Rating Endorsements using Generalized Linear Models

Edward W. Frees and Gee Lee

University of Wisconsin – Madison

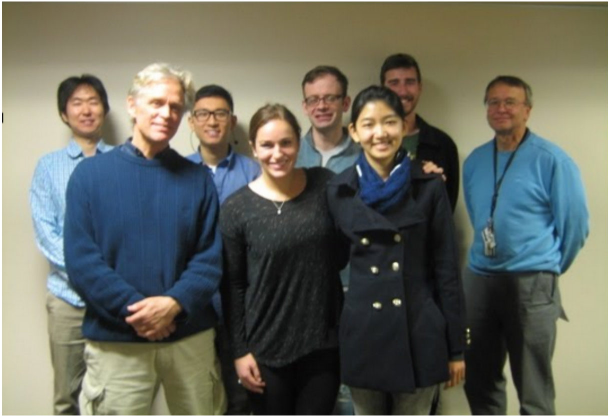
March, 2015

Outline

1. Introduction
2. LGPIF Data
3. Claims Modeling
4. Results from the Claims Modeling
5. Out of Sample Performance
6. Concluding Remarks

Fall 2014 Team

Introduction
LGPIF Data
Fund Claims and Rating Variables
Endorsements
Claims Modeling
Results from the Claims Modeling
Out of Sample Performance
Concluding Remarks



- Back Row: Gee Lee, Henry Chao, James Farley, Jordan Paszek, Greg Wanner
- Front Row: Jed Frees, Whitney Foss, Frances Lei
- Edward W. (Jed) Frees is the Principal Investigator

Why Endorsements?

- An endorsement, or a rider,
  - provides optional insurance coverage
  - may include alternative deductibles and coverage limits
  - also provides extensions to the type of peril (e.g., stolen jewelry in homeowners insurance) covered
- If there were no charge, it is not optional
- How do we charge for an endorsement in a generalized linear model setting?
  1. Endorsements form a relatively small fraction of the premium base and so only informal, ad hoc, approaches are needed.
  2. Use information from an external agency for this set of relativities
  3. Treat endorsements as another type of coverage and use GLM techniques to determine this set of prices.
     - Requires a substantial amount of data
     - Requires claims that are identified by type of endorsement.

Local Government Property Insurance Fund

- We seek other alternatives
- Local Government Property Insurance Fund (LGPIF)
  - was established to provide property insurance for local government entities that include counties, cities, towns, villages, school districts, and library boards.
  - covers over a thousand local government entities who pay approximately $25 million in premiums each year and receive insurance coverage of about $75 billion.
  - offers three major groups of insurance coverage: building and contents (BC), inland marine (construction equipment), and motor vehicles.
  - acts as a stand-alone insurance company, charging premiums to each local government entity (policyholder) and paying claims when appropriate.
  - not permitted to deny coverage for local government entities. Thus, the LGPIF acts as a "residual" market to a certain extent

Determining Effective Relativities

- Because of the size of the fund, there will be little difficulties using GLM to determine relativities/rates for the basic variables
- Endorsements are more difficult
  1. Fund is undergoing a major rate restructuring, politically sensitive
  2. Information from external agencies is expensive and not particularly relevant
  3. LGPIF data for optional coverages is limited
- We employed GLM techniques with restrictions on the coefficients through shrinkage using well-known penalized likelihood methods. Advantages:
  1. We provide relativities for endorsements in a disciplined manner, mitigating ad hoc adjustments
  2. Because we use GLM techniques, our approach is naturally multivariate and the introduction of endorsements accounts for the presence of other rating variables.

Building and Contents Claims Summary

- Number of policyholders is declining yet coverage is steady

Year	Average Frequency	Average Severity	Average Coverage	Number of Policyholders
2006	1.015	17,729	32,498,186	1,154
2007	1.235	15,158	35,275,949	1,138
2008	1.041	10,728	37,267,485	1,125
2009	1.277	9,934	40,355,382	1,112
2010	1.285	33,026	41,242,070	1,110
2011	1.036	20,554	42,503,989	1,094

- Going forward, 2006–2010 is for training, 2011 is for validation

Summary of Claim Frequency and Amount, Deductibles, and Coverages

	Minimum	Median	Average	Maximum	Spearman Correlation Frequency	Claim*
Claim Frequency	0	0	1.165	263	–	0.413
Claim Amount	0	0	17,258	12,922,218	0.413	–
Deductible	500	1,000	3,365	100,000	0.073	0.324
Coverage (000's)	9	11,354	37,281	2,444,797	0.438	0.243

Note: *The claim correlations are based on 1,679 observations with at least one claim using the average claim (amount divided by frequency).

- All variables are right skewed
- Frequency and severity are related

Description of Base Rating Variables

Variable	Description
`EntityType`	Categorical variable that is one of six types: (Village, City, County, Misc, School, or Town)
`LnCoverage`	Total building and content coverage, in logarithmic millions of dollars
`LnDeduct`	Deductible, in logarithmic dollars
`NoClaimCredit`	Binary variable to indicate no claims in the past two years
`Fire5`	Binary variable to indicate the fire class is below 5 (The range of fire class is $0 \sim 10$)

Claims Summary by Entity Type, Fire Class, and No Claim Credit

Variable	Number of Policies	Average Frequency	Average Claim
`EntityType`			
Village	1,341	0.529	11,869.75
City	793	2.042	39,177.27
County	328	4.973	95,832.87
Misc	609	0.204	40,011.89
School	1,597	1.500	70,606.31
Town	971	0.118	18,449.46
`Fire5`-No	2,508	0.563	18,346.54
`Fire5`–Yes	3,131	1.655	68,798.51
`NoClaimCredit`–No	3,786	1.571	53,283.61
`NoClaimCredit`–Yes	1,853	0.349	32,666.54
Total	5,639	1.169	49,358.53

- There is substantial variation in the claims distribution by each rating variable
- By itself, `Fire5` is counter-intuitive. We anticipate lower claims when the fire class is below 5 (`Fire5`=Yes)

Description of Endorsements

Variable	Description
Business Interruption	Reimburses an insured for business interruption (lost profits and continuing fixed expenses)
Accounts Receivable	Adds coverage for money owed by its debtors during business interruption due to a covered loss.
Pier and Wharf	Loss of watercraft, by the pressure of ice or water on piers and wharves
Fine Arts	Adds coverage (agreed value) on fine arts, either per item or per exhibit
Golf Course Grounds	Adds coverage to golf course type property such as greens, tees, fairways, etc.
Special Use Animal	Adds coverage for police enforcement animals, such as dogs and horses
Zoo Animals	Adds coverage for zoo animals. Animal mortality is specifically excluded.
Vacancy Permit	Allows claims from covered losses arising from vacant property
Monies and Securities	Adds coverage for monies and securities for loss by theft, disappearance, or destruction (A: loss inside premise, B: loss outside premise).
Other Endorsements	Other additional endorsements, including ordinance & law, and extra expenses

Summary of Claim Frequency and Severity by Endorsement

Endorsements	Num of Obs	Average Frequency	Average Claim	Average Endorsement Coverage
Business Interruption	225	6.493	228,393	2,679,595
Accounts Receivable	172	5.360	283,634	853,966 7
Pier and Wharf	312	2.599	41,262	245,445
Fine Arts	67	13.537	419,083	12,160,956
Golf Course Grounds	28	18.036	469,986	237,500
Zoo Animals	10	73.900	1,615,405	1,102,790
Special Use Animal	256	5.617	95,790	21,903
Vacancy Permit	225	4.902	158,402	1,779,212
Monies and Securities	2,137	2.071	60,868	58,928
Other Endorsements	53	5.000	40,819	4,763,019
All Policies				
Total	5,639	1.169	49,359	

- There is substantial variation in the claims distribution by each endorsement

Shrinkage Estimation

- Begin with classic linear model shrinkage estimation, minimize

$$\sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{k} x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^{k} \beta_j^2 .$$

  - Values of λ control the complexity of the model; smaller values mean less shrinkage
- Can write this in terms of classical "ridge regression"

$$\hat{\beta}_{shrink} = \left(\mathbf{X}'\mathbf{X} + \lambda \mathbf{I} \right)^{-1} \mathbf{X}'\mathbf{y}$$

  - appealing in instances of collinearity
- For (nonlinear) GLMs, we use a penalized likelihood of the form

$$l(\beta) = \sum_{i=1}^{n} \log f(y_i) - \lambda \left\| \mathbf{R}\beta - \mathbf{r} \right\|^2 ,$$

Offset

- We used the offset variable

$$\texttt{offset} = \ln(0.95)\texttt{AC05} + \ln(0.90)\texttt{AC10} + \ln(0.85)\texttt{AC15}.$$

- AC05 represents a binary variable to indicate the presence of a 5% alarm system (meaning that automatic smoke alarms exist in some of the main rooms)
- similarly for AC10 and AC15
- Also included an offset for vacancy permits

Endorsements

- Summary statistics suggest that *endorsement coverage amount* may influence claims outcomes
- To capture this, using GLMs
  - y_B represents claims from a base coverage, mean $\mu_B = \exp(\mathbf{x}'\beta)$
  - Let y_E be the claims from an endorsement, mean μ_E.

$$\mu = \mathrm{E}\ y = \begin{cases} \mu_B = \exp(\mathbf{x}'\beta) & \text{endorsement not present} \\ \mu_B + \mu_E = \exp(\mathbf{x}'\beta + \beta_E x_E) & \text{endorsement present} \end{cases}.$$

  - Let $Coverage_E$ and $Coverage_B$ represent amount of coverage for the endorsement and base (building and contents)
  - Define

$$x_E = \ln\left(1 + \frac{Coverage_E}{Coverage_B}\right).$$

  - We will use x_E in our GLM specifications

Endorsements

- Base mean $\mu_B = \exp(\mathbf{x}'\beta)$, Endorsement mean μ_E.

$$\mu = \mathrm{E}\ y = \begin{cases} \mu_B = \exp(\mathbf{x}'\beta) & \text{endorsement not present} \\ \mu_B + \mu_E = \exp(\mathbf{x}'\beta + \beta_E x_E) & \text{endorsement present} \end{cases}.$$

$$x_E = \ln\left(1 + \frac{Coverage_E}{Coverage_B}\right).$$

- With this specification, we have

$$\begin{array}{rcl} \mu_E & = & \exp(\mathbf{x}'\beta + \beta_E x_E) - \mu_B \\ & = & \mu_B \left[\left(1 + \frac{Coverage_E}{Coverage_B}\right)^{\beta_E} - 1 \right] \\ & \approx & \mu_B \left[\left(1 + \beta_E \frac{Coverage_E}{Coverage_B}\right) - 1 \right] \\ & = & \beta_E \times Coverage_E \times \left(\frac{\mu_B}{Coverage_B}\right), \end{array}$$

using the approximation $(1+z)^b \approx 1 + bz$.

- Endorsement Price μ_E is a factor times the endorsement coverage, rescaled by the overall cost per unit coverage. The factor, β_E, is estimated from the data.

Claims Modeling

- For frequency-severity, we fit Poisson for frequency and gamma for severity
- For pure premiums, we used a Tweedie model. I focus on the frequency-severity results
- Used the logarithmic link throughout
- Investigated shrinking *endorsement* parameter estimates, although not base model estimates

Poisson Frequency Model Using Shrinkage Estimation

	$\lambda = 0$		$\lambda = 5$		$\lambda = 1,000$	
	Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error
Basic Rating Variables						
Intercept	-1.542	0.127	-1.519	0.127	-1.669	0.127
LnCoverage	0.703	0.045	0.704	0.045	0.716	0.046
LnDeduct	-0.061	0.011	-0.064	0.010	-0.043	0.010
TypeCity	0.154	0.142	0.163	0.142	0.142	0.142
TypeCounty	-0.449	0.200	-0.679	0.194	-1.637	0.164
TypeMisc	-0.214	0.173	-0.174	0.171	-0.105	0.170
TypeSchool	-6.008	0.173	-6.009	0.173	-5.987	0.173
TypeTown	-0.460	0.160	-0.461	0.160	-0.446	0.161
Fire5	-0.154	0.037	-0.159	0.037	-0.176	0.037
NoClaimCredit	0.000	0.104	0.004	0.104	0.018	0.104
LnCoverage*TypeCity	0.010	0.049	0.007	0.049	0.009	0.049
LnCoverage*TypeCounty	0.123	0.056	0.177	0.055	0.385	0.050
LnCoverage*TypeMisc	-0.257	0.059	-0.262	0.058	-0.280	0.058
LnCoverage*TypeSchool	1.241	0.051	1.242	0.051	1.228	0.051
LnCoverage*TypeTown	0.166	0.092	0.165	0.092	0.154	0.092
LnCoverage*NoClaimCredit	-0.194	0.025	-0.194	0.025	-0.197	0.025
Endorsements						
LnBusInterCovRat	0.188	0.046	0.189	0.045	0.036	0.020
LnAccRecCovRat	0.132	0.093	0.018	0.088	0.008	0.022
LnAddInsCovRat	0.295	0.045	0.284	0.045	0.044	0.021
LnPierWarfCovRat	0.039	0.078	0.045	0.075	0.003	0.022
LnSpecialAnimalCovRat	0.348	0.599	0.067	0.289	0.001	0.022
LnZooAnimalCovRat	3.491	0.810	0.430	0.295	0.004	0.022
LnFineArtsCovRat	0.321	0.054	0.392	0.048	0.068	0.020
LnGolfCourseCovRat	1.019	0.330	0.257	0.273	0.002	0.022

Gamma Severity Model for Average Claim

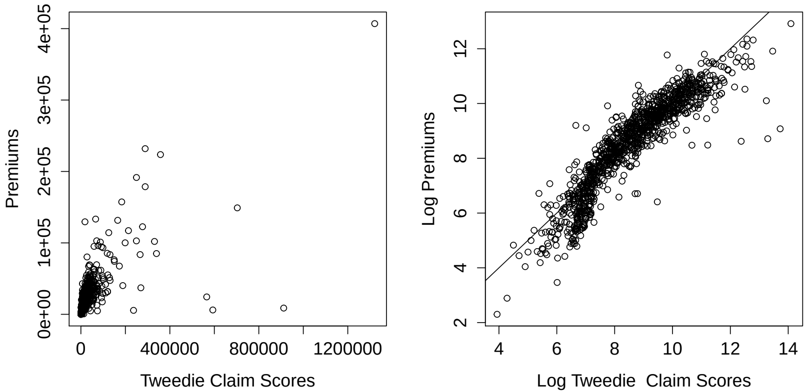
	Estimate	Standard Error
Base Rating Variables		
Intercept	8.396	0.105
LnCoverage	0.404	0.033
TypeCity	-0.286	0.119
TypeCounty	-0.562	0.146
TypeMisc	0.606	0.187
TypeSchool	0.941	0.118
TypeTown	1.622	0.195
Endorsements		
LnMoneySecCovRat	0.153	0.261
LnMoneySecLimitedCovRat	0.244	0.165
ϕ (dispersion)	2.143	

- Did not shrink the severity model

Relativities for Base Rating Variables and Endorsements

	$\lambda = 0$	$\lambda = 5$	$\lambda = 1,000$
Basic Rating Variables			
LnCoverage	3.024	3.029	3.065
LnDeduct	0.941	0.938	0.958
TypeCity	0.885	0.891	0.874
TypeCounty	0.414	0.348	0.166
TypeMisc	1.129	1.169	1.229
TypeSchool	0.023	0.023	0.023
TypeTown	3.804	3.797	3.814
Fire5	0.857	0.853	0.839
NoClaimCredit	0.815	0.818	0.827
Endorsement			
LnBusInterCovRat	1.207	1.208	1.036
LnAccRecCovRat	1.141	1.018	1.008
LnAddInsCovRat	1.343	1.329	1.045
LnPierWarfCovRat	1.040	1.046	1.003
LnSpecialAnimalCovRat	1.416	1.069	1.001
LnZooAnimalCovRat	32.816	1.538	1.004
LnFineArtsCovRat	1.379	1.481	1.071
LnGolfCourseCovRat	2.771	1.293	1.002
LnMoneySecCovRat	1.165	1.165	1.165
LnMoneySecLimCovRat	1.276	1.276	1.276

Comparison of Tweedie model scores to external agency premium scores



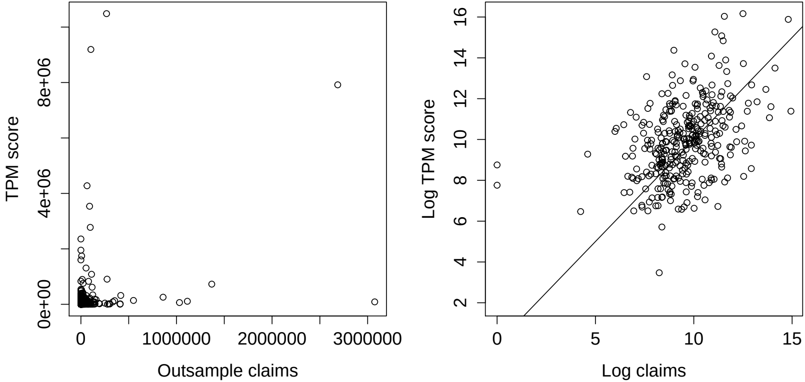
- The Spearman correlation is 94.17%.

Spearman Correlations among Scores and Out of Sample Claims

	Freq Sev Model		Tweedie		Premiums
	$\lambda = 0$	$\lambda = 1,000$	$\lambda = 0$	$\lambda = 1,000$	
Frequency Severity Model, $\lambda = 1,000$	0.9980				
Tweedie Model, $\lambda = 0$	0.9033	0.9011			
Tweedie Model, $\lambda = 1,000$	0.9126	0.9120	0.9767		
Out of Sample Premiums	0.9008	0.8999	0.9417	0.9477	
Out of Sample Claims	0.4269	0.4217	0.4128	0.4152	0.4218

- Frequency-severity relativities easier to interpret than one's from Tweedie
- Tweedie did a little better in reproducing premiums from an external agency
- Both Frequency-severity and Tweedie were comparable to external premiums in anticipating held-out claims

Comparison of Frequency Severity Scores to Out of Sample Claims for 2011



- The Spearman correlation coefficient is 42.69%.

Contributions of this Work

1. Detailed analysis of the Wisconsin Local Government Property Insurance Fund.
   - There is little in the literature on government property and casualty actuarial applications.
   - The LGPIF is similar to small commercial property insurance, making our work of interest to a broad readership.
2. Detailed analysis in the manner of a case study so that other analysts may replicate parts of our approach.
   - We provide relativities not only for our primary rating variables but also for endorsements.
   - Introduce an approach for handling these optional coverages when it is not known whether or not a claim is due to an endorsement.

Contributions of this Work

1. Detailed analysis of the Wisconsin Local Government Property Insurance Fund.
2. Detailed analysis in the manner of a case study so that other analysts may replicate parts of our approach.
3. Explored the use of shrinkage estimation in ratemaking
   - Shrinkage is particularly appealing in the case of endorsements.
   - Little predictive ability was lost by using shrinkage methods and they gave much more intuitively appealing relativities.
   - Helpful to have relativities that can be calibrated in a disciplined manner and are consistent with sound economic, risk management, and actuarial practice.

Introduction

LGPIF Data
- Fund Claims and Rating Variables
- Endorsements

Claims Modeling

Results from the Claims Modeling

Out of Sample Performance

Concluding Remarks

Conclusion

More information about our efforts to model the LGPIF at:

`https://sites.google.com/a/wisc.edu/local-government-property-insurance-fund/`

Thank you for your kind attention.