Institute of Actuaries of Australia

Understanding Competitor Premiums

Prepared by **Colin Priest BEc FIAA**

Presented to the Institute of Actuaries of Australia
16th General Insurance Seminar 9-12 November 2008
Coolum, Australia

The Institute of Actuaries of Australia
Level 7 Challis House 4 Martin Place
Sydney NSW Australia 2000
Telephone: +61 2 9233 3466 Facsimile: +61 2 9233 3446
Email: actuaries@actuaries.asn.au Website: www.actuaries.asn.au

Abstract

The insurance cycle is at a turning point. Insurers who do not change their rates with the market may end up either losing business or with lower profit margins than their peers. There is also the risk of adverse selection if an insurer is caught out pricing materially different to their competitors. There is a clear need for insurers to quickly discover when their competitors have changed rates, and for a detailed understanding of how their rates sit against their competitors. This paper explains the practical difficulties in measuring competitor rates – the complexity of rating systems, and the dynamic features of the assets being insured. The paper then examines the effectiveness of some techniques that can be used to overcome these difficulties.

Keywords: premium rating; quantitative market analysis.

The Motivation for This Paper

My day to day work as a general insurance actuary requires me to set prices, comment on underwriting results, forecast results and reserve for claims. All of these things require me to understand the market environment within which the insurer is operating. When I read the newspaper, industry news, talk to sales staff, agents and brokers I get told anecdotes about what is happening with premium rates, but quite often one person's anecdotes contradict another's.

Some of the questions that I have been asking myself over the past 12 months are:

- Will insurer rates react to the greater incidence of weather claims we have been experiencing since La Niña began?
- When some of the listed insurers tell share investors that they are increasing rates, are they telling the complete truth, or just trying to talk up their share prices?
- Where are we in the insurance cycle?
- When my mix of business changes, is it because of anti-selection?
- Am I priced differently to the market, and if so, who's right?
- Can I increase premium rates without materially impacting sales?
- Where am I able to price most competitively, and can I define that segment so that the marketing team can target market?

These questions are often difficult to answer. General insurance pricing is complex, and prices are usually not published. A simple list of possible home building premium rating factors would include the sum insured and post code, and assuming you insured every post code and every sum insured from $100,000 to $1,000,000 you would have 236,800,000,000 possible combinations of sum insured and post code. If you obtained one competitor quote per minute, then it would take you 450,376 years to cover all of those combinations. I'm too impatient to wait that long, and that's even before you consider other rating factors such as construction type…

Types and Purposes of Competitor Premium Analysis

There are three different types of competitor premium analysis that can be done:

1. Profile Comparisons: Ranking insurers
2. Premium Indices: How premium rates move through time
3. Premium Structures: Comparing how different insurers charge for each different risk

Profiling

The simplest way to compare insurer premiums is to create one or more hypothetical consumers, determine a risk profile for each consumer, and compare the prices that different insurers quote for each profile. A risk profile will include answers to the underwriting and rating questions that each insurer asks e.g. the location of the item to be insured. One can see profiling in use with:

- consumer organisations (e.g. Choice and Cannex)
- market analysts (e.g. Macquarie Equities)

Profiling is also used internally by insurers. The aim of profiling is to rank insurers by their quoted premiums (and maybe their policy wording features) to understand which insurers have the most competitive product offerings for different market segments.

Profiling is the simplest of the three approaches that I outline in this paper. But while it is simple, profiling has some traps that are easy to fall into:

1. insurer rankings can be chaotic – changing the value for a single rating factor can arbitrarily change the rankings
2. some rating factor values belong together and some don't – the items insured by customers from different socio-economic and geographic areas can be quite different
3. some insurers market into niches and / or avoid certain types of risks – use of profiles that include risks that those insurers are not interested in can lead to the misleading conclusion that those insurers are uncompetitive, or can invalidate the profiling approach if the insurers do not insure that profile at all
4. profiles need to be representative of the market mix of risks – however many market surveys give equal weight to each of their profiles. Giving equal weight to profiles from different sized market segments can lead to too much importance being given to smaller market segments.
5. different insurers use different rating factors – this can lead to the chaotic rankings mentioned in the first point above, and it also raises the more practical problem of how to determine which rating factor value to use when comparing quotes between insurers
6. profiling is open to abuse – if insurers know in advance what profiles are to be used, then they have the opportunity to fine tune their rating so that the particular profiles rates well, in a way that is not indicative of their general pricing. However I am not aware of any instances so far of this occurring.

How can something that sounds so simple go wrong so easily? Insurance is often treated as a commodity product by policyholders (despite TV advertising in which each insurer claims to be different to the rest), and economic theory tells us that a commodity product should result in a type of market in which competitive forces result in a tight spread of prices. The problem is that the product usage isn't homogenous nor is that individual usage perfectly understood by either supplier or purchaser. Because of this variation in usage, insurance is more like mobile phone plans than like bread or milk – there are a bewildering range of prices and packages and therefore comparisons between suppliers are not trivial. The lack of perfect information on the part of both consumers and insurers results in a bewildering range of rating factors and rating factor relativities.

Figure 1: Profiling used to compare iPhone plans

Rank	Phone	Product	Price/month	Total Value	Estimated Usage	Total Cost
1		**Optus** $49 iPhone Cap	$62	$300		$1,488
2		**Optus** $59 iPhone Cap	$68	$350		$1,632
3		**Vodafone Online Store** $49 Maxi Cap	$76	$310		$1,826 + Bonus
4		**Virgin Mobile** $70 iPhone Cap with 1GB Data	$74	$520		$1,776
5		**Vodafone** $49 Maxi Cap	$76	$310		$1,826
6		**Optus** $79 iPhone Cap Unlimited SMS	$83	$550		$1,992
7		**Vodafone** $69 iPhone Cap	$82	$310		$1,965
8		**Optus** $89 iPhone Cap	$89	$600		$2,136
9		**Vodafone Online Store** $79 Super Cap	$103	$550		$2,473 + Bonus
10		**Virgin Mobile** $100 iPhone Cap with 5GB Data	$100	$520		$2,400

Any insurance market analyst who has been asked to compare premiums between insurers can't help but have a feeling of déjà vu when looking at the comparison of iPhone plans shown above (source http://mobile-phones.smh.com.au). These should be commodity products because they are all offering the same hardware and the same type of data service. But since iPhone applications are more likely to use data than a normal phone, and iPhone users are likely to have differing data requirements, the telcos have developed pricing schemes that offer different prices for different usage. The complexity of the plans means that no simple statement can be made about which provider is most competitive, so the best that can be done is to use profiles and say which plan is best for which profile.

Consider a hypothetical niche insurer that is risk averse to high sums insured because it doesn't want the volatility of loss ratios that come from occasional total losses on these vehicles. If it chooses to quote premiums 25% higher than market for those vehicles with sums insured greater than $30,000 and 2% lower than market for the other vehicles then an industry survey that used five vehicle profiles with sums insured of ten, fifteen, twenty, thirty and fifty thousand dollar values could conclude from its average premium quotes that it was uncompetitive. The bias comes not only from the implicit assumption that 20% of vehicles have sums insured greater than $30,000, but also from the fact that vehicles with a higher sum insured have a higher premium, and therefore gain a higher weighting when calculating the average premium across all profiles.

A similar effect can occur for another hypothetical insurer that has a strategy of marketing to regional areas. If it prices 5% below market in regional areas, and 5% above market in suburban areas, then a profile approach would conclude that the insurer's prices were uncompetitive because its average premiums across all profiles was high. This would be due to the lower average premiums generally seen in regional areas leading to a lower weight being given to the competitive premiums.

Finally, consider an insurer that writes a niche motor insurance product that is pitched at people who do not drive their vehicles as much as the norm (such as Pay as You Drive). The use of an average profile to compare insurers will disadvantage this product because the product is not pitched to be attractive to an average driver.

In order to understand what types of problems and solutions there are in real life, I will compare the profiling practices of two market surveys:

	Survey A	Survey B
Number of Profiles	10	4
Number of suburbs	5 for each profile (50 different ones in total), 5 states tested	16 suburbs, repeated for each profile
Gender	random 50/50 split	Works on profiles: 1 single male, 2 couples, 1 family with young driver
Age	randomly selected 16 to 60	Profiles 21, 38, 48, 60
Make	make determined randomly by selecting from a cumulative frequency table*	Hyundai, Holden, Mazda, Toyota
Model	The most popular model is chosen	
Manufacture Year	year determined randomly by selecting from a cumulative frequency table*	2004, 2005, 2007, 2007
Sum Insured	$2,150 - $77,000	$10,000 – $31,900
Other insurance provider	determined randomly by selecting from a cumulative frequency table*	
	*Stats based on market share /score /population	
Other Assumptions	All drivers have a good driving record.	Same
	Each driver attained their license at the age of 18.	See profile
	One driver, single.	Allows for other drivers in the family
	Car purchased new	Same
	Standard features: AirCon, Driver and Passenger Air Bag, Central Locking.	Cars have standard manufacturer's accessories. No metallic paint if that's an optional extra.
	No modifications	Same
	Vehicle is owned, registered and used privately.	Same
	If a "Value to be insured" is required, go to www.carsales.com.au and use the midpoint of the private sale value range.	Same
	Garaged at night.	Street during the day and on a driveway on a private property at night.
	No other cars are owned.	Same
	Assume 15000km per year.	Same
	No extra options: no hire car, no windscreen, no NCD protection.	Same
Excess	No multi-policy discounts.	Same
NSW	500	Uses standard excess
VIC	450, 500	Uses standard excess
QLD	400	Uses standard excess
SA	300, 350	Uses standard excess
WA	250	Uses standard excess

Survey A and Survey B take different approaches to profiling. Each has different strengths. Survey A's strengths are:

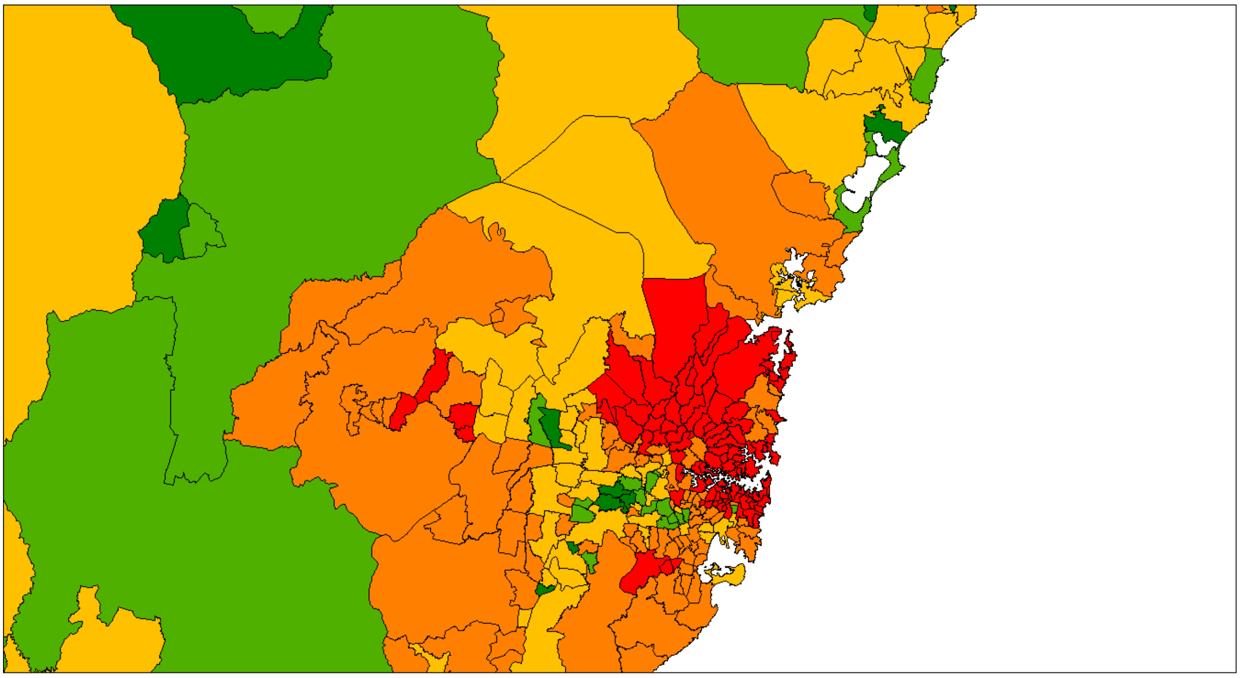
- the driver age, vehicle make and manufacture year, and current insurer used in the sample profiles are implicitly weighted to the population mix via the cumulative frequency approach. Compared to Survey B this is a strength because Survey B profiles
  - use only relatively new vehicles, which biases the ranking to insurers that are competitive on newer vehicles
  - have a young driver for one quarter of their profiles, which is not representative of the population mix, and which does not allow for insurers that choose to price themselves away from these risks
- the vehicle model, sum insured, modification, private usage, car ownership, driving record and distance driven use the most common or most likely values. While this doesn't cover the full range of possible values, it covers the greatest proportion of the population that is possible with a single quote. This is a way of dealing with complexity and with insurers that use different rating factors to one another.
- the excess is standardised so that different insurers are compared against the same excess
- sums insured span a wider range of values than Survey B, which is inclusive of the lower sums insured for older vehicles

Survey B's strengths are:

- the driver genders are linked to the secondary driver profiles and the driver age profiles, which allows for the way these rating factors tend to go together
- uses a narrower range of sums insured, which avoids bias from outliers from quotes for high sums insured
- the vehicle model, sum insured, modification, private usage, car ownership, driving record and distance driven use what the surveyor considers likely values for that state. While this doesn't cover the full range of possible values, it covers the greatest proportion of the population that is possible with a single quote. It also means that comparisons are done on the types of risks for which insurers are competing the most. This is a way of dealing with complexity and with insurers that use different rating factors to one another.
- The excess uses the standard excess, which is the most likely excess for the insured to choose. While this doesn't cover the full range of possible values, it covers the greatest proportion of the population that is possible with a single quote. On the other hand it may mean that it is not comparing like with like between insurers that have different standard excesses.

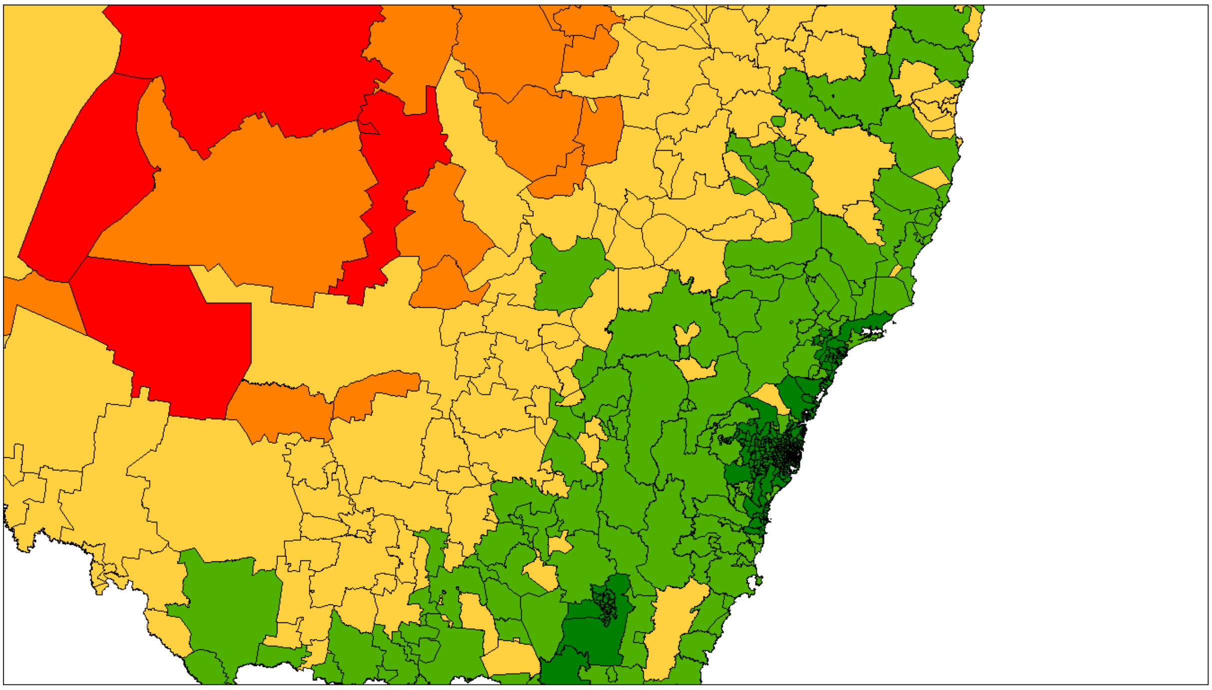
When looking at post codes and suburbs it is useful to reduce the number of dimensions. With thousands of post codes and an even greater number of suburbs in Australia, the number of possibilities to consider for comparison is impractical. The Australian Bureau of Statistics (ABS) produces some use statistics about geographic areas. These statistics allow the user to rank post codes on different measures, which in turn allows the profiler to use stratified sampling. This paper uses a very simple example of this analysis that uses two indices: the socio-economic index and the remoteness index. In practice one would need to consider the risk factors that actually affect the type of risk being considered, and choose a range of measures that are appropriate to measuring those risks.

Figure 2: Socio-Economic Indices



The ABS produces four socio-economic indices. This paper uses the "Advantage and Disadvantage" index, which is "a continuum of advantage (high values) to disadvantage (low values) which is derived from Census variables related to both advantage and disadvantage, like household with low income and people with a tertiary education" (ABS 2039.0). Looking at the map above, we see that the index scores highest for wealthy post codes.

Figure 3: Remoteness Indices



The Australian Standard Geographical Classification (ASGC) "was developed by the ABS in response to a demand for a statistical geography that allows quantitative comparisons between 'city' and 'country' Australia where the defining difference between 'city' and 'country' is physical remoteness from goods and services." (ABS Census Geography Paper 03/01).

Looking at the map above, we see that it scores highest for urban areas. This is a useful way of quickly separating urban suburbs from regional suburbs.

Figure 4: Comparison of geographic profiles used in two surveys

Remoteness & SEI	Survey A	%	Survey B	%
Major Cities of Australia	44	88%	7	44%
Inner Regional Australia	4	8%	3	19%
Outer Regional Australia	2	4%	5	31%
Remote Australia	-	0%	1	6%
Total Postcodes	50	100%	16	100%
Average SEI	1,057		1,013	
SEI 25th percentile	994		923	
SEI 75th percentile	1,116		1,081	
Max SEI	1,209		1,198	
Min SEI	819		835	

Survey A uses suburbs that better represent the mix and importance of socio-economic and remoteness characteristics in Australia. It has chosen suburbs from a broader range of index values, and it has weighted them according to the population. This means that Survey B's results are biased towards insurers that are competitive in regional areas, and its use of a smaller number of suburbs leaves it more exposed to sample error. It is quite possible that these two surveys would come to significantly different conclusions to the ranking of different insurers.

The weaknesses of both surveys are:

- The suburbs chosen do not represent a balanced mix of socio-economic status and remoteness
- The sample sizes are too small to
  - obtain a stable estimate of insurer ranking – this is point discussed later in this section
  - understand the rating relativities – this is especially so for Survey B which doesn't have enough samples in major cities to cover the number of suburb rating districts that one typically sees
  - understand the implications of changes in rating factor values – when the sample size is small there is insufficient coverage of possible combinations of rating factor values to be able to determine which rating factor relativity rerate led to a premium change
- Rating factor values are considered to be independent, but many rating factors belong together
  - Lower socio-economic regions are less likely to have policyholders with a new or high value vehicle
  - Lower age groups are less likely to have policyholders with a new or high value vehicle
  - Different gender choose different types of vehicles
  - Different socio-economic and remoteness regions have different age profiles
  - People drive their vehicles differently in inner city suburbs versus outer regional areas e.g. someone living in Mosman is likely to catch public transport to work and leave their vehicle locked in a garage all day, so using a profile of being parked on a street during the day is not the most likely scenario

- The choice of make and model of a vehicle should linked to the year of manufacture

So what can profilers do to minimise the problems inherent in the profiling approach?

1. Choose rating factor values that have regard for sensitivity testing of the rating factor values being used
2. Choose rating factor values that represent the market segment that you are interested in
3. Allow for natural groupings of rating factor values
4. Weight the profiles depending on their importance

Sensitivity testing the rating factor values involves varying the rating values used, and noting which variations cause a material change in the insurer rankings. For example, you may discover that insurers have different rating relativities for suburbs – some insurers may be more competitive in particular suburbs, and other insurers may be competitive in different suburbs. Your profiles need to include more samples for the rating factors that are flagged by the sensitivity testing. You do not need as many samples where the premium rankings are not sensitive to the rating factor values.

Figure 5: Sensitivity test of insurer premium rank

	Insurer						
	A	B	C	D	E	F	G
Suburb 1	6	4	7	2	5	3	1
Suburb 2	5	2	6	1	7	4	3
Suburb 3	7	3	5	1	4	6	2
Suburb 4	4	2	7	1	5	6	3
Suburb 5	2	6	7	5	3	4	1

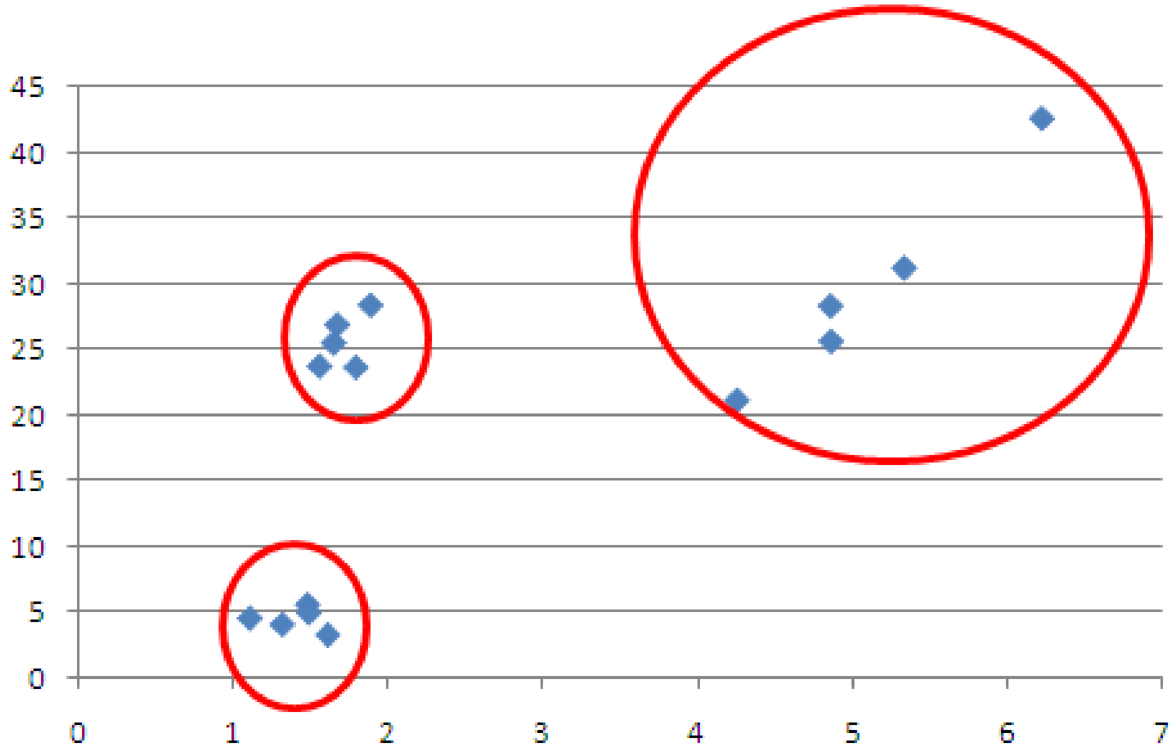
The table above shows the effect on insurer rankings of changing the suburb in an actual profile used by a survey. The rankings of insurers in this survey are extremely sensitive to the choice of suburbs. Insurer D's ranking changes from second to worst, depending on the choice of suburb. Insurer C is consistently ranked towards the worst ranking. This type of sensitivity testing shows how chaotic insurance pricing can be, but can also confirm that some insurers are more competitive than others.

Also, consider the situations for rating factors that are only used by one or two insurers , whether those rating factors make a large difference to the premiums and whether there is likely to be much variation in the value of those rating factors. For example some insurers rate new business using the person's existing no claim bonus (from another insurer) and some do not. There is a large difference in premium between a 60% no claim bonus and a 0% no claim bonus. Profilers typically assume that the person has full no claim bonus, which is the most likely possibility, but this may not be an appropriate comparison for some niche markets. Sometimes one is safe choosing the most likely rating value for comparison. For example, one insurer loads motor premiums by 4% if the policyholder is unemployed. In such a case one knows from publicly available statistics that approximately 95% of the adult population is not unemployed, and that the rating differential is not all that large, so it would be reasonable to only use comparison profiles in which the insured is employed.

If you are looking at the whole market, then your profiles need to consider the whole market, and you need to use rating factor values that represent the broad range of values that exist across the market. For example if you are profiling domestic motor premiums, then you should consider older and newer vehicles, higher and lower value vehicles, and suburbs with a wide range of socio-economic indices. On the other hand, if you are an insurer comparing yourself to competitors, then your profiles should be chosen with regard to your marketing strategy. For example, if you are a state based insurer, then your profiles should be restricted to the state that you operate in, and the range of insured items and policyholders that are seen in your market. One way to do this (if you have the data) is to use your premium quote database for selecting the profiles because your premium quotes reflect the market that you are operating in. if you do not have access to a quote database, then you may use your policy database, but this is considered suboptimal because it reflects the policies that you are winning, which may be a biased sample of the entire market.

Rating factor values belong together, sometimes in surprising ways. I have found that cluster analysis is very useful for understanding how rating factors belong together. For those who haven't come across cluster analysis before, it is a statistical technique for segmenting a set of data into groups of data that are similar to one another within a group, but where each group is dissimilar to the other.

Figure 6: Cluster analysis



Consider the graph above. The data points naturally fit into three groups. Cluster analysis would identify those groups and which data points belong in each group. You would then know that very low values in the y-axis can only occur for a narrow range of values in the x-axis. The same type of analysis can be applied to rating factor values to find groups of rating factor values that belong together.

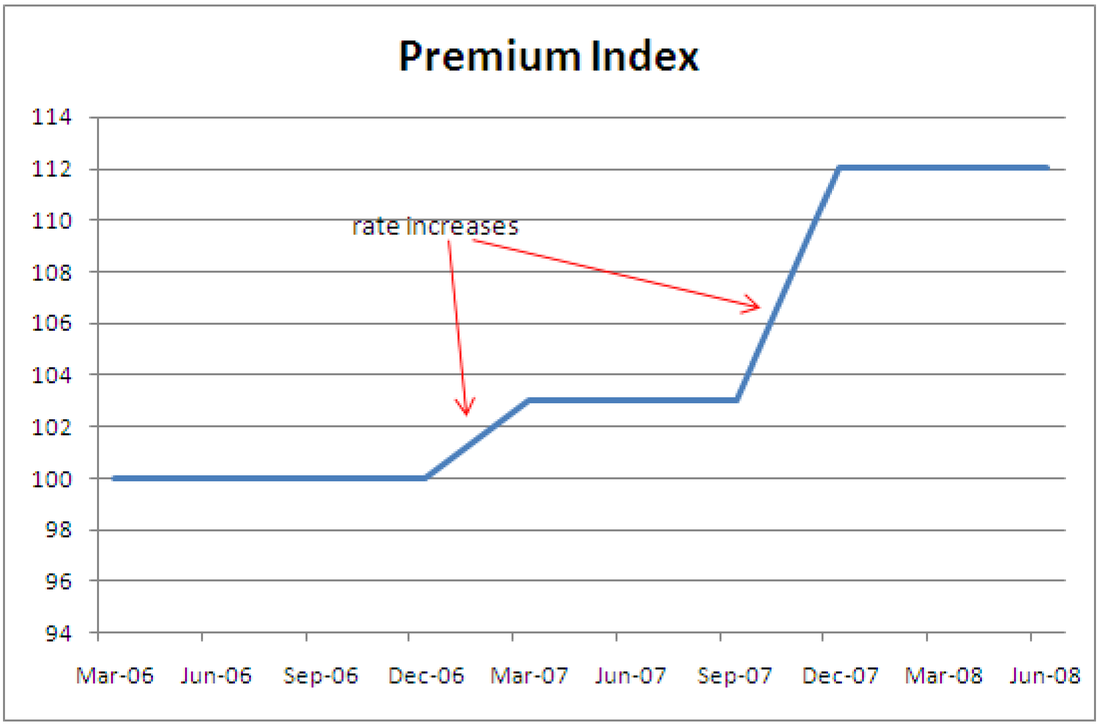
Figure 7: Cluster analysis of mix of business

The output above shows part of a cluster analysis output. You can see that each cluster has a different profile and that some rating factors belong together. For example you can see that cluster 5 is people who are single, and that single people are less likely to have any nominated drivers, less likely to drive a Falcon, and less likely to garage their vehicle. While this is just a sample output, it demonstrates the point that the choice of values for different rating factors should not be considered independently of one another.

Profiles should be chosen to represent the range of different rating factor values, then the profiles should be weighted to represent their relative importance. If most of the population has the same set of rating factor values, then only one profile is needed for that market segment, but that single profile should then be given a higher weight to allow for its importance.

Premium Indices

We appear to be at a turning point in the insurance cycle. I see news such as "Renewals reveal premium rises" (Cover Note Issue 1595) which leads me to think that rates are starting to harden. Some insurers are announcing that they intend to increase rates such as the news that "CGU was increasing premium rates 2% to 12%" (Cover Note 1564). But I also see news such as "many insurers say they intend to increase premiums in these classes, they are reluctant to give up hard-won market share, based on price alone" (Cover Note Issue 1577). Then we have the time invariant anecdotal evidence of sales staff claiming that the rates are too high…



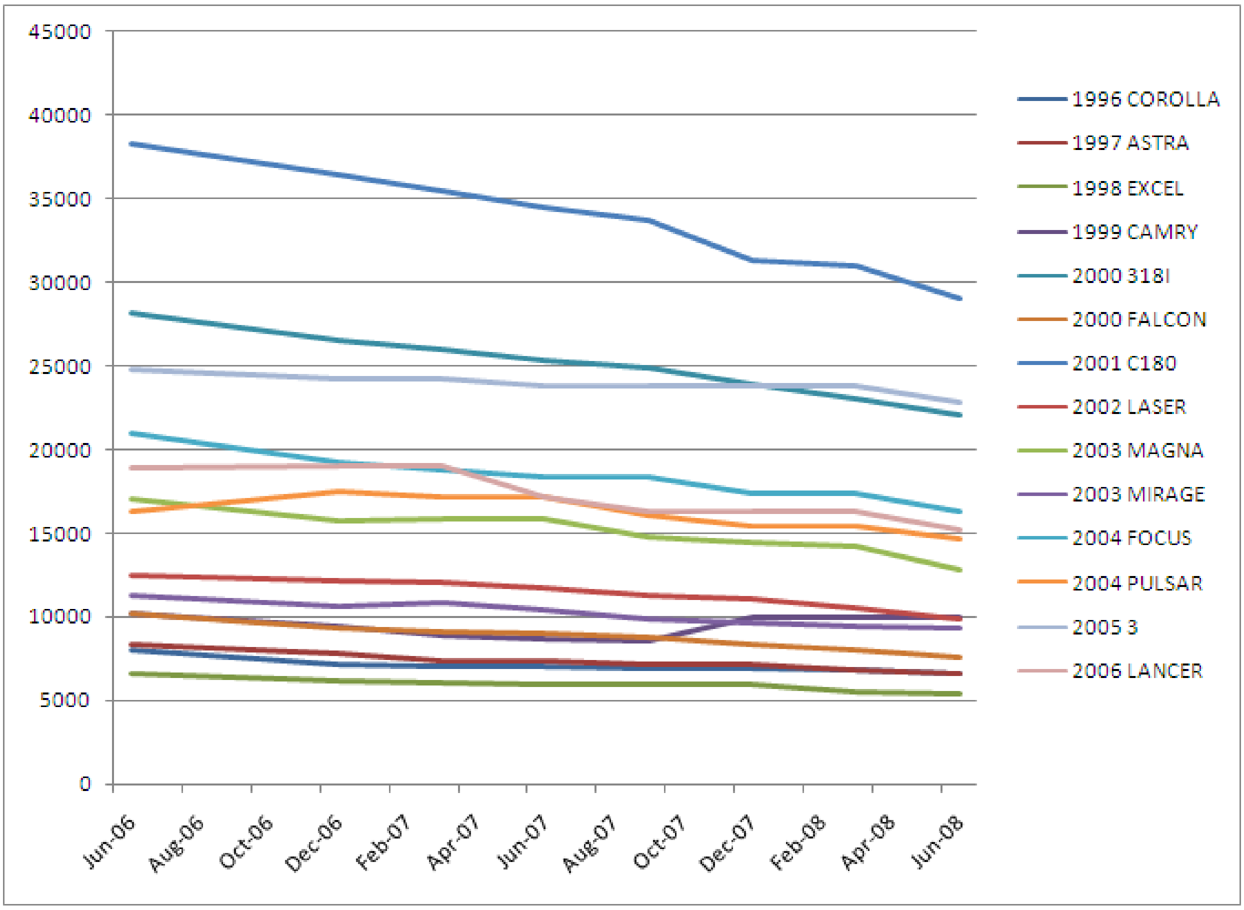
Some market commentators (e.g. Macquarie Equities) track premium rates through time to measure whether rates are hardening or softening. The approach that is taken is to use the profiling technique across a period of time, and to measure how the premiums for each profile change across time. When premium rates are increased by an insurer, the premiums for the affected risk profiles will increase (everything else being equal). So one should be able to determine when an insurer changes their rates by watching for movements in the premium index.

Rather than repeating the issues with profiling, this section will focus only on the issues related to tracking of premiums through time. The problems that one encounters are:

1. Inflation: many assets appreciate in value over time, so the required sum insured increases
2. Depreciation: many assets depreciate in value over time, so the required sum insured decreases. To make things even more difficult, the rate of depreciation can vary by asset type and age.
3. Asset age is often a rating factor: in some classes of business e.g. householders, and domestic motor, the age of the asset being insured is a rating factor. The premium quotes on an asset therefore changes with time even if the premium rates remain unchanged.
4. Obsolescence: some assets become obsolete, and different products replace them. This means that you sometimes cannot get identical assets of different ages to directly compare e.g. a vehicle manufacturer may stop manufacturing one model, and instead release a brand new model
5. Change in mix over time: the market mix of risks changes over time e.g. the proportion of fibro buildings is reducing over time. This needs to be allowed for, or your profiling weighting will be biased towards older risks.

The issue of changing sums insured affects almost all classes of business. Premiums are usually not directly proportional to sum insured i.e. a 10% increase in sum insured usually results in premium increase of less than 10%. That is because some policy costs are the same no matter how much the sum insured. One therefore cannot correctly adjust for changes in sum insured over time without understanding the sum insured rating relativities. While the most accurate solution is to find out the full rating structure of the insurer (such as discussed in the next section of this paper), this is not usually practical in terms of time or effort. A simpler solution is to get pairs of quotes on sums insured that are a few percent different, and to assume a linear relationship between sum insured and premium, to approximate the effect of change in sum insured. The other simple solution is to quote against the same sum insured in consecutive periods, but this may not solve the problem if the asset has a market value e.g. domestic motor because the sum insured variation from the market value may in itself be a rating factor. To do this properly you will need to understand the insurer's rating relativities for sum insured.

Figure 8: Depreciation varies by vehicle make and model

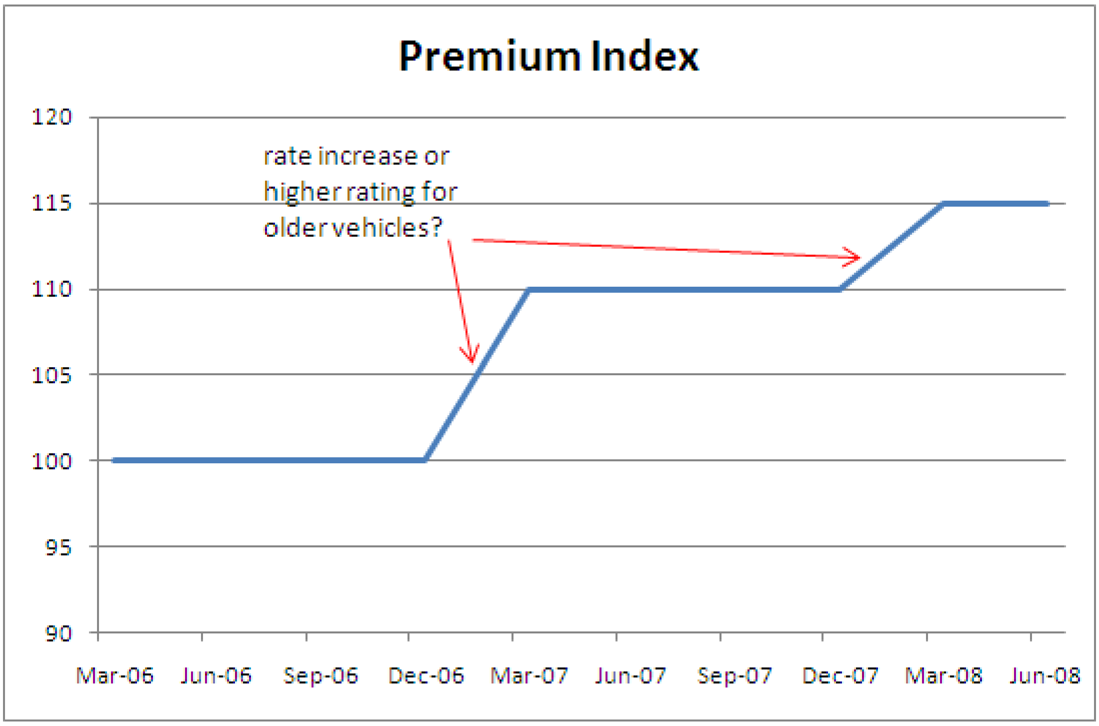


Depreciation is not constant over time, nor is it constant across different assets. In the graph above we can see that different types of vehicles have experienced different rates of depreciation, and some have even had periods of appreciation in value. Similarly buildings tend to appreciate in value over time, but at different rates, and some may depreciate in value at times. One can track the amount of appreciation or depreciation by tracking average sums insured. When tracking sums insured, one should ensure that changes are not simply due to changes in the mix of the portfolio e.g. writing more high value risks. The best way to track movements in sum insured is to compare the same assets across time.

For asset or risk profile substitution there are two possible solutions:

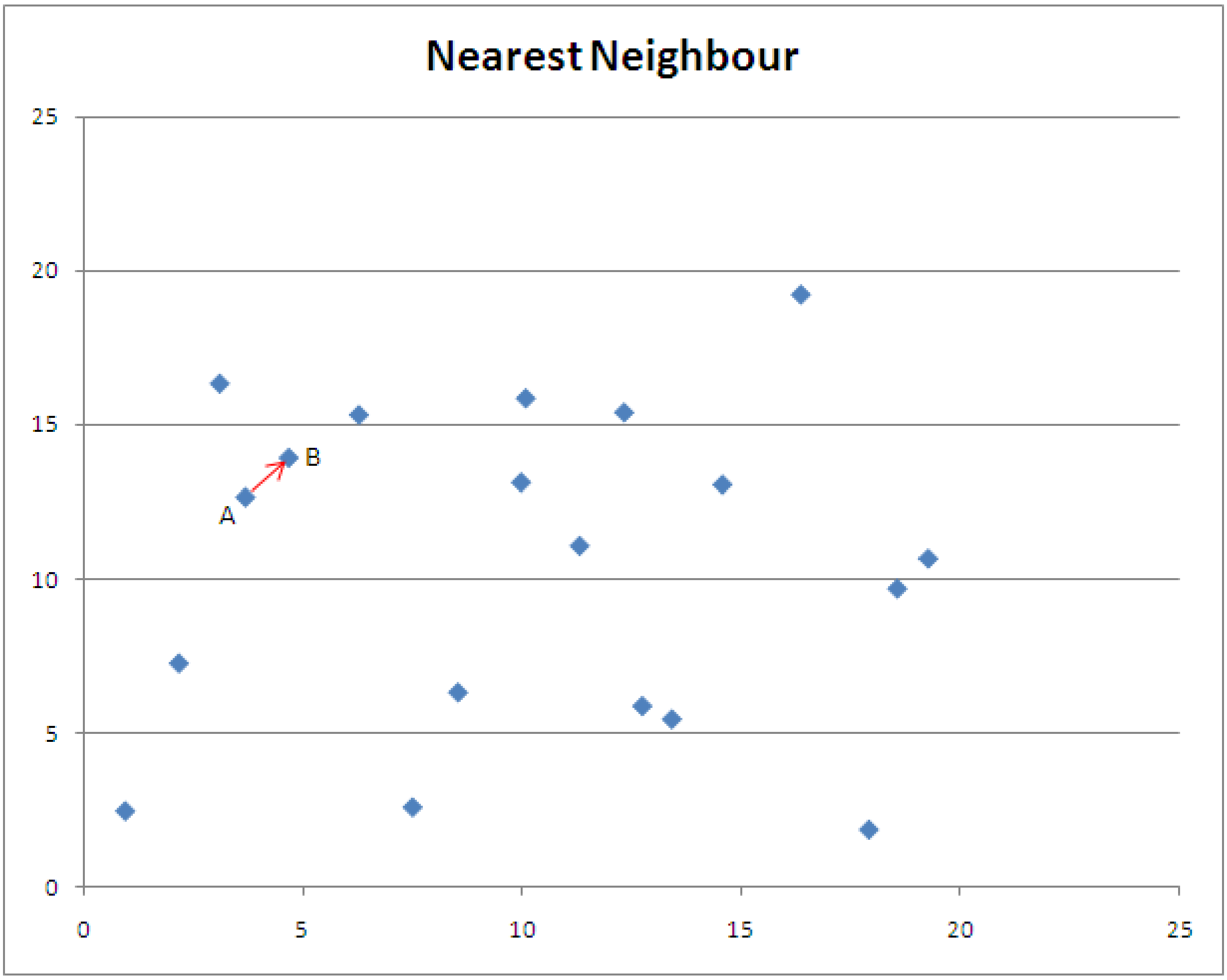
1. removing the substituted asset / risk profile from the period on period comparison
2. finding a substitute asset / risk profile using nearest neighbour search

The first solution, removing the obsolete asset from the comparison, is practical only if a small number of assets became obsolete during the period. But there are cases where this is not going to be true e.g. going from a December quarter to a March quarter all vehicles cease to be current models!



In this case one must find a way to compare the two quarters. The underlying asset will not have changed much over three months, but the way it is treated by an insurer may change. In the case of motor, the age of a vehicle is usually a rating factor, and that age is usually calculated as the current calendar year less the year of manufacture. So the rating factor value of a vehicle changes from December to the following January. This may either increase or decrease the premium charged, depending upon the insurer's rate relativities. It is therefore necessarily to understand the insurer's year of manufacture rate relativities in order to adjust for the aging of a vehicle. If you do not adjust for this, you will come to the wrong conclusion about rate movements.

Figure 9: Nearest neighbour search



Nearest neighbour search is simple in concept. In the graph above each data point has a nearest neighbour – the data point that is closest to that point. The red arrow signifies the distance from point A to point B, where point B is the nearest neighbour to point A. The concept can be extended to rating factors but need some judgement as to the relative weight to give to different distances within rating factors, and how to find a distance for non-numeric rating factors.

With non-numeric rating factors, one can take two approaches:

1. find substitute numeric variables e.g. for post code you can't use the difference between the post code numeric values because post code 3996 in Victoria is nowhere near postcode 4000 in Queensland, so instead you can use the latitude and longitude of the centroid of the postcode to calculate the physical distance between the postcodes
2. subjectively determine a scoring system e.g. for parking you may determine that garaged and carport are not much different, so only score a "distance" of 1, while garaged and parked on a street may have quite different risks and score a "distance" of 10

In practice premium indices cannot be created without some understanding of premium structure because a risk profile's rating factor values change with time, so its premium changes with time, even if an insurer has not changed its rates. This leads us to the next section in this paper.

Premium Structure

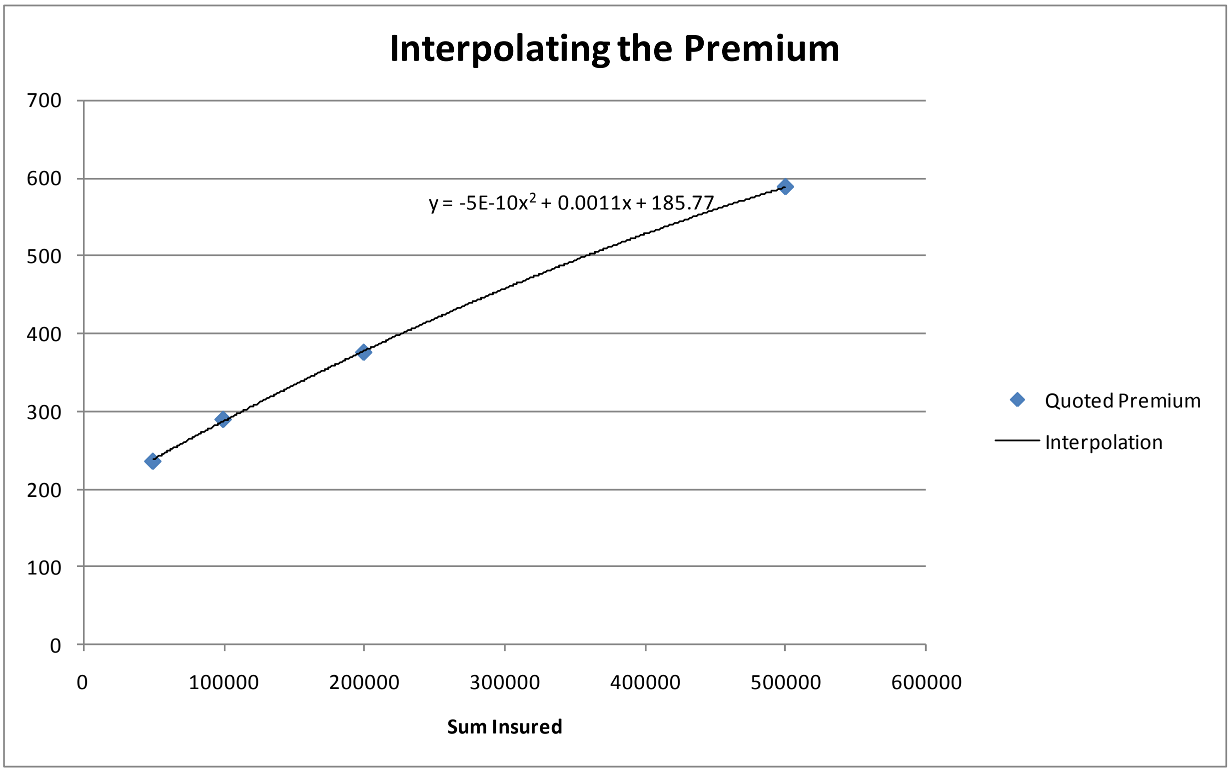
The analysis of a competitor's premium structure can be used by insurers to understand where their rating is competitive, to help determine appropriate rates when entering a new market, and to understand the price elasticity of their customers. One does not approach this analysis lightly because it is usually time consuming and complex.

As discussed earlier in this paper, there are usually too many possible rating factor values to fully enumerate all of the possible combinations when quoting prices. So a more practical approach is required. The steps are:

1. determine sample values for continuous rating factors
2. determine sample values for geographic rating factors
3. one way analysis
4. informed use of two way and even three way analysis
5. interpolation
6. random samples to test the accuracy of the interpolated structure

The rating factors with the greatest number of possible values are those with continuous values e.g. sum insured, or with a large number of non-ordered values e.g. suburb. For sum insured we can make a simple assumption that reduces the dimensionality – we can assume that the higher the sum insured, the higher the premium (all other things being equal). This assumption makes sense for rational economic decision makers, because if an insurer offered a lower premium for a higher sum insured, then the policyholder would just choose the higher sum insured, get more cover and pay less for that extra cover.

Figure 10: Interpolating the premium for different sums insured



This assumption helps us because it sets limits around the premium that can be charged for sums insured that haven't been sampled e.g. because we know that if we have sample quoted premiums for sums insured of $50,000 and $55,000, that the premium for a sum insured of $54,000 must lie between these two premium values, and is likely to be nearer to the premium for a sum insured of $55,000. A similar principle applies for excess – the higher the higher, the lower the premium. Application of this principle enables interpolation of the premium.

This means that for continuous rating factor values, like sum insured, we only need to obtain enough premium quotes to allow us to fit the shape of the curve.

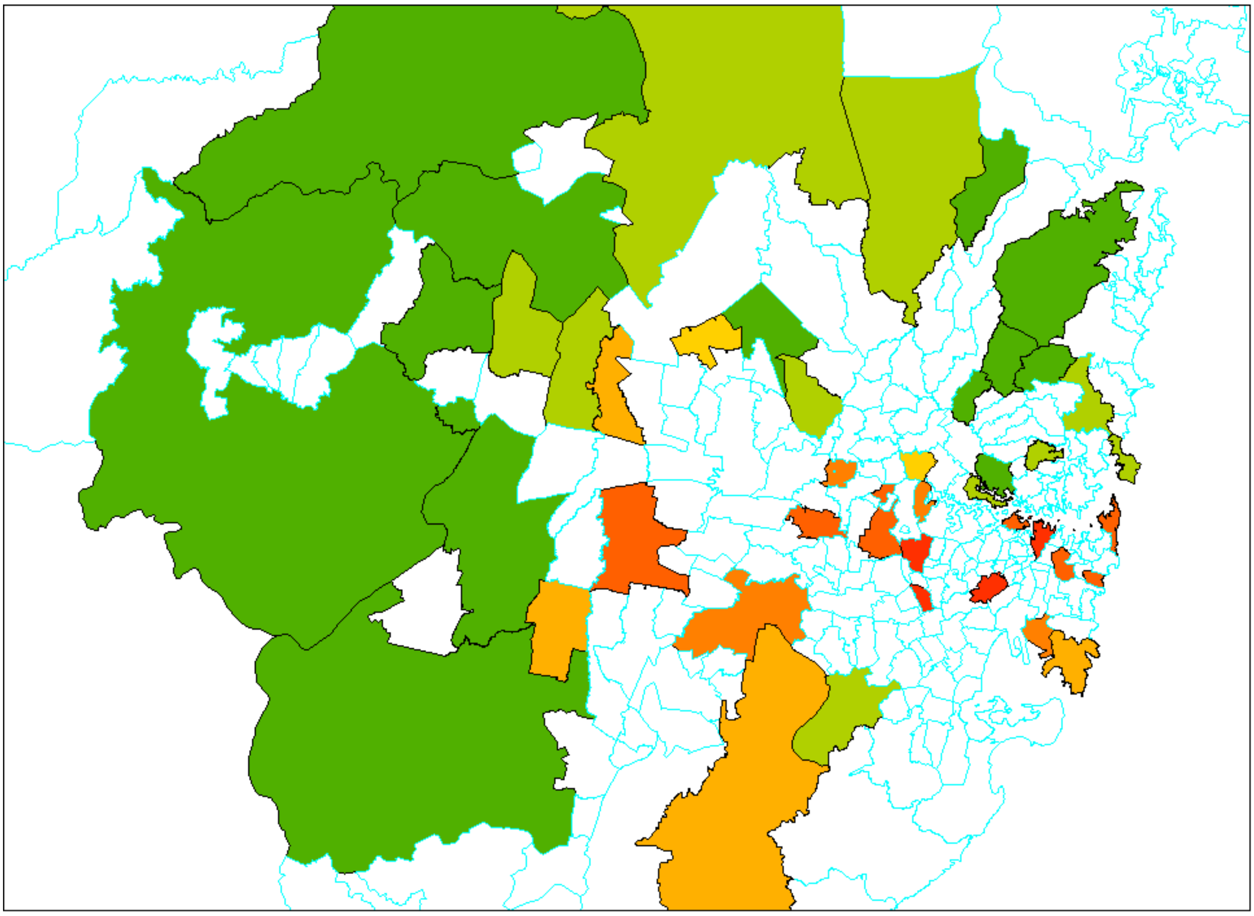
This concept can be extended into two dimensions – into suburb and post code rating. But there are differences from what we were able to do with sum insured:

- a greater sum insured required a higher premium, but there is no such thing as an absolute ranking of suburbs (which suburb is greater than the other?)
- geographic patterns are more complex than a smooth curve because of the complexities of physical geography and demographic patterns e.g. two suburbs can be separated by water around Sydney, and although they are close they are not connected and have different characteristics

Interpolation can still be applied to geographic rating. It just needs to be boosted by actuarial judgement and the use of some socio-economic and geographic measures.

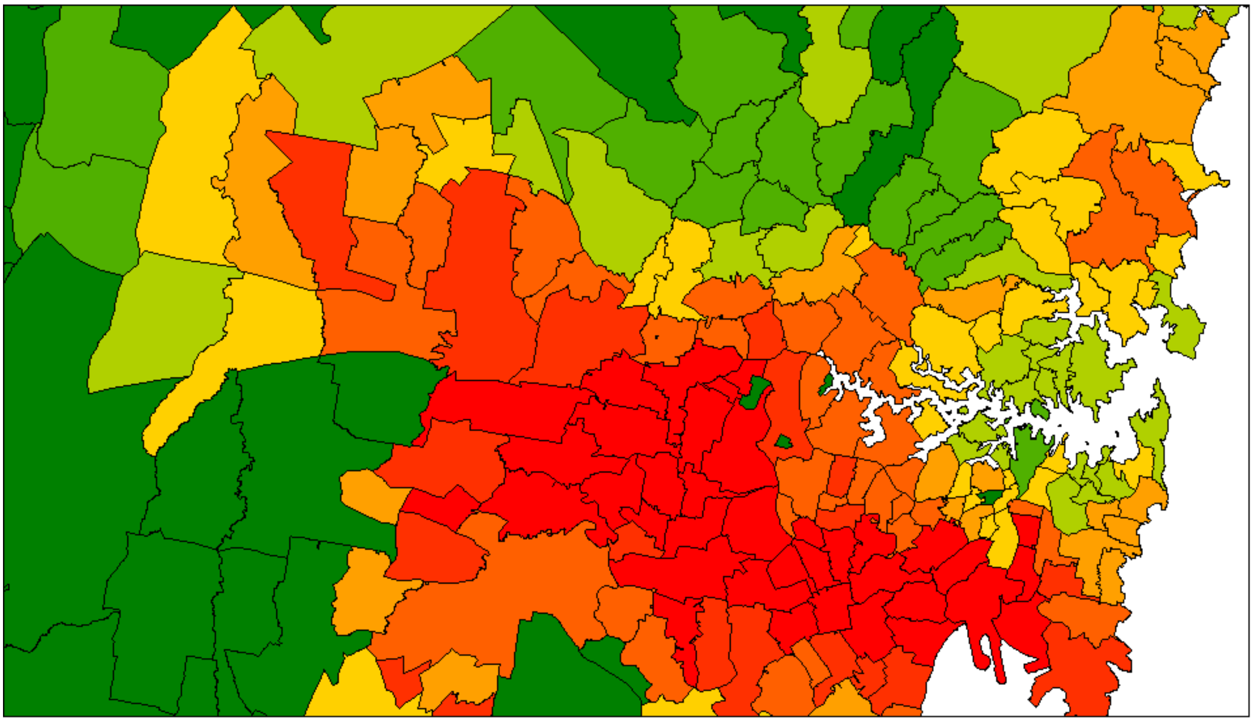
The first step is to choose some numeric measures for each suburb or post code. These can be chosen by hypothesising about what factors are likely to affect the nature of the risk. For the purpose of this paper I am using socio-economic index, but this is just for the purpose of example and should not be taken as a recommendation of the most appropriate index to use.

Figure 11: Diverse locations of sample premium quotes



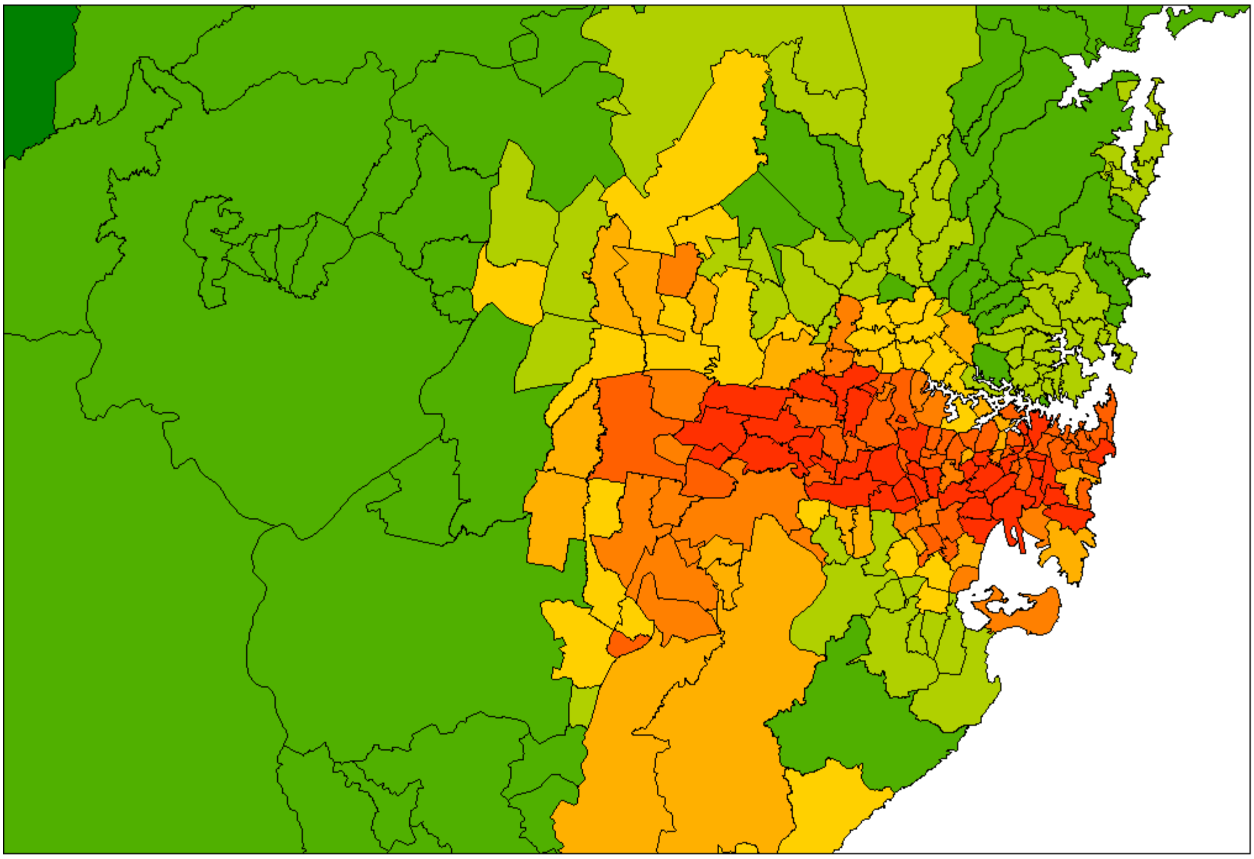
The second step is to choose some sample locations for obtaining quotes. Once you have chosen your factors, map them and look for regional patterns. Then choose a small number of samples that give a range of combinations of values from these underlying factors. The map above shows some sample premium quotes selected to give a range of different socio-economic indices.

Figure 12: Interpolated post code rating (first iteration)



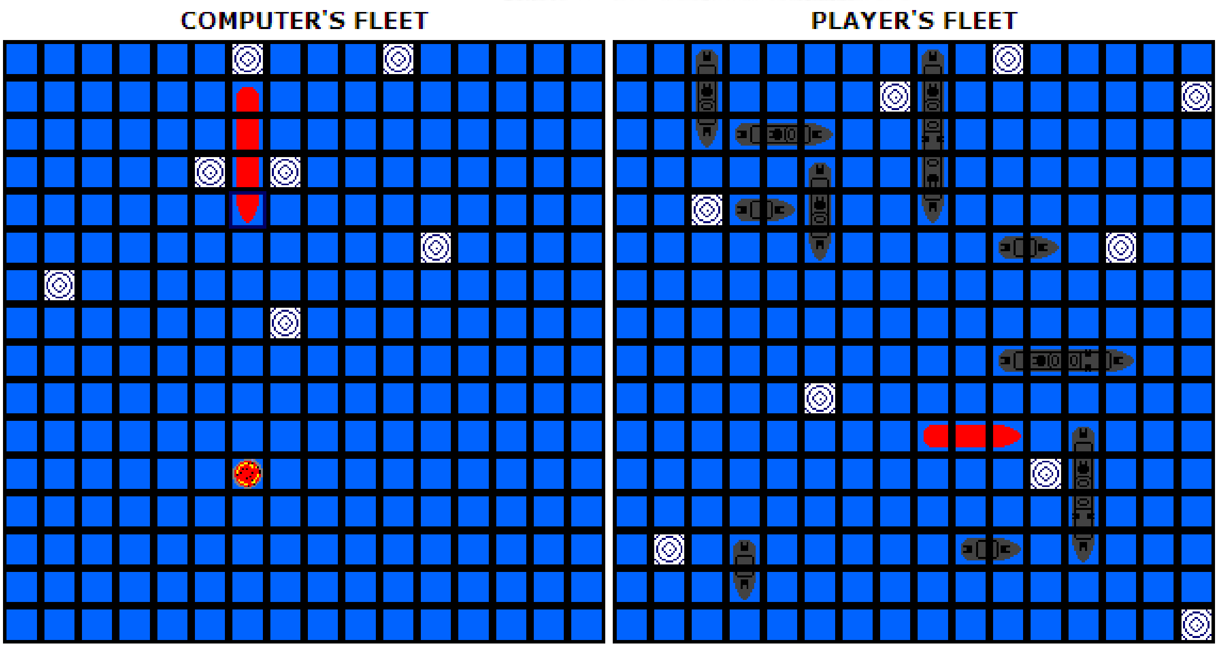
The third step is to fit a function that explains those sample quotes in terms of the underlying geographic factors that you chose back in step one. The underlying geographic factors tend to interact when explaining the premium, so a simple linear function may not always work. This type of problem is well suited to neural networks. The map above shows an interpolated rating that was achieved via fitting a neural network to the sample quotes.

Figure 13: Interpolated post code rating (final iteration)



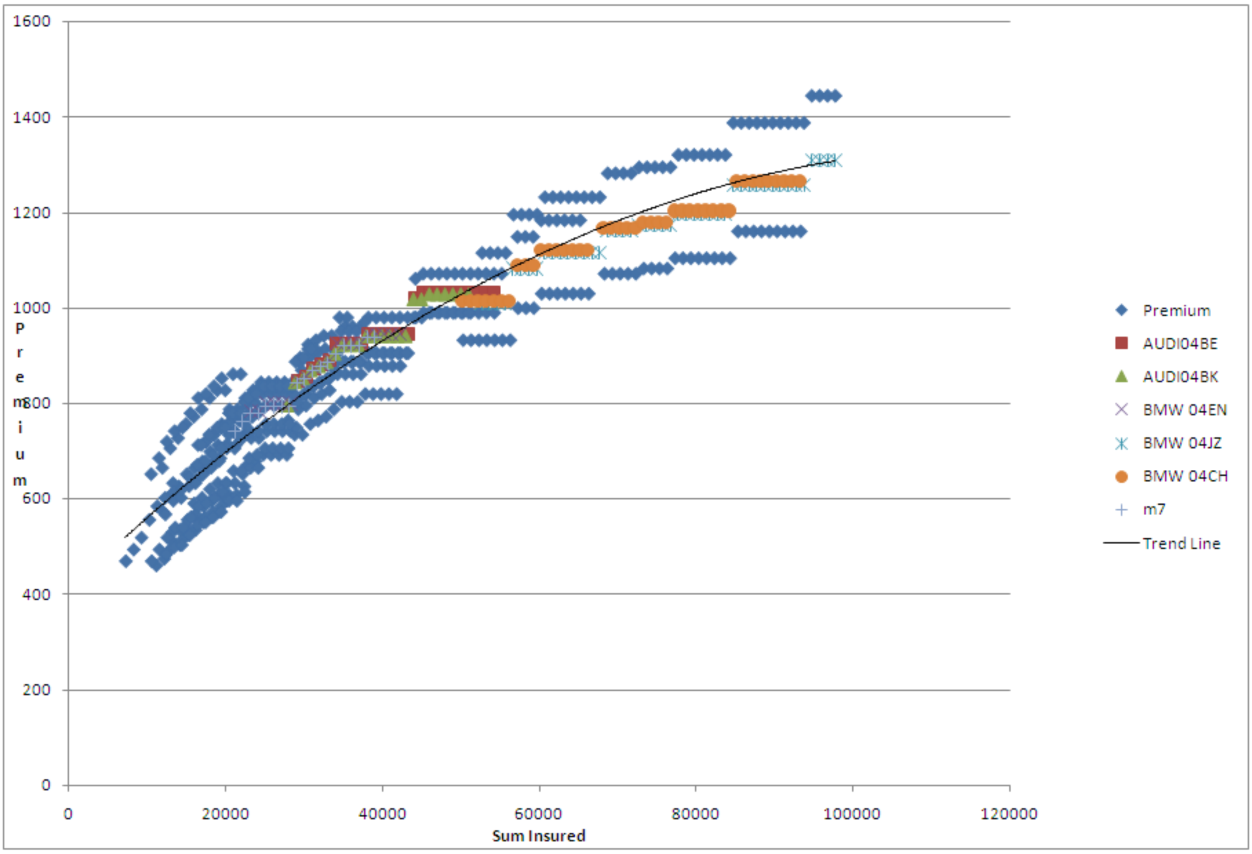
Finally, you need to test and fine tune the interpolation. Choose some suburbs that stand out as being different to the surrounding suburbs, and choose some other suburbs randomly and obtain quotes for those suburbs. If the quotes fit your interpolated values, then your model works well and you are finished. Otherwise you may need to add more samples, or add more underlying numeric factors that explain the discrepancy. The map above shows a rating that has been fine tuned to account for all discrepancies.

Figure 14: Partway through a game of battleship



You can think of this fine tuning step as being much like playing a game of battleship. You start by selecting some seemingly random (but not truly random – they need to be spread apart) positions, and assessing their worth. The results from these selections start to inform you – if you hit a ship then you know that you need to test some extra positions around that location but you don't know the exact size or location of the ship. By intelligently adding samples you find out more about the size of location of the ship until you have it destroyed. Similarly with geographic rating you start with some geographically diverse samples, and then you respond to results that are unusual and fine tune until you understand the patterns.

Figure 15: Vehicle make, model and year of manufacture with varying sum insured / premium relationship



Some rating factors interact. One of the most different interactions to allow for is the combinations of vehicle make, model, year of manufacture and sum insured. When this happens, you need to put more effort into allowing for these interactions, otherwise you will not be comparing like with like e.g. how is a $30,000 year 2000 BMW 323 rated compared to a new $30,000 Mazda 323? This is a situation where multivariate regression can be helpful. In the graph above you can see how different vehicle makes and models have different premiums for different years of manufacture and different sums insured. The premiums generally increase with sum insured, but also vary by vehicle make and model and year of manufacture. The premium needs to be explained in terms of all of these factors.

Conclusion

Profiling requires some thought to the number, composition and weighting of quote profiles that are used. One must choose profiles that cover the range of variations that occur in profiles. Then those profiles must be weighted in accordance with the market size each profile represents. Sensitivity testing of the values that are used will identify which changes in rating factors produce arbitrary changes in insurer rank. Rating factor values are often not independent of one another. Cluster analysis can assist the choice of sample profiles.

Premium indices are subject to the same problems as profiling, plus additional problems relating to the difficulty of comparing prices across time. Some assets appreciate in value over time while some depreciate. Premium rates are not usually directly proportional to the sum insured, so one must understand the sum insured rating relativities in order to adjust for this effect. The age of an asset is sometimes a rating factor, so once again one must understand the rating relativities in order to adjust for this effect.

It is not a trivial task to understand insurers' rating relativities, but you need to do it in order to correctly calculate a premium index. Rates are usually not published, are often quite complex. The sheer volume of combinations of rating factor values precludes a brute force approach. By considering which rating factors should have monotonic relationships with premium, once can selectively use interpolation for some numeric rating factors. This concept of interpolation can be extended to geographic rating by linking suburbs to underlying risk factors for the location.

Understanding insurers' premium rating is not as simple as comparing petrol prices. There are traps in the analysis that require an understanding of insurance and strong numeric skills to overcome.

References

Australian Bureau of Statistics, 2001, *Census Geography Paper 03/01 - ASGC Remoteness Classification - Purpose and Use*, Australian Bureau of Statistics

Australian Bureau of Statistics, 2006, *2039.0 - Information Paper: An Introduction to Socio-Economic Indexes for Areas (SEIFA),* http://www.abs.gov.au/ausstats/abs@.nsf/mf/2039.0/

Choice, 2008, *Test: Home and Contents Insurance Online,* http://www.choice.com.au/

Cover Note, 2008, *Cover Note*, http://services.thomson.com.au

Macquarie Research, 2008, *2Q08 Macquarie Motor Pricing Survey*, http://www.macquarie.com/research

Pay As You Drive, 2008, *Pay as You Drive*, http://www.payasyoudrive.com.au