duration claims.

Chart 21: Raw Recovery Rate Data – Age at Disability



5.4 Model Fitting

We began by fitting an initial model that included a selection of single variables and variable interactions that we expected would have strong impacts on recovery rates based on our prior experience working with group LTD claims, as well as our review of the raw data. Based on results from that initial model, we made decisions to refine the model by modifying variable segmentation, removing variables that appeared to be redundant, and adding additional single variables and variable interactions. The model refinements were based on statistical significance as well as judgment. Charts 22 through 33 show results from that process along with commentary on the modeling decisions we made.

Charts 22 and 23 show the claim exposures and modeled recovery rate relativities by Claim Duration and Age at Disability, respectively. The relativities represent the effect on recovery rates of these two variables alone. As we expected, there are strong negative correlations between recovery rates and both Claim Duration and Age at Disability, indicating that both are important variables. Both variables were included in the final model.

The relativities by Claim Duration do not exhibit the spike at the end of the own occupation period (months 26 to 30) that we saw in the raw recovery rates shown in Chart 20. The reason for the lack of this spike in the Claim Duration relativities is that the impact of the change in definition is captured by the Own Occ to Any Occ Transition variable, which identifies when a claim is near the end of the own occupation period.

Chart 22: Recovery Relativities by Claim Duration

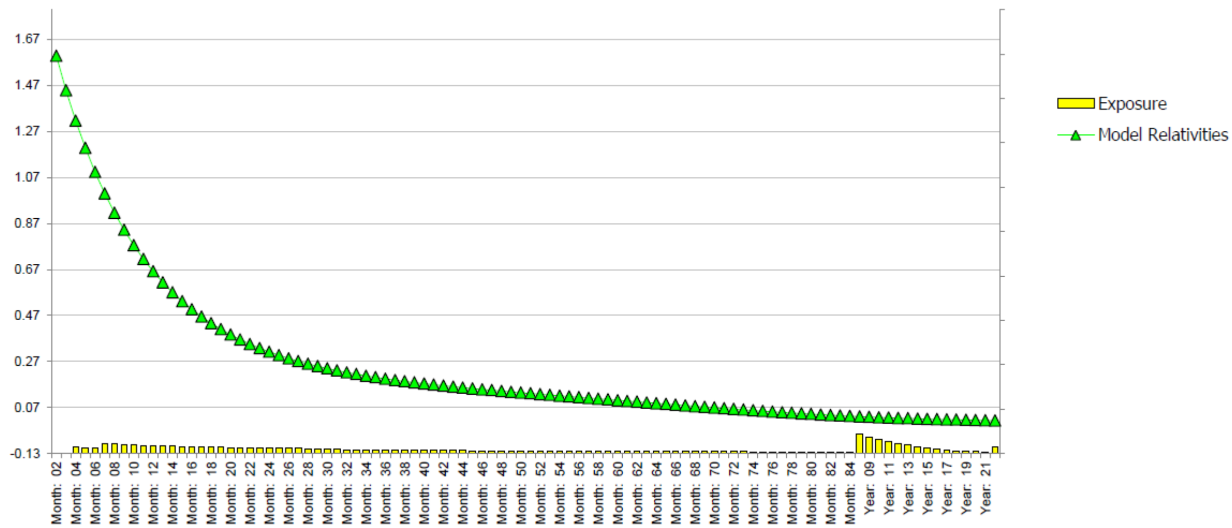


Chart 23: Recovery Relativities by Age at Disability

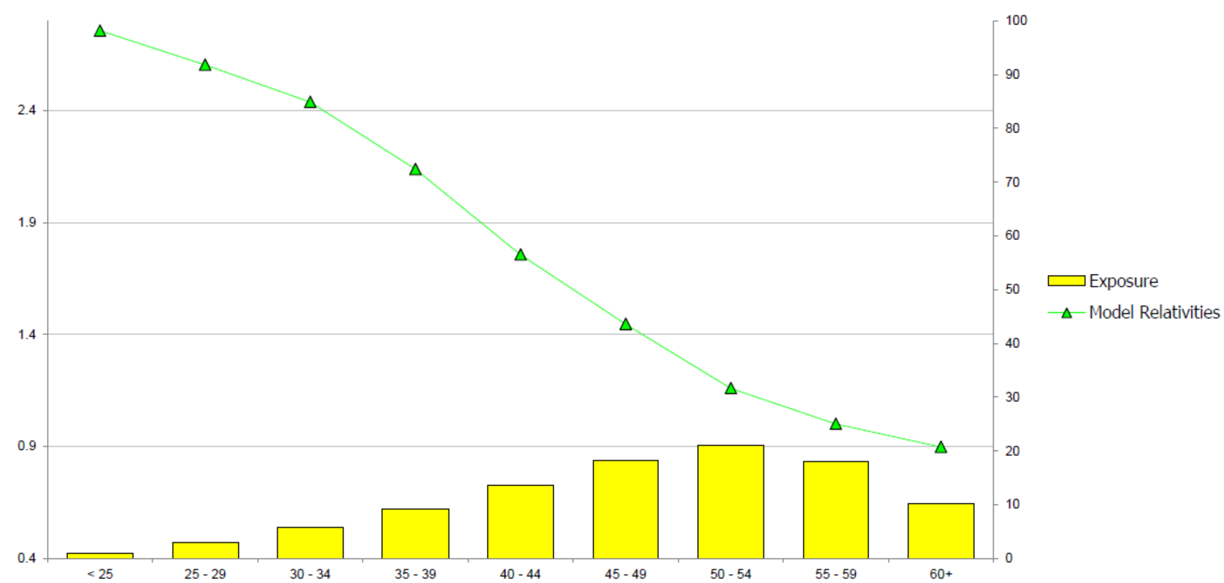


Chart 24 shows the interaction between Age at Disability and Claim Duration. The rates for each age group are shown relative to the age 55-59 value, which is the pink line. As Claim Duration increases (through Year 7) the impact of Age at Disability on recovery rates increases significantly. Then as Claim Duration increases beyond Year 7, the impact of Age at Disability declines steadily. We have included an interaction between Age at Disability and Claim Duration in the final model.

Chart 24: Age at Disability and Claim Duration

Rates shown relative to base level of Age at Disability (55-59)

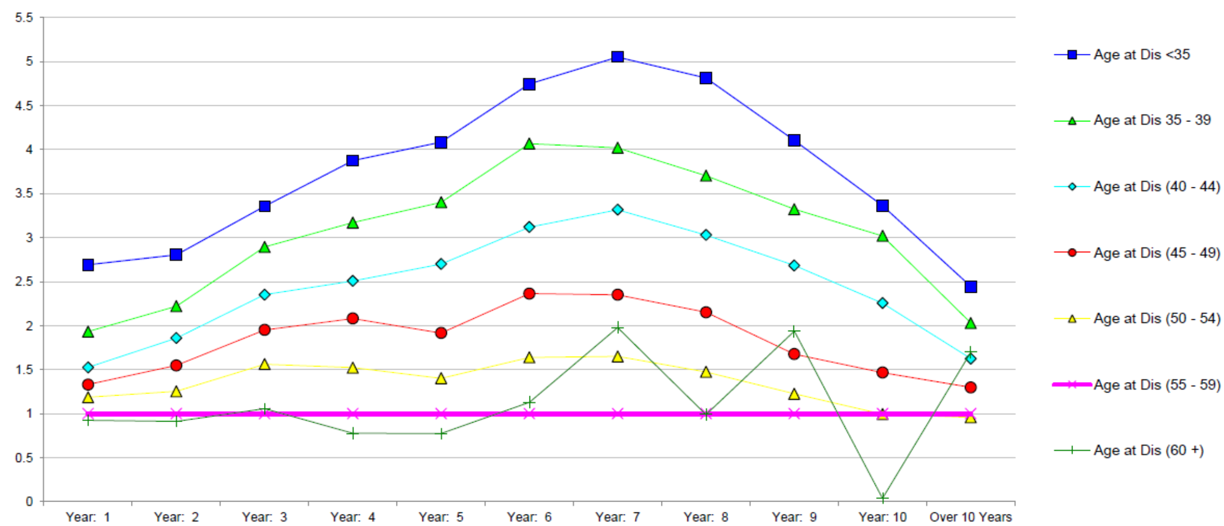


Chart 25 shows the claim exposures and modeled recovery rate relativities by Gross Indexed Benefit Amount (GIB). As for the Claim Duration and Age at Disability relativities described above, the GIB recovery relativities represent the effect on claim recovery rates of GIB alone. The strong negative correlation between recovery rates and GIB was something we expected to see based on our experience working with LTD claims experience. GIB is included in the final model.

Chart 25: Recovery Relativities by Gross Indexed Monthly Benefit Amount

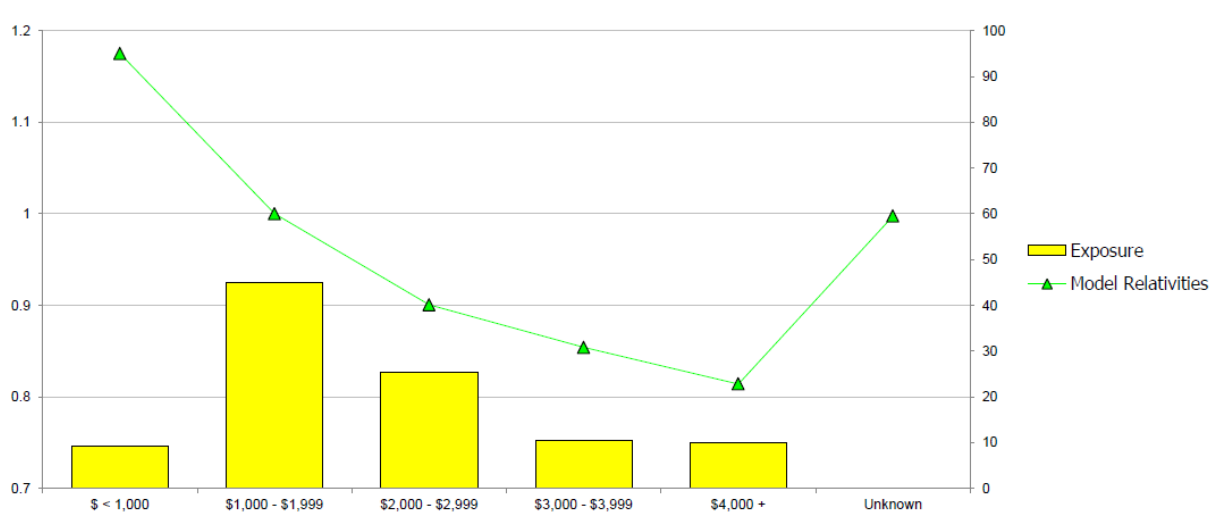


Chart 26 shows the interaction between GIB and Claim Duration. The rates for each age group are shown relative to the $1,000-$1,999 value, which is the green line. As claim duration increases from Year 1 to Year 3, the impact of GIB on recovery rates increases significantly. Then as Claim Duration increases beyond Year 4, the impact of GIB declines. We have included an interaction between GIB and Claim Duration in the final model.

Chart 26: Recovery Rates by Gross Indexed Benefit Amount (GIB) and Claim Duration
Shown Relative to Base GIB of $1,000 - $1,999

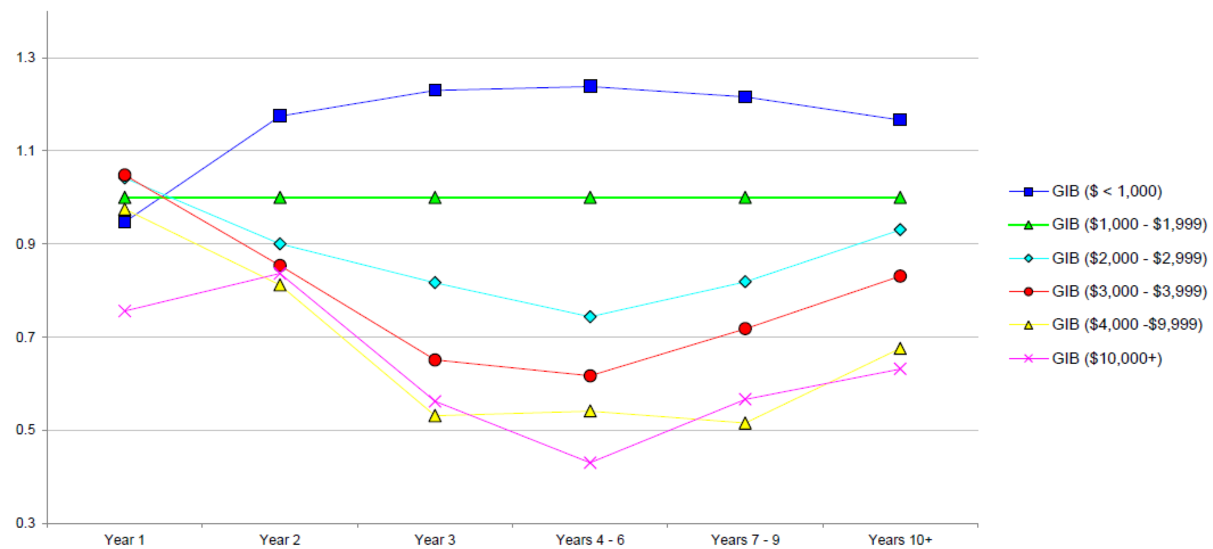


Chart 27 shows the claim exposures and modeled relativities by the Own Occupation Any Occupation Transition variable, which is the claim duration relative to the end of the own occupation period. For example, the value "Own + 1" means the claim duration is equal to the month after the end of the Own Occupation period. As we expected, values of "Own +0" and "Own +1" have a large positive impact on claim recovery rates. This variable is included in the final model.

Chart 27: Recovery Relativities by Own Occupation Any Occupation Transition

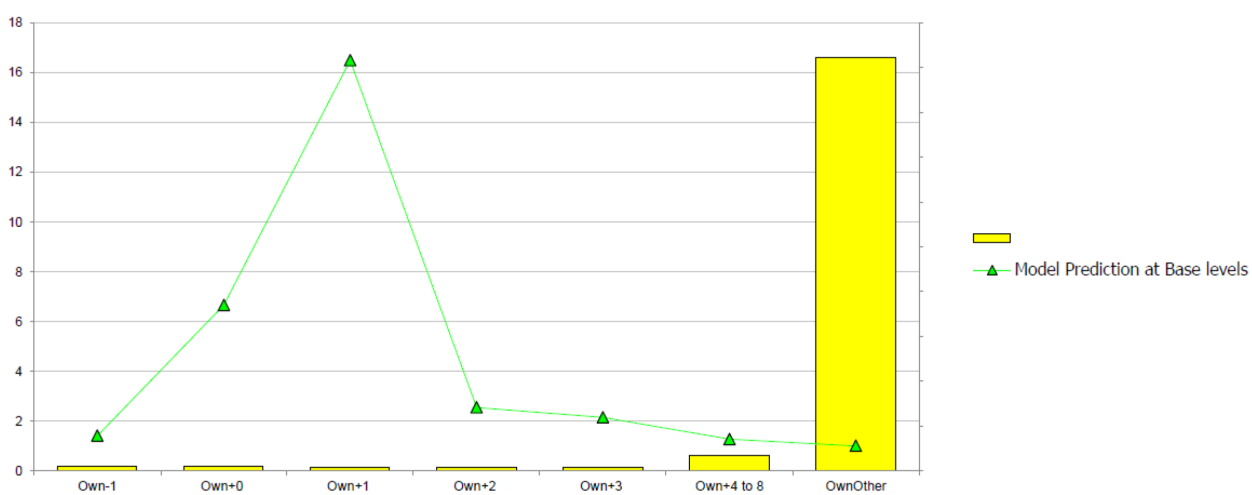


Chart 28 shows the interaction between Age at Disability and Own Occupation Any Occupation Transition. The rates for each age group are shown relative to the ages 55+ value, which is the pink line. The impact of Age at Disability on recovery rates is significantly lower around the end of the own occupation period than at other times during the life of a claim. This interaction was included in the final model.

Chart 28: Recovery Rates by Age at Disability and Own Occupation Any Occupation Transition
Shown Relative to Base Age at Disability (55+)

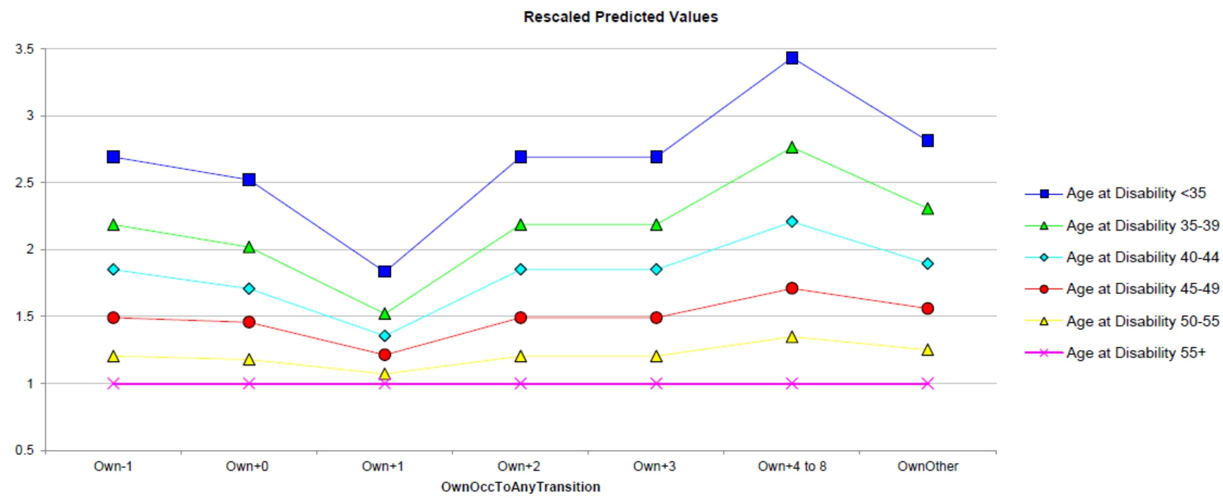


Chart 29 shows the claim exposures and recovery rate relativities by Disability Category. Disability Category has a strong effect on claim recovery rates. However, as is shown in the next three charts, the effect of Disability Category varies significantly depending on the values of Claim Duration and Age at Disability. Disability Category is included in the final model.

Chart 29: Recovery Relativities by Disability Category

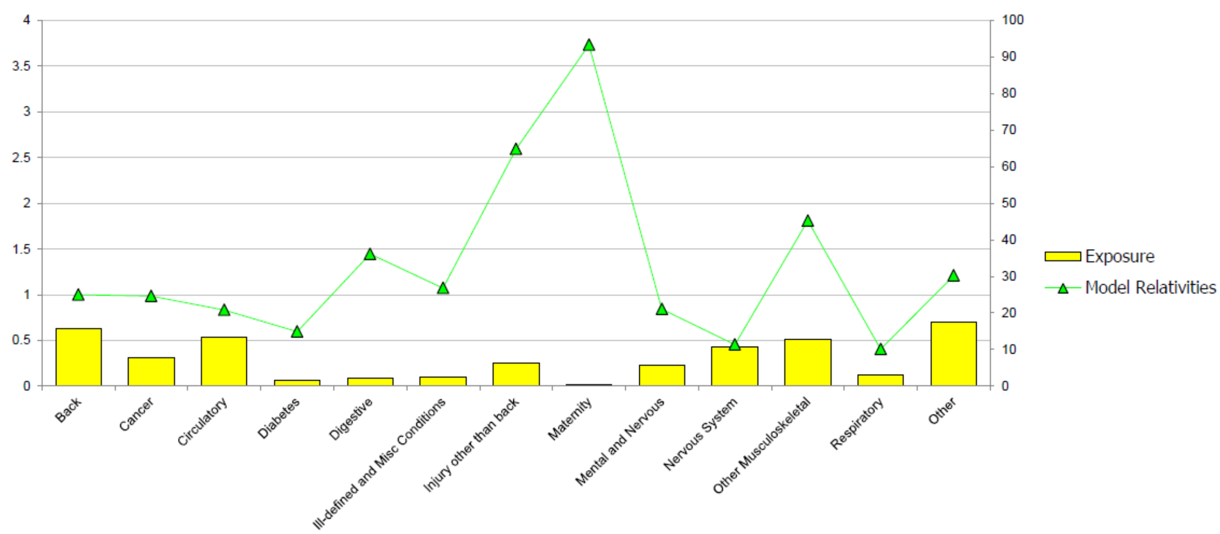


Chart 30 shows the interaction between Disability Category and Age at Disability. The rates for each Disability Category are shown relative to the "back" value, which is the blue line. The impact of Disability Category on recovery rates generally increases as Age at Disability increases. This variable was included in the final model.

Chart 30: Recovery Rates by Disability Category and Age at Disability
Shown Relative to Base Category of "Back"
(Maternity Excluded)

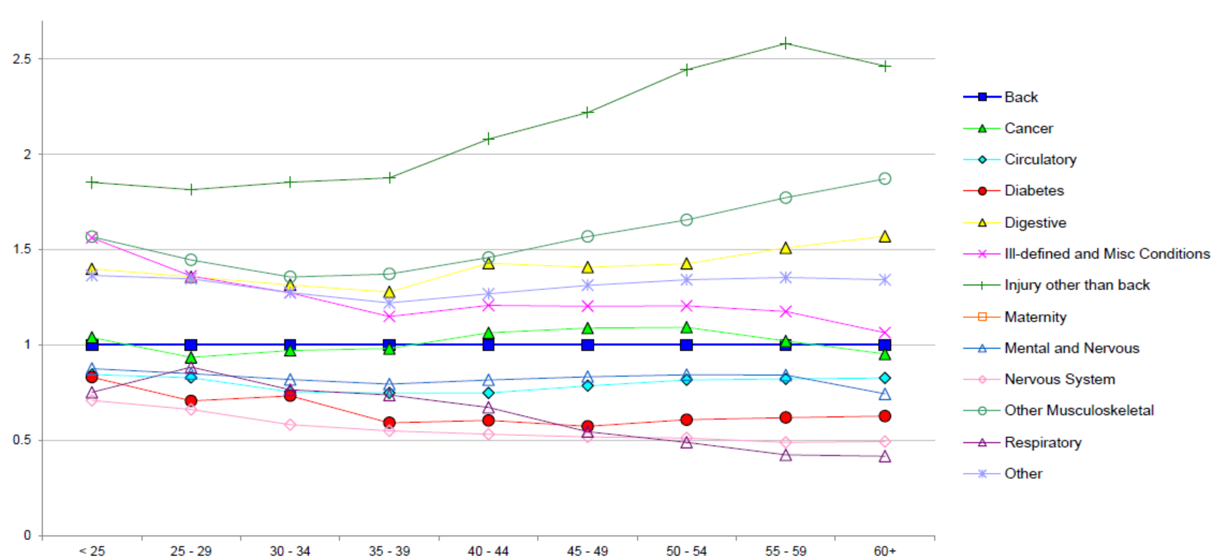
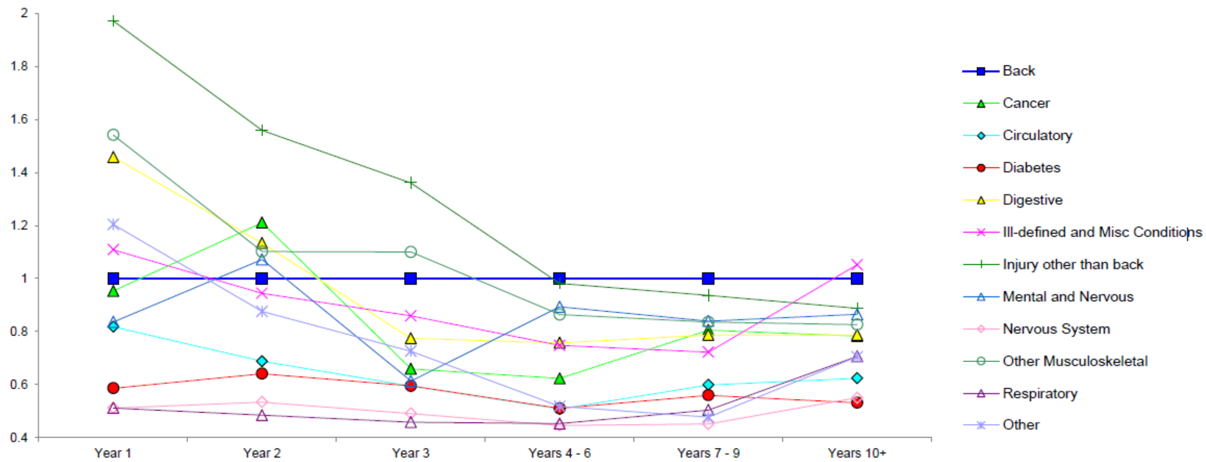


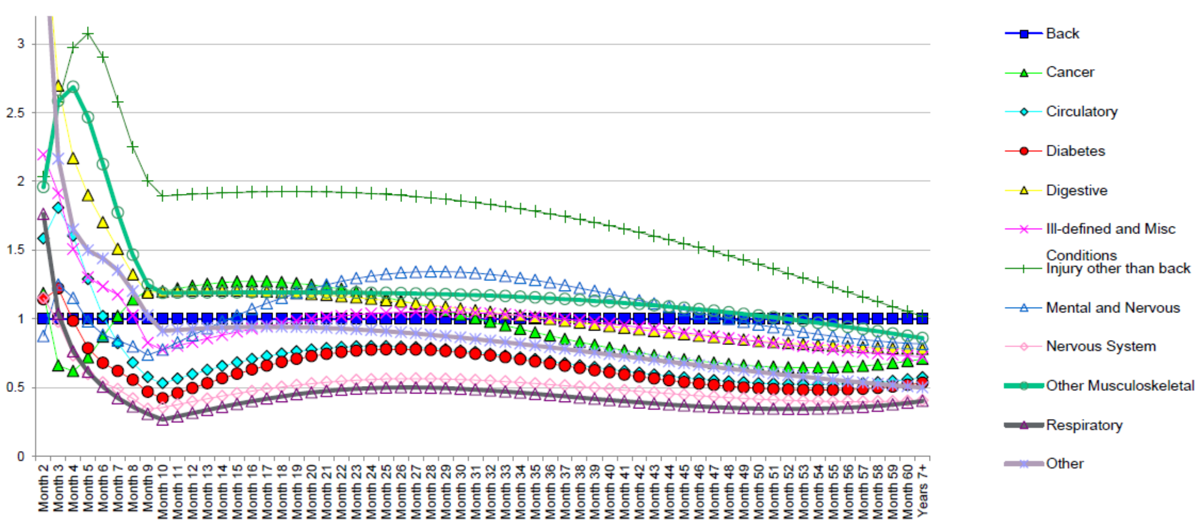
Chart 31 shows the interaction between Disability Category and Claim Duration. The rates for each Disability Category are shown relative to the "back" value, which is the blue line. The impact of Disability Category on recovery rates generally decreases as Claim Duration increases. We did include an interaction between Disability Category and Claim Duration in the final model. However, rather than use the annual claim duration variable shown here, we used a more granular version of the claim duration variable as shown in Chart 31 on the next page.

Chart 31: Recovery Rates by Disability Category and Duration
Shown Relative to Base Category of "Back" (Maternity Excluded)



As shown in Chart 32, the impact of Disability Category varies significantly during the first year of claim duration, which led us to use the monthly claim duration variable for this interaction in the final model.

Chart 32: Recovery Rates by Disability Category and Duration (Monthly)
Shown Relative to Base Category of "Back"
(Maternity Excluded from the chart)



5.5 Model Validation

We validated the model by comparing the predicted (fitted) recovery rates to the actual recovery rates for the exposure data that was not used to fit the model (the hold-out sample).

Table 33 below shows that at a high level the model fits the hold-out sample well.

Table 33 - Actual and Modeled Recovery Rates by Calendar Year (Using Hold-Out Data)

Calendar Year	2004	2005	2006	2007	2008	2009	2010	2011	2012	**Total**
Actual Recovery Rate	0.0158	0.0169	0.0167	0.0171	0.0173	0.0164	0.0168	0.0175	0.0175	**0.0169**
Expected Recovery Rate	0.0175	0.0171	0.0170	0.0169	0.0172	0.0169	0.0173	0.0167	0.0163	**0.0170**
A / E	90.2%	98.6%	98.3%	101.2%	100.6%	96.6%	97.5%	104.5%	107.2%	**99.7%**

Charts 34 through 37 show examples of more detailed actual and fitted comparisons we used to review the fit of the model. Overall the comparisons show that the model fits the hold-out sample data well. However, Chart 37 shows an example where there may be an opportunity to improve the model by adding an additional interaction variable.

Chart 34: Model Validation – Actual and Fitted Values by Claim Duration
Disability Category: Other Musculoskeletal
Using data sample withheld from modeling

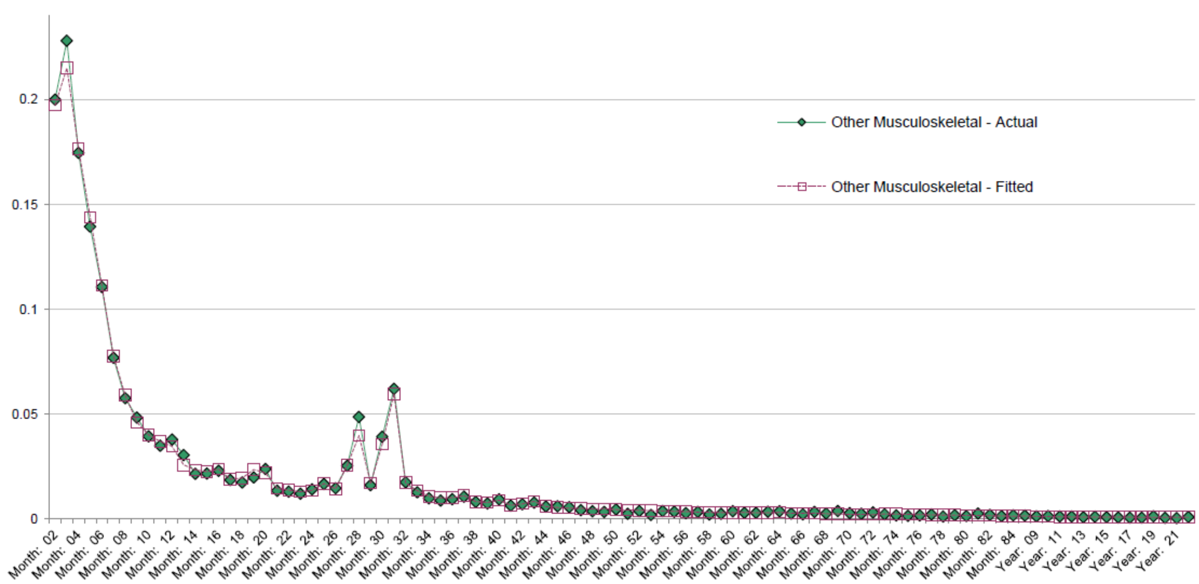


Chart 35: Model Validation – Actual and Fitted Values by Claim Duration
Disability Category: Cancer
Using data sample withheld from modeling

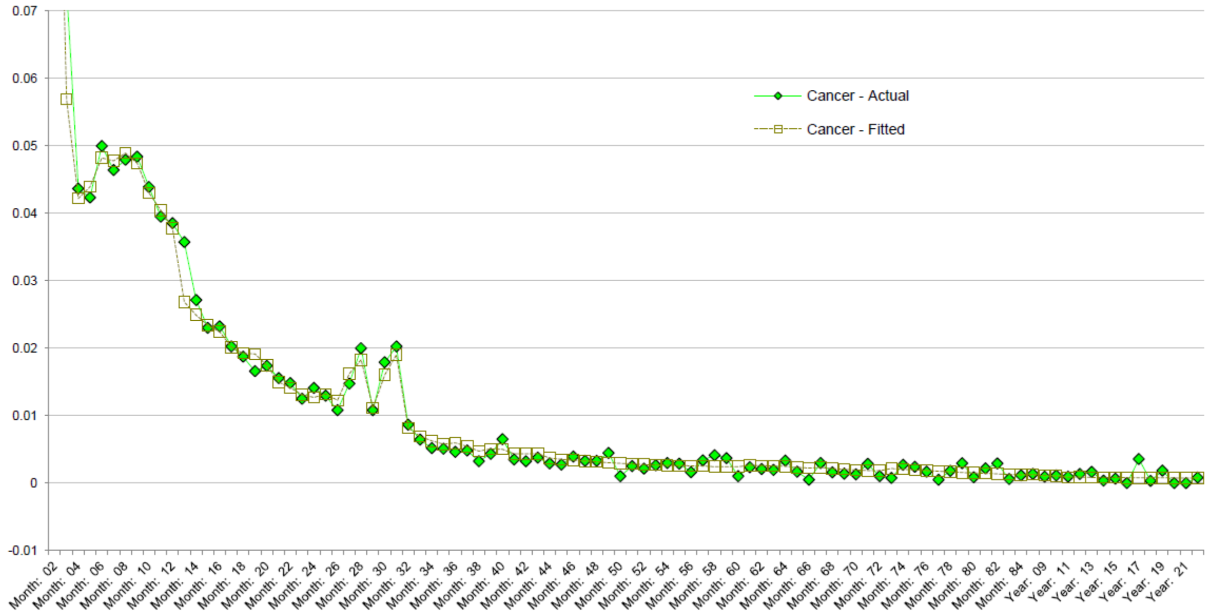


Chart 36: Model Validation – Actual and Fitted Values by Claim Duration
Disability Category: Digestive
Age at Disability: 50 - 59

Using all data

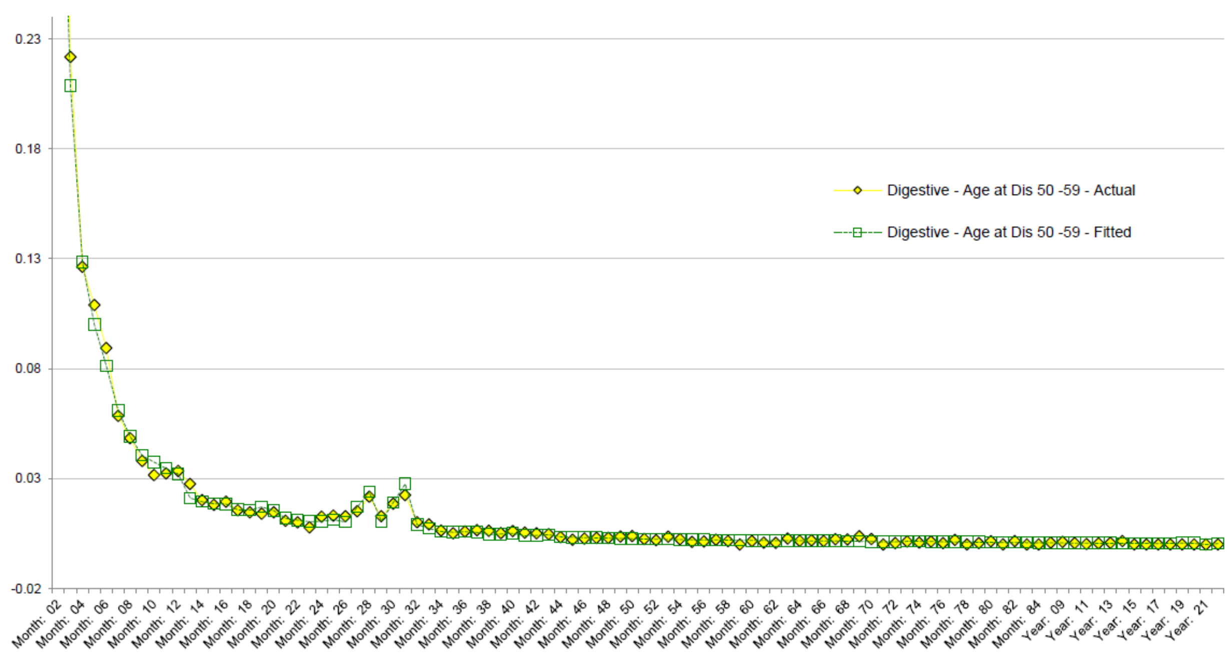
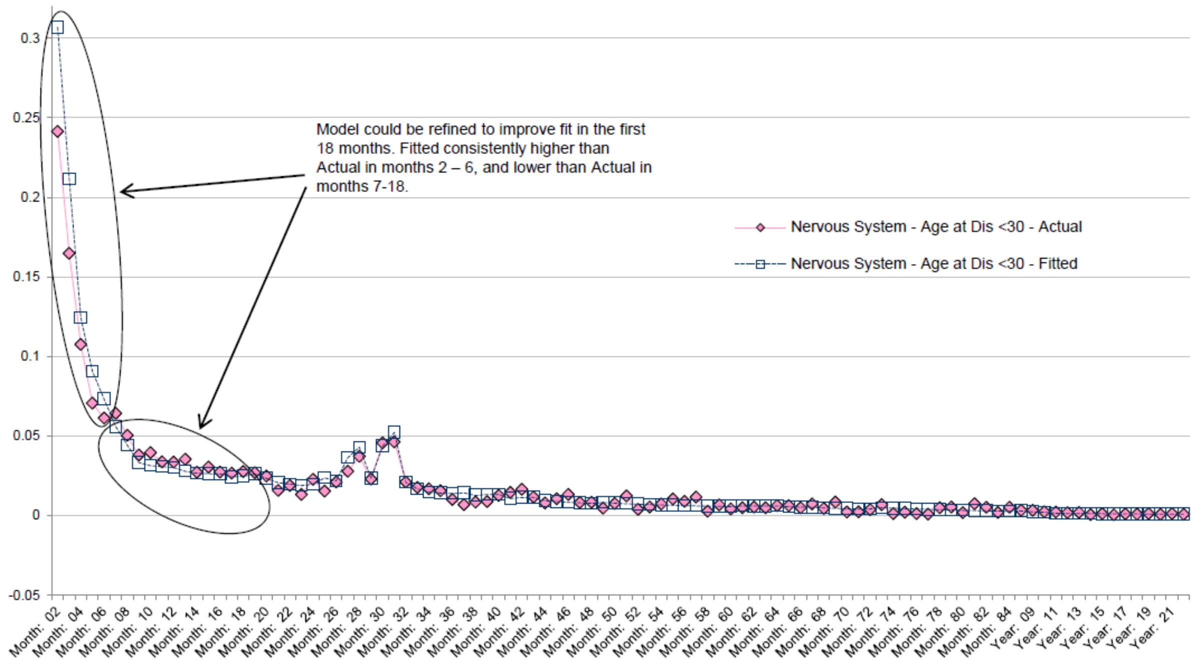


Chart 37: Model Validation – Actual and Fitted Values by Claim Duration

Disability Category: Nervous System

Age at Disability: <30

Using all data due to limited data for disability ages less than 30



Section 6: Predictive Analytics Trend Attribution Analysis

The 2016 study actual-to-expected ratios (A/Es) for recoveries (using the 2008 Experience Table as E) exhibited noticeable increases over the observation period (as well as modest decreases in mortality A/Es). From the analyses in Section 4, we determined that change in exposure mix by STD Integration Status appeared likely to explain a significant portion of the trend in recovery rates during the study period, while changes in mix for other variables appeared likely to have only minor influences on the recovery trend. In this section, we have used a predictive analytics (PA) approach to further analyze the drivers of the trend.

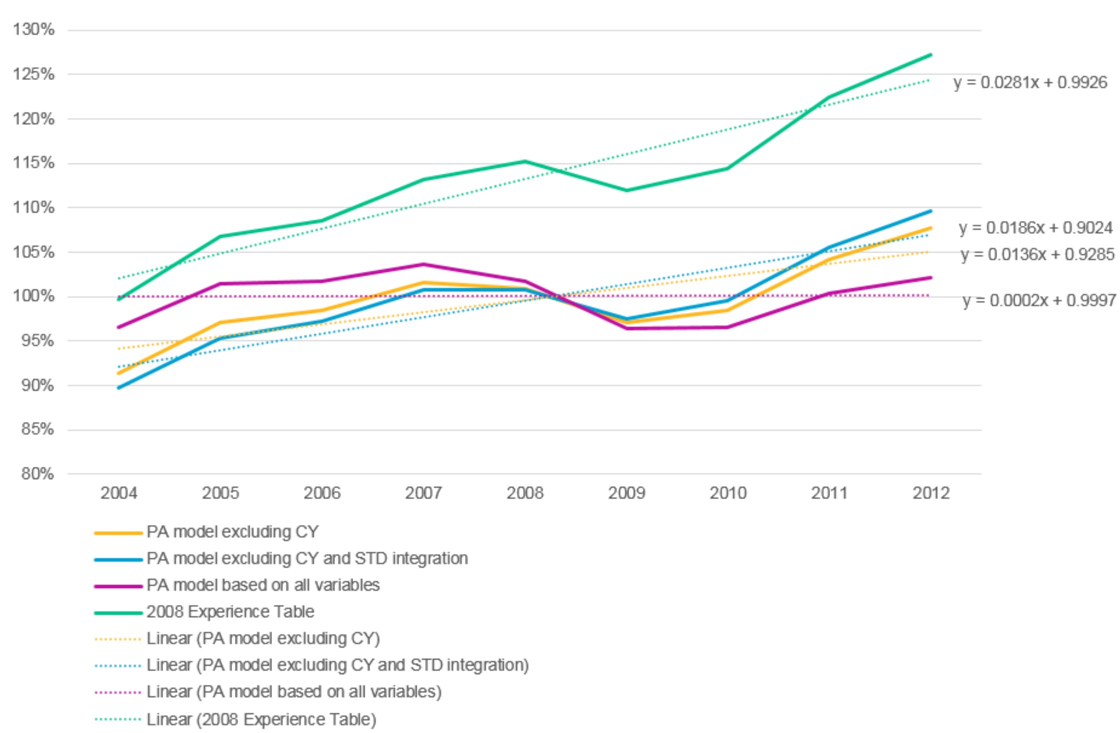
For our PA-based attribution analysis, identifiers for specific contributors were not available (not in CDB), so we used Company Size as a proxy.

We started by constructing a generalized linear model (GLM) (as described in Section 5) that incorporated 15 of the variables included in the 2016 study as well as 10 interaction variables. For the purposes of the analysis described here, we also fit a second version of the GLM described in Section 5 that added the Calendar Year (CY) variable to the model.

Our general measure of the importance of a specific variable to the trend was to measure A/Es using the 2016 terminations as "A" and different GLM versions as the "E"s; i.e., keeping the numerator constant while varying the denominator (We note that this is different than the traditional approach to A/E studies, which looks at different views of the numerator (A) while keeping the denominator (E) constant).

Our GLM based versions of the E involved re-calibrating the GLM after individual variables were removed. Thus, each E, or set of expected results, represents a different GLM that was developed separately.

Chart 38: Actual / Expected Recoveries by Calendar Year



The recovery A to E ratios using the actual data for the 2016 study and various iterations of the E are shown in chart 38 above:

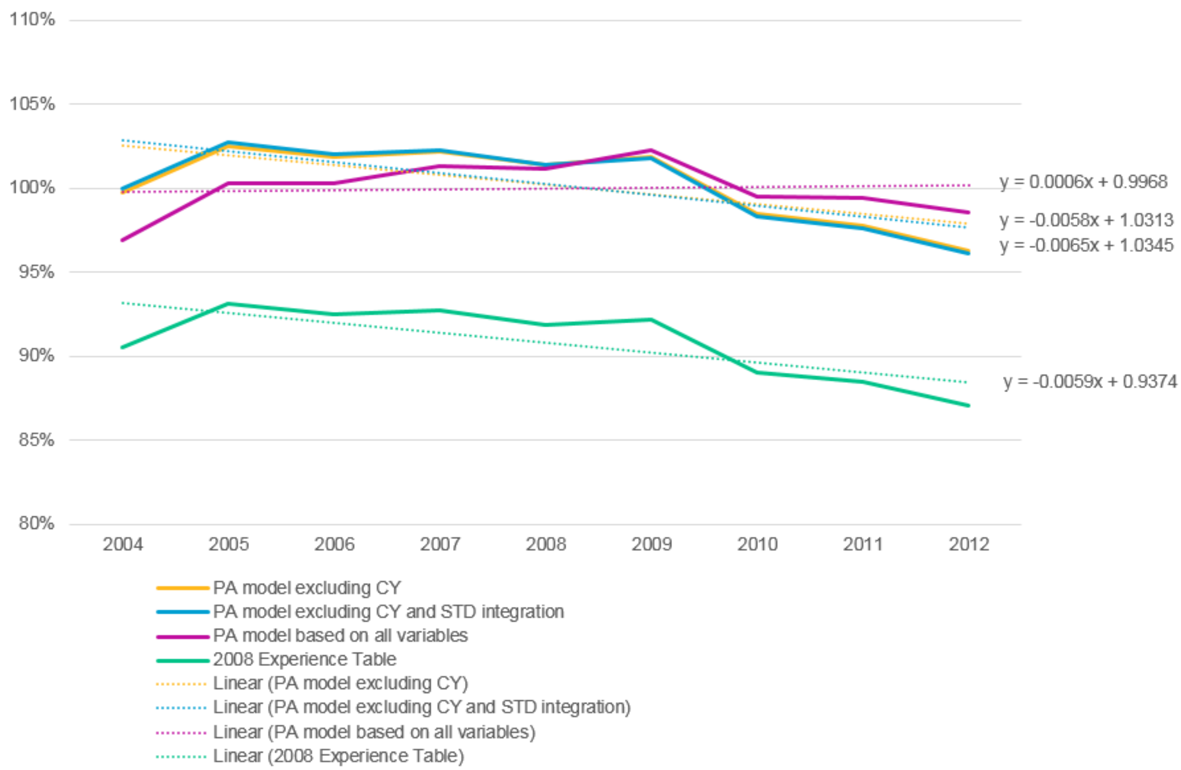
- 2008 Experience Table = 2016 study actual / 2008 Experience Table (trend = 2.81%)
- GLM based on all variables = 2016 study actual / GLM based of all variables including CY (trend = 0.0%)
- GLM excluding CY = 2016 study actual / GLM including all variables except CY (trend = 1.86%)
- GLM excluding CY and STD Integration = 2016 study actual / GLM including all variables except GLM except CY & STD integration (trend = 1.36%)

Our observations are:

- The GLM based on all variables (including CY) shows 0.00% trend – which implies a good fit to the data
- The GLM based on all variables excluding CY implies that the recovery trend is about 1.36%. (Note: This assumes that the caveats for the GLM approach listed below did not have a major impact on the results.)
- The 2016 Experience study data shows a trend of 2.81% when based on expected using the 2008 Experience Table. If we assume that the GLM excluding CY (1.36%) is the better measure of trend (all the caveats for the GLM would appear to apply equally to the 2008 Experience Table), this would imply that variables not reflected in the 2008 Experience Table expected essentially doubled the apparent trend. We scanned the 2016 data for variables that exhibited conspicuous changes in distribution over the 2016 study period and (as noted above) STD integration stood out. STD integration is not a variable included in the construction of the 2008 Experience Table.
- The GLM excluding CY and STD integration shows a trend of 1.86%, implying that the changes in proportion of business with integration could drive about one third of the difference between the 2.81% and the 1.36%. This implies that inclusion of the other 5 new variables plus any changes in the proportions of other variables, may account for the remaining 0.95%As mentioned above, the GLM and the 2008 Experience Table are subject to potentially material caveats:

- The trend line approach is potentially volatile given the limited number of years
- The analysis is limited to variables that were captured in the 2016 LTD study. Movement in other (non-captured) variables (e.g. Social Security approval) could also be contributors
- The GLM does not capture any contributions to trend that changes in specific company mix over time may have had. (Note: it does include the large/medium/small company splits, which was one of the study variables.)

Chart 39: Actual / Expected Deaths by Calendar Year



We applied the same approach to the mortality data for the 2016 study using various iterations of the GLM, with the results shown in chart 39 above:

- 2008 Experience Table = 2016 study actual / 2008 Experience Table (trend = -0.590%)
- GLM based on all variables = 2016 study actual / GLM based of all variables including CY (trend = 0.06%)
- GLM excluding CY = 2016 study actual / GLM including all variables except CY (trend = -0.58%)
- GLM excluding CY and STD Integration = 2016 study actual / GLM including all variables except GLM except CY & STD integration (trend = -0.65%)

Our observations are:

- The GLM based on all variables (including CY) shows a very low 0.06% trend – which implies a good fit to the data
- The GLM based on all variables excluding CY implies that the mortality trend is about -0.58%. (Note: This assumes that the caveats for the GLM approach listed above did not have a major impact on the results.)

- The 2016 Experience study data shows a trend of -0.59% when based on expected using the 2008 Experience Table, which implies that neither the new variables, nor the changes in mix for old variables had much net impact.
- The GLM excluding CY and STD integration shows a trend of -0.65%, which reinforces the above statement.

Caveats for mortality are the same as for recoveries.

About The Society of Actuaries

The Society of Actuaries (SOA), formed in 1949, is one of the largest actuarial professional organizations in the world dedicated to serving more than 30,000 actuarial members and the public in the United States, Canada and worldwide. In line with the SOA Vision Statement, actuaries act as business leaders who develop and use mathematical models to measure and manage risk in support of financial security for individuals, organizations and the public.

The SOA supports actuaries and advances knowledge through research and education. As part of its work, the SOA seeks to inform public policy development and public understanding through research. The SOA aspires to be a trusted source of objective, data-driven research and analysis with an actuarial perspective for its members, industry, policymakers and the public. This distinct perspective comes from the SOA as an association of actuaries, who have a rigorous formal education and direct experience as practitioners as they perform applied research. The SOA also welcomes the opportunity to partner with other organizations in our work where appropriate.

The SOA has a history of working with public policymakers and regulators in developing historical experience studies and projection techniques as well as individual reports on health care, retirement and other topics. The SOA's research is intended to aid the work of policymakers and regulators and follow certain core principles:

Objectivity: The SOA's research informs and provides analysis that can be relied upon by other individuals or organizations involved in public policy discussions. The SOA does not take advocacy positions or lobby specific policy proposals.

Quality: The SOA aspires to the highest ethical and quality standards in all of its research and analysis. Our research process is overseen by experienced actuaries and nonactuaries from a range of industry sectors and organizations. A rigorous peer-review process ensures the quality and integrity of our work.

Relevance: The SOA provides timely research on public policy issues. Our research advances actuarial knowledge while providing critical insights on key policy issues, and thereby provides value to stakeholders and decision makers.

Quantification: The SOA leverages the diverse skill sets of actuaries to provide research and findings that are driven by the best available data and methods. Actuaries use detailed modeling to analyze financial risk and provide distinct insight and quantification. Further, actuarial standards require transparency and the disclosure of the assumptions and analytic approach underlying the work.

Society of Actuaries
475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
www.SOA.org