watsonwyatt.com

CAS Predictive Modeling Seminar

C13: Homeowners Modeling

James Tanser, FIA
October 6 & 7, 2008

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

US History

- 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

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

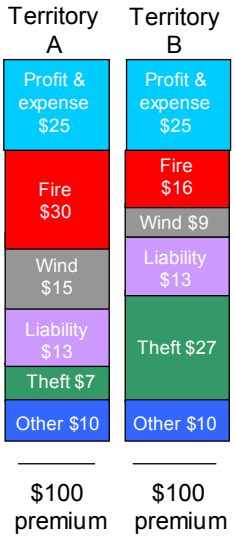
Why unbundle?

- 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

Not all segments are the same

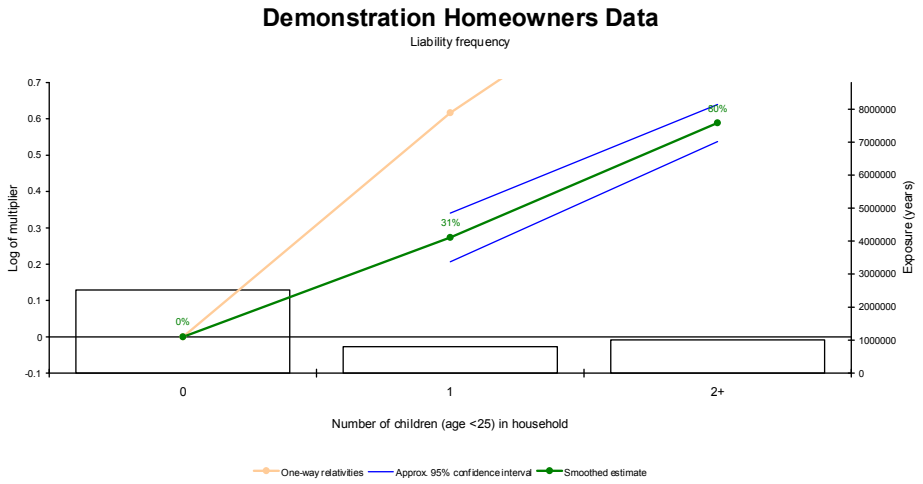
- Consider different locations:
  - Rural house: Low theft, high fire
  - Urban apartment: High theft, low fire
  - Suburban sprawl: Medium theft, medium fire
- Consider different occupants:
  - Young professional: 1 or 2 occupants, 1 PC
  - Student: 4 occupants, 4 PCs

Percent of losses by peril varies across territories

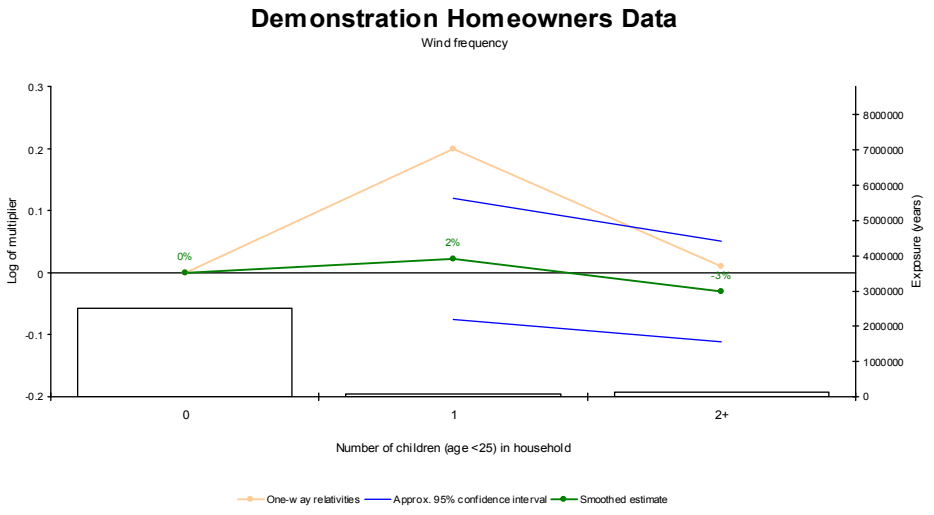


- Flat discounts for anti-theft measures (eg 10% reduction in total premium) may be larger than actual theft cost
- Cross subsidy between segments / perils is not necessarily clear to the company
- If you don't understand your rates, how do you know they are correct?

Inhabitant information: Effect of children on Liability



Inhabitant information: Effect of children on Wind



Best practice pricing

- Model dwelling and contents separately
  - A given building costs the same to rebuild regardless of the contents when it burnt down
- Separate territories by peril
  - Theft and weather happen in different places!
  - Level of needed granularity may differ by peril
  - Consider external flood mapping data
- Variable categorization by peril
  - AOI granularity may differ by peril (but see later)
  - deductible options may differ by peril
- Large loss thresholds by peril

Best practice pricing

- 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

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 accurate pricing of options and changes
- Helps with reserving

Agenda

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

Practical Issues

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

Practical Issues

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

Know your product

- Rating Algorithm
  - How much flexibility do you have to change it?
  - Can you add new variables or interactions?
- IT issues
  - How good is the allocation of claims to peril?
  - How much data is missing?
  - Can IT systems cope with by peril pricing?
- Perils
  - How much data is in each peril?

Know your product
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

Group together where necessary

Practical Issues for by-peril analysis

- Know your product
- 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

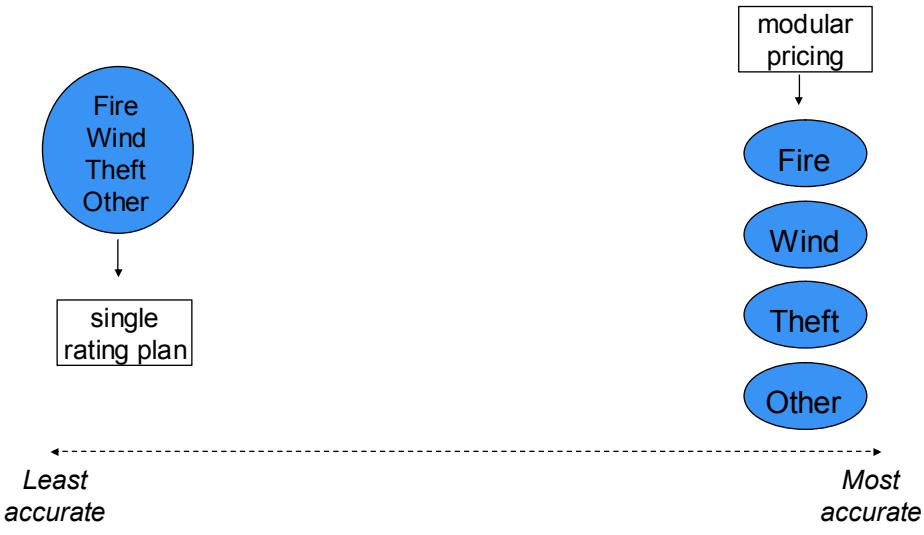
Volume required

- Minimum (by claim type):
  - 1,000 claims
  - 4 or 5 variables with few levels
- Target (by claim type):
  - 10,000+ claims
  - 20 to 30 variables with many levels
- Minimum (by categorical level)
  - 100 claims per level
- Target (by categorical level)
  - 500+ claims per level

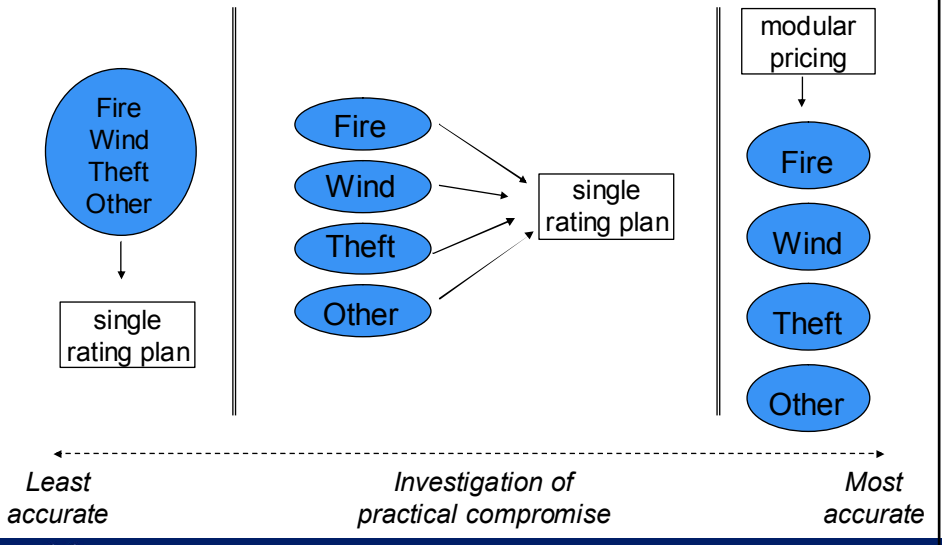
Practical Issues for by-peril analysis

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

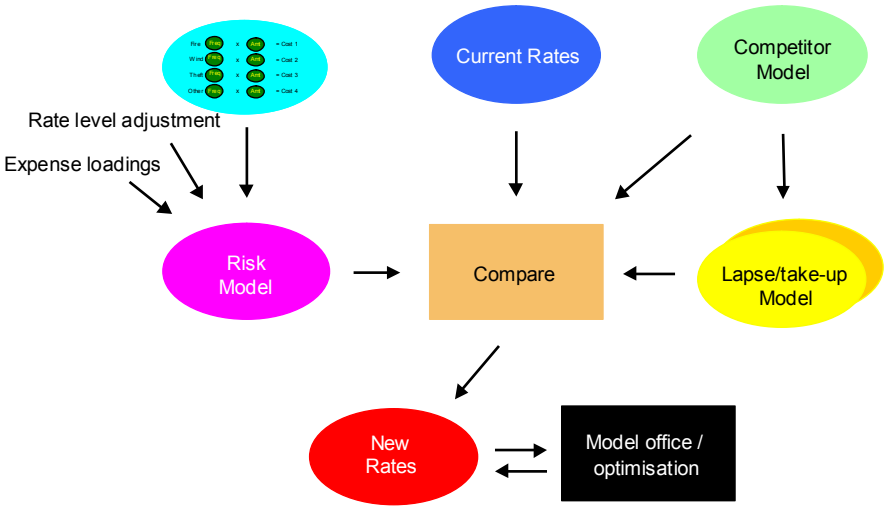
Point of sale options



Point of sale options



Back to school…



Modelling the cost of claims

Fire	Freq	x	Amt	= Cost 1
Wind	Freq	x	Amt	= Cost 2
Theft	Freq	x	Amt	= Cost 3
Other	Freq	x	Amt	= Cost 4

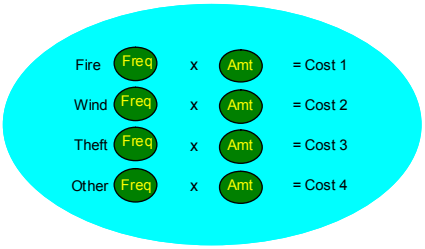
Combining claim elements - I

- Multiply factors for frequencies and amounts
- Calculate risk premium as sum of claim elements

Risk premium for one claim type

- The multipliers are simply the product of the multipliers in the underlying two models
- But because we use a log link, we can simply add the parameter estimates
- The models' standard errors can also be calculated simply
- Risk premium standard errors here are the square root of the sum of the squares of the standard errors from the model for any given factor level

Combining claim elements - II



- Consider current exposure
- Ignore claims
- Calculate expected frequency and amount for each claim type
- Combine to give expected total cost of claims
- Fit model to this expected value

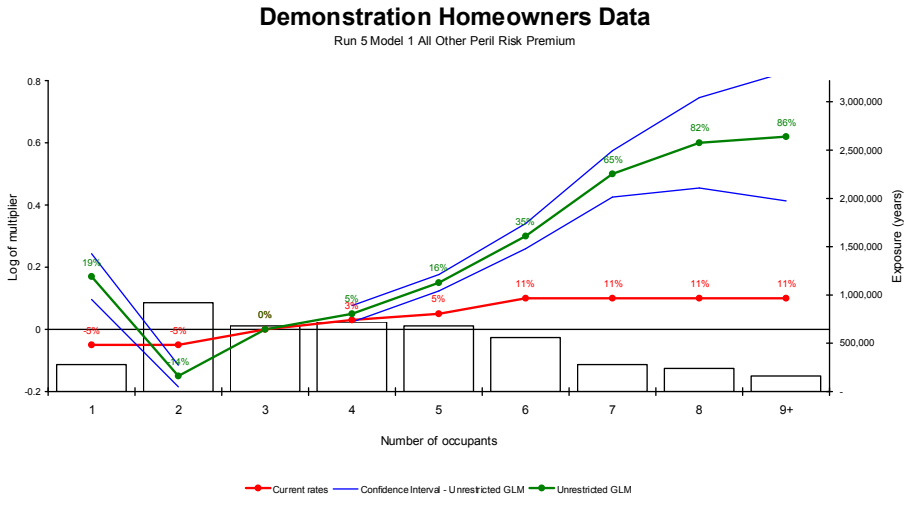
Risk premium relativities

- Fit a model to the global risk premium
- This calculates the best multiplicative approximation to the theoretical result
- Standard errors from this model are small as there is no random error, just model error

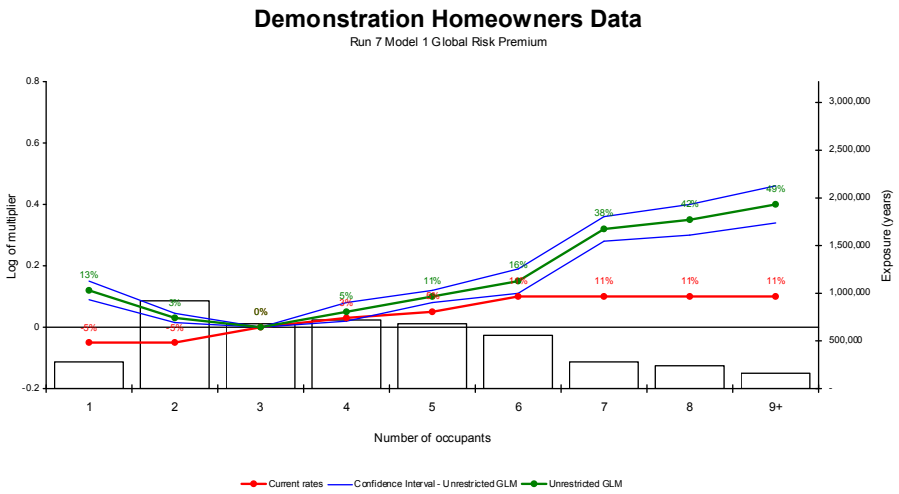
Investigating practical compromise

- 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

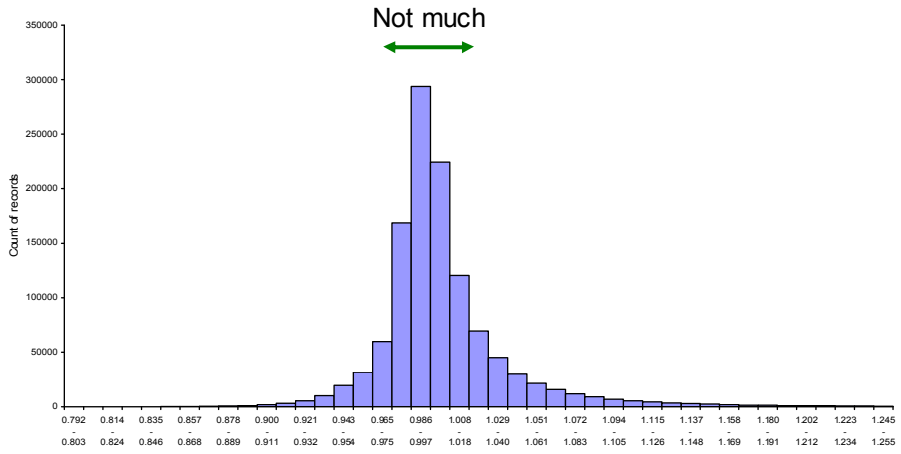
Sample output - risk premium by peril



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

Practical Issues for by-peril analysis

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

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
- New rating variables

Traditional rating variables

- Policy form
- AOI
- Deductible

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)

- 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

Amount of insurance (AOI)

- Treat AOI as continuous variable
  - Allows use of full information
  - But may not reflect way levels are selected
- Take care with multiplicative models
  - Straight lines in log space are exponential in $ space
  - Use log(AOI) to allow for this
  - AxBxCx…x(AOI^p)
- Fit a regression spline to log(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

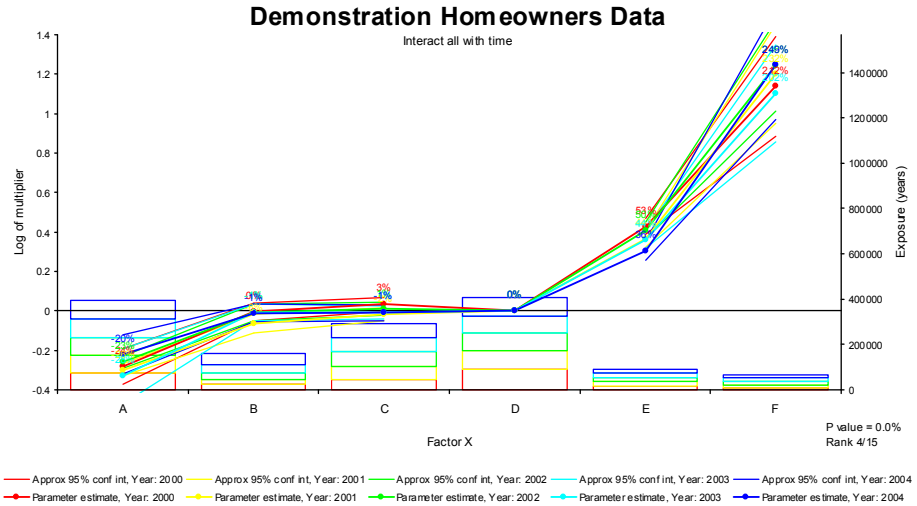
New rating variables

- Missing levels
- Investigate consistency over time
- Internal information (eg inhabitant info)
- External information (eg geodemographics)

Missing levels

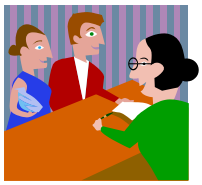
- Information may not be collected on every exposure
- Ensure "missing" is not base level
- Investigate exposure distribution of missing level with other factors
  - consider altering data to alleviate problem (eg use more recent years)
  - consider changing order of factors in the model to force aliasing 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?
  - Risk rarely proportional to AOI

Score based on property characteristics

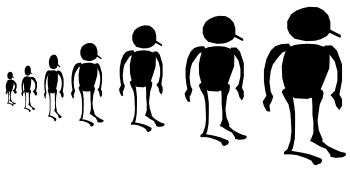
- Fit GLM with traditional rating factors and several property characteristics (eg $R_1xR_2xR_3xP_1xP_2xP_3$)
- Transform model results for property variables ($P_1xP_2xP_3$) into points-based score variable = R_4
- Categorize score variable appropriately
  - consider # of categories & proportion of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

External information

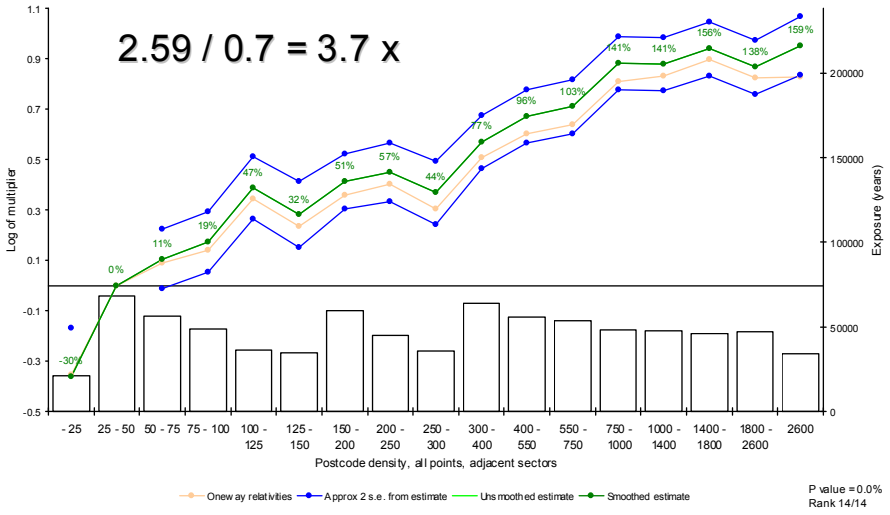
- Geodemographics (avg characteristics in an area)
  - population density
  - length of home ownership
  - average age of residents
  - financial information
- Weather data per area (relating to vulnerability of buildings)
  - max wind speed
  - avg temperature
  - avg high/low temperature
  - avg rainfall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



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

Aside: PCA and clustering suggestion

- Use PCA to identify main “directions” of data set
  - Convert categorical into 0/1 for this purpose
  - Insurance risk not part of this calculation
- Use clustering on PCA output to derive categorical factor
  - This uses insurance risk

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

- Work with other areas of the company
  - underwriting, legal, marketing, IT, …
- Review integrity of data
- 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

Questions

- James Tanser
  - +44 1737 274249
  - james.tanser@watsonwyatt.com

watsonwyatt.com

CAS Predictive Modeling Seminar

Homeowners Modeling

James Tanser, FIA
October 6 & 7, 2008

Watson Wyatt Worldwide