Our Agenda

- **Evolution of Predictive Modeling**   Bill Carpenter

    – Modeling claim propensity

    – Incorporating rate comparisons

- **State of the Art – Price Optimization**   Thomas Bayley

- **A Regulatory Perspective**   Chuck Romberger

Milliman

What is Predictive Modeling?

- Mathematical and statistical techniques to predict future outcomes based on data (variables) at the individual policy level

- How is it used in insurance?
  - Underwriting – both new business and renewal
  - Pricing / tiering
  - Customer acquisition
  - Customer service
  - Claims, e.g. fraud detection
  - Customer Retention
  - Agency Management

Milliman

Once Upon a Time …

- The expense and/or lack of processing power and electronic storage did not support large scale predictive modeling

	1980	1985	1990	1995	2000	2008
Storage Cost per Megabyte	$190	$ 70	$ 10	$0.90	$0.05	<$0.001
Microprocessor Speed, MHz	5-8	16	33	120	800	3,200

Milliman

Once Upon a Time …

- Easily accessible internal transaction level data did not exist

- The ability to append external data did not exist

- Management had to rely on instinct or "gut" feel

- Data was an "after the fact" verification of the accuracy of those intuitions

Milliman

What Changed ...

- Predictive modeling in the form of credit scoring dates back to the 1960s in the banking industry
  - An automated decision process was a must for the booming credit card business
- In the 1980s, other predictive modeling applications gained wide acceptance in direct marketing
- By the late 1980s, the first application of the extensive credit database started to be applied to auto insurance

Milliman

Steps in Predictive Modeling

- Data Mining: knowledge finding in large databases
  - Data exploration techniques (some brute force)
  - e.g. discover strength of credit variables
- Model Building: apply statistical techniques (like regression) after data mining phase is completed.
  - Quantify & synthesize relationships found in data mining exercise
  - e.g. build a credit model

Milliman

Data Mining: Sources of Data

- Moved from solely credit score to 1,000s of variables

- Company's internal data
  - Policy-level records
  - Loss & premium transactions
  - Agent database
  - Billing transactions (e.g. late payments, payment method …)
  - Vehicle information

- Externally purchased data
  - Credit, D&B, …
  - CLUE, MVR
  - Census
  - Marketing datasets
  - Geo-coding
  - ….

Milliman

Data Mining: Exploring the Data

- 1-way analyses of predictive variables

- Data Visualization

  – E.g. Decision trees, relationship maps …

- Create new predictive variables

  – Opportunity to incorporate underwriting judgment

  – E.g. # drivers vs. # vehicles

  – E.g. single vehicle / liability only …

Milliman

Data Mining

- Key step to building a successful model

- Key step in getting "buy-in" from others in your organization

  – Use this step to communicate and get input from others that will create the "buy-in" when you get to model building

- Can yield significant benefits, even before you get to the model building state

  – *Insight* tool (complete segmentation, lift model)

Milliman

Segmentation Lift Curve

Lift Curve Example
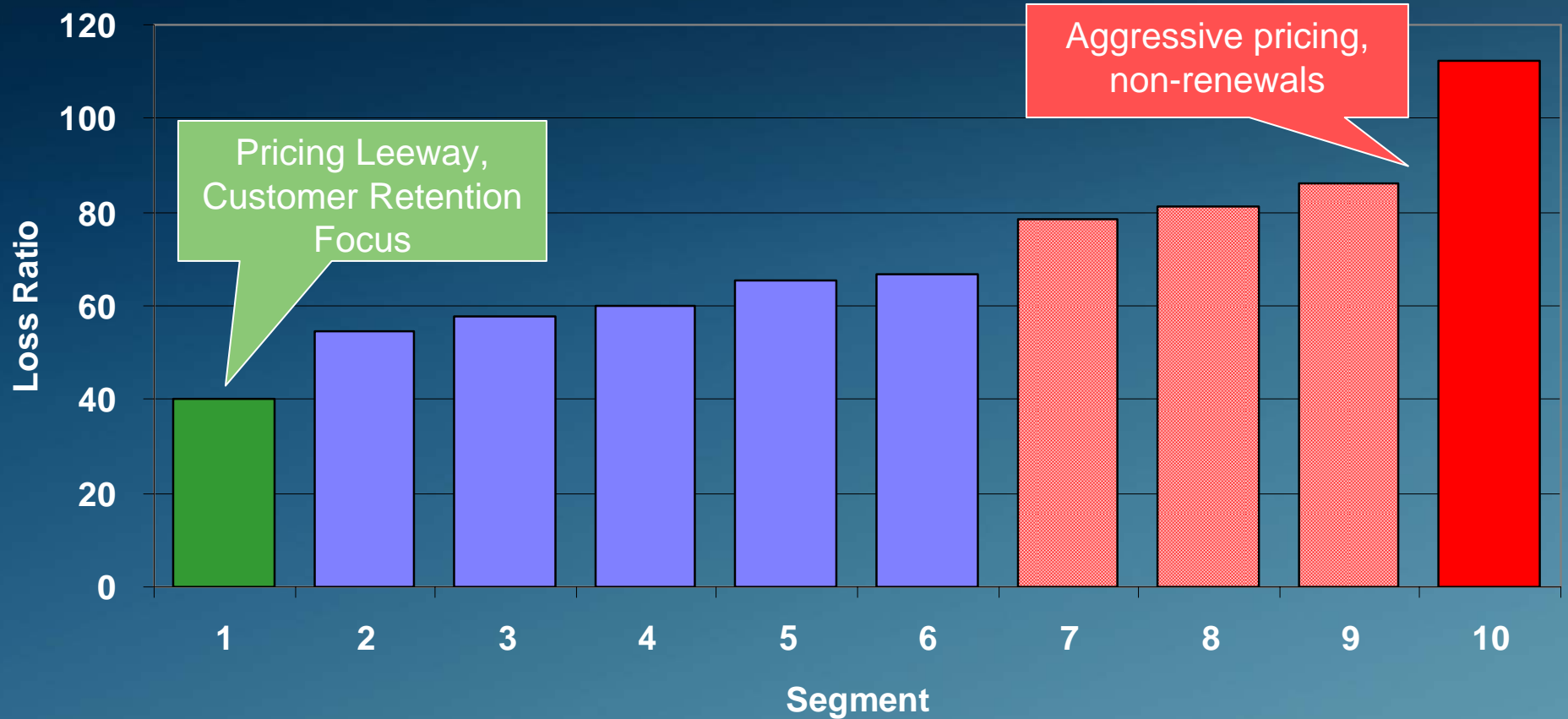
Building Models – Important Concepts

- **Scoring engine is a form of predictive model**

- **Not all scoring models are credit models**
  - Models can be built independent of outside credit data (in fact, this is required in some states)

- **Why build a scoring model after data mining**

  - Going for further lift

  - Ease of application … compare a score to a set of rules, speeding up and enforcing consistency of decision making

Milliman

Traditional Application

- **<u>Scoring engine</u>**: formula that segments the policy base by relative profitability
- An equation *f*( ) of several predictive variables
  - Both linear and non-linear flavors
- The result is a continuous range of scores
  - *score = f(X$_1$, X$_2$, …, X$_N$)*

Milliman

Elements of a Scoring Engine $f(X_1, X_2, \ldots, X_N)$

- **The scoring function or algorithm**
  - Regression/GLM
  - Decision Trees/MARS
  - Neural Networks
  - Test several

- **The inputs $X_1, X_2 .. X_N$: as important as the algorithm**
  - A large part of the process consists of variable creation and selection … the data mining steps
  - Not unusual to generate hundreds of variables
 - Examine correlations / eliminate redundant, weak, poorly distributed variables

Milliman

Testing the Results

- ## Important to divide data (randomly) into 3 pieces

  - **Training** data: to fit the model

  - **Test** data: create a lift curve (an iterative process until you are happy with the resulting model)

  - **Validation** data: to produce a final lift curve, giving an unbiased estimate of future performance

Lift Curve Example

- ## Lift Curve

  - Lift = highest / lowest segment (usually deciles for models)

 - Lift = segmentation power of your model

Milliman
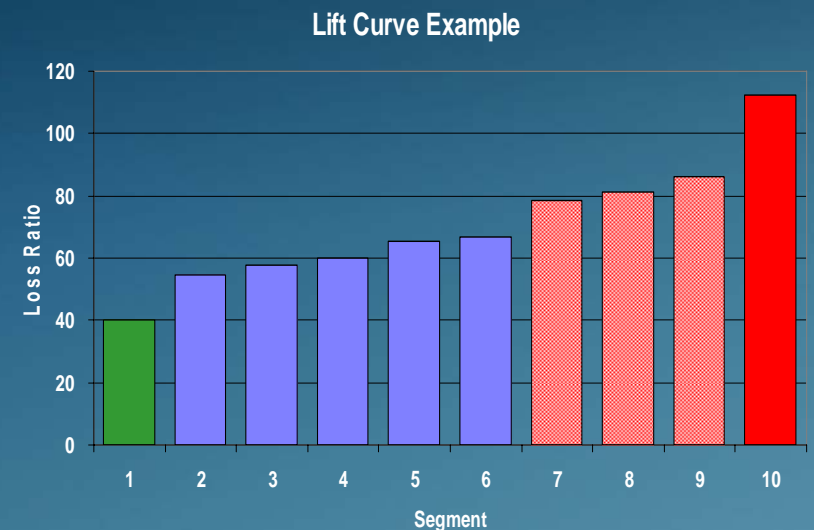
Summary of the Predictive Model Process

- Organize your data

- Data mining

  – Intense process to identify the candidate variables to go into your predictive model

  – Pare variables down to a manageable set

- Modeling building

  – Try several different models / algorithms

  – Test different combinations of variables

- Test the model

  – Lift curves

  – Hold back data for testing

  – Key step in getting "buy-in" from your users

- Monitor performance

  – Distribution of scores over time, predictive performance, usage of model...

  – Plot model maintenance / update strategy

Milliman

Incorporating Competitive Pricing Data

- So far … our discussion has focused on modeling claim propensities, usually measured as loss ratios or pure premiums

- But … focusing on profitability is only half on the Portfolio Management mantra of "profitable growth"

- Modeling had to evolve … the first surrogate for growth was competitive position

- The next step in the evolution of modeling has been to combine these claim propensity indications with competitive indications

Milliman

Profit / Competitiveness Analysis
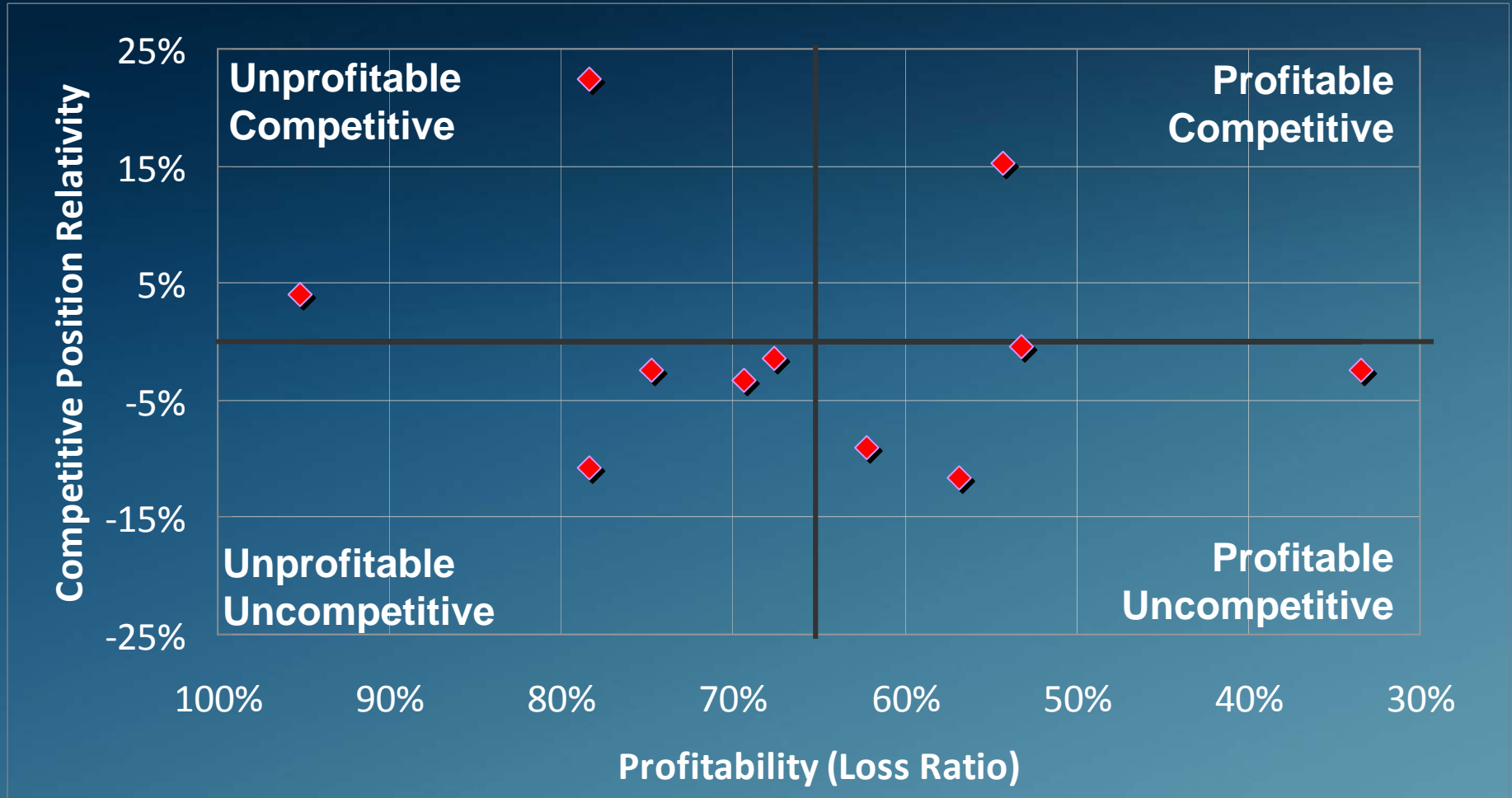
Achieving Profitable Growth

- Comparative rate analysis alone is no guarantee for achieving profitable growth
- Another tool is needed to know what competitive position to establish for different segments in order to achieve desired growth rates
  - Price optimization

Milliman

Some Evidence That We've Entered a New Stage

Ad spending by P/C insurers is at a record high

Year	$ Billions
99	$1.74
00	$1.74
01	$1.80
02	$1.71
03	$1.88
04	$2.11
05	$2.98
06	$3.70
07E	$4.32

Source: Insurance Information Institute from consolidated P/C Annual Statement data.

Milliman

Why The Increase in Advertising Spending?

- Back to Those Savings Comparisons
- Liberty Mutual – $328 a year
- AIG Auto – $366 a year
    - Switch from State Farm - $331
    - Switch from GEICO - $375
    - Switch from Allstate - $443
- Esurance – $550 a year
- GEICO – over $500 a year
- Hartford AARP – $303 a year

Which Company Do You Want To Be ??

Milliman

Which Company Gets Your Business?

Savings	Company (description)
$500	Brand X (unbranded)
$425	"I Heard of Them, Don't Know Much About Them" Company
$350	Gold Standard, A+ rated
$300	Top Brand AND endorsed by your favorite affinity group
---	I'm not switching

Milliman

Why Might Esurance Make Branding a Priority?

- Still new … aggressive growth plans
- But … imagine if they could attract customers with only $450 annual savings (instead of $550)
  - More growth at the current price level
  - Vs. more profit (at a higher price) for the same growth level

- Hmm … sounds like an optimization challenge

Milliman

Predictive Modeling - Next Steps in Pricing

For more information, see the CAS website for presentations from past Predictive Modeling seminars

www.casact.org/education/specsem/f2008/handouts/
www.casact.org/education/specsem/f2007/handouts/

Milliman