



# Antitrust Notice

- **The Casualty Actuarial Society is committed to adhering strictly to the letter and spirit of the antitrust laws. Seminars conducted under the auspices of the CAS are designed solely to provide a forum for the expression of various points of view on topics described in the programs or agendas for such meetings.**
- **Under no circumstances shall CAS seminars be used as a means for competing companies or firms to reach any understanding – expressed or implied – that restricts competition or in any way impairs the ability of members to exercise independent business judgment regarding matters affecting competition.**
- **It is the responsibility of all seminar participants to be aware of antitrust regulations, to prevent any written or verbal discussions that appear to violate these laws, and to adhere in every respect to the CAS antitrust compliance policy.**



**CL-4: Predicting the unpredictable commercial line  
business – predictive modeling applications for  
excess loss and specialty lines**

CAS Ratemaking and Product Management  
Seminar

12 March 2013

# Agenda

---

- ▶ Excess loss claims overview and considerations
- ▶ A predictive modeling application
- ▶ Details of excess loss models
- ▶ Predictive model data sources
- ▶ Examples of predictor variables
- ▶ Model performance metrics
- ▶ Use of excess loss models
- ▶ Summary

---

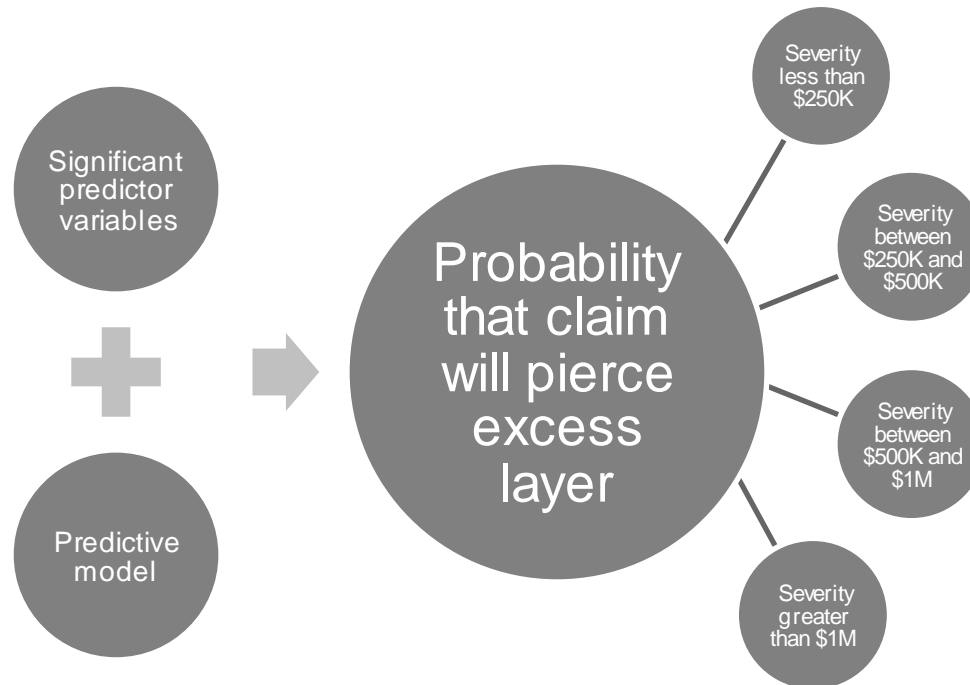
# Excess loss claims overview and considerations

---

- ▶ Claim severity data often has insufficient volume to make credible predictions of excess losses due to a limited number of claims that actually pierce the excess layer
- ▶ Claims that pierce the excess layers usually take several years to settle
- ▶ It is challenging to build predictive models that estimate the actual claim severity with data that is not fully developed
- ▶ Alternatively, predictive models can estimate the likelihood of a claim piercing the excess layer and a potential range of severity values
- ▶ These types of models are less sensitive to data limitations and the long time claims could take to fully develop

# How predictive modeling can help

- ▶ A predictive model can estimate the probability that a claim will pierce the excess layer
- ▶ The model can be further refined to predict the probability that a claim will reach or exceed a certain incurred loss value



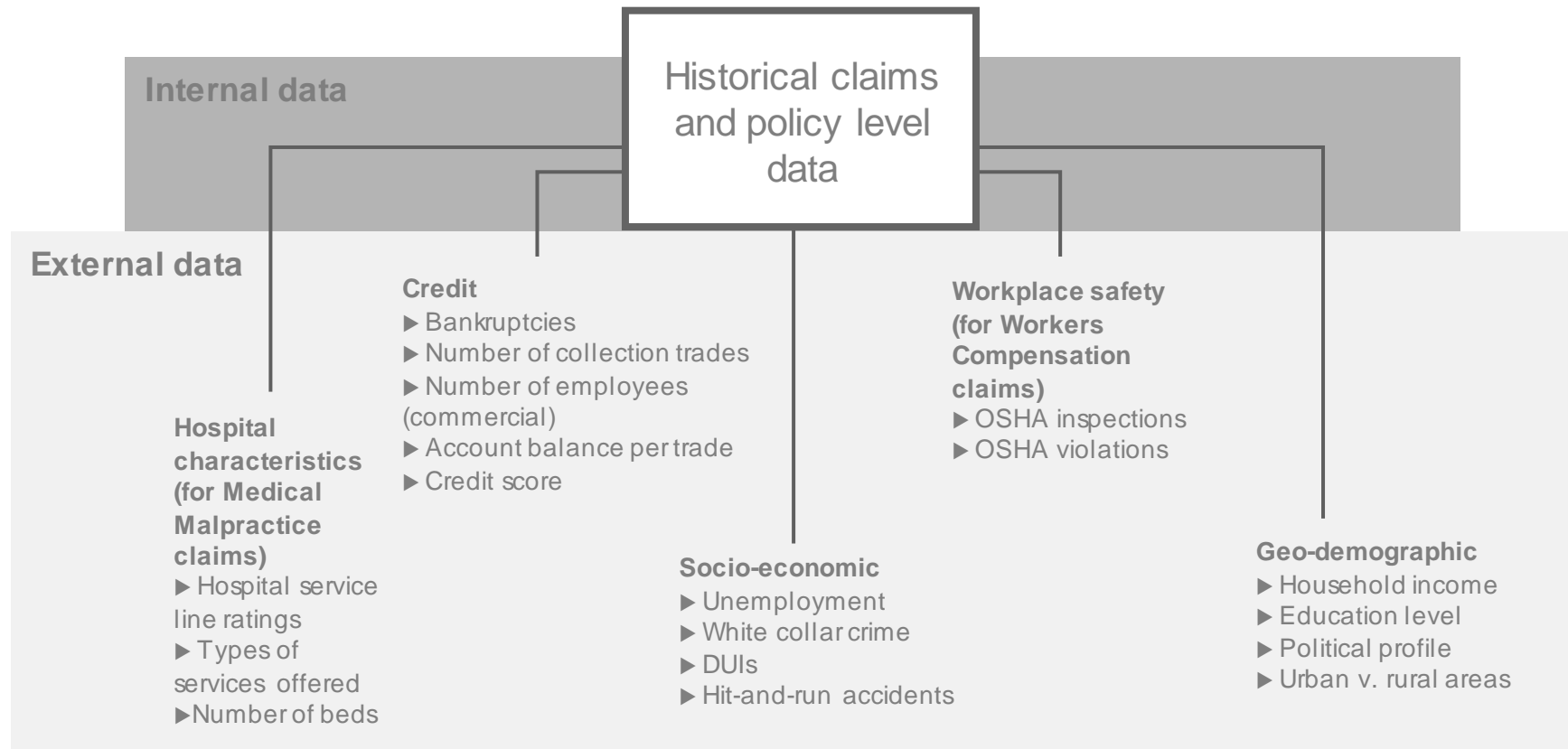
# How does an excess loss model work

---

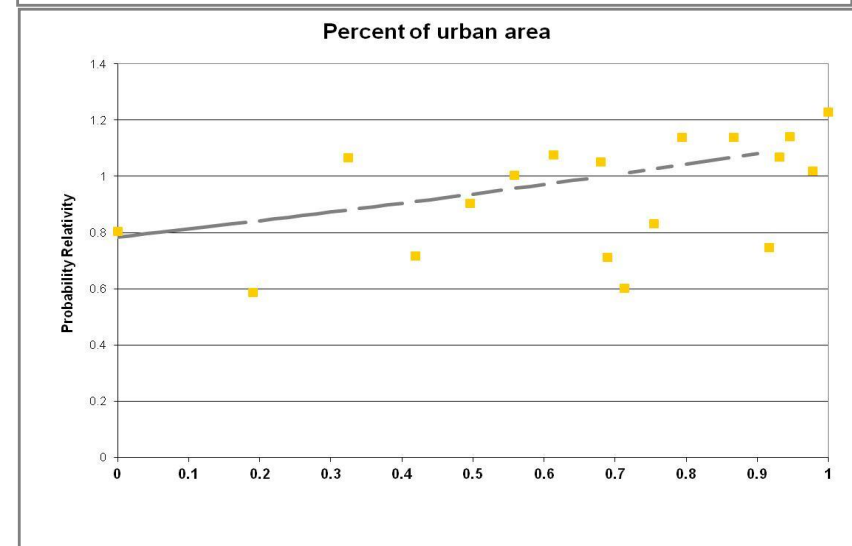
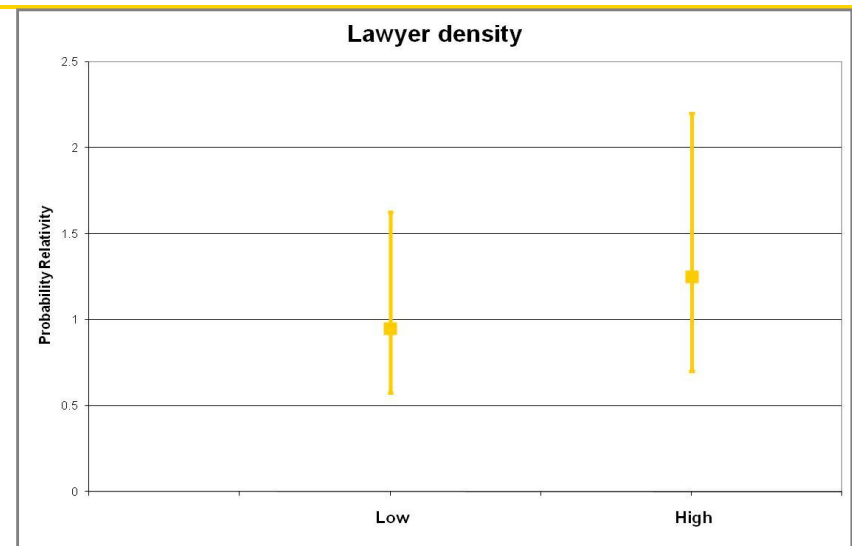
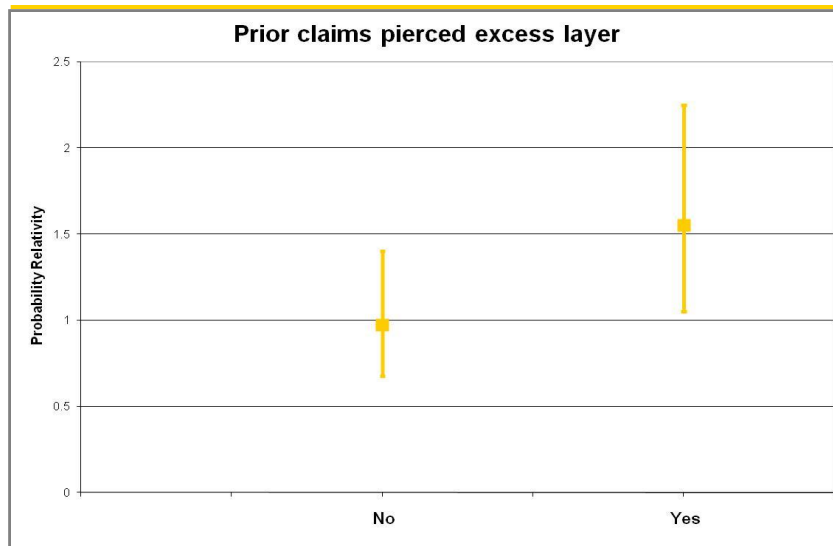
- ▶ This type of model belongs in the family of generalized linear models, which are widely used in the insurance industry for the development of underwriting models, the management of claims, the development of marketing strategies, etc
- ▶ The actual form of the model is a logistic regression where the response variable is a Yes or a No for whether or not a particular claim will pierce an excess layer
  - ▶ The response can be broken down further to ranges of severity values (as shown in the prior slide) and then a multinomial logistic regression model would be developed
- ▶ The model would then produce a probability for each value of the response
  - ▶ For example, in the Yes/No scenario, the model could predict that for a specific claim the probability of the claim piercing the excess layer is 0.85 and the probability of the claim not piercing the excess layer is 0.15
  - ▶ A probability greater than 0.5 is usually translated to a certain outcome, meaning that in the example above, the model prediction would indicate that the claim would pierce the excess layer

# What types of data sources have been found to be predictive?

- ▶ A variety of internal and external data sources have been found to be predictive



# Selected examples of variables found to be predictive

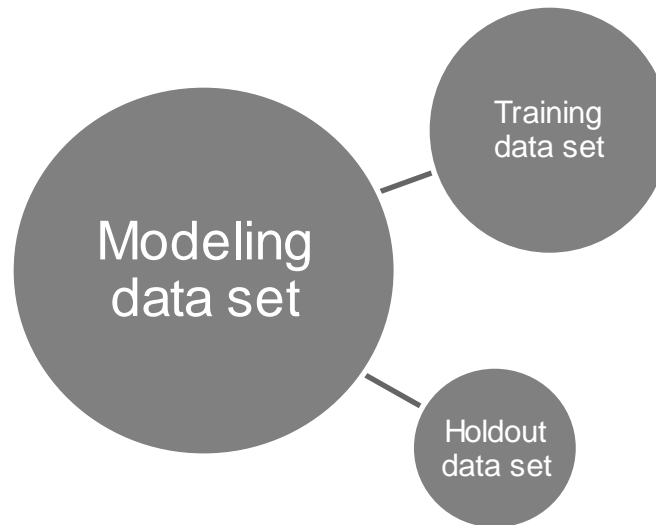




# Model performance metrics

---

- ▶ The modeling data set is usually split into a training data set to develop the model and a holdout data set to test the model performance
  - ▶ A reasonable distribution for training versus holdout data is a 2:1 ratio



- ▶ A few different metrics exist to measure the model performance using the holdout data set including the following:
  - ▶ Concordance metric comparing pairs of observations
  - ▶ Alternative concordance metric assessing individual observations
  - ▶ Liftchart
- ▶ These metrics are discussed in the following slides

# Concordance metrics

- ▶ Concordance metrics capture the percentage of cases that the model “gets it right”
- ▶ The concordance metric that compares pairs of observations is based on the percentage of concordant pairs of observations from the modeling dataset

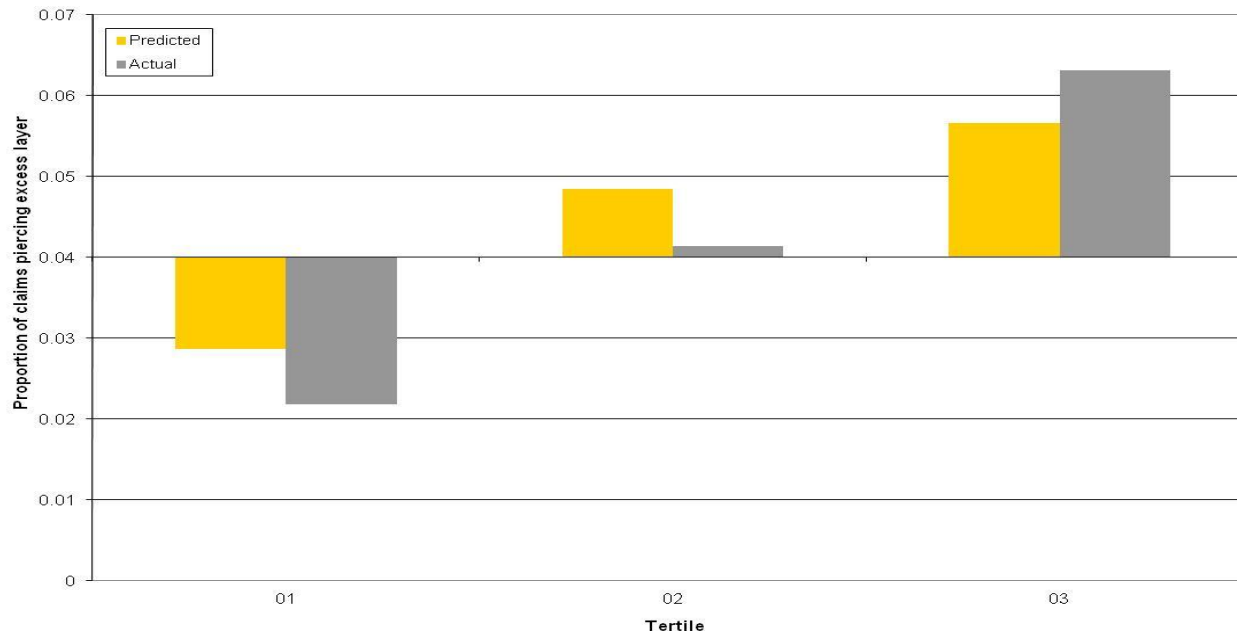
	Claim		Claim	
	A	B	C	D
Pierced excess layer	Yes	No	Yes	No
Model predicted probability of piercing excess layer	0.85	0.56	0.42	0.65
Pair concordant	Yes		No	

- ▶ The concordance metric that assesses individual observations is based on the percentage of cases where the model accurately predicts the observed outcome

	Claim		Claim	
	A	B	C	D
Pierced excess layer	Yes	No	Yes	No
Model predicted probability of piercing excess layer	0.85 (Yes)	0.56 (Yes)	0.42 (No)	0.65 (Yes)
Model predicted observed outcome accurately	Yes	No	No	No

# Liftchart

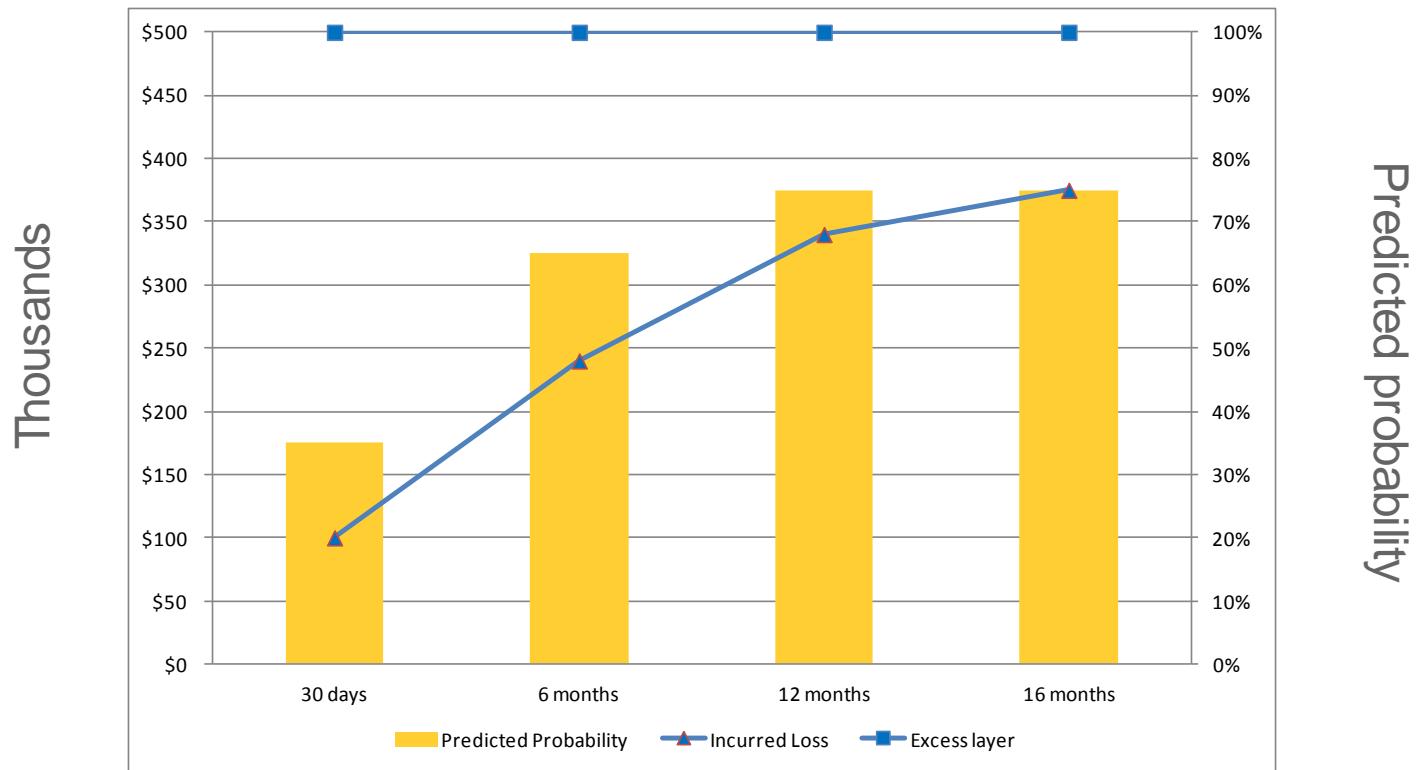
- ▶ A liftchart can be created from the holdout data set using the following process:
  - ▶ Order the claims from lowest to highest predicted probability of a claim piercing the excess layer
  - ▶ Split the claims into a credible number of groups (based on claim volume) with equal numbers of claims
  - ▶ Calculate the average predicted probability of claims piercing the excess layer and the average proportion of claims piercing the excess layer in each group
- ▶ A liftchart provides a visual representation of the comparison between the model predictions and the actual claim severity outcomes
- ▶ An example is shown below with three groups of claims where the average proportion of claims piercing the excess layer is 4% (this model example is based on a working layer where claims would pierce the layer every year):



# How can excess loss models be used?

## Claim management process

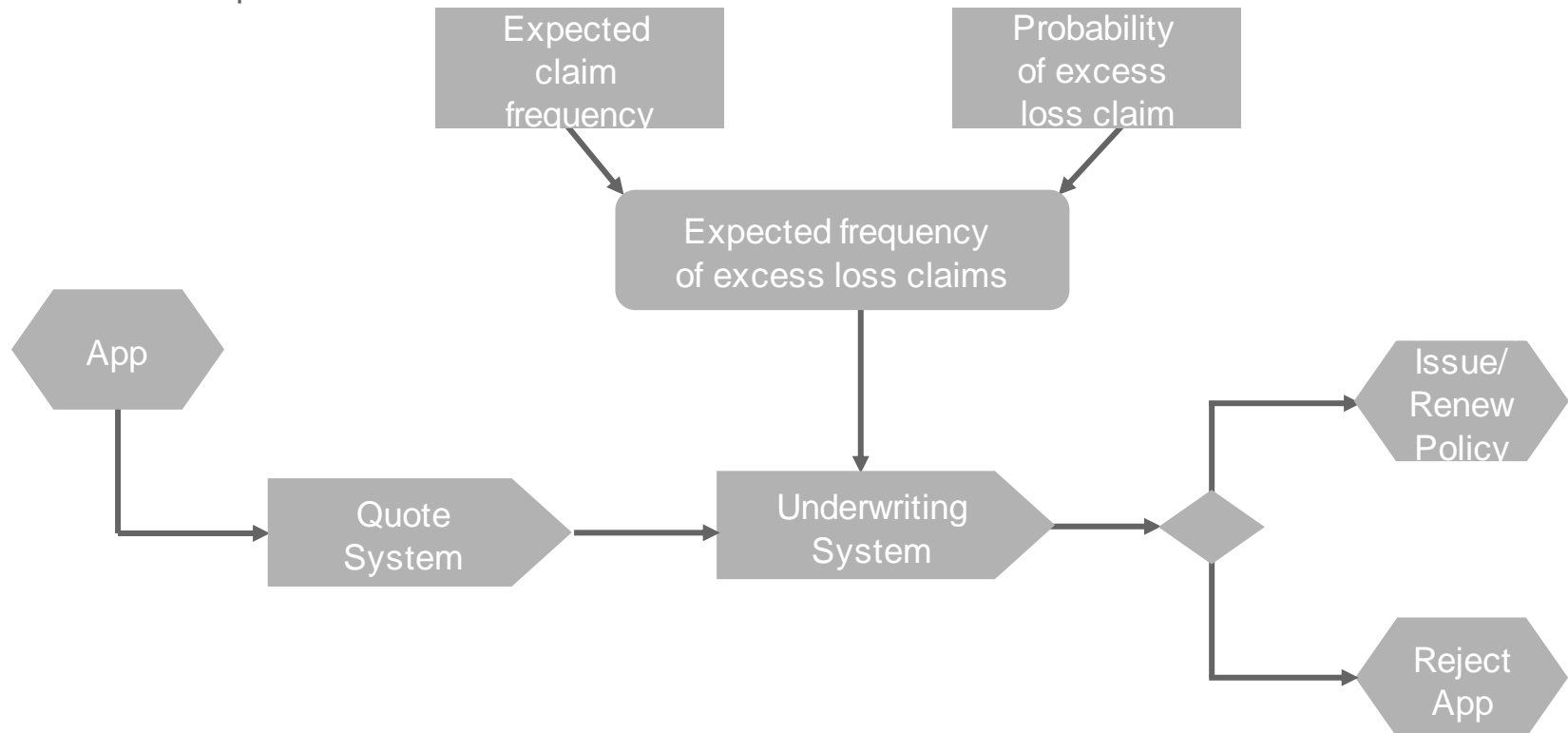
- ▶ Models that predict the probability that a claim will pierce the excess layer can be used to enhance the claim management process
  - ▶ After the claim is reported and periodically when new information is gathered the model would be used to calculate a predicted probability that an individual claim will pierce the excess layer
  - ▶ The example below shows the timeline of an individual claim in terms of the predicted probability and the incurred loss in relation to the excess layer



# How can excess loss models be used?

## Underwriting process

- ▶ Models that predict the probability that a claim will pierce the excess layer can also be used to enhance the existing underwriting process
  - ▶ Expected # of claims \* Predicted probability of piercing excess layer = Expected # of claims piercing excess layer
  - ▶ The excess loss model would need to be modified to adjust for predictor variables that are claim-specific since no individual claim information would be known



---

# Summary

---

- ▶ Excess loss models can be successfully developed after adjusting for data limitations and claim development timelines
- ▶ A form of logistic regression can be used to develop excess loss models
- ▶ Internal and external predictor variables should be considered
- ▶ Excess loss models can be deployed as a step in the claim management or the underwriting process