watsonwyatt.com

CAS Predictive Modeling Seminar

Homeowners Modeling

Gaétan Veilleux, FCAS
October 11, 2007

Watson Wyatt Worldwide

Agenda

- Case for unbundling the perils
- Practical Issues
- Traditional rating variables
- New rating variables

Agenda

- Case for unbundling the perils
- Practical Issues
- Traditional rating variables
- New rating variables

Legacy of indivisible premium for residential property lines

- Package policy formed when Fire was major % of total losses (1950s)
- ISO issues simplified (standardized) policy form (1970s)
- Remnant of paper manuals and inflexible quoting systems
- Lack of attention to specific cause of loss trends
- Comfort in status quo

In contrast, personal auto premium

- Coverages are priced with modular approach
- Accepted by customers, agents, regulators, etc.
- In general, more pricing segmentation than homeowners
- More responsive trend detection (eg liability trends vs parts/labor trends)
- Matches how experience is monitored

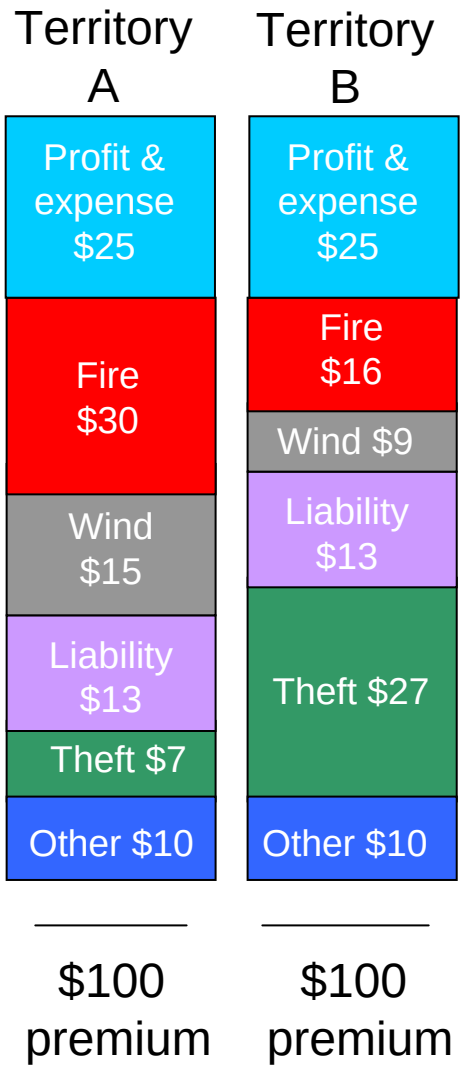
Why unbundle?

- Improved rating accuracy
  - rate classification equity
  - favorable selection
  - better competitive position
  - improved profitability
- Improved ability to monitor and respond to trends and emerging causes of loss

More detailed reasons...

- The changing landscape - claims
- The share of loss costs by peril varies considerably by geography
- Effect of rating factors varies considerably by peril
  - traditional rating factors
  - territory
  - inhabitant info
  - external info

Percent of losses by peril varies across territories



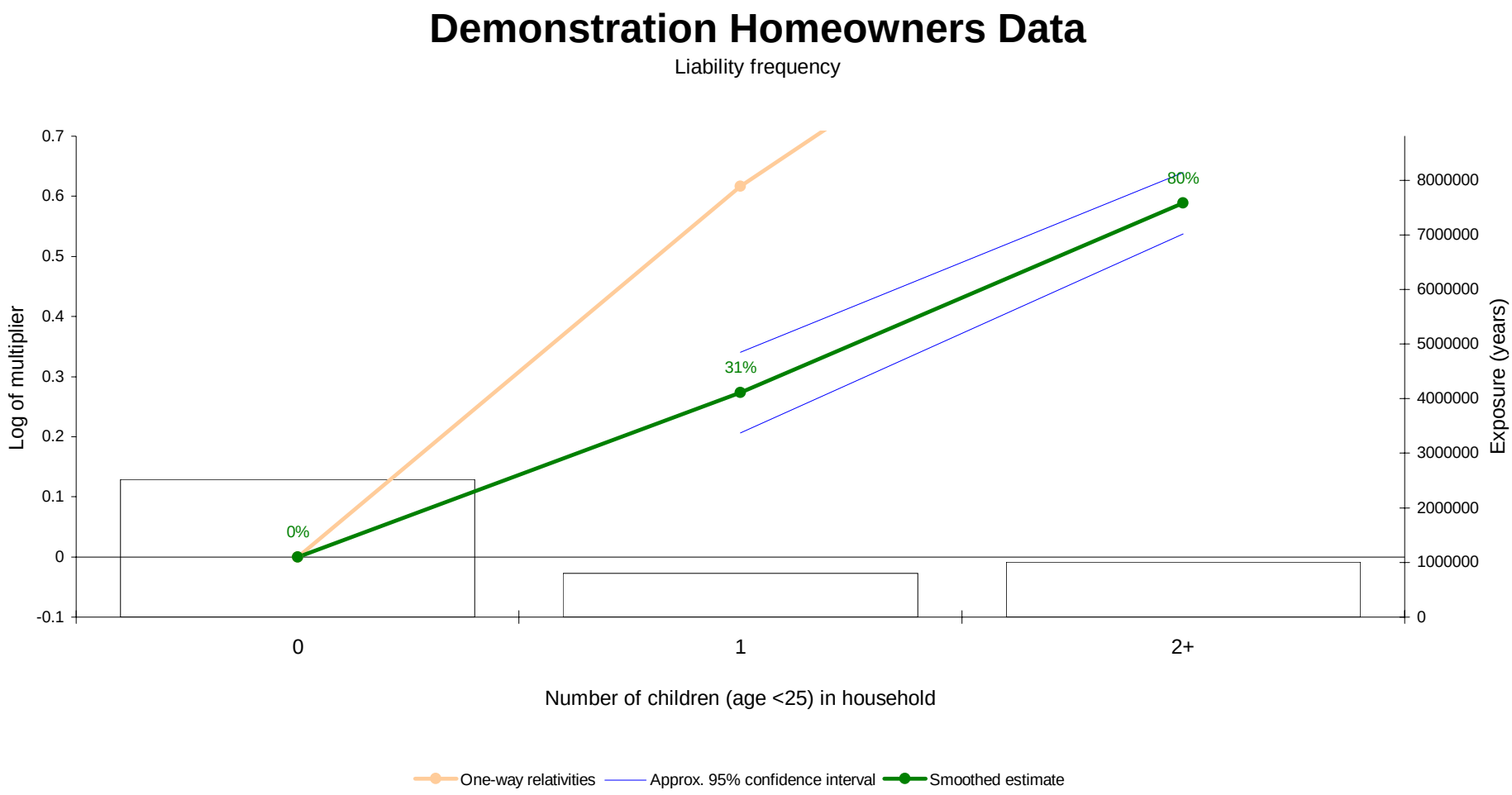
10% Theft system discount

- Territory A: credits more premium ($10) than losses expected from theft ($7)
- Territory B: credit ($10) may represent appropriate amount of savings in total theft losses ($27)

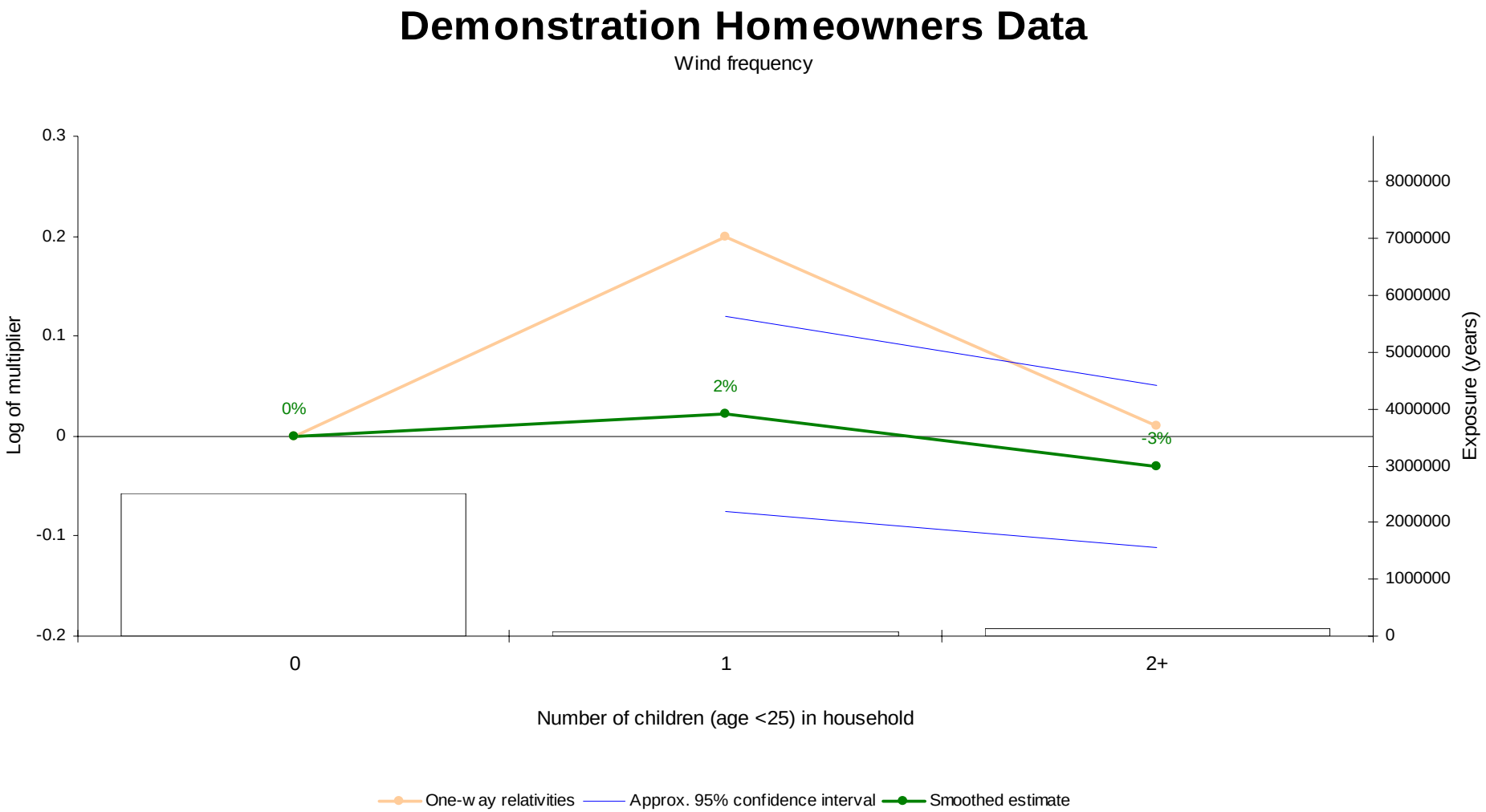
More detailed reasons...

- The changing landscape - claims
- The share of loss costs by peril varies considerably by geography
- Effect of rating factors varies considerably by peril
  - traditional rating factors
  - territory
  - inhabitant info
  - external info

Inhabitant information: Effect of children on Liability



Inhabitant information: Effect of children on Wind



More detailed reasons (cont'd)

- Model dwelling and contents separately
- Separate territories by peril
  - liability affected by demographics, but sinkhole affected by meteorological and geological phenomena
  - level of needed granularity may differ by peril
- Variable categorization by peril
  - AOI granularity may differ by peril
  - deductible options may differ by peril
- Large loss thresholds by peril

More detailed reasons (cont'd)

- Different ratemaking analysis methods to be applied to each peril
  - loss trends and development
  - data used (eg company experience for non-cat and simulated data for cat)
  - expenses allocation
  - cost of capital considerations

More detailed reasons (cont'd)

- Facilitates separation of liability for loss reserving and monitoring
- Facilitates endorsement pricing (for those tied to specific peril)

Agenda

- Case for unbundling the perils
- Practical Issues
- Traditional rating variables
- New rating variables

Practical Issues for by-peril analysis

- Know your product
- Frequency by peril
- Volume required
- Point of sale algorithm
- Others issues

Practical Issues for by-peril analysis

- Know your product
- Frequency by peril
- Volume required
- Point of sale algorithm
- Other issues

Know Your Product

- Rating Algorithm
- Changes?
- IT issues
- Perils

Contract Review - Perils

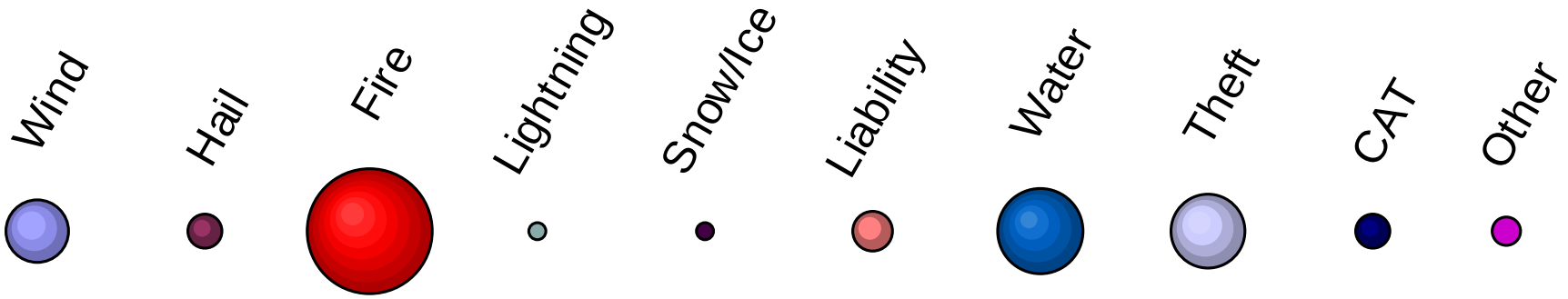
- Fire
- Lightning
- Water
- Wind
- Hail
- Weight of Snow and Ice
- Freezing
- Catastrophes
- Theft
- Vandalism and Malicious Mischief
- Liability
  - Bodily Injury
  - Property Damage
- Identity Theft
- All Other Perils

Practical Issues for by-peril analysis

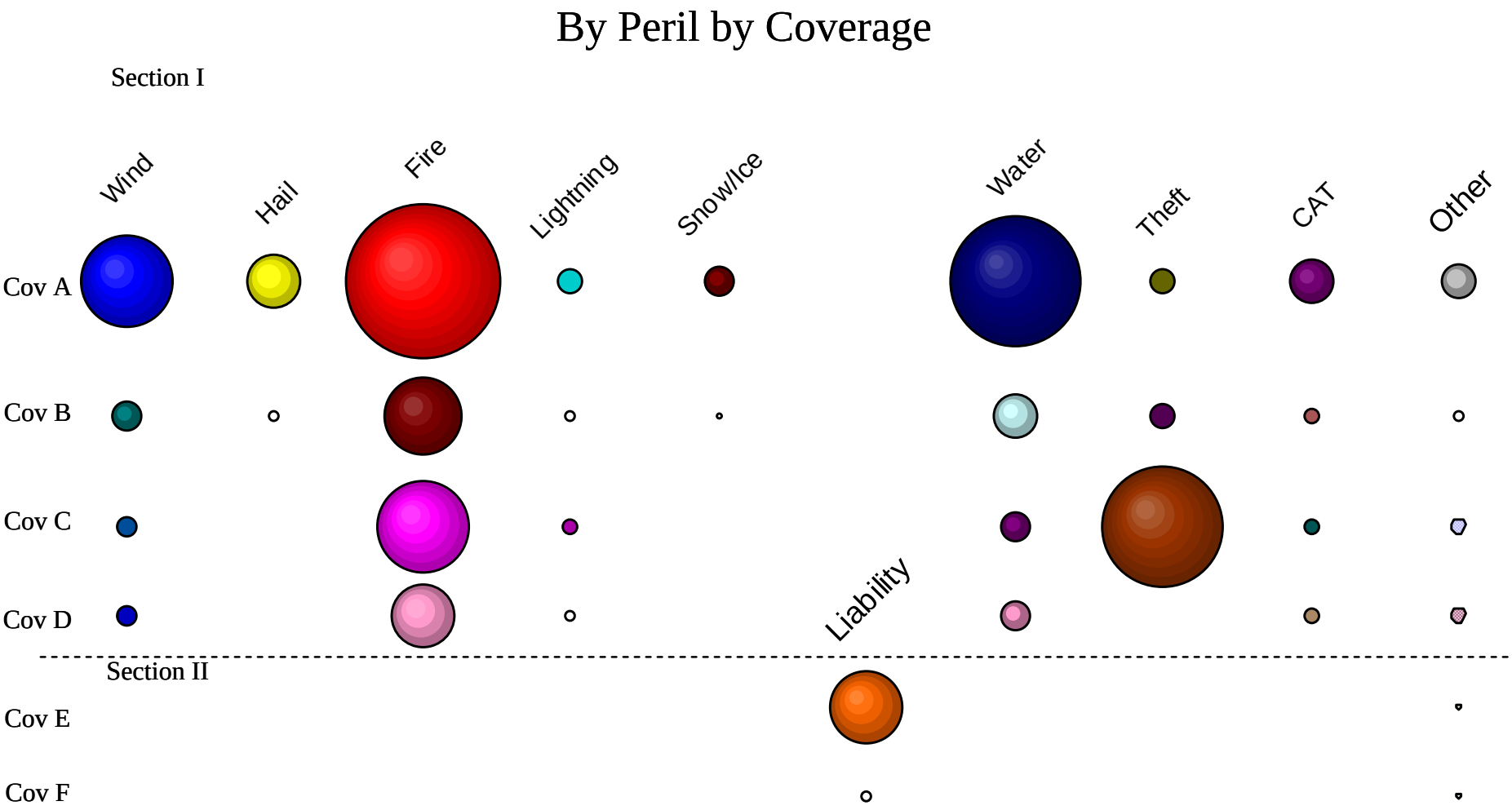
- Know your product
- Frequency by peril
- Volume required
- Point of sale algorithm
- Other issues

Frequency

By Peril

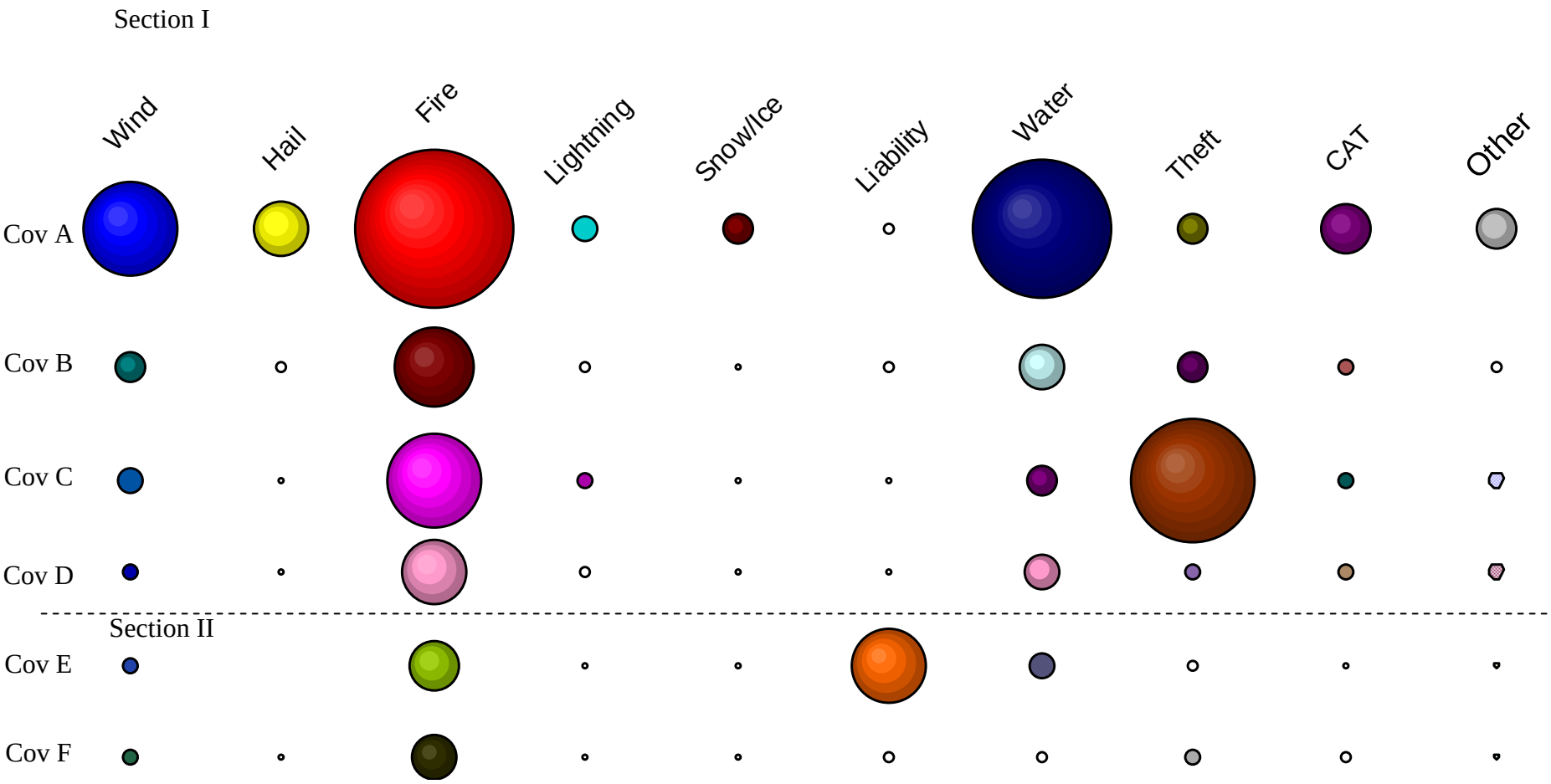


Frequency


By Peril by Coverage
Section I
Wind
Hail
Fire
Lightning
Snow/Ice
Water
Theft
CAT
Other
Cov A
Cov B
Cov C
Cov D
Liability
Section II
Cov E
Cov F

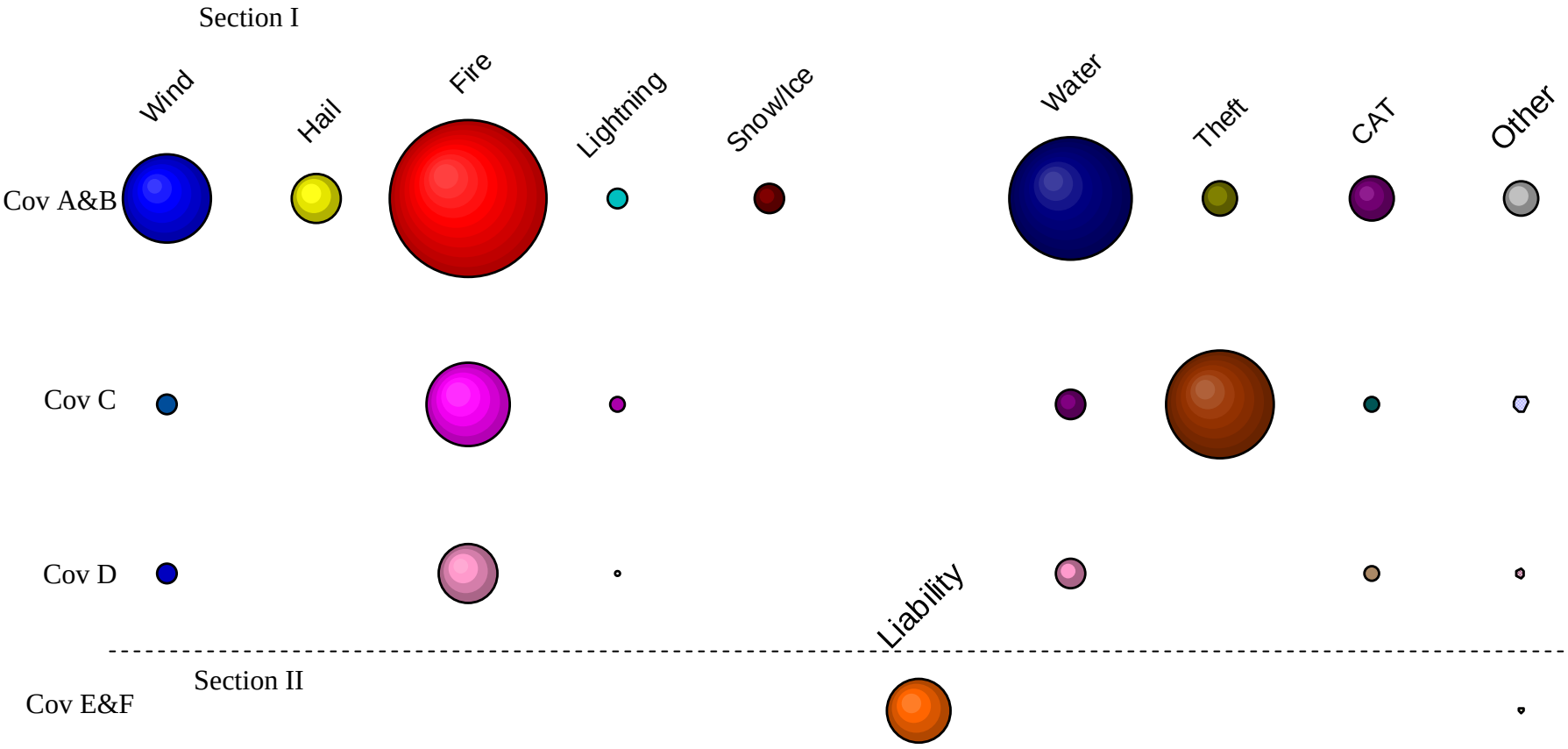
Frequency

Data Issues - By Peril by Coverage



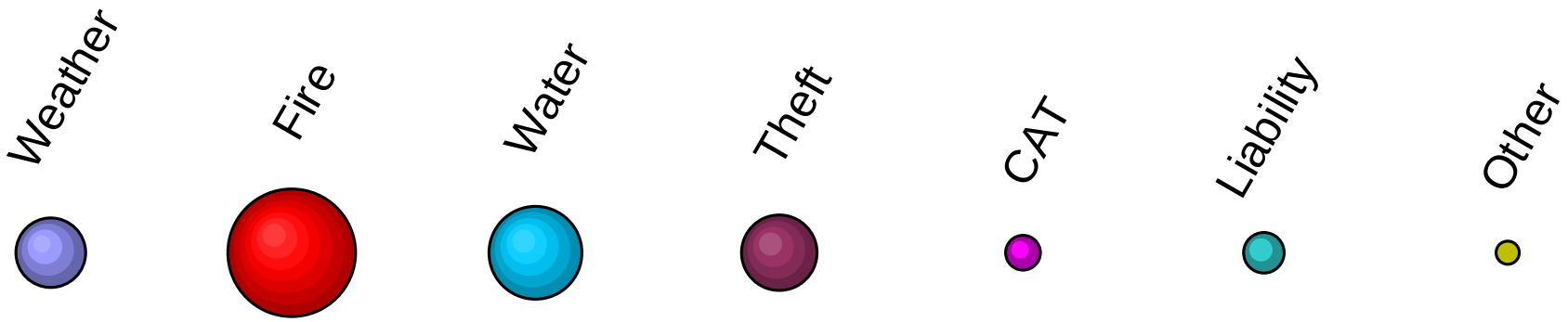
Frequency

Grouping Coverages - By Peril by Coverage



Frequency

A Selected Peril List



Practical Issues for by-peril analysis

- Know your product
- Frequency by peril
- Volume required
- Point of sale algorithm
- Other issues

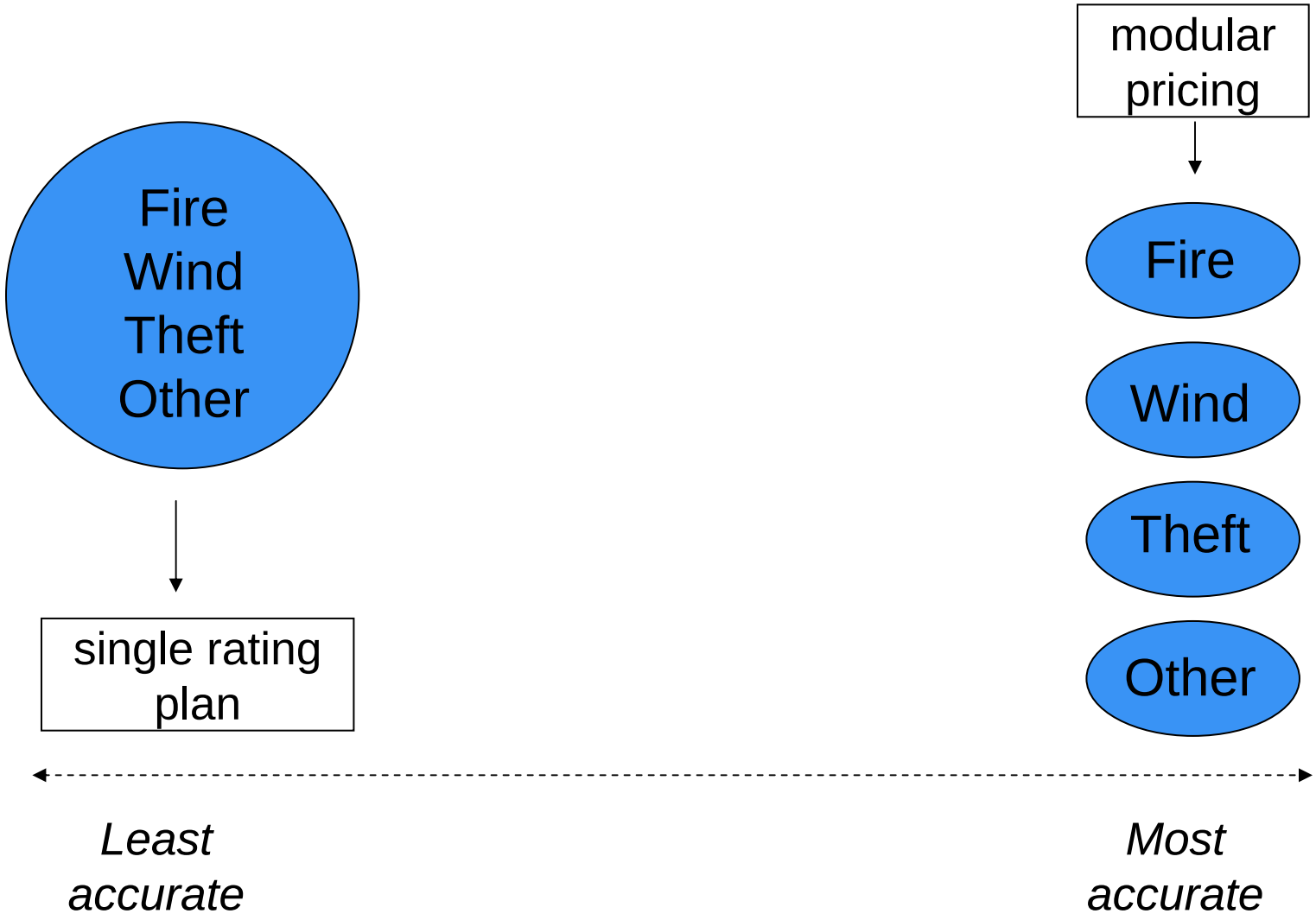
Volume

- Generally seek a few thousand claims per claim type to attain meaningful models
- Depends on the number of variables to be examined

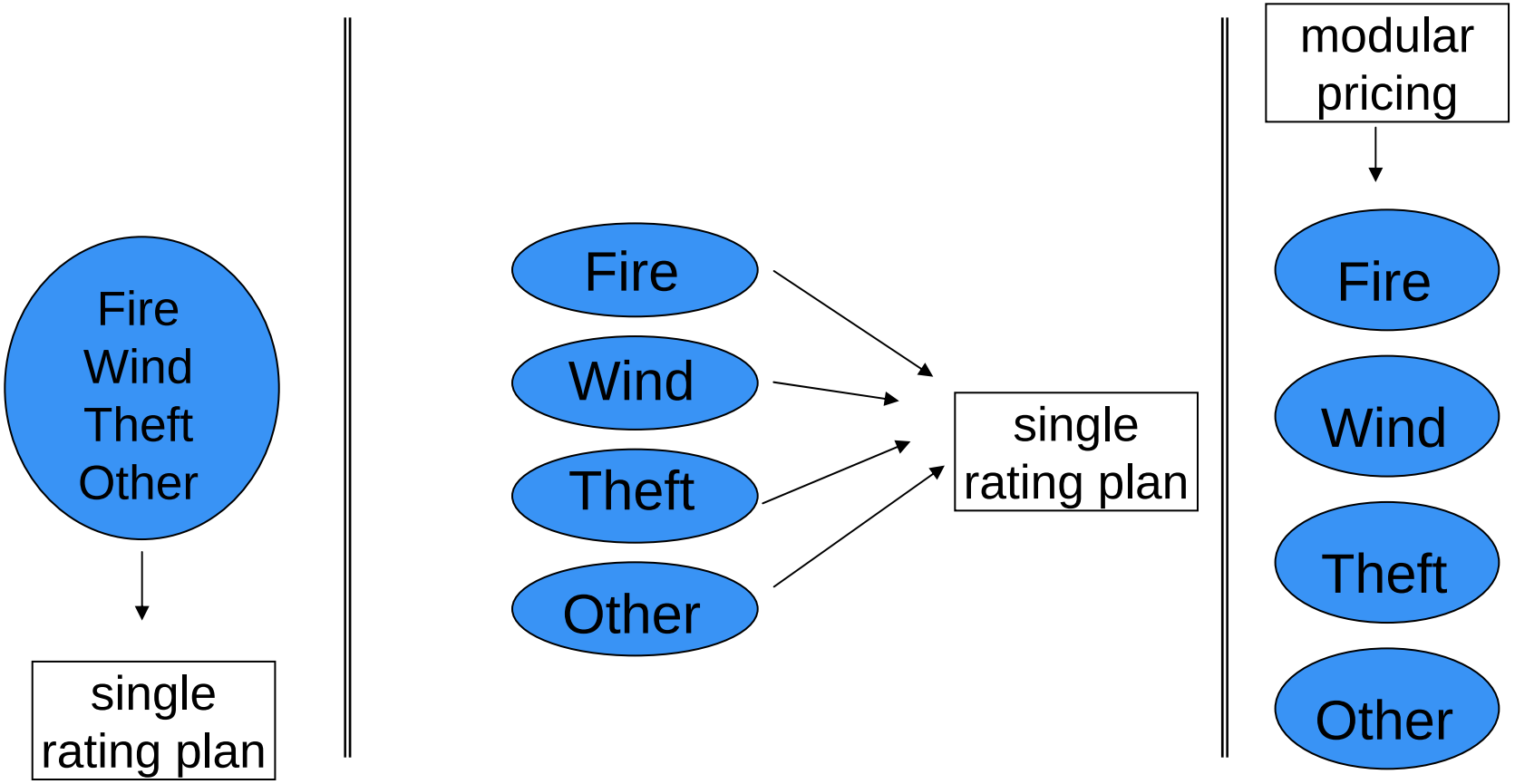
Practical Issues for by-peril analysis

- Know your product
- Frequency by peril
- Volume required
- Point of sale algorithm
- Other issues

Point of sale options



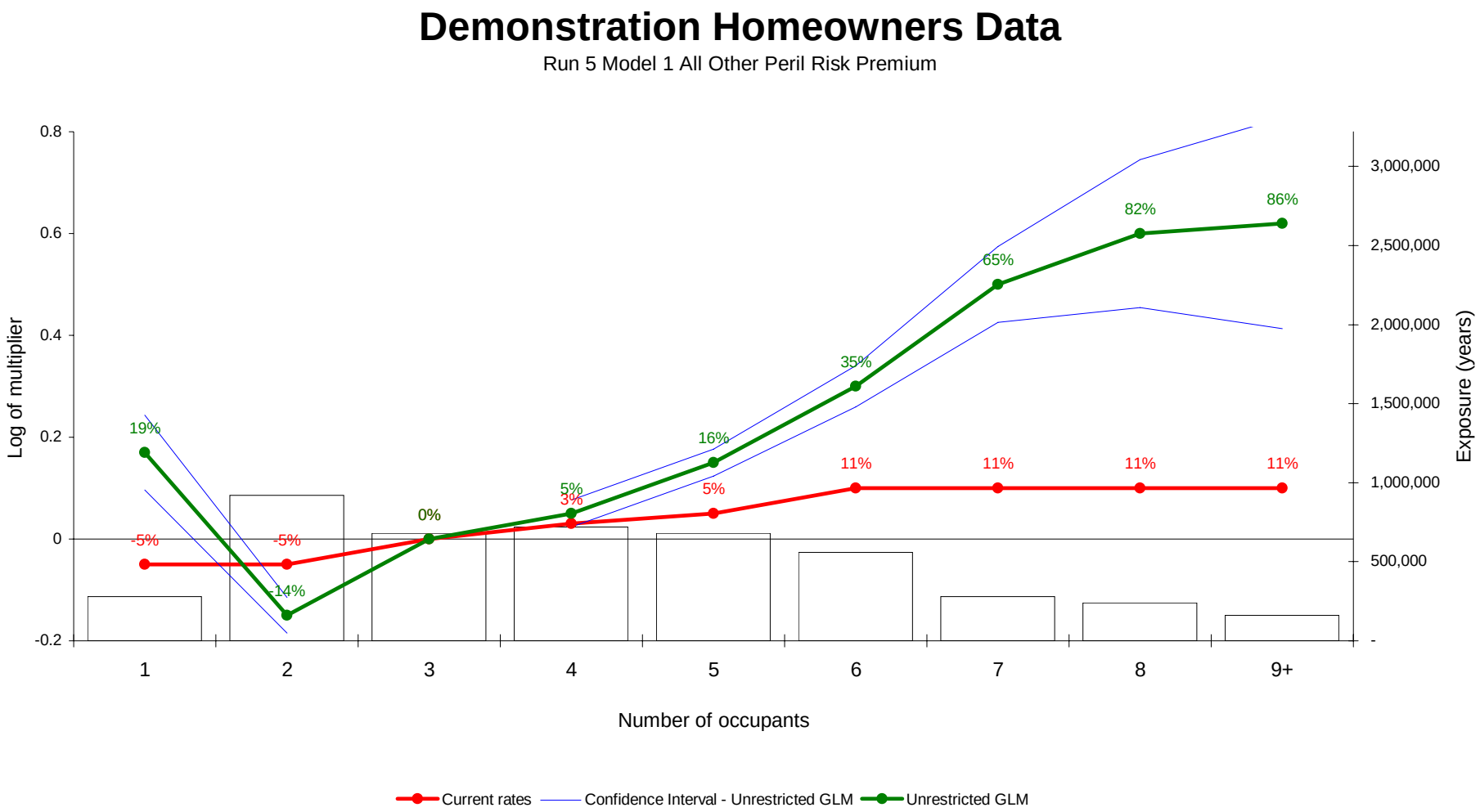
Point of sale options



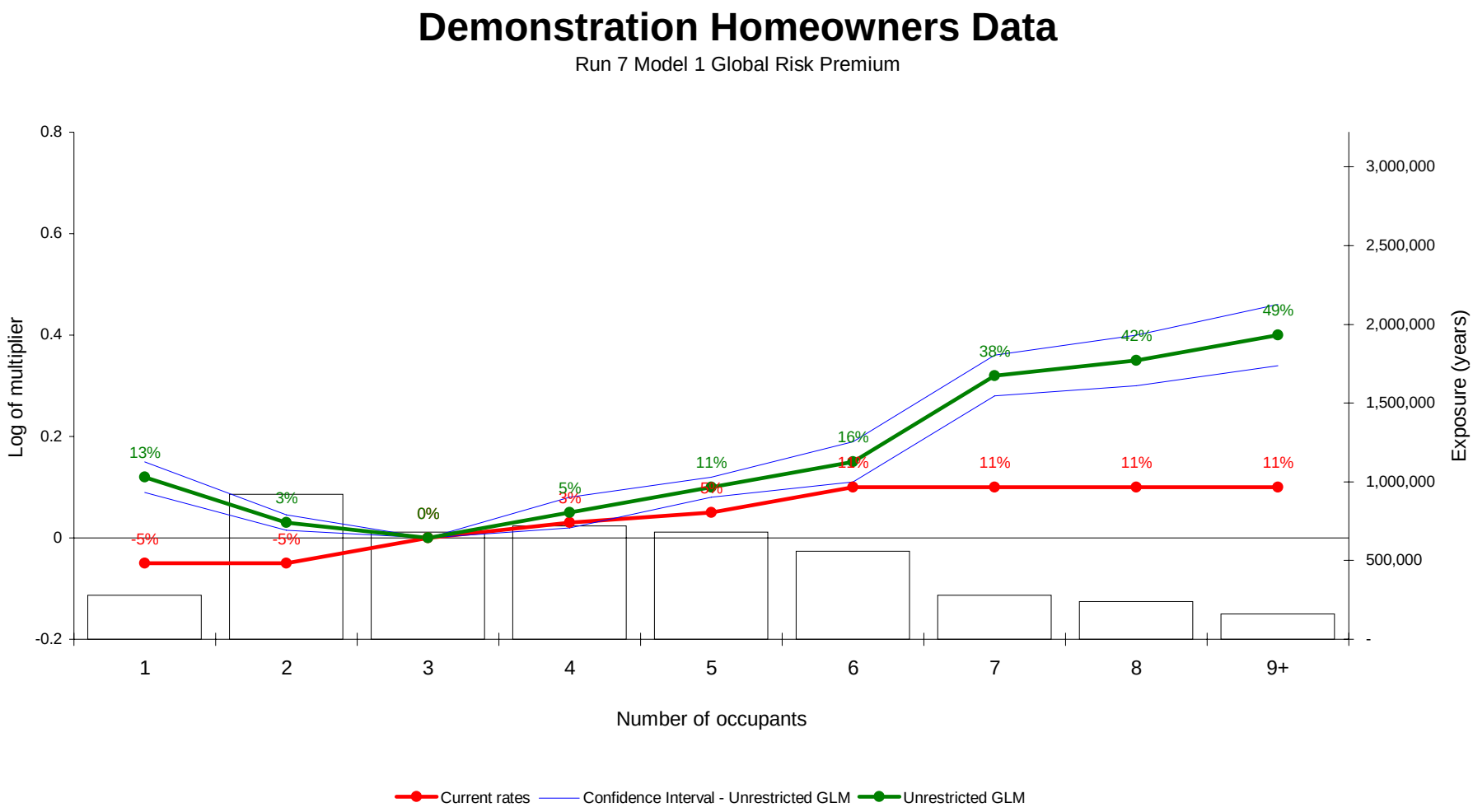
Investigating practical compromise

- Global risk premium across all perils
  - populate fitted values by peril for each individual record
  - calculate the sum of the by peril fitted loss costs
  - fit model to this modeled data
  - somewhat analogous to a single loss-weighted average of underlying by-peril models
- Investigate loss of accuracy in global risk premium model

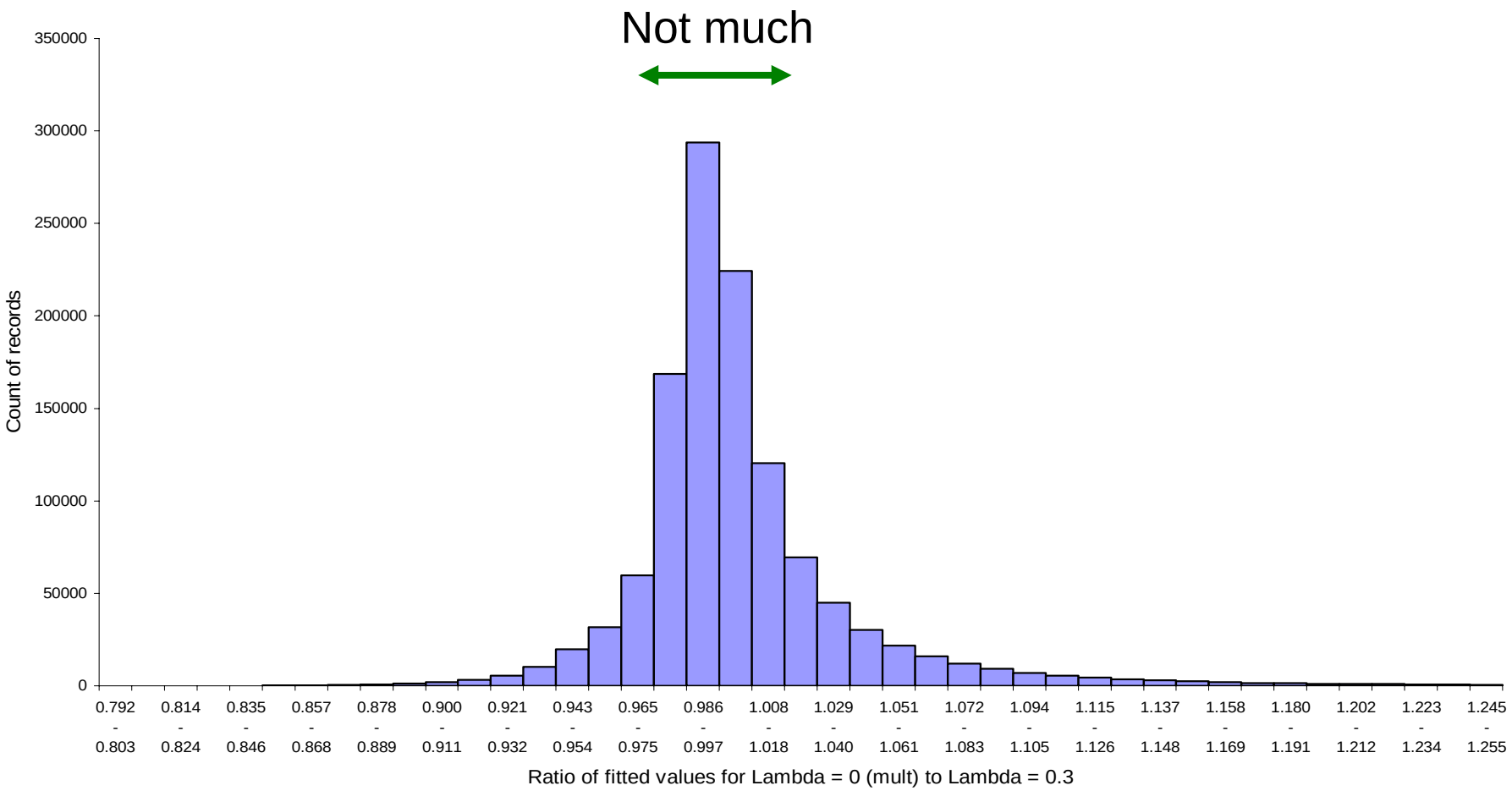
Sample output - risk premium by peril



Sample output - global risk premium



Investigating loss of accuracy



Ratio of fitted values from global risk premium to sum of fitted values from individual by-peril models

Other practical considerations for by-peril analysis

- Know your product
- Frequency by peril
- Volume required
- Point of sale algorithm
- Other issues
  - IT concerns (eg separate territory definitions by peril)
  - Lack of competitive benchmarks by peril
  - Complication by policy form
  - Endorsements priced as % of base premium
  - Incorporating catastrophe loads
  - Statistical plan requirements

Agenda

- Case for unbundling the perils
- Practical Issues
- Traditional rating variables – for example:
  - policy form
  - AOI
  - deductible
- New rating variables

Policy form

- Model separately by form allows
  - different variable categorization by form (eg amount of insurance)
  - different large loss thresholds
  - understanding loss cost effects by form
- Model home and renters/condo separately and include form as an independent variable
- Model all combined with form as an independent variable
- Consider interactions by form

Amount of insurance (AOI)

- Could model AOI as a categorical factor with many levels (consider categories that straddle common AOIs eg $98.5-101.5K)
  - this allows the true effect to be seen for both frequency and amounts models
  - smooth the relativities carefully so that the risk premium result for AOI shows a sensible progression
  - either charge a premium based on interpolated banded AOI, or perform simple interpolation between exposure weighted mid points of the bands to get a continuous scale
- Alternatively fit a regression spline to AOI and incorporate in rating algorithm or use to populate a detailed table

Deductible

- Model incurred losses net of deductible
- Include in underlying frequency and severity models
- If results counter-intuitive, may need to remove factor and offset model by log of relativities from external study (eg current relativities or results from LER)
- Careful of changing selection behavior in future
- Changing deductibles by AOI

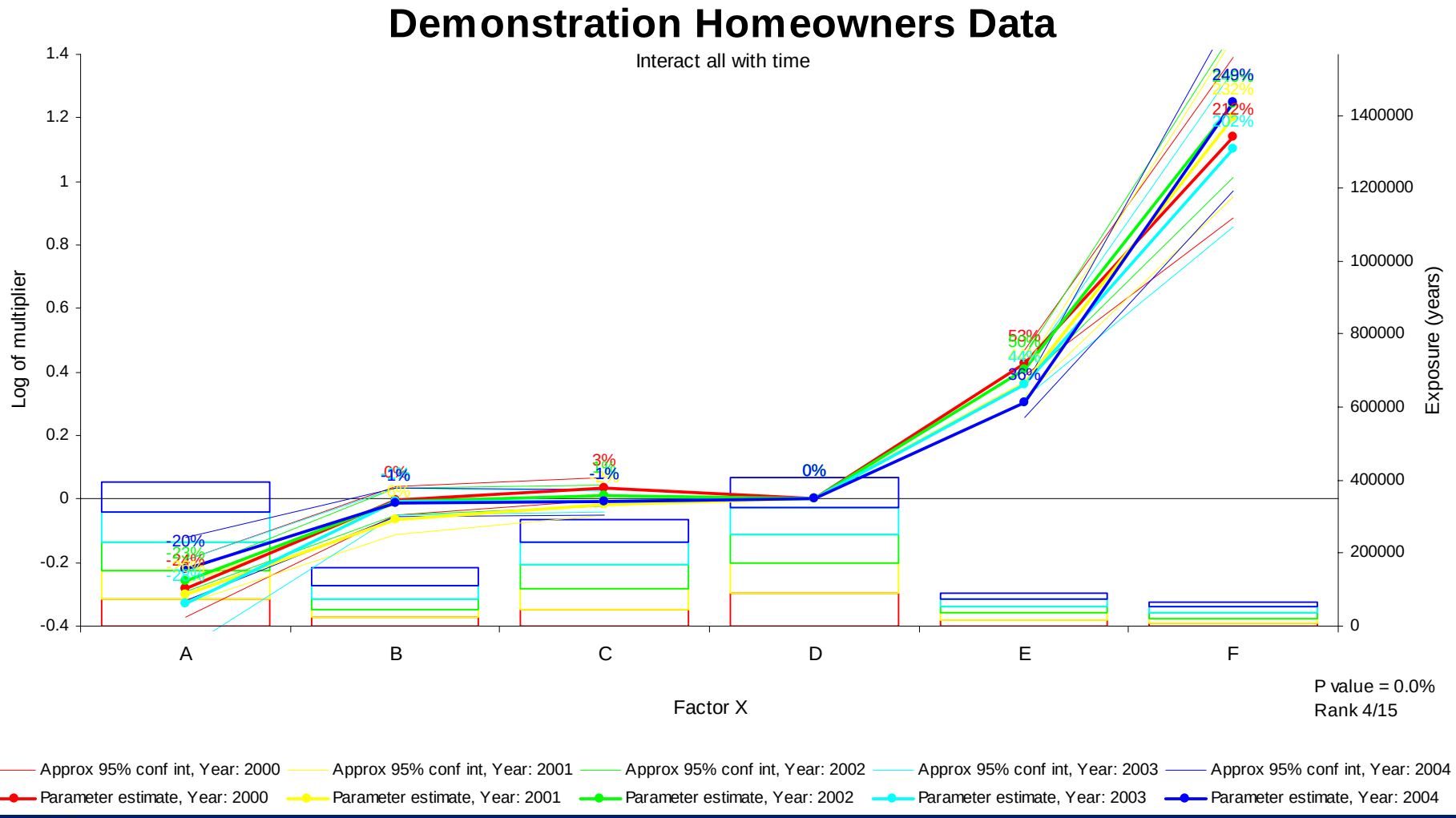
Agenda

- Case for unbundling the perils
- Practical Issues
- Traditional rating variables
- New rating variables
  - concern over missing levels
  - investigate consistency over time
  - internal information (eg inhabitant info)
  - external information (eg geodemographics)

Factors with missing levels

- Common problem as information may not be collected on every exposure
- Do not choose "missing" as base level
- Investigate exposure distribution of missing level with other factors - eg does missing occur only on older years or older houses?
  - consider altering data to alleviate problem (eg use more recent years)
  - consider changing order of factors in the model to force alias in another variable
- Model with and without factor to understand effect

Consistency over time



Internal variables

- Inhabitant information
  - # occupants
  - age, gender, marital status
  - unusual exposure (eg dogs)
- Relationship with company
  - optional endorsements
  - products held
  - # years with company
  - affinity membership

Internal variables

- Detailed information on property
  - square feet
  - number of rooms
  - foundation shape
  - roof attributes (age, shape, covering)
  - interior construction materials
  - pool/spa

Property characteristics

- Consider correlation with AOI – ie could something inherent to AOI algorithm actually predict risk better than AOI?
- Could you live without AOI?

Score based on property characteristics

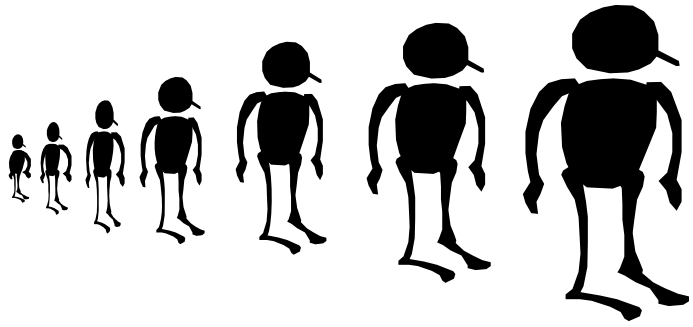
- Fit GLM with traditional rating factors and several property characteristics (eg $R_1xR_2xR_3xP_1xP_2xP_3$)
- Transform model results for property variables ($P_1xP_2xP_3$) into points-based score variable = R_4
- Categorize score variable appropriately
  - consider # of categories & proportion of business in each
- Include new score variable in claims model (ie $R_1xR_2xR_3xR_4$) and consider interacting with other variables

External information

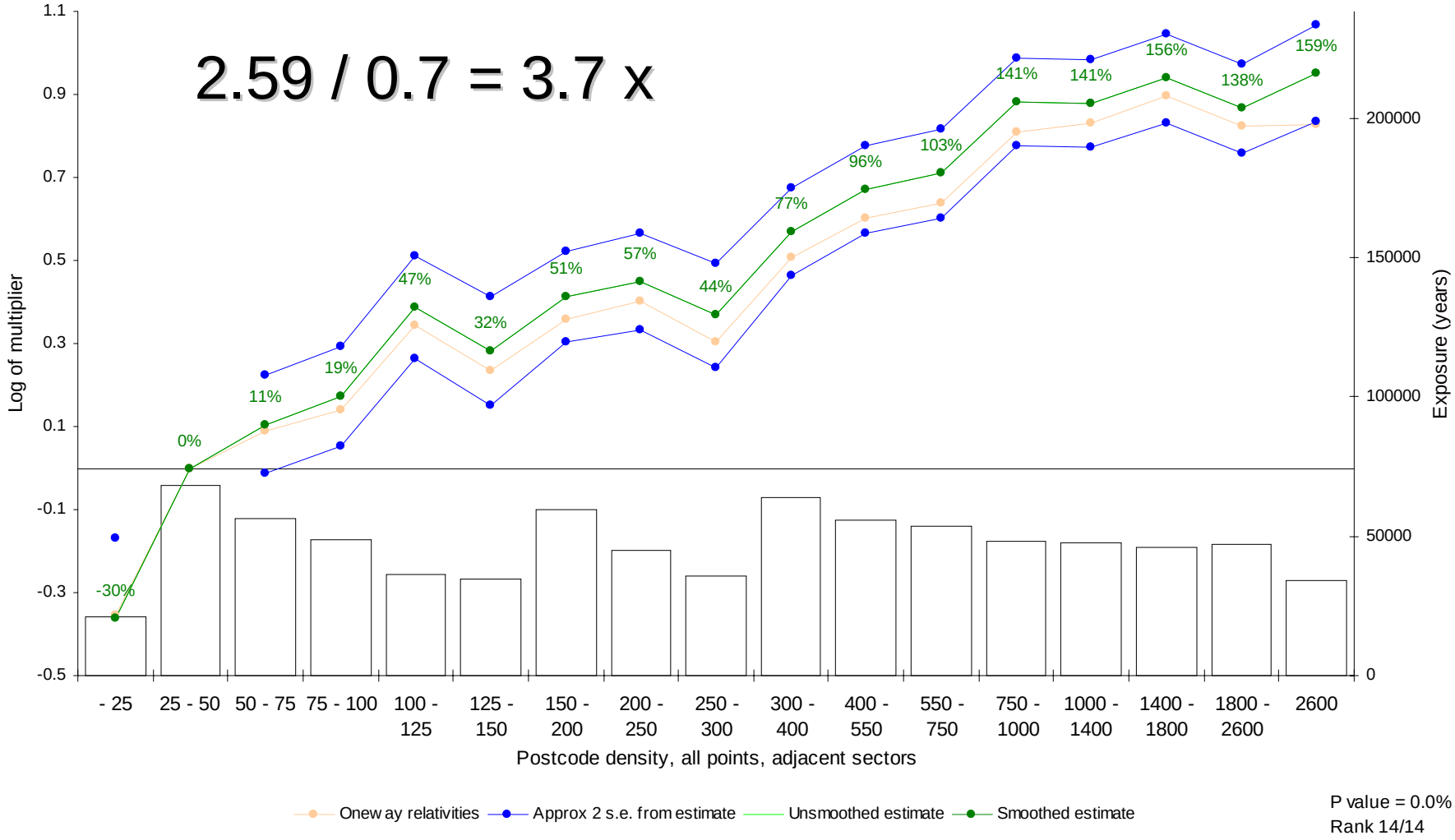
- Geodemographics (avg characteristics in an area)
  - population density
  - length of home ownership
  - average age of residents
  - financial information
- Weather data per area (relating to vulnerability of buildings)
  - max wind speed
  - avg temperature
  - avg high/low temperature
  - avg rainfall
  - soil type

Geodemographic data

- Often designed for marketing retail products
- Attaches to zip code therefore easy to use at point of sale
- Marketing segment types often not predictive
- Underlying data often more interesting
- Simple measure of urban density often predictive



Example effect of urban density on homeowners theft frequency



P value = 0.0%
Rank 14/14

Effect of density varies

- Effect of increasing density on risk:

	Frequency	Severity
Theft	↑	↑
Fire	↓	↑
"Other"	↑	↑

Geodemographics can be rather related!

	R1	R2	R3	R4	G1	G2	G3	G4	G5	G6
R1										
R2	11%									
R3	32%	3%								
R4	17%	7%	58%							
G1	8%	2%	57%	16%						
G2	8%	2%	53%	15%	49%					
G3	7%	3%	44%	14%	33%	33%				
G4	5%	4%	21%	8%	30%	30%	30%			
G5	3%	2%	31%	6%	36%	35%	34%	31%		
G6	8%	2%	65%	16%	37%	35%	31%	29%	34%	
G7	8%	2%	65%	16%	36%	34%	30%	30%	34%	71%

Cramer's V for a selection of standard rating factors (R1, .., R4) and geodemographic factors (G1, …, G4)

Coping with related factors

- Can be hard to interpret output from a GLM that includes a very large number of related characteristics
- Options
  - test related factors (within "families") one at a time to find most predictive member (eg # of late pays in 60 days may be most predictive of "late pay" family)
  - apply principal components analysis first

Example of geodemographic factors

Real GLM output cannot be disclosed in handouts

Graph in presentation showed strong multivariate effect of geodemographic factor related to average life-stage of an area

Example of geodemographic factors

Real GLM output cannot be disclosed in handouts

Graph in presentation showed strong multivariate effect of another geodemographic factor

Example of geodemographic factors

Real GLM output cannot be disclosed in handouts

Graph in presentation showed strong multivariate effect of average type of building in area

Same graph, rescaled to show one-way effect

Real GLM output cannot be disclosed in handouts

Graph in presentation showed strong multivariate effect of average type of building in area

External information

- Geodemographics (avg characteristics in an area)
  - population density
  - length of home ownership
  - average age of residents
- Weather data per area (relating to vulnerability of buildings)
  - max wind speed
  - avg temperature
  - avg high/low temperature
  - avg rainfall
  - soil type

Examples of geophysical data

Real GLM output cannot be disclosed in handouts

Graph in presentation showed strong multivariate effect of weather-related geophysical data item

Examples of geophysical data

Real GLM output cannot be disclosed in handouts

Graph in presentation showed strong multivariate effect of weather-related geophysical data item

Examples of geophysical data

Real GLM output cannot be disclosed in handouts

Graph in presentation showed strong multivariate effect of weather-related geophysical data item

External data

- Can add predictive power and thus give competitive pricing edge
- Can improve speed and accuracy of quotation process
- Can help assess risk when own data insufficient
- New philosophy for agents, regulators, etc.
- May complicate ability to compare to existing rates on factor by factor basis (eg comparing "old" territory to "new" territory plus population density)

Must balance accuracy with model parsimony and point of sale concerns.

Example Homeowners Rating Factors UK

- Post code (so geodemographic and geophysical factors can be derived)
- Amount of insurance
- Number of rooms / bedrooms
- Wall type
- Roof type
- State of repair
- Extensions
- Ownership status (rent/own)
- Occupancy in day
- Neighborhood watch scheme
- Approved locks, alarms, smoke detectors
- Deductibles
- Riders purchased, value> £x
- How long held insurance / when last claimed
- Policyholder details
  - Age
  - Sex
  - Marital status
  - Number of children
  - Occupation
  - Residency
  - Criminal convictions
  - Claims in past 2/5 years
- Smokers present in house
- Non family members sharing house
- Length of time living at property
- Use (principal/ second / business / let)
- Cover selected (buildings/contents/both)
- Source business (eg internet)

Organizational advice

- Review/discuss variables in advance with other areas of the company (underwriting, legal, marketing, IT)
- Review integrity of data (especially if can't explain effects)
- Aim for visual aids (including maps)
- Address what matters most to the organization (removal of cross-subsidy, change in competitive position, policyholder dislocation, etc)
- Examine effect of business decisions (i.e. penalty to theoretical)

watsonwyatt.com

CAS Predictive Modeling Seminar

Homeowners Modeling

Gaétan Veilleux, FCAS
October 11, 2007

Watson Wyatt Worldwide