The Minimum Bias Procedure— A Practitioner's Guide

Sholom Feldblum, FCAS, FSA, MAAA and
J. Eric Brosius, FCAS, MAAA

The Minimum Bias Procedure

A Practitioner's Guide

prepared by

Sholom Feldblum
and
Dr J Eric Brosius

[January 2002]

{The authors are indebted to Robert A. Bailey, LeRoy J. Simon, Thomas G. Myers, and Dr. Ernesto Schirmacher for review of and corrections to the final draft of this document. Dr. Schirmacher, in particular, made extensive corrections to both the substance and the style of this *Practitioner's Guide*. Any remaining errors are the authors' own and should not be attributed to Mrssrs. Bailey, Simon, or Myers.}

The Minimum Bias Procedure

This *Practitioner's Guide* is geared to the practicing actuary who would like to optimize classification relativities. It provides the intuition underlying the minimum bias procedure and the alternative methods that have been proposed subsequently. It uses a simple illustration to show the computations required for each method, and to evaluate their advantages and drawbacks.

All the procedures discussed here can be easily coded in modern spreadsheets using the built-in functions provided by vendors. Practicing actuaries should be able to quickly implement minimum bias procedures, and the intuition here should enable students to readily master the methods.

BACKGROUND

The minimum bias procedure was first introduced in a 1960 *Proceedings* paper by Robert Bailey and LeRoy Simon, "Two Studies in Automobile Insurance." Bailey and Simon examined models with two types of arithmetic functions (multiplicative and additive), two types of bias functions (balance principle and χ-squared), and two data types (loss costs and loss ratios).[1]

Bailey and Simon used their procedure (i) to judge the merits of an additive versus a multiplicative classification model for Canadian private passenger automobile business and (ii) to choose optimal rate relativities.[2] The 1960 Bailey and Simon paper discusses the rationale for the minimum bias procedure, the characteristics of a suitable rating model, and the rating scenarios that fit the various types of models.

The 1960 Bailey and Simon paper concluded that: (i) the additive model fits the Canada private passenger automobile data better than the multiplicative model, and (ii) the χ-squared function is the optimal bias function. The first conclusion was based on a goodness-of-fit test; the second conclusion was based on the credibility assigned by the χ-squared function.

In a 1963 *Proceedings* paper, "Insurance Rates with Minimum Bias," Robert Bailey summarized the minimum bias theory, outlining the considerations that support the use of the balance principle as the bias function and explaining when loss ratios serve better than loss

---

1 References to the *Proceedings* are to the *Proceedings of the Casualty Actuarial Society.*

2 The minimum bias procedure deals with loss cost relativities, which we refer to here as classification relativities. In practices, actuaries determine rate relativities. The two types of relativities may differ if expenses are not a fixed percentage of premiums. These issues are discussed in a subsequent section of this *Guide.*

costs. Bailey's 1963 *Proceedings* paper was on the CAS examination syllabus for many years, serving as a teaching text for a generation of actuaries.

In a 1988 *Proceedings* paper, "Minimum Bias with Generalized Linear Models", Robert Brown expanded the minimum bias method to use two additional types of bias functions. Brown retained the balance principle and χ-squared functions from the Bailey and Simon papers. He added a least squares function, which is similar to the χ-squared function, and a maximum likelihood function, which assumes certain distributions of claim frequency or claim severity in the insured population.

Brown also examined generalized linear models (GLM), which have potential statistical advantages and may accomplish the same objectives as the minimum bias procedures. He did not find that the generalized linear models produced more accurate results. For the Canadian private passenger automobile business, Brown found the multiplicative model superior to the additive model.

In 1990, Gary Venter published a thorough discussion of Brown's paper, in which he introduced several extensions of the existing procedures, such as a combined additive and multiplicative model, along with an analysis of credibility consideration and other modeling issues. Brown's *Proceedings* paper, along with Venter's discussion, replaced the 1963 Bailey paper on the actuarial syllabus in the mid-1990's.

Bailey's 1963 paper, Brown's paper, and Venter's discussion have proved difficult for practicing actuaries to understand and for actuarial candidates to master. These authors wrote for experienced actuaries who were familiar with the ratemaking issues and proficient with the mathematical methods.

Dr J. Eric Brosius and Sholom Feldblum have taught the minimum bias procedure to several hundred actuarial candidates since 1995. They have developed heuristic illustrations and intuitive explanations that clarify the theory.

This paper is a practitioner's guide to the minimum bias procedure. It combines the theory of the original actuarial papers with the teaching material prepared by Brosius and Feldblum. It explains the rationale for the procedure and it shows its applications. It teaches the method to new actuaries and it gives them the background to read the original *Proceedings* papers.

The title of this paper is the "Minimum Bias Procedure," since that name is now common in the U.S. actuarial profession. The subject of this paper should more properly be described as the development of multi-dimensional classification systems. This subject is broad. The paper covers part of this subject, of which one component is the minimum bias procedure and the alternative methods discussed here.

This *Practitioner's Guide* does not cover generalized linear models. Generalized linear models are commonly used in the United Kingdom and in continental Europe for multi-dimensional classification ratemaking. We treat generalized linear models in a companion *Practitioner's Guide.*

THE PRACTITIONER'S GUIDE

Practicing actuaries are unique professionals. Their goal is to manage business endeavors, not simply to provide statistical advice. Yet their expertise rests on a large body of theoretical knowledge, not just on experience. The role of this *Practitioner's Guide* is to transform theoretical knowledge into practical business situations.

The pure actuary concentrates on fundamental theory, confident that sound theory will find its way into multiple applications. Many actuarial papers are written from this perspective.

The practicing actuary, in contrast, begins with the business problem and works backward to find theoretical solutions. Similarly, this Practitioner's Guide begins with the business need for multi-dimensional classification relativity systems. It unveils the intuition underlying the statistical methods and examines their suitability for the business scenarios.

THE WORLD BEFORE BAILEY AND SIMON

Before Bailey and Simon introduced the minimum bias procedure, classification relativities were determined one dimension at a time. The appendices to Charles McClenahan's "Ratemaking" chapter and Robert Finger's "Risk Classification" chapter in the CAS *Foundations of Casualty Actuarial Science* textbook illustrate the procedure. This remains the dominant classification ratemaking system for many lines of business in the United States. In other countries, such as Great Britain, actuaries have made more use of generalized linear models (GLM) to develop classification relativities; see the companion paper on GLM.

If a line of business has a one-dimensional classification system, the minimum bias procedure adds nothing to the traditional calculation. Workers' compensation, for example, uses industry as the only dimension in the classification system. Insurers are now examining other classification dimensions for workers' compensation; the minimum bias procedure and generalized linear models may prove valuable in this analysis.

The minimum bias procedure is useful when the classification system has two or more dimensions. Throughout this paper, we use examples of two dimensions. The extension to three or more dimensions is straight-forward, but the arithmetic becomes cumbersome and it is more difficult to format the arrays on a two-dimensional page. These problems are eliminated by spreadsheet implementations of the procedure.

There are several reasons for using a procedure which looks at all dimensions of the classification system simultaneously; we provide the intuition further below. The primary statistical issue relates to the optimal bias function. The 1960 Bailey and Simon paper emphasizes the credibility argument, from which the authors concluded that the χ-squared function is the optimal bias function. The 1963 Bailey paper emphasizes the bias argument, from which Bailey inferred that the balance principle is the optimal bias function. Some statisticians argue that the maximum likelihood function is inherently superior to the other bias functions; this method is discussed in Brown's paper.

We define the terms, explain the statistical procedures, and review the intuition underlying each method further below. It is hard for some readers to grasp the intuition until they have a working knowledge of the methods. We provide the explanations alongside a series of heuristic illustrations.

We are setting pure premiums. We do not deal with expenses by classification or with gross premiums.[3] We base the pure premiums upon the empirical observations in each cell of an

---

[3] For expense loadings by classification and development of gross premiums in competitive markets, see S. Feldblum [1996: PAP].

array. For a two dimensional classification system, this means each cell in a matrix. The observations can be average loss costs (i.e., pure premiums), loss frequencies, or loss ratios. In practice, the data would consist of losses and exposures (for loss costs), claim counts and exposures (for loss frequencies), or losses and premiums (for loss ratios).

ILLUSTRATIONS

A series of illustrations forms the backbone of this *Practitioner's Guide*. The basic illustration has two dimensions with two values in each dimension. This prevents the intuition from getting submerged under tedious mathematics. The illustrations are constructed so that they are not conceptually different from real scenarios. For practical work, the minimum bias procedure is most important for multi-dimensional classification systems that have multiple entries in each dimension.

Most of the illustrations show only one iteration. (The meaning of an "iteration" is provided below.) In practice, multiple iterations are needed for convergence, since the procedures do not have closed form solutions. The work would be tedious were it done by hand. With current spreadsheet applications, the work is elementary. Some spreadsheets have built-in iterative functions, such as "goal-seek" and "solver" in Excel. Some software packages, such as SAS, have built-in routines for generalized linear models. Once the intuition is clear, the required programming is not difficult.

Illustration: A classification system for private passenger automobile insurance has two dimensions: (i) urban vs rural and (ii) male vs female. A company insures exactly four drivers, one in each cell, with the following observed loss costs:[4]

	Urban	Rural
Male	\$800	\$400
Female	\$400	\$200

We seek to determine pure premium relativities.[5] We compare all males with all females, or \$1,200 for two exposures compared to \$600 for two exposures. This gives a pure premium relativity of

Male to Female = 2 to 1

We do the same for urban versus rural, and we get the same relativity. We choose "rural female" as the base class, and we get the following set of relativities:

Male:	2.00	=	s_1
Female:	1.00	=	s_2
Urban:	2.00	=	t_1
Rural:	1.00	=	t_2

An urban male driver has an indicated pure premium of the base rate times the urban relativity times the male relativity, or \$200 × 2.00 × 2.00 = \$800. More generally, the pure premium in cell (i,j) is \$200 × s_i × t_j.

[4] We defer for the moment our discussion of credibility. If the observed loss costs were fully credible – that is, if they were fully accurate and unbiased estimators of the expected pure premiums within each cell of the matrix – we would use the observed loss costs as the pure premiums, and we could dispense with the classification ratemaking problem.

[5] We sometimes refer to the pure premium relativities as rate relativities, since this is the more common actuarial term. For the conversion of pure premium relativities into rate relativities, see Feldblum [1996: PAP, pages 231-237].

In this illustration, the indicated pure premiums match the observed loss costs. The minimum bias method is not needed for this case.[6]

Assumptions

Two assumptions underlie this analysis.

1. Instead of using the observed loss costs as the indicated pure premiums, we convert them into a system of classification relativities. There are both practical and theoretical reasons for using classification relativities.

   a. *Practical:* In many lines of business, there are several classification dimensions with numerous classes in each. Using separate rates for each cell of the array is unwieldy.
   b. *Theoretical:* Using relativities improves the accuracy of the rate indications, since we use all the information regarding each cell's expected pure premium. This reason comes under the general rubric of credibility considerations.

   The practical reason (reason "a") was once compelling, though the development of on-line premium quoting system has reduced its importance. If a cell has low volume, credibility considerations justify basing future rates more heavily on classification relativities than on observed data in that cell. We discuss this further below.

2. We assumed that the relativities model is multiplicative. A multiplicative model means that the relativity for a given cell is the product of the relativities in its row and column. Robert Finger [1976] justifies this assumption (in another context) by stating that several independent factors interact multiplicatively to determine the liability claim size. For automobile insurance, these factors include

   - the speed of the vehicles before impact
   - the health of the injured party
   - the protection (e.g., with seat belts, interior padding) of the victim
   - the income of the victim
   - the skill of the plaintiff's attorney, and
   - the skill of the defendant's claims adjusters.

   This reasoning would be more persuasive if these factors were also the classification dimensions. They are not; instead, the classification dimensions are driver characteristics, garaging territory, and use of the vehicle.

---

[6] The McClenahan and Finger chapters of the *Foundations of Casualty Actuarial Science* textbook uses this procedure to determine rate relativities.

The multiplicative model implicitly assumes that the classification dimensions are independent. It is less appropriate when classification dimensions are correlated. The pricing actuary must determine – both empirically and logically – what model is best. The minimum bias procedure aids this determination.

THE ADDITIVE MODEL

The indicated pure premiums may differ from the observed loss costs for two reasons:

- The model structure may be incorrect.
- Random loss fluctuations influence the observed loss costs.

We treat the first reason, the model structure, in this section. We are determining pure premium relativities, but the observed loss costs are as shown below.

	Urban	Rural
Male	\$700	\$500
Female	\$400	\$200

We begin in the same fashion as we did before. We compare all males to all females, giving a pure premium relativity of \$1,200 to \$600, or 2 to 1. We compare all urban to all rural, giving a pure premium relativity of \$1,100 to \$700, or 1.571 to 1. We choose rural females as the base class.

The indicated pure premium relativities no longer match the observed loss costs. The indicated pure premium for rural males is \$200 × 2.000 = \$400, but the observed loss costs are \$500. The indicated pure premium for urban females is \$200 × 1.571 = \$314, but the observed loss costs are \$400. The differences are significant.

No multiplicative factors work perfectly. In urban territories, the relationship of male to female is \$700 to \$400, or 1.75 to 1. In rural territories, the relationship of male to female is \$500 to \$200, or 2.50 to 1. A male to female relativity appropriate for the urban territories is not optimal for the rural territories.

Similarly, for male drivers, the urban to rural relativity is \$700 to \$500, or 1.4 to 1. For female drivers, the urban to rural relativity is \$400 to \$200, or 2 to 1. An urban to rural relativity appropriate for male drivers is not optimal for female drivers.

The discussion in the paragraphs above assumes that the rating model is multiplicative. In this illustration, an additive is more appropriate. We add or subtract a dollar amount for each class instead of multiplying by a factor.

Let the base class be rural females, with a base rate of $200, and let the relativities be as shown below:

- Male: +$300
- Female: +$0
- Urban: +$200
- Rural: +$0

The rate for any cell is the base rate plus the male/female relativity plus the territory relativity.

The indicated pure premiums now match the observed loss costs. Rural male = $200 + $300 + $0 = $500. Urban male = $200 + $300 + $200 = $700.

The additive method provides an exact match because

1. The dollar difference between males and females is the same for the rural column as for the urban column. In both columns, the dollar difference is $300.
2. The dollar difference between urban and rural is the same for the male row as for the female row. In both rows, the dollar difference is $200.

ADDITIVE INTUITION

Some casualty actuaries implicitly assume that pure premium relativities should be multiplicative, not additive. Current multi-dimensional classification systems for the casualty lines of business generally use multiplicative factors.

Regulators sometimes castigate insurers for using multiplicative factors that "unduly" increase the rates for high-risk insureds. Some actuaries assume that this criticism is purely political, not actuarial. This is often true, but it is not always correct. When two or more dimensions of the classification system are correlated, multiplicative systems may be biased. For some types of insurance, multiplicative systems may be biased even when classification dimensions are not correlated.

Life insurance rating systems provide an example. If smokers have twice the mortality of non-smokers, and persons with high-blood pressure have twice the mortality of persons with average blood pressure, should high-blood pressure smokers have four times the mortality of average blood pressure non-smokers? Life insurance underwriters employ judgment to assess the rating for applicants with multiple causes of high mortality. A pure multiplicative rating system would not be appropriate.

We discuss multiplicative and additive models throughout this *Practitioner's Guide.* It is useful to understand the circumstances that justify the use of each type of system.

We use the illustration presented in the 1960 Bailey and Simon paper. We have two rating dimensions: (i) class of driver and (ii) merit rating class.

1. Class of driver refers to the driver characteristics, such as age, sex, and marital status, as well as use of the vehicle, such as pleasure use or business use.

2. Merit rating class refers to the number of immediately preceding accident free years, ranging from 0 to 3.

These two rating dimensions are partially correlated. Young, unmarried male drivers have a high average class relativity. Because these drivers either are new drivers or (if not new) are more likely to have had an accident in the past year, they have relatively few accident free years.

Mature female drivers have a low class relativity. Because they are more experienced drivers with fewer past accidents, they also have (on average) merit rating class credits. The two rating dimensions are not independent.

To choose between a multiplicative model and an additive model, we first find an optimal model of each type. We use a minimum bias procedure to select the optimal multiplicative factors for the multiplicative model and the optimal additive factors for the additive model. We then compare the goodness-of-fit of the indicated pure premiums from each model to the observed loss costs. Both Rob Brown and Bailey and Simon did this analysis for the Canadian merit rating plan factors. As Venter notes in his discussion of Brown's paper, some private passenger automobile insurance carriers used a combined additive and multiplicative model; see below.

This teaching guide does not advocate any particular rating method. Once readers are comfortable with the procedures described here, they should be well equipped to optimize the classification relativities for any scenario.

BIAS FUNCTIONS

In practice, the indicated pure premiums do not perfectly match the observed loss costs for either an additive model or a multiplicative model. We illustrate with the same 2 by 2 classification system. The observed loss costs are shown in the table below:

	Urban	Rural
Male	$800	$500
Female	$400	$200

Neither an additive model nor a multiplicative model provides a perfect match. If we use a model that does not perfectly match the observed data, we must determine how to minimize the mismatch between the observed loss costs and the indicated pure premiums. A "bias function" is a means of comparing two or more models to see which fits the data with the smallest degree of mismatch.[7] To choose the optimal model, we proceed along three steps:

1. We choose a rating method, such as an additive model, a multiplicative model, or a combined model.
2. We select a bias function to optimize the rating method. This *Practitioner's Guide* discusses the balance principle, least squares, χ^2, and maximum likelihood bias functions.
3. For each optimized rating method, we examine the goodness-of-fit of the indicated pure premiums with the observed loss costs.

For models using a maximum likelihood bias function, we must also choose a probability distribution function for losses within each cell.

We begin with the balance principle, since it is the bias function most commonly used in practice. The 1963 Bailey paper provides a compelling justification for the balance principle.[8]

---

[7] The "bias function" is not a standard statistical term, and the balance principle is not a standard statistical principle. As used in this *Practitioner's Guide*, the bias function is the means of determining how "close" the indicated pure premiums are to the observed loss costs or how great the "mismatch" is between these two sets of data. The sum of the squared deviations and the χ-squared deviation are common statistical bias functions. The balance principle, which was introduced by Bailey and Simon in 1960 and then strongly endorsed by Bailey in 1963, minimizes the bias along the dimensions of the classification system, thereby leading to the term "minimum bias."

[8] In contrast, Brown [1988] and Mildenhall [1999] argue that the standard statistical functions, such as least squares, χ-squared, and maximum likelihood, along with generalized linear models, should be considered in place of the balance principle.

THE BALANCE PRINCIPLE

The balance principle means that

the sum of the *indicated pure premiums* = the sum of the *observed loss costs*

along *every row* and *every column*.

We examine the balance principle for both the additive and the multiplicative models in our simplified illustration. On the left are the observed loss costs; on the right are the indicated pure premiums. We begin with the multiplicative model.[9]

	Urban	*Rural*		*terr₁*	*terr₂*
Male	\$800	\$500	sex_1	$200 \times s_1 \times t_1$	$200 \times s_1 \times t_2$
Female	\$400	\$200	sex_2	$200 \times s_2 \times t_1$	$200 \times s_2 \times t_2$

To balance along the first row (the "male" row), we must have

$$800 + 500 = 200 \times s_1 \times t_1 + 200 \times s_1 \times t_2$$

To balance along the second row (the "female" row), we must have

$$400 + 200 = 200 \times s_2 \times t_1 + 200 \times s_2 \times t_2$$

To balance along the first column (the "urban" column), we must have

$$800 + 400 = 200 \times s_1 \times t_1 + 200 \times s_2 \times t_1$$

To balance along the second column (the "rural" column), we must have

$$500 + 200 = 200 \times s_1 \times t_2 + 200 \times s_2 \times t_2$$

We have two rows and two columns, for a total of four equations. We have four variables, so we can solve the equations.

---

9 To keep the notation simple, we use rating dimensions of male vs female and urban vs rural throughout this *Practitioner's Guide*. For the formulas in the illustrations, we use $sex_1 = s_1$ = male, $sex_2 = s_2$ = female, $terr_1 = t_1$ = urban, and $terr_2 = t_2$ = rural. The recursive equations use variable names of *x*, *y*, and *z*, and rating dimensions of *i* and *j*.

Although we have four equations in four unknowns, we do not have a unique solution. There are two special considerations we must be aware of. These two considerations offset each other in such a way as to yield a unique set of indicated pure premiums.

1. Dependence among the equations: These equations are related by a totality constraint. Using any three of these equations we can derive the fourth, since the sum of the rows equals the sum of the columns. For instance, the fourth equation equals the first equation plus the second equation minus the third equation.

More generally, the equation for any column equals the sum of the equations for the rows minus the sum of the equations for the other columns. The equation for any row equals the sum of the equations for the columns minus the sum of the equations for the other rows.

2. Invariance under reciprocal scalar multiplication: We can set one of the variables arbitrarily, and we can still solve the system of equations. To see this most clearly, suppose that we have solved these equations for values of the four variables s_1, s_2, t_1, and t_2. Another solution is $2s_1$, $2s_2$, $\frac{1}{2}t_1$, and $\frac{1}{2}t_2$. We could use any constant in place of "2." No matter which set of relativities we pick, the values in the cells remain the same. The values in the cells are the product of an "s" relativity and a "t" relativity, so the additional constant cancels out.

We have an additional variable. The pure premium in each cell depends on the "base rate" (actually, the "base pure premium"). If the relativities s_1, s_2, t_1, and t_2 optimize the rating model for a base pure premium of $200, the relativities $2s_1$, $2s_2$, t_1, and t_2 optimize the rating model for a base pure premium of $100.[10]

The minimum bias procedure optimizes the relationship of the rating variables along each dimension of the classification system. If $s_1 = 2 \times s_2$ for a given base rate and a given set of territorial relativities, then $s_1 = 2 \times s_2$ for any other base rate and territorial relativities.

By convention, we choose a base class in each classification dimension. This is often the largest class, though any class may be used. The base class in each classification dimension is given a relativity of unity. This determines the values of all other rating variables as well as the value of the base rate.

[10] With so much leeway in choosing the classification relativities, one might ask what we are "optimizing." We are optimizing the indicated pure premiums. Each set of classification relativities give the same indicated pure premiums. The optimization is relative to the bias function. The optimal pure premiums have the least bias, the least sum of squared deviations, the least χ-squared value, or the greatest likelihood.

Solving the Equations

The equations are *not* linear, so there is no closed form solution. We begin with an arbitrary (but reasonable) choice of relativities for one dimension, and we solve the equations iteratively.[11]

Illustration: We choose a set of relativities for urban and rural. Suppose we choose 2.00 for urban and 1.00 for rural.

	Urban	*Rural*		$terr_1 = 2$	$terr_2 = 1$
Male	\$800	\$500	sex_1	$200 \times s_1 \times 2$	$200 \times s_1 \times 1$
Female	\$400	\$200	sex_2	$200 \times s_2 \times 2$	$200 \times s_2 \times 1$

The balance equation for the first row (the "male" row) says that

$$800 + 500 = 200 \times s_1 \times 2 + 200 \times s_1 \times 1$$

This give us: $1{,}300 = 600 \times s_1$, or $s_1 = 13/6$.

Balancing along the second row (the "female" row) gives

$$400 + 200 = 200 \times s_2 \times 2 + 200 \times s_2 \times 1,$$
$$\text{or } s_2 = 600/600 = 1.$$

We now have intermediate values for the male and female relativities of 13/6 and 1. We discard the initial values for the urban and rural relativities of 2.00 and 1.00, and we solve for new intermediate values by balancing along the columns.

The balance equation for the first column (the "urban" column) says that

$$800 + 400 = 200 \times (13/6) \times t_1 + 200 \times 1 \times t_1$$

This implies that $1{,}200 = 633.33\ t_1$, or $t_1 = 1.895$.

Balancing along the second column (the "rural" column) gives

---

[11] Iterative methods were originally adopted because there are no closed form solutions. They are ideal for spreadsheet applications, which have eliminated the hand calculations.

$$500 + 200 = 200 \times (13/6) \times t_2 + 200 \times 1 \times t_2,$$
$$\text{or } t_2 = 1.105.$$

We continue in this fashion. We discard the previous male and female relativities, and we solve for new ones:

Balancing along the first row (the "male" row) gives

$$800 + 500 = 200 \times s_1 \times 1.895 + 200 \times s_1 \times 1.105$$

and balancing along the second row (the "female" row) gives

$$400 + 200 = 200 \times s_2 \times 1.895 + 200 \times s_2 \times 1.105$$

We solve these two equations for new values of the male and female relativities, we discard the previous values of the urban and rural relativities, and we balance along the columns for new values of the urban and rural relativities.

We continue in this fashion until the relativities converge. Convergence means that the change in the relativities from an additional iteration is not material. Pencil and paper calculation of the minimum bias relativities is tedious, particularly if there are many classes within each dimension. The built-in iterative functions in standard spreadsheet packages eliminate this problem.

In practice, we begin with starting values determined by the simple rate relativities procedure. In this illustration, the urban to rural relativity is 12 to 7. If we choose a pure premium relativity of 1.000 as the starting value for the rural class, we would choose a starting value of $12 \div 7 = 1.714$ for the urban class. The starting values have no effect on the final rates in each cell.

Once the series converges, the common practice is to normalize the base class relativities to unity. We normalize by changing the base rate. For instance, suppose that the series above converged after a single iteration. (It does not actually converge after a single iteration; we simply want to show the normalization technique.)

A. The territorial relativities are 1.895 for urban and 1.105 for rural. If the rural territory is the base class, we change the rural relativity to 1.000, we change the urban relativity to $1.895 \div 1.105 = 1.715$, and we change the base rate to $\$200 \times 1.105 = \$221$.

B. The male/female relativities are 13/6 for males and 1.000 for females. If females are the base class, these relativities are already normalized. If males are the base class, we change the male relativity to 1.000, the female relativity to 6/13, and we multiply the base rate by 13/6.

There are several equivalent formulas for the additive model. The rate in cell x_iy_j, or row "i" and column "j," is

A. Base rate + $x_i + y_j$,
B. Base rate × $(1 + u_i + v_j)$, or
C. Base rate × $(p_i + q_j)$

To see the equivalence of these formulas, suppose the base rate in formula "A" is $10.

- In formula "B," the base rate is also $10, each "u" value is one tenth the corresponding "x" value in formula "A," and each "v" value in formula "B" is one tenth the corresponding "y" value in formula "A": $u_i = 0.1 \times x_i$ and $v_j = 0.1 \times y_j$.
- Formula "C" is equivalent to formula "B," except that either the "p" values are all increased by 1, the "q" values are all increased by 1, or the "p" values are increased by a constant (*c*) and the "q" values are increased by the complement of that constant (*1–c*): $p_i = 1 + u_i$ *or* $q_j = 1 + v_j$ (both not both) or p_i = "c" + u_i *and* q_j = "1–c" + v_j.

We use the first form – formula "A" – for our illustrative example, since it shows the additive method most clearly. In practice, formula "C" might be preferred, since only the base rate need be increased for inflation. In formula "A," the base rate and all the relativities must be increased for inflation. The inflation adjustments would necessitate new rate pages each year.

We choose initial values for the urban and rural relativities: say, $250 and $0.[12] These initial values are based on the simple rate relativities procedure, since the average differential between the urban and rural observed loss costs is $250.[13]

	Urban	*Rural*		$terr_1 = 250$	$terr_2 = 0$
Male	$800	$500	sex_1	$200 + s_1 + 250$	$200 + s_1 + 0$
Female	$400	$200	sex_2	$200 + s_2 + 250$	$200 + s_2 + 0$

Balancing along the first row (the "male" row) gives

[12] The term "relativities" is more appropriate for a multiplicative model, where the relativities are multiplicative factors. An additive model uses dollar relativities that are added to the base rate.

[13] The average differential is ½ × [(800 – 500) + (400 – 200)] = 250.

$$800 + 500 = 200 + s_1 + 250 + 200 + s_1 + 0,$$
$$\text{or } s_1 = 650/2 = 325.$$

Balancing along the second row (the "female" row) gives

$$400 + 200 = 200 + s_2 + 250 + 200 + s_2 + 0,$$
$$\text{or } s_2 = -50/2 = -25.$$

We discard the initial values for the urban and rural relativities, and we balance along the columns, using the intermediate values of the male and female relativities, to get new values for the urban and rural relativities. We continue the iterative process until the series converges.

The negative relativity of –$25 for females seems odd at first. In truth, the relativity for female drivers is not inherently negative; this is an artifact of the base rate and the starting values.

We could make the relativity for females positive by adding a constant to the male and female relativities and subtracting the same constant from the rural and urban relativities. For instance, we could add $75 to the male and female relativities to get relativities of $400 and $50, and we would subtract $75 from the rural and urban relativities.

We can make all the relativities positive or negative by adjusting the base rate. For instance, by choosing a base rate of $1,000, we obtain negative relativities for male drivers, female drivers, rural drivers, and urban drivers. Companies may do this for marketing reasons. All drivers get discounts from the base rate, so all drivers feel they are gaining from the classification system.

Alternatively, we may set a relativity of $0 for the base class in each classification dimension. This determines the base rate and all other relativities.

EXPOSURES

The illustrations assume one driver in each cell or the same number of drivers in each cell. In practice, there may be different numbers of risks in each cell.

Two types of adjustments are needed: an adjustment to the bias function and an adjustment for credibility.

- We adjust the bias function for the relative volume of business in each cell (see below), not for the absolute volume of business.
- We may make a credibility adjustment based on the absolute volume of business in a cell.

Illustration: Insurer A has 100 exposures in each cell; insurer B has 10,000 exposures in each cell. Insurer A may rely more heavily on the minimum bias procedure. Insurer B may give greater weight to the empirical observations.

We deal with the adjustment to the bias function in this section. We defer the adjustment for credibility until later.

We deal with multiple exposures in a cell as though there were multiple cells laid on top of each other. We set the sum of the observed loss costs in each row or column equal to the sum of the indicated pure premiums in the corresponding row or column. If there are two drivers in a given cell, we double both the observed loss costs and the indicated pure premiums in that cell. If there are "n" drivers in a given cell, we multiply both the observed loss costs and the indicated pure premiums in that cell by "n."

In the illustrations above, we used two matrices: a 2 by 2 matrix of observed loss costs and a 2 by 2 matrix of indicated pure premiums. This is sufficient when there is exactly one driver in each cell, or when the number of drivers in each cell is the same. When the number of drivers varies by cell, we need a matrix of the number of drivers in each cell.

For the multiplicative model, suppose that the number of drivers were as follows:

- male urban: 1200
- male rural: 600
- female urban: 1000
- female rural: 800

We include the number of drivers in the matrices:

	Urban	*Rural*		*$terr_1$*	*$terr_2$*
Male	$1200 \times \$800$	$600 \times \$500$	*sex_1*	$1200 \times 200 \times s_1 \times t_1$	$600 \times 200 \times s_1 x t_2$
Female	$1000 \times \$400$	$800 \times \$200$	*sex_2*	$1000 \times 200 \times s_2 \times t_1$	$800 \times 200 \times s_2 x t_2$

We modify the balance equations to include exposures.

To balance along the first row (the "male" row), we must have

$$1200 \times 800 + 600 \times 500 = 1200 \times 200 \times s_1 \times t_1 + 600 \times 200 \times s_1 \times t_2$$

To balance along the second row (the "female" row), we must have

$$1000 \times 400 + 800 \times 200 = 1000 \times 200 \times s_2 \times t_1 + 800 \times 200 \times s_2 \times t_2$$

To balance along the first column (the "urban" column), we must have

$$1200 \times 800 + 1000 \times 400 = 1200 \times 200 \times s_1 \times t_1 + 1000 \times 200 \times s_2 \times t_1$$

To balance along the second column (the "rural" column), we must have

$$600 \times 500 + 800 \times 200 = 600 \times 200 \times s_1 \times t_2 + 800 \times 200 \times s_2 \times t_2$$

This paper proceeds simultaneously along three paths: illustrations, rigor, and intuition. Readers new to this topic should focus on the illustrations. Readers seeking to implement these methods should focus on the rigor. The rigor presents the equations which must be incorporated into a spreadsheet or program as well as the goodness-of-fit tests that indicate which models are superior.

More experienced actuaries should focus on the intuition. The intuition explains why we are using these procedures and it provides the rationale for the methods. The skilled actuary must understand not only the mechanics of a procedure but also what the procedure attempts to accomplish and why a particular procedure is appropriate for a given scenario.

We begin with observed loss costs. One might wonder: *Why can't we use these figures, appropriately developed and trended, as the indicated pure premiums for the coming policy period? Instead of fitting either multiplicative or additive models to the observed data, let us use $800 as the indicated pure premium for urban male drivers, $400 as the indicated pure premium for urban female drivers, $500 as the indicated pure premium for rural male drivers, and $200 as the indicated pure premium for rural female drivers.*

The common answer is that the individual cells are "not fully credible." This answer is correct, but the terminology is not ideal. The term "credible" has a vague connotation. Let us be more precise, so that we grasp the intuition.

Credibility is a relative concept. No cell is inherently credible or not credible. The cell's credibility depends on the value of its own experience versus the information provided by the values in other cells.

Consider the basic illustration with the following observed loss costs:

	Urban	*Rural*
Male	$800	$500
Female	$400	$200

The urban male observed pure premium of $800 is a mixture of expected losses and random loss fluctuations. How might we judge whether it is biased upwards or downwards?

Let us suppose first that the rating values combine additively to generate the expected losses. The urban male observed loss cost of $800 is $300 more than the rural male observed loss cost of $500. This suggests that the urban attribute of the vehicle's location adds about $300 to the expected loss costs.

However, the urban female observed loss cost of $400 is only $200 more than the rural female observed loss cost of $200. This suggests that the extra cost associated with the urban attribute is only $200, not $300. This implies that the urban male loss cost of $800 might be biased upwards.

We perform a similar analysis for male versus female. The urban male observed loss cost of $800 is $400 more than the urban female observed loss cost of $400. This suggests that the male attribute adds about $400 to the expected loss costs.

However, the rural male observed loss cost of $500 is only $300 more than the rural female observed loss cost of $200. This suggests that the extra cost associated with the male attribute is only $300, not $400. In other words, the urban male loss cost of $800 might be biased upwards.

The $800 observed loss cost in the urban male cell does not tell us how much of this observed loss cost is expected and how much is distorted by random loss fluctuations. If we know the mathematical function linking the cells – that is, if the characteristics of the driver and the vehicle have some additive or multiplicative relationship – we can use additional cells to provide information about the true expected costs for urban male drivers. In this illustration, the other cells imply that the $800 observed loss cost might be biased upwards.

If we assume that the cells are linked by a multiplicative relationship, our inferences change. The urban male observed value of $800 is 160% of the rural male observed value of $500. This suggests that the urban attribute of the vehicle's location adds about 60% to the expected loss costs.

The urban female observed loss cost of $400 is twice the rural female observed loss cost of $200. This suggests that the extra cost associated with the urban attribute is +100%, not +60%. The urban male loss cost of $800 might be biased downwards.

We use a similar analysis for male versus female. The urban male observed loss cost of $800 is twice the urban female observed loss cost of $400. This suggests that the male attribute adds about 100% to the expected loss costs.

The rural male observed loss cost of $500 is 250% of the rural female observed loss cost of $200. This suggests that the extra cost associated with the male attribute is +150%, not +100%. The urban male loss cost of $800 might be biased downwards.

If the cells are linked additively, we infer that the urban male observed loss costs of $800 might be biased upwards. If the cells are linked multiplicatively, we infer that the urban male observed loss costs of $800 might be biased downwards.[14]

With a 2 by 2 matrix, there are 4 cells in total. If the exposures are evenly distributed among the cells, each cell contains 25% of the total exposures, whether there is 1 car in cell or 10,000 cars in each cell. We give much credence to the observed value in that cell compared to our inferences from other cells. With a larger array, such as a 10 by 10 by 10 array, there are many more cells. The average cell contains only 0.1% of the total exposures. We give less credence to the observed loss costs in that cell compared to our inferences from other cells.

This is the intuition for classification ratemaking in general and for the minimum bias procedure in particular. The rating model – such as additive, multiplicative, or combined – tells us the type of relationship joining the cells. The bias function – such as balance principle, χ-squared, least squared error, or maximum likelihood – provides a method of drawing inferences from one cell to another.

CREDIBILITY

The original papers on the minimum bias procedure discuss credibility considerations, but they come to differing conclusions. Insurers and rating bureaus using versions of the minimum bias procedure sometimes add credibility enhancements.

We do not discuss credibility in depth until we cover additional bias functions, goodness-of-fit tests, and classification models. We introduce here the general credibility issues.

The 1960 Bailey and Simon paper uses credibility considerations to pick a bias function. The authors' view of the appropriate credibility for each cell led them to choose the χ-squared bias function over the balance principle.

The 1963 Bailey paper, which advocates the balance principle, has no explicit discussion of credibility. The balance principle has an implicit credibility component. The credibility of a cell is proportional to the exposures in that cell. If we double the exposures in a cell, we double the credibility of that cell.

We conceive of a cell with multiple exposures as a stack of cells each containing a single exposure. These cells are all at the same location in the array, but there is nothing special joining these cells into a group. Instead of one cell with two exposures and an average loss

[14] In practice, the direction of the bias rarely depends on the type of rating model. The more common scenario might show an observed loss cost of $600, an additive model indicated pure premium of $550, and a multiplicative model indicated pure premium of $530. We might infer that the observed loss costs are biased upwards by random loss fluctuations.

cost of \$500, we conceive of two cells with one exposure in each. The observed loss costs in the two cells have an average value of \$500.

Venter looks at credibility from a different angle. We said above that the \$800 observed loss cost for urban male drivers might be overstated for an additive model or understated for a multiplicative model. The over- or understatement stems from random loss fluctuations. If there is a single exposure in each cell, an overstatement or understatement is likely. If there are 10,000 exposures in each cell, the degree of overstatement or understatement is likely to be smaller.

If there is single exposure in each cell, we might attribute the\$800 observed loss cost to random loss fluctuations. If there are 10,000 exposures in the urban male cell, we might attribute the observed loss cost to the special hazards of being male and living in a city. The number of exposures in the cell may suggest how much credence to give to the observed value in that cell versus the inferences from other cells.

ITERATIVE FORMULAS

We have presented simple illustrations and intuitive explanations. To program this procedure on a spreadsheet or in source code, we need general formulas.

We derive the iterative formulas for the multiplicative balance principle model. For the balance principle, we balance along the rows and the columns until we achieve convergence. Convergence means that an additional iteration does not materially change the relativities. With modern spreadsheets, the speed of convergence is not a concern. If the insurance premium is expressed in whole dollars, we might define convergence such that the indicated pure premiums stay the same to the nearest dollar.

We designate the number of exposures in row i and column j by n_{ij}, and the observed pure premiums in row i and column j by r_{ij}. The balancing equations for the multiplicative model are

$$\sum_j (n_{ij} r_{ij}) = \sum_j (n_{ij} x_i y_j)$$

In this equation, x is a row relativity and y is a column relativity.[15] The equation assumes a base rate of unity. We solve this equation for x_i to get

[15] In the illustrations, we use s for the row relativity and t for the column relativity as acronyms for the classification dimensions (sex and territory). The variables x and y are commonly used in the literature.

$$x_i = \frac{\sum_j n_{ij} r_{ij}}{\sum_j n_{ij} y_j}$$

We sum over the *j* subscript when we balance along the rows (the *i* subscripts). We do this separately for each *i*. When we balance along the columns, we sum over the *i* subscripts and we do this separately for each *j*. When the series converges, we set the relativity for the base class in each classification dimension to unity, and we adjust the base rate to offset this.

For readers who wish to proceed to the original actuarial papers, this formula is the fifth expression on page 10 in the 1963 Bailey paper. In Brown's 1988 paper, this is formula (1.1) on page 189, formula (2.3) on page 192, and formula (3.3) on page 195. It is also formula (5.20) on page 201, which is Brown's multiplicative Poisson model with maximum likelihood estimation. This is also the final equation on page 44 of the 1999 Holler, Sommer, and Trahair paper (*CAS Forum*, Winter 1999), which follows Brown's derivation.

We used two dimensions in this formula. One might assume that the two dimensions correspond to the two variables *x* and *y*. That is not correct. The two dimensions correspond to the two subscripts *i* and *j*. The *x* and *y* variables correspond to two sets of relativities. A dimension can have two or even more sets of relativities.

Illustration: The classification system has two dimensions: male vs female and territory A vs territory B. Territory A has more attorneys than territory B has, resulting in a higher propensity to sue and higher loss costs. Territory B has several blind intersections, leading to more accidents. We might presume that the higher attorney involvement in territory A increases the cost of all claims, so a multiplicative factor is appropriate, whereas the blind intersections in territory B adds additional hazards, so an additive factor is appropriate. The rating model might take the form

$$\textit{indicated pure premium} = x_i \times y_j + z_j,$$

where *i* represents the male/female classification dimension and *j* represents the territory dimension. To optimize this rating model, the balance principle is not sufficient; we would have to employ one of the other bias functions.[16]

[16] The balance principle provides *i+j* equations, but we have *i+2j* variables. The other bias functions discussed in this *Practitioner's Guide* provide *i+2j* equations.

The arithmetic is similar for any number of dimensions. The multiplicative model has one set

$$x_i = \frac{\sum_{j,k} n_{ijk} r_{ijk}}{\sum_{j,k} n_{ijk} y_j z_k}$$

of relativities for each dimension. With three dimensions, the iterative formula is

Similarly, we develop the general formula for the balance principle additive model by assuming a base rate of \$0. The balance principle equation is

$$\sum_j (n_{ij} r_{ij}) = \sum_j n_{ij}(x_i + y_j)$$

and the iterative formula is

$$x_i = \frac{\sum_j n_{ij}(r_{ij} - y_j)}{\sum_j n_{ij}}$$

MULTIPLICATIVE MODEL

Illustration: A multiplicative rating model with two dimensions and two classes in each dimension. The observed loss costs and exposures in each class are shown below. We use the balance principle to optimize the pure premium relativities.

Using a base rate of 100 and starting values of 1.00 for y_1 and 1.500 for y_2, we compute the first iterative values of y_1 and y_2.

Loss Costs		y_1	y_2
	x_1	300	300
	x_2	200	400

Exposures		y_1	y_2
	x_1	100	150
	x_2	100	100

Iterations

We are given starting values for y_1 and y_2. To compute the first iterative values of y_1 and y_2, we must first compute the intermediate values for x_1 and x_2.

- From the starting values for y_1 and y_2, we compute values for x_1 and x_2.
- From the computed values of x_1 and x_2, we compute new values for y_1 and y_2.

Since the base pure premium is \$100, the indicated pure premiums are \$100 $\times x_i \times y_j$. To simplify the mathematics, we compute all values in units of \$100. The indicated pure premiums are $x_i \times y_j$, and the observed loss costs are \$3, \$3, \$2, and \$4.

We form a matrix of observed loss costs and indicated pure premiums:

	y_1	y_2		y_1	y_2
x_1	3	3	x_1	$x_1 \times y_1$	$x_1 \times y_2$
x_2	2	4	x_2	$x_2 \times y_1$	$x_2 \times y_2$

We multiply each of the figures by the number of exposures in the cell:

	y_1	y_2		y_1	y_2
x_1	100×3	150×3	x_1	$100 \times x_1 \times y_1$	$150 \times x_1 \times y_2$
x_2	100×2	100×4	x_2	$100 \times x_2 \times y_1$	$100 \times x_2 \times y_2$

The starting values for y_1 and y_2 are 1.00 and 1.50. We use the balance principle to obtain values for x_1 and x_2:

$$100 \times 3 + 150 \times 3 = 100 \times x_1 \times 1.00 + 150 \times x_1 \times 1.5,$$
$$\text{or } 300 + 450 = 100 \times x_1 + 225 \times x_1,$$
$$\text{or } x_1 = 2.308.$$

and

$$100 \times 2 + 100 \times 4 = 100 \times x_2 \times 1.00 + 100 \times x_2 \times 1.5,$$
$$\text{or } 200 + 400 = 100 \times x_2 + 150 \times x_2,$$
$$\text{or } x_2 = 2.400.$$

We now have intermediate values for x_1 and x_2. We discard the initial values for y_1 and y_2, and we balance along the columns.

$$100 \times 3 + 100 \times 2 = 100 \times 2.308 \times y_1 + 100 \times 2.400 \times y_1,$$
$$\text{or } 300 + 200 = 230.8 \times y_1 + 240 \times y_1,$$
$$\text{or } y_1 = 1.062.$$

and

$$150 \times 3 + 100 \times 4 = 150 \times 2.308 \times y_2 + 100 \times 2.400 \times y_2,$$
$$\text{or } 450 + 400 = 346.2 \times y_2 + 240 \times y_2,$$
$$\text{or } y_2 = 1.450.$$

This completes the mathematics. We comment on several items in this exercise.

Data and Assumptions

The number of exposures in each cell may be viewed as a credibility item. We give 50% more credence to the observed loss costs in the x_1/y_2 cell. Intuitively, we believe the relationships involving cell x_1/y_2 more than we believe other relationships.

- The observed loss costs in the x_1 row indicate that there is no difference between y_1 and y_2. The observed loss costs in the x_2 row indicate that the y_2 class should have twice the pure premium as the y_1 class. We give more credence to the first of these two relationships.

- The observed loss costs in the y_1 column indicates that the x_2 class should have a pure premium 33% *lower* than the x_1 class. The observed loss costs in the y_2 column indicates that the x_2 class should have a pure premium 33% *higher* than the x_1 class. We give more credence to the second of these two relationships.

Since there are only four cells in the array, the indicated pure premiums should be close to the observed loss costs. If the relationships implied by the various rows and columns were consistent, the number of exposures in each cell would have little effect on the indicated pure premiums.[17] In this exercise, the relationships implied by the different rows and columns are not consistent. The higher exposures in the x_1/y_2 cell tilts the indications toward the relationships involving that cell.

We see this in the computed values for x_1 and x_2. If the exposures were the same in each cell, we would have no reason to rate x_1 differently from x_2: the y_1 column suggests that the x_2 class should have a 33% lower pure premium and the y_2 column indicates that the x_2 class should have a 33% higher pure premium. Since we give more credence to the second of these two relationships, the x_2 relativity is slightly higher than the x_1 relativity.

Rating Model

In practice, the pricing actuary is optimizing the pure premium relativities within a given rating model, such as the multiplicative model in this exercise. In theory, the pricing actuary might use the minimum bias procedure to determine the optimal rating model.

The general rule is that a multiplicative model is indicated when the observed loss costs are more dispersed, and an additive model is indicated when the observed loss costs are less dispersed.[18] More precisely, a multiplicative model is indicated when the high rated classifications stem from multiple high relativities. An additive model is indicated when the combination of high relativities does not result in very high observed loss costs.

Sometimes the type of dispersion is evident, such as high pure premiums for young unmarried urban male drivers and low pure premiums for mature suburban female drivers. In this exercise, the degree of dispersion is not evident.

---

[17] By "consistent," we mean that the relationship among the column relativities are the same in each row.

[18] One can phrase this in statistical terms as a multiplicative model works better when the coefficient of variation is high and an additive model works better when the coefficient of variation is low. The generalizations are intended to guide practicing actuaries in choosing a model. In truth, the actuary must test the goodness-of-fit of each model to see which works best.

ADDITIVE MODEL

An additive model with two dimensions has the observed loss costs shown below. Each cell has 1000 exposures. The base lost cost is 100. The formula for loss costs by cell is Loss $Cost_{ij}$ = (Base Loss Cost) x ($x_i + y_i$). We use the starting values shown below to compute intermediate values for y_1 and y_2.

- Loss Costs:

	y_1	y_2
x_1	500	750
x_2	250	475
x_3	150	400

- Starting Values:

x_1	4.500
x_2	3.000
x_3	2.000

Computations

The number of exposures is the same in each cell. To simplify the computations, we may assume that there is a single exposure in each cell, since the factor of "1,000" cancels out of every equation.

The base rate is $100. To simplify the mathematics, we use units of $100 and a base pure premium of unity. The indicated pure premiums are $x_i + y_i$, and the observed loss costs are divided by 100.

The matrix of observed loss costs and indicated pure premiums is shown below:

	Observed Values			Indicated Values	
	y_1	y_2		y_1	y_2
x_1	5	7.5	x_1	$x_1 + y_1$	$x_1 + y_2$
x_2	2.5	4.75	x_2	$x_2 + y_1$	$x_2 + y_2$
x_3	1.5	4	x_3	$x_3 + y_1$	$x_3 + y_2$

We balance along the columns. For the first column, we have

$$5.00 + 2.50 + 1.50 = (x_1 + y_1) + (x_2 + y_1) + (x_3 + y_1)$$

We substitute the starting values of the "x"s to get

$$5.00 + 2.50 + 1.50 = (4.50 + y_1) + (3.00 + y_1) + (2.00 + y_1),$$
$$\text{or } 3\,y_1 = 9.00 - 9.50, \text{ or } y_1 = -0.167.$$

For the second column, we have

$$7.50 + 4.75 + 4.00 = (x_1 + y_2) + (x_2 + y_2) + (x_3 + y_2)$$

We substitute the starting values of the "x"s to get

$$7.50 + 4.75 + 4.00 = (4.50 + y_2) + (3.00 + y_2) + (2.00 + y_2),$$
$$\text{or } 3\,y_2 = 16.25 - 9.50, \text{ or } y_2 = 2.25.$$

We have finished balancing along the columns. The next step is to balance along the rows. We take the new *y* values, $y_1 = -0.167$ and $y_2 = +2.25$, and we compute new values for x_1 and x_2 by balancing along each row. We continue this process – alternately balancing along rows and columns – until the new values at the end of an iteration do not differ significantly from the old values at the beginning of that iteration along both dimensions.

Balance

It is helpful to see the convergence of the iterative equations.

- During the iterative process – before convergence – the plan is alternately balanced along the rows or along the columns, but not along both.
- Once the series converges, the plan is balanced along both the rows and the columns.

We have just balanced along the columns. To see that we are not yet balanced along the rows, we examine the first row:

$$5.00 + 7.50 = (x_1 + y_1) + (x_1 + y_2).$$

Substituting the starting values of the "x"s and the first iterative values of the "y"s, we get

$$12.50 = 4.50 + (-0.167) + 4.50 + 2.25 = 11.083.$$

The equality does not hold, since the plan is not yet balanced. Since we are still far from convergence, balancing along the columns distorts any balance along the rows.

Other Classification Dimensions

The basic illustrations use the minimum bias procedure to set pure premium relativities simultaneously for the male/female dimension and the urban/rural dimension. There may be other dimensions to the classification plan as well, such as age of driver, marital status, type of vehicle, use of the car, driver education, prior accident history, and so forth.

How do these other dimensions affect our analysis?

Ideally, we would use a multi-dimensional minimum bias procedure to set all classification relativities simultaneously. In practice, this may not be possible. Some rate relativities may be set on a statewide basis, whereas other rate relativities may be set on a countrywide basis. Some rate relativities may be analyzed each year, whereas other rate relativities may be analyzed every several years.

Territorial analyses must be done on a state by state basis. Certain driver characteristics and vehicle characteristics may be analyzed on a countrywide basis, for two reasons:

1. The relativities are not expected to vary by state, as long as the states use the same insurance compensation system.[19] The effects of type of vehicle, such as sedan, SUV, or sports car, on insurance loss costs should not vary by state. The same is true for the age, sex, and marital status of the driver.

2. Some classification cells would have few exposures in a state analysis, and the results may be distorted by random loss fluctuations. The countrywide analysis uses more data, providing more credible results. For example, we may wish to analyze driver age in yearly increments: age 17, age 18, age 19, and so forth. Single state data may be too sparse to give credible results.

Some classification dimensions, such as driver education, have a minor effect on overall loss costs. We may analyze these classification dimensions every five years or so, not every year.

Suppose that we analyze the male/female dimension and the urban/rural dimension on a statewide basis, while relativities for other classification dimensions are set on a countrywide basis. We use a minimum bias method for the statewide analysis.

---

[19] The countrywide analysis may actually be done on all tort liability states or all no-fault states. The type of vehicle may have different effects, depending on the compensation system. The bodily injury rate relativities may be higher for SUV's (sports utility vehicles) than for sedans in tort liability states. The reverse may be true in no-fault states.

If all the classification dimensions are independent, the analysis should work well. If one or more of the other classification dimensions is correlated with the male/female or urban/rural dimensions, the rating analysis may be distorted.

Illustration: Suppose that young people migrate to urban areas, for university education, work opportunities, and the glamor of urban social activities. Older people move to the suburbs and rural areas, to buy homes and raise families away from the vices of urban areas. The age of the driver is correlated with the urban/rural garaging location.

The statewide analysis may indicate an urban to rural relativity of 2 to 1. The countrywide analysis may indicate a relativity for young unmarried male drivers of 3 to 1 when compared to adult drivers.

The relativity for young unmarried urban male drivers is not 6 to 1, even if a multiplicative model is appropriate for automobile insurance. Many of the young unmarried male drivers in the countrywide analysis live in urban areas, and many of the urban drivers are young and unmarried.

The optimal solution is to use a complete multi-dimensional statewide analysis, including age and marital status of the driver. This is not always practicable, since there are too many possible inter-relationships. For instance, the effect of vehicle type on bodily injury loss costs would probably be analyzed only on a countrywide basis, since there are too many vehicle types to give credible results on a statewide basis. But suburban married women with children are more likely to drive SUV's, and young unmarried urban male drivers are more likely to drive sports cars.

Loss Ratios

The common solution is to use loss ratios instead of loss costs in the minimum bias procedure. More precisely, we use loss ratios adjusted to the base rates for the classification dimensions included in the minimum bias analysis.

We use the basic illustration to develop the intuition. The first illustrations clarify the concepts. We then use a more involved setting to show the application of the procedure for private passenger automobile rating.

Suppose the empirical experience consists of loss ratios by classification, not loss costs. We observe the following loss ratios for four drivers for a multiplicative model:

	Urban	*Rural*
Male	75%	85%
Female	90%	80%

We could take either of two approaches:

First Approach: We treat the unadjusted loss ratios as though they were loss costs. Instead of using pure premium relativities, we use loss ratio relativities. These relativities are in addition to whatever pure premium relativities are embedded in these loss ratios.

In this scenario, the minimum bias procedure will indicate about equal loss ratio relativities for urban vs rural and slightly higher loss ratio relativities for females than for males. This does not mean that urban risks are similar to rural risks, or that female drivers have more accidents than male drivers have. If the current rate relativities are reasonable, we would expect the loss ratios in all cells to be about equal. In this scenario, the current male to female rate relativity might be 2.4 to 1. Since the average female loss ratio of 85% is higher than the average male loss ratio of 80%, the loss ratio relativities would indicate that we should slightly reduce the male to female rate relativity.

Second Approach: We convert the raw loss ratios to base class loss ratios. Suppose the current rate relativities are 2.4 to 1 for male to female and 1.8 to 1 for urban to rural. We must divide the male premiums by 2.4 and the urban premiums by 1.8. This is equivalent to multiplying the male loss ratios by 2.4 and the urban loss ratios by 1.8. In sum, we multiply the raw loss ratios by the current classification relativities, as shown in the table below.

	Urban	*Rural*
Male	75% × 2.4 × 1.8 = 324%	85% × 2.4 × 1.0 = 204%
Female	90% × 1.0 × 1.8 = 162%	80% × 1.0 × 1.0 = 80.0%

We apply the minimum bias procedure to the adjusted loss ratios. The resulting loss ratio relativities would be the same as the indicated rate relativities.

To see this, suppose that the base rate is $100. We determine the observed loss costs in each cell:

- For the male/urban cell, the premium is $100 × 2.4 × 1.8 = $432. The observed loss ratio is 75%, so the loss cost is 75% × $432 = $324.
- For the male/rural cell, the premium is $100 × 2.4 × 1.0 = $240. The observed loss ratio is 85%, so the loss cost is 85% × $240 = $204.

- For the female/urban cell, the premium is $100 × 1.0 × 1.8 = $180. The observed loss ratio is 90%, so the loss cost is 90% × $180 = $162.
- For the female/rural cell, the premium is $100 × 1.0 × 1.0 = $100. The observed loss ratio is 80%, so the loss cost is 80% × $100 = $80.

As we mentioned above, the common practice is to set the rate relativity to unity for the base class in each classification dimension. To facilitate this procedure, we divide each adjusted loss ratio in the matrix by the adjusted loss ratio for the base class. The resulting loss ratios in this illustration are shown below.

	Urban	Rural
Male	324% / 80% = 405.0%	204% / 80% = 255.0%
Female	162% / 80% = 202.5%	80.0% / 80% = 100.0%

LOSS RATIO INTUITION

We have shown how to convert loss ratios to reflect the loss costs in each cell. This might be useful if the observed data were loss ratios and we wanted to use loss costs for the minimum bias procedure. But the observed data are loss costs, not loss ratios. We must first convert the observed loss costs to loss ratios before converting back to loss costs.

The purpose of this conversion from loss costs to loss ratios and then back to loss costs is to eliminate the potentially distorting effects of classification dimensions that are not being analyzed in the minimum bias procedure.

We explain by illustration. We have average observed bodily injury loss costs for four groups of drivers, with 1000 drivers in each cell.

	Urban	Rural
Male	$800	$500
Female	$400	$200

There are other dimensions in the classification system.

Type of Vehicle: For bodily injury, cars are subdivided into various groups. SUV's (sports utility vehicles) and similar vehicles, such as light trucks, are larger and sturdier. They provide better protection for their occupants, but they cause greater damage to others. Sedans and small cars cause less damage to others. Sedans and small cars are more common in urban areas; SUV's and light trucks are more common in rural areas.

The distribution of vehicles types between urban and rural areas, along with the appropriate surcharge or discount for each type of vehicle, affects the observed loss costs. The pricing actuary may not have this distribution for the state under review. This is not necessary; the use of loss ratios instead of loss costs corrects for the effects on vehicle type.

It is hard to follow abstract intuition. To keep the mathematics clear, we assume that there are only two types of vehicles – SUV's and sedans – and that SUV's receive a +20% surcharge for bodily injury coverage. In this state, SUV's comprise 40% of the rural vehicles and 10% of the urban vehicles.

Age of Driver: The male/female rate relativity applies to all male and female drivers. Unmarried male drivers under the age of 21 receive additional surcharges, ranging from +25% for 20 year old drivers to +125% for 16 year old drivers. There is no corresponding surcharge for unmarried female drivers under the age of 21.[20] These surcharges are determined from a countrywide analysis.

The pricing actuary performing the minimum bias analysis does not have a distribution of male drivers by age and marital status. This is not necessary; the loss ratios are sufficient. To clarify the mathematics, however, we assume that 10% of male drivers are unmarried and under the age of 21. The average surcharge for these drivers is +50%.

DOUBLE COUNTING AND OFFSETTING

If we do not take vehicle type and driver age into account, we will overcharge male drivers and rural drivers.

Male Drivers: The male/female relativity is determined from the statewide analysis. The surcharges for young unmarried male drivers is determined from a separate countrywide analysis. The poor driving experience of young unmarried male drivers is counted twice: once at the countrywide level for the surcharges and once at the statewide level for the male/female

---

[20] For many years, it was unclear why young males are such hazardous drivers. Maturity or temperament were often given as vague explanations. For life insurance, the higher mortality rates of males is assumed to be physiological, but there seemed to be no similar relationship between male drivers and auto accidents.

Advances in biological science have removed much of the puzzle. Young males have high levels of testosterone, which often leads to aggressive, risk-taking behavior. Much of the adventurous and often dangerous activities of young males, which previous generations ascribed to social acculturation, may have the same biological roots as facial hair and deep voices. Auto insurance rating by testosterone level may be far too intrusive; rating by age, sex, and marital status is a more acceptable proxy.

Lest readers misunderstand our comments, we note that young female drivers have higher loss costs than adult female drivers, but the differences are much smaller than they are for male drivers. Driving experience, maturity, and temperament have an effect, though they account for a relatively small portion of young unmarried male driver loss costs.

relativity. To accurately determine the male/female relativities, we must remove the hazardous effects of being young and unmarried from the male driver classification.

Rural Drivers: Rural drivers are less hazardous than urban drivers, but they drive more dangerous vehicles. The vehicle surcharge is determined in the countrywide analysis. To properly determine the urban/rural relativities, we must remove the effects of vehicle type from the statewide experience.

Removing these effects is not easy. It might seem reasonable to include additional dimensions in the statewide minimum bias analysis. We could include vehicle type as a third rating dimension and avoid its potential distorting effects on the minimum bias procedure.

For two reasons, this is not a practical or a reasonable solution.

A. Credibility: The experience for certain types of vehicles may be sparse in the statewide data. Rate relativities determined from sparse data may reflect random loss fluctuations.
B. Consistency: The relativity for a given vehicle type should be the same in all states. If separate statewide analyses are done, the relativities by type of vehicle will vary by state.

The countrywide analysis gives more credible relativities that are consistent from state to state.

To remove the effects of vehicle type and age of driver from the statewide analysis, we assume that the countrywide relativities are accurate. We examine each risk in the minimum bias procedure. We divide the actual loss costs by the type of vehicle relativity and by the driver age relativity. This gives the loss costs that we would have expected to see were the vehicle evenly distributed over all other rating dimensions.

Illustration: Suppose a four door sedan is the base vehicle type and age 21+ is the base age. A two door compact has a bodily injury discount of 10% and an unmarried 20 year old male driver has a surcharge of 25%.

Suppose the observed loss costs for a 20 year old unmarried male driver of a two door compact car are \$450. The loss costs adjusted for age and vehicle type equal

$$\$450 / (0.90 \times 1.25) = \$400.$$

It is not practicable to make these adjustments car by car. There may be several classification dimensions that might distort the statewide analysis.

Using loss ratios adjusts for all classification dimensions simultaneously. Using observed loss ratios instead of observed loss costs adjusts for driver age, driver sex, territory, vehicle types, and all other rating dimensions. We add back in the current rating relativities for classification dimensions that we are analyzing, or male/female and urban/rural in this illustration.

We show the effects for this illustration. The calculations below are heuristic. We need not perform them for the minimum bias analysis. They reveal the intuition underlying the use of loss ratios instead of loss costs.

Illustration: The average observed loss costs for the 1000 drivers in each of four classes are displayed above. The current relativities are 2.4 to 1 for male to female and 1.8 to 1 for urban to rural. The average SUV to sedan relativity is 1.2 to 1. SUV's comprise 40% of rural cars and 10% of urban cars. Unmarried males under the age of 21 comprise 10% of male drivers, and their average surcharge is +50%. We convert the observed loss costs to adjusted loss costs for the minimum bias analysis.

Let us suppose that the pure premium for a female driving a sedan in a rural territory is $200. The choice of the base rate does not affect the results, since the same multiplicative factor affects all four cells. We work out the premium in each cell.

- Rural/female: We adjust for vehicle type with a multiplicative factor of 1 + 20% × 40% = 1.08. The average pure premium is $200 × 1.08 = $216.00.

- Urban/female: We adjust for vehicle type with a multiplicative factor of 1 + 20% × 10% = 1.02. The average pure premium is $400 × 1.02 = $408.00.

- Rural/male: We adjust for vehicle type with a multiplicative factor of 1 + 20% × 40% = 1.08 and for driver age with a multiplicative factor of 1 + 10% × 50% = 1.05. The average pure premium is $500 × 1.08 × 1.05 = $567.00.

- Urban/male: We adjust for vehicle type with a multiplicative factor of 1 + 20% × 10% = 1.02 and for driver age with a multiplicative factor of 1 + 10% × 50% = 1.05. The average pure premium is $800 × 1.02 × 1.05 = $856.80.

We divide the average loss costs in each cell by the average pure premium in that cell to get the observed net loss ratios in the cell.[21]

---

[21] These loss ratios are net of expenses; they are losses divided by pure premiums. (These net loss ratios are sometimes called "experience ratios.") In practice, we have gross premiums, not pure premiums, so we use traditional loss ratios, not net loss ratios. The traditional loss ratios are slightly distorted by expense flattening procedures. Loss costs show pure premium relativities, whereas traditional loss ratios show rate relativities. In many cases, the distortion is not material. When the potential distortion is material, offsetting adjustments must be made. These adjustments depend on the expense flattening procedure; a full explanation would take us too far afield. On rate relativities versus pure premium relativities, see Feldblum [1996: PAP].

	Urban	Rural
Male	\$800 / \$856.80 = 93.37%	\$500 / \$567.00 = 88.18%
Female	\$400 / \$408.00 = 98.04%	\$200 / \$216.00 = 92.59%

We have removed the effects of all classification dimensions from the observed loss costs. We multiply by the current pure premium relativities for male/female and urban/rural to restore these effects to the observed data.

	Urban	Rural
Male	93.37% x 2.4 x 1.8 = 403.36%	88.18% x 2.4 x 1.0 = 211.64%
Female	98.04% x 1.0 x 1.8 = 176.47%	92.59% x 1.0 x 1.0 = 92.59%

The necessary adjustments are now done. If we want the base class in each dimension to have a relativity of unity, we divide by the base class adjusted loss ratio to get relative loss ratios, as shown below.

	Urban	Rural
Male	403.36% / 92.59% = 435.63%	211.64% / 92.59% = 228.57%
Female	176.47% / 92.59% = 190.59%	92.59% / 92.59% = 100.00%

If we wish, we can convert the relative loss ratios to adjusted loss costs by multiplying the cells by the base rate.

	Urban	Rural
Male	435.63% x \$200 = \$871.26	228.57% x \$200 = \$457.14
Female	190.59% x \$200 = \$381.18	100.00% x \$200 = \$200

LOSS RATIO APPROACH

A classification system with two classes in each of two dimensions shows the observed data in the four matrices below.

- The relativities matrix shows the current pure premium relativities for male to female and urban to rural.
- The loss costs matrix shows the average loss costs per driver.
- The exposures matrix shows the number of cars in each cell. The company writes predominantly rural business.
- The premium matrix shows the premium collected in each cell.

The base pure premium is$66.67 for rural females.

Relativities	Urban	Rural
Male	4	2
Female	2	1

Loss Costs	Urban	Rural
Male	180	120
Female	100	40

Exposures	Urban	Rural
Male	100	1000
Female	100	1000

Earned Premium	Urban	Rural
Male	25000	125000
Female	13333	66667

We use the loss ratio approach to adjust the observed loss costs for the effects of other rating factors, and we calculate the first iteration for the minimum bias analysis, using the balance principle and a multiplicative model. We start with initial relativities of 1.5 for urban and 0.75 for rural.

Intuition

The use of loss ratios instead of pure premiums adjusts for potential distortions resulting from an uneven mix of business by classification. Let us review the rationale for the adjustments. We explain why we don't just use the given loss costs for the minimum bias procedure.

The loss cost approach (or the pure premium approach) implicitly assumes that the distribution of all other classification dimensions is homogeneous across class and driving record. When the distribution is not even, the loss ratio approach corrects the problem, as long as relativities for the unanalyzed classification dimensions are accurate.

We reason through the exercise. The matrix of relativities shows that the relativity for male drivers is twice the relativity for female drivers The loss costs per car in the first column of the loss costs matrix for male drivers are 1.8 times those for female drivers.

If we had only the matrix of observed loss costs, we would presume that the relativities should be adjusted. The male to female relativity should be 1.8 to 1, not 2 to 1.

The exposures and premium matrices show the error in this reasoning. Two cells in the first column (urban) have the same number of exposures: 100 for each cell. If there were an even distribution along other classification dimensions, then the premium for urban males would be twice the premium for urban females. The actual premiums are 25,000 and 13,333, for a ratio of 1.875 to 1.

To aid the intuition, let us assume that the only other classification dimension is vehicle type. We expect premium in the ratio of 2 to 1 for urban males compared to urban females. We actually have premium in the ratio of 1.875 to 1. That means that more females than males are driving high rated vehicles. If we evened out the distribution among territories, then the ratio of average loss costs between these males and females would increase from 1.8 to something greater.

Let us consider how much greater. Our first impulse is to say that the observed loss costs should be 2 to 1, not 1.8 to 1, since the pure premium relativities are 2 to 1, not 1.8 to 1. This is not correct, for two reasons.

- The pure premium relativities are the current relativities, not the indicated relativities. We do not know if these relativities are correct. Perhaps they were correct several years ago, but they are no longer correct now. Perhaps they were never exact, but they were chosen as round numbers.
- Even if the relativities are correct, random loss fluctuations may distort the observed loss costs. From these data alone, we can not determine if the relativities are correct and the

observed loss costs are distorted by random loss fluctuations or the observed loss costs are correct and the current relativities are not accurate.

We must adjust the observed loss costs from the experience data, not from our current relativities. The adjustment factor is based on the relativities, the exposures, and the premium. If the 100 urban males drove the same vehicles as the 100 urban females, their expected loss costs would increase by 2.000/1.875. The observed loss cost relativity would be 1.800 × (2.000 ÷ 1.875) = 1.920.

To clarify the intuition, let us revise the figures in the problem. Suppose that we change the premium for urban females to $25,000. Urban males and urban females both have premiums of $25,000, both have exposures of 100, yet males have a relativity twice that of females.

Let us further suppose that all urban males drive sedans and all urban females drive SUV's. There are no other classification dimensions besides sex of driver, territory, and vehicle type.

If sedans and SUV's had the same rates, the premium for urban males would be twice the premium for urban females. Since the urban males and the urban females have the same premium, we infer that the premium rate for SUV's is twice the premium rate for sedans.

If we assume that the rates for vehicle type are indicative of expected loss costs, part of the observed loss cost relativity between urban males and urban females stems from the different types of cars that these two groups drive. If all the urban females exchanged their SUV's for sedans, their observed loss costs should drop in half. The observed loss cost relativity between urban males and urban females would be 3.6, not 1.8.

Similarly, if all the urban males exchanged their sedans for SUV's, their observed loss costs should double. The observed loss cost relativity between urban males and urban females would be 3.6, not 1.8. This is what we mean when we say: "What would be the observed loss cost relativity if the effects of other classification dimensions were eliminated?"

We restate this as follows. The premium for urban females ($13,333) is higher than expected ($12,500) compared to the premium for urban males based on the number of exposures and the male/female relativity. This implies that the urban females are driving more hazardous vehicles than the urban males are driving. More generally, this implies that the group of urban female drivers have other characteristics that are raising their premiums and loss costs relative to those of the urban male drivers.

The current male to female relativity is 2.00. In the formula

$$1.80 \times (2.00 \div 1.875) = 1.92$$

the rate relativity for the classification dimension which we are reviewing appears in the numerator. The denominator is the rate relativity of the two groups (urban males and urban females), based on the premium ratio divided by the exposures ratio.

Let us revise the figures again to clarify the intuition for the denominator. Suppose that the male to female rate relativity were 1 to 1, the exposures in the two cells were equal, but the premiums were $25,000 and $13,333. Suppose also that the only other classification dimension is vehicle type, and that all the urban males drive sedans and all the urban females drive SUV's.

If sedans and SUV's had the same premium rate, then the premiums for urban males and for urban females would be the same. Since the premium for urban males is 1.875 × the premium for urban females, we infer that the premium rate for sedans is 1.875 × the premium rate for SUV's.

Part of the observed loss cost difference between urban males and urban females stems from the different vehicle types. If all the urban females drove sedans, their premiums and their expected losses would increase by a factor of 1.875. The ratio of observed loss costs for urban males compared with urban females would be reduced by a factor of 1.875.

The general procedure is straightforward. We remove the effect of *all* classification dimensions from the array of observed loss costs by converting it into an array of loss ratios. In our intuitive reasoning, this is accomplished by dividing by 1.875, since

loss ratios = average loss costs × exposures / premiums, or
loss ratios = average loss costs / (premiums / exposures)

The average effect of all rate relativities in a cell is proportional to the premium in that cell divided by the exposures in that cell. This is equivalent to saying that the premium for any car is the base rate times the product of the rate relativities for each classification dimension. To remove the effects of all dimensions from the array of loss costs, we divide by the ratio of premiums to exposures; that is, we multiply by the ratio of exposures to premium.

We restore the effects of the classification dimensions which we are reviewing. In our intuitive reasoning, this is accomplished by multiplying by the male to female relativity of 2.000.

Each cell contains loss costs × exposures ÷ premiums. These are the loss ratios in the cells. We now multiply by the array of rate relativities for the classification dimensions that we are analyzing. This restores the effect of these classification dimensions.

This leaves us with the loss costs relativities that would be observed were there complete homogeneity in the distribution of exposures of the other classification dimensions.

An additional simplification is to form relative loss ratios to the base class and then multiply by the base rate. This additional step is not essential. In a multiplicative system, multiplying all figures by a constant does not affect the final rate relativities.

Let us keep an eye on this relationship. We want to revise the matrix so that the loss cost relationship is not 1.80 to 1 but $1.80 \times 2.00 \div 1.875 = 1.92$ to 1. The same is true for all other relationships in the matrix of loss costs. Since we have many relationships, we use a systematic method for adjusting them. We use the following sequence of computations.

Form loss ratios: We form loss ratios, using the (i) loss costs per exposure, (ii) the exposures in each cell, and (iii) the earned premium in each cell, as

$$\text{loss ratio} = \text{loss costs per exposure} \times \text{exposure} \div \text{premium}.$$

This gives us a matrix of loss ratios, as shown below:

Absolute Loss Ratios:	*Urban*	Rural
Male	0.72	0.96
Female	0.75	0.6

We want to form a matrix of adjusted loss costs. These are the expected loss costs were there no other classification dimensions affecting the experience. We multiply by the base rate of the base class, which is the class of rural females.

It is easiest to do this with relative loss ratios (not absolute loss ratios), with the relativity for the base class being 1.000. We divide the matrix of loss ratios by the base class loss ratio of 60% to get the matrix of relative loss ratios.

Relative Loss Ratios:	*Urban*	Rural
Male	1.2	1.6
Female	1.25	1

We want the expected loss costs in each cell. We multiply the relative loss ratios by the classification differentials for each cell to get the relative loss costs by cell. These are relative loss costs by cell, not absolute loss costs by cell.

We are given the current differentials in the relativities matrix.

Current Relativities:	*Urban*	Rural
Male	4	2
Female	2	1

We multiply the relative loss ratios by the current class relativities to get the relativities implied by the observed data. These relativities are like the observed loss costs in the earlier illustrations in this *Practitioner's Guide.*

Observed Relativities	*Urban*	Rural
Male	4.8	3.2
Female	2.5	1

Let us check our work. $4.80 \div 2.50 = 1.92$, which equals $1.80 \times 2.00 \div 1.875$. We have now adjusted the figures so that we can perform the minimum bias calculation.

We can proceed in either of two fashions.

(i) We can multiply by the pure premium for the base class and proceed with the pure premium approach to the minimum bias procedure.

(ii) We can do the minimum bias analysis, and then multiply by the base class pure premium.

The two methods are mathematically equivalent. We use the latter method here.

We set up the matrices of observations and of indicated relativities, as shown below.

	Observed Relativities			*Indicated Relativities*	
	Urban (t_1)	Rural (t_2)		Urban (t_1)	Rural (t_2)
s_1	4.8	3.2	s_1	$s_1 \times t_1$	$s_1 \times t_2$
s_2	2.5	1	s_2	$s_2 \times t_1$	$s_2 \times t_2$

The initial values of urban (t_1) and of rural (t_2) are 1.5 and 0.75. We balance along the rows, and we multiply each cell by the corresponding number of exposures:

$$100 \times 4.80 + 1{,}000 \times 3.20 = 100 \times s_1 \times 1.5 + 1{,}000 \times s_1 \times 0.75,$$
$$\text{or } s_1 = 3680 / 900 = 4.089.$$

$$100 \times 2.50 + 1{,}000 \times 1.00 = 100 \times s_2 \times 1.5 + 1{,}000 \times s_2 \times 0.75,$$
$$\text{or } s_2 = 1250 / 900 = 1.389.$$

We use these relativities for male and female (s_1 and s_2), and we discard the starting values for urban/rural (t_1 and t_2). We balance along the columns.

$$100 \times 4.80 + 100 \times 2.50 = 100 \times t_1 \times 4.089 + 100 \times t_1 \times 1.389,$$
$$\text{or } t_1 = 730 / 547.8 = 1.333.$$

$$1000 \times 3.20 + 1000 \times 1.00 = 1000 \times t_2 \times 4.089 + 1000 \times t_2 \times 1.389,$$
$$\text{or } t_2 = 4200 / 5478 = 0.767.$$

These are the pure premium relativities after one iteration. If expenses are proportional to premiums, they are also the rate relativities.[22] To convert to actual rates, we multiply by the base rate of $66.67.

New Rates	Urban	Rural
Male	$363.39	$209.09
Female	$123.44	$71.03

[22] If all expenses are proportional to the gross premium, as is true for state premium taxes, the expenses are also proportional to the pure premium, and the pure premium relativities equal the rate relativities.

Loss Ratio Intuition

The loss ratio method is commonly used by practicing actuaries. The mathematics is not complex, but the intuition sometimes seems elusive We show another illustration, where we focus on the intuition, not the arithmetic.

The previous illustration showed the loss costs and exposures in each cell. We needed only the incurred losses and earned premiums by cell to determine the loss ratios. We used the loss costs by cell only to show the intuition of the method. In practice, the practicing actuary may have only incurred losses and earned premiums by cell, but not exposure counts by cell. For the exposition in this section, we speak of loss ratios and rate relativities, not loss costs.

We are using loss ratios for their theoretical benefits, though the practical benefits may be equally great. Accurate exposure counts by cell are not always available to the pricing actuary. When we lack exposure counts by cell, we can not use a straight-forward loss costs method, even if the distribution of insureds by other rating dimensions is even.

We focus here on transforming the data into loss cost relativities by cell. Once we have the loss cost relativities by cell, we can use any of the models in this *Practitioner's Guide* to determine rate relativities. We explain the intuition in steps.

Illustration: Loss Ratio Intuition

Incurred Losses		
	Urban	*Rural*
Male	$2,700	$2,000
Female	$1,500	$1,200

Earned Premium		
	Urban	*Rural*
Male	$3,000	$4,000
Female	$2,400	$1,600

The current relativities by sex of driver and by garaging location are as follows:

Male:	1.50	Urban:	1.20
Female:	1.00	Rural:	1.00

Three Causes

To correct for potential distortions caused by an uneven distribution of insureds by other classification dimensions, we use loss ratios instead of loss costs. The underlying principle is that if the current premiums are actuarially correct – that is, if the current rate relativities are actuarial proper – the loss ratios in each cell should all be equal, except for random loss fluctuations.

The reader might wonder: *If all the current rate relativities were actuarially correct, we would not need to perform a rate relativities analysis. We perform the analysis because we want to examine whether the rate relativities are correct. What exactly are we assuming here?*

We restate this question as follows:

- We need data with no uneven distribution of insureds by classification dimension to determine accurate rate relativities.
- We must assume that the rate relativities are accurate to correct for uneven distributions of insureds by classification dimension.

The logic in these two statements seem circular at first. The exposition below explains what we are assuming and provides the justification for the assumption.

We assume that the rate relativities are correct for all classification dimensions besides those being examined now. We make no assumptions about the current rate relativities for the classification dimensions being examined. We are forming new rate relativities for the classification dimensions being examined now, and we ignore the current rate relativities.

We examine this in the numerical illustration. The observed loss ratios, or the ratios of the incurred losses to the earned premiums in each cell, are shown in the matrix below.

Loss Ratios		
	Urban	*Rural*
Male	90.00%	50.00%
Female	62.50%	75.00%

If all current rate relativities were actuarially proper and there were no random loss fluctuations, all cells would have the same loss ratio. They do not have the same loss ratios in this illustration. There are three possible causes.

Cause 1: The differences may be caused by random loss fluctuations. The importance of random loss fluctuations is a credibility issue. The credibility of the data increases with the volume of data and decreases with the dispersion of the loss distribution of these data. Credibility issues are important, but they are distinct from the minimum bias issues.

This *Practitioner's Guide* includes a short section on credibility. Most of the exposition assumes either that the data are fully credible or that the pricing actuary has already made (or will make) whatever adjustments are warranted by credibility considerations.

Cause 2: The differences are caused by improper rate relativities in other classification dimensions and there is an uneven distribution of insureds by these other classification dimensions. For example, perhaps the rates for a certain type of vehicle are too low, and the proportion of urban males driving that type of vehicle is greater than the proportions of the insureds in the other cells driving that type of vehicle.

If this is the cause of the differences, we are stymied. However, as long as the uneven distribution of insureds by the other classification dimension is not too serious, an inaccuracy in the rates will not distort our analysis too greatly. We may restate our assumption as follows:

> *For other classification dimensions, either the current rate relativities are accurate or the distribution of the insureds that we are examining is relatively even across these other dimensions.*

We may rephrase this assumption to fit the illustration as follows. Either the along other classification dimensions, such as vehicle type, are actuarially correct, or if a certain vehicle types has too high or too low a classification relativity, the proportion of males and females driving that type of vehicle is the same.

In many instances, this assumption is not perfect. Nevertheless, even if the use of loss ratios does not perfectly correct distortions caused an uneven distribution of insureds along other classification dimensions, it corrects the distortions at least partially. This assumption may not be perfect, but it makes our analysis better.

Cause 3: The final cause of the differences in the loss ratios by cell is inaccuracies in the rate relativities for the two classification dimensions that we are examining: sex and territory. This can be corrected by the minimum bias procedure, since the loss ratios by cell times the current relativities by cell equal the relative loss costs by cell.

Illustration: Suppose the loss ratio for male drivers is 90% and the loss ratio for female drivers is 75%. If the current male to female rate relativity is 2 to 1, the male to female loss cost relativity is 2 × 90% to 1 × 75% = 2.4 to 1.

For the illustration in this section, we form a matrix of relativities by sex and territory:

Current Rate Relativities		
	Urban	*Rural*
Male	1.80	1.50
Female	1.20	1.00

The relative loss costs by sex and territory are the product of the matrix of relativities and the matrix of loss ratios:

Loss Cost Relativities		
	Urban	*Rural*
Male	1.62	0.75
Female	0.75	0.75

We now proceed to determine optimal rate relativities by any of the minimum bias models discussed in this *Practitioner's Guide.*

Combined Models

Throughout this *Practitioner's Guide*, we have used simple multiplicative and additive models. In part, this reflects insurance practice, since most lines of business use simple multiplicative and additive models.

In truth, the business practice reflects ratemaking capabilities. Actuaries did not have simple procedures to optimize combined models, so these models did not gain wide acceptance.

The rationale for combined models is strong. Since the least squares and χ-squared bias functions provide simple recursive equations for many combined models, they may become more popular in the future.[23]

Illustration: Rating territory may have a variety of effects on insurance loss costs.

1. High crime areas may have a greater incidence of car theft and claim fraud. Thefts would raise comprehensive pure premiums, and fraud would raise liability pure premiums.
2. Areas with more sophisticated medical facilities may have higher loss costs for bodily injury claims.
3. Territories with a higher incidence of attorneys per capita often experience a higher incidence of bodily injury claims per physical accident.[24]

The first of these three effects argues for an additive model; the last of these three effects argues for a multiplicative model; and the second of these three effects may have both additive and multiplicative components.

Intuition alone is rarely sufficient to optimize a rating model. The minimum bias method allows the pricing actuary to determine the optimal rating structure from the observed loss costs.

COMBINED MODEL

Suppose a classification system has two dimensions: sex of driver and rating territory. Each classification dimension has two values: male vs female and urban vs rural. The male/female rating dimension has a multiplicative effect on loss costs. The rating territory dimension has both a multiplicative and an additive effect on loss costs. We show the structure of this rating model, and we explain how it can be optimized.

---

[23] Generalized linear models allows the optimization of even more complex rating models. We hope to provide a companion *Practitioner's Guide* on the use of generalized linear models for classification ratemaking.

[24] See Conners and Feldblum [1998] for the effects on private passenger automobile insurance.

For the male/female classification dimension, we use rate relativities of s_1 and s_2. For the urban/rural dimension, each class has two relativities: a multiplicative relativity denoted by t_1 and t_2, and an additive relativity denoted by z_1 and z_2. We denote the base rate as B.

The indicated pure premium for any class is $B \times (s_i \times t_j + z_j)$. The subscripts "i" and "j" denote the classification dimension. The indicated pure premiums are shown in the table below.

Observed Loss Costs			*Indicated Pure Premiums*		
	Urban	*Rural*		t_1, z_1	t_2, z_2
Male	r_{11} = \$800	r_{12} = \$500	*Male*	$B \times (s_1 \times t_1 + z_1)$	$B \times (s_1 \times t_2 + z_2)$
Female	r_{21} = \$400	r_{22} = \$200	*Female*	$B \times (s_2 \times t_1 + z_1)$	$B \times (s_2 \times t_2 + z_2)$

If we use the balance principle as the bias function, we balance along the two rows and the two columns. This gives four equations, of which only three are independent, since there is a totality constraint. We must solve for six classification relativities.

We can not solve for the optimal solution by straight-forward iterative methods. We show the procedure used earlier to highlight the problems.

Using the methods explained earlier, we choose starting values for the "t" relativities and the "z" relativities. We balance along the first row to determine the intermediate value for the s_1 (male) relativity, and we balance along the second row to determine the intermediate value for the s_2 (female) relativity.

We discard the values for t_1 and t_2, but we retain the starting values of z_1 and z_2. We balance along the first column to obtain the intermediate value for t_1 (urban), and we balance along the second column to obtain the intermediate value for t_2 (rural).

We discard the values for z_1 and z_2, but we retain the intermediate values of t_1 and t_2. We balance again along the first column to obtain the intermediate value for z_1 (urban), and we balance along the second column to obtain the intermediate value for z_2 (rural).

We have no problem doing the calculations. We show the general equations below, where $r_{i,j}$ is the observed loss cost for class (i,j) and $n_{i,j}$ is the number of exposures for class (i,j).

We multiply both the observed loss costs and the pure premiums by the number of exposures. Balancing along the first row gives us

$$\sum_j n_{1j}\, r_{1j} = \sum_j n_{1j}\, (s_1 \times t_j + z_j)$$

We are solving for s_1, going across the row, and using all the *j* values. We transpose this equation to give

$$s_1 = \sum_j n_{1j} (r_{1j} - z_j) \div \sum_j n_{1j} t_j$$

For the general equation, we substitute the *i* subscript for the "1" subscript in the *s, n,* and *r* variables, to give

$$s_i = \sum_j n_{ij} (r_{ij} - z_j) \div \sum_j n_{ij} t_j$$

We balance along the columns, using the first column as our example:

$$\sum_i n_{i1} r_{i1} = \sum_i n_{i1} (s_i \times t_1 + z_1)$$

We transpose this equation to solve for either *t* or for *z*. Solving the equation for t_j yields

$$t_j = \sum_i n_{ij} (r_{ij} - z_j) \div \sum_i n_{ij} s_i.$$

This looks like the equation for the *s* variables, with one difference. When we solve for *s*, we have z_j in the formula. When we solve for s_i, the formula uses all the *z* values. When we solve for *t*, we have a particular *z* value in the formula (e.g., when we solve for t_2, we have z_2 in the formula). Similarly, we can solve for the *z* variables to get

$$z_j = \sum_i n_{ij} (r_{ij} - s_i t_j) \div \sum_i n_{ij}$$

The series will not necessarily converge. If the series does converge, it does not necessarily have a unique limit. In a multi-dimensional combined multiplicative and additive model, there are many more relativities than there are equations when the balance principle is used as the bias function.

If we use a least squares or a χ-squared bias function, the combined model is not conceptually different from a simple multiplicative or additive model. We form the least squares or χ-squared expression in the same manner as did above. We set the partial derivative with respect to each rating variable equal to zero. We have the same number of equations as we have rating variables.

Caveats

The use of the minimum bias procedure with combined models is a powerful rating tool. But as the rating models grow more complex, the potential rating errors become more serious.

If there are a large number of exposures in each class, the optimization procedure is less likely to be distorted by random loss fluctuations. As the number of exposures in each cell decreases, the effects of random loss fluctuations become more serious.

OUTLIERS

The least squares and χ-squared bias functions are sensitive to outliers. Outliers are observed values that differ substantially from the expected values because of random loss fluctuations. Distortions stemming from random loss fluctuations can be controlled in several ways.

- Losses can be capped at basic limits or similar retentions.
- Low volume classes can be assigned limited credibility.
- The data in each cell can be examined for unusual values.

The use of low retentions or low credibility conflicts with the objective of basing rates on observed experience as much as possible. The examination of the observed data for unusual values is too time consuming for the exigencies of practical work.

Rather, the bias function should be chosen so that the results are not too sensitive to outliers.

Illustration: The classification system has two dimensions: male/female along one dimension and ten territories along the other dimension. The current driver relativities are 1.000 for female and 2.000 for male. The current territorial relativities are 1.000, 2.000, . . ., 10.000 for the ten territories, labeled (01, 02, . . . , 10). The base rate is $100, and a multiplicative model is used.

Scenario A: The observed loss costs are shown below, in units of 100 dollars.

Territory:	01	02	03	04	05	06	07	08	09	10
Male	$2	$4	$6	$8	$10	$12	$14	$16	$18	$20
Female	$1	$2	$3	$4	$5	$6	$7	$8	$9	$10

The observed loss costs exactly match the indicated pure premiums in the current rating system. All three bias functions discussed so far – balance principle, least squares, and χ-squared – would indicate retention of the current relativities.

Scenario B: Because of a random large loss, the observed loss costs for the males in territory 10 are $10,000 instead of $2,000. The "territory 10 / male" cell shows $100 instead of $20. This type of random loss fluctuation is common in classification analysis for small populations.

We have starting values of (1.000, 2,000, . . ., 10.000) for the ten territories. We determine the intermediate value for the male relativity.

The balance principle selects the male relativity "s_1" such that

$$(s_1 \times t_1) + (s_1 \times t_2) + \ldots + (s_1 \times t_{10}) = r_{1,1} + r_{1,2} + \ldots + r_{1,10}$$
$$s_1 \times \$55 = \$190$$
$$s_1 = 3.455$$

The least squares bias function selects the male relativity to minimize the squared error

$$\sum (r_{1,i} - s_1 \times terr_i)^2$$

We set the partial derivative with respect to s_1 equal to zero:

$$\sum (r_{1,i} - s_1 \times terr_i) \times (-terr_i) = 0$$

$$s_1 = \sum (r_{1,i} \times terr_i) / \sum terr_i^2 =$$

$$[(1\times2) + (2\times4) + (3\times6) + \ldots + (9\times18) + (10\times100)] / [1^2 + 2^2 + 3^2 + \ldots 9^2 + 10^2] = 4.078$$

Compared with the balance principle, the least squares bias function exacerbates the distortion caused by random loss fluctuations. In this instance, the χ-squared bias function magnifies the distortion less than the least squares bias function does. In other instances, the χ-squared bias function magnifies the distortion more than the least squares bias function does. Since combined models are more sensitive to random loss fluctuations than simple models are, and since the least squares or χ-squared bias function must be used, the pricing actuary must be particularly careful to exclude outliers from the data.

Other Bias Functions

Summary: We examine other bias functions, beginning with the χ-squared function and the squared error function. We continue with our simple 2 by 2 illustration for both additive and multiplicative models using these bias functions.

We review arguments for and against specific bias functions. We examine two goodness-of-fit tests – average absolute error and χ-squared – and we consider the relationship between the bias function chosen and the goodness-of-fit test.

We review the maximum likelihood bias function and the distributions commonly used with it. We discuss some of the potential advantages and drawbacks of the more sophisticated models compared to the balance principle.[25]

SQUARED ERROR AND X-SQUARED

Let us return to the simple illustration with which we began, as reproduced below.

	Urban	Rural		$terr_1$	$terr_2$
Male	$800	$500	sex_1	$200 \times s_1 \times t_1$	$200 \times s_1 \times t_2$
Female	$400	$200	sex_2	$200 \times s_2 \times t_1$	$200 \times s_2 \times t_2$

The left-hand side of the matrix shows the observed loss costs; the right-hand side shows the indicated pure premiums. Our objective is to pick classification relativities such that the indicated pure premiums are "as close as possible" to the observed loss costs.

We used the balance principle earlier to fit the classification relativities. We did not attempt to justify the balance principle; we provide its justification further below.

Statisticians would normally fit the classification relativities using other methods, such as:

1. Minimize the average absolute error between the indicated and observed figures.

---

[25] This *Practitioner's Guide* summarizes three seminal papers on the minimum bias procedure:

- Bailey and Simon [1960] recommends the χ-squared bias function, based on a credibility argument.
- Bailey [1963] recommends the balance principle, based on a bias argument.
- Brown [1988] investigates other bias functions, based on statistical arguments.

This *Guide* highlights the implications of the various bias functions, though it leaves conclusions to the reader.

2. Minimize the sum of the squared errors between the indicated and observed figures (i.e., the least squares bias function)
3. Minimize the sum of the relative squared errors between the indicated and observed figures (i.e., minimize the χ-squared error).
4. Maximize the likelihood of obtaining the observations given the classification relativities.

Minimizing the average absolute error makes sense to practicing actuaries. Minimizing the average absolute error is rarely used in statistics, perhaps because it was once thought to be mathematically intractable.[26] The three other methods result in relatively simple iterative equations for the minimum bias procedure.

We use the average absolute error as one of the goodness-of-fit tests. Given a set of classification relativities, it is easy to calculate the average absolute error. It is not easy to determine the set of classification relativities that minimizes the average absolute error.

Methods 2 and 3 – least squares and χ-squared are similar. We first show the procedures, and then we discuss the intuition for each.

Squared Error

The squared error for each cell is the square of the difference between the observed pure premium and the indicated pure premium. For urban male drivers in our basic illustration, this number is $(\$800 - \$200 \times s_1 \times t_1)^2$.

We sum the squared errors for the four cells to get (SE = sum of squared errors):

SE	$= (\$800 - \$200 \times s_1 \times t_1)^2$	urban male
	$+ (\$500 - \$200 \times s_1 \times t_2)^2$	rural male
	$+ (\$400 - \$200 \times s_2 \times t_1)^2$	urban female
	$+ (\$200 - \$200 \times s_2 \times t_2)^2$	rural female

To minimize the sum of the squared errors, we take partial derivatives with respect to each variable and set them equal to zero. For the "male" classification relativity ("s_1"), we have

$$0 = \partial SE/\partial s_1 = 2 \times (\$800 - \$200 \times s_1 \times t_1) \times (-\$200 \times t_1) + 2 \times (\$500 - \$200 \times s_1 \times t_2) \times (-\$200 \times t_2)$$

---

[26] Compare Cook [1967], page 200: "Why then do we use the method of least squares? Simply because absolute values are alleged to be mathematically inconvenient." Cook provides an algorithm for minimizing the average absolute error, which is simple to compute by pencil and paper and even easier to program in a spreadsheet. It would be useful to compare a minimum bias procedure based on Cook's algorithm with the methods in this *Practitioner's Guide*. [Charles Cook, "The Minimum Absolute Deviation Trend Line," Proceedings of the CAS, vol 80 [1967], pages 200-204.

We need to consider the cells only in the male (s_1) row. The other cells do not have an s_1 term in the squared error, so the partial derivative with respect to s_1 is zero.

Taking partial derivatives with respect to each of the classification relativities gives us four equations in four unknowns. The equations are not linear, so we use iteration to solve them.

Let us choose the same starting values for the squared error bias function as we chose for the balance principle (namely $t_1 = 2$ and $t_2 = 1$):

	Urban	Rural		$terr_1 = 2$	$terr_2 = 1$
Male	\$800	\$500	sex_1	$200 \times s_1 \times 2$	$200 \times s_1 \times 1$
Female	\$400	\$200	sex_2	$200 \times s_2 \times 2$	$200 \times s_2 \times 1$

Using the squared error bias function, we solve for the male relativity (s_1):

$$0 = \partial SE/\partial s_1 = 2 \times (\$800 - \$200 \times s_1 \times 2) \times (-\$200 \times 2) + 2 \times (\$500 - \$200 \times s_1 \times 1) \times (-\$200 \times 1)$$

To avoid dealing with multiples of 10, we choose a base rate of \$2 and we evaluate the observed pure premiums in multiples of \$100.

$$0 = \partial SE/\partial s_1 = 2 \times (\$8 - \$2 \times s_1 \times 2) \times (-\$2 \times 2) + 2 \times (\$5 - \$2 \times s_1 \times 1) \times (-\$2 \times 1)$$

$$-64 + 32s_1 - 20 + 8s_1 = 0$$
$$40s_1 = 84$$
$$s_1 = 2.1$$

Similarly, we solve for the female relativity (s_2):

$$0 = \partial SE/\partial s_2 = 2 \times (\$400 - \$200 \times s_2 \times 2) \times (-\$200 \times 2) + 2 \times (\$200 - \$200 \times s_2 \times 1) \times (-\$200 \times 1)$$

Simplifying as before, we get

$$0 = \partial SE/\partial s_2 = 2 \times (\$4 - \$2 \times s_2 \times 2) \times (-\$2 \times 2) + 2 \times (\$2 - \$2 \times s_2 \times 1) \times (-\$2 \times 1)$$

$$-32 + 32s_2 - 8 + 8s_2 = 0$$
$$-40 + 40s_2 = 0$$
$$s_2 = 1$$

We now discard the starting values of $t_1 = 2$ and $t_2 = 1$. Using the intermediate values of $s_1 = 2.1$ and $s_2 = 1$, we set the partial derivatives of the sum of the squared errors with respect to t_1 and t_2 equal to zero and we solve for new values of t_1 and t_2. We continue in this fashion until the series converges.

Squared Error Intuition

The properties of squared error minimization in the minimum bias procedure are unlike the properties of squared error minimization in other statistical problems, as explained below. We note first that the bias function makes a difference, even in this simple illustration.

A. The balance principle bias function gives $s_1 = 13/6 = 2.167$ and $s_2 = 1$.
B. The squared error bias function gives $s_1 = 2.100$ and $s_2 = 1$.

The balance principle looks at the errors; the squared error bias function looks at the square of the errors. The χ-squared bias function looks at the square of the errors relative to the expected value. The squared error and χ-squared bias functions place more weight on outlying cells, where the squares of the errors are large. The balance principle and the squared error bias function place more weight on the cells with large dollar values. The χ-squared bias function weights all cells more evenly.

Illustration: A classification system with two dimensions has male vs female in one dimension and territories 1, 2, and 3 in the other dimension. The starting relativities are 1.00, 2.00, and 3.00 for territories 1, 2, and 3. The observed loss costs for the three territories in the male row are \$2, \$4, and \$12, with equal exposures in each cell.

	territory 1 (1.00)	territory 2 (2.00)	territory 3 (3.00)
male	\$2.00	\$4.00	\$12.00
female	—	—	—

We want to determine the indicated relativity for males. Our concern here is not to solve this problem but to understand the effects of the different bias functions. We examine the effects of different choices for the male relativity.

- If the male relativity is 2.00, the indicated pure premiums are \$2, \$4, and \$6. The first two cells have a perfect fit, and the third cell is too low by \$6.

- If the male relativity is 4.00, the indicated pure premiums are \$4, \$8, and \$12. The first two cells are too high by a total of \$6, and the third cell has a perfect fit.

The balance principle considers the first power of the errors. To achieve balance, we choose a male relativity of 3.00. The indicated pure premiums are \$3, \$6, and \$9. The first two cells are too high by a total of \$3, and the third cell is too low by \$3.

The squared error bias function is more concerned with large errors than with small errors. We are more concerned with the error for territory 3, which is relatively large, than with the errors for territories 1 or 2, which are relatively small. To minimize the sum of squared errors, we increase the male relativity slightly, thereby reducing the error in territory 3 and increasing the errors in territories 1 and 2.

To solve this problem using the squared error bias function, we minimize the following expression:

$$(2-x)^2 + (4-2x)^2 + (12-3x)^2.$$

Taking the partial derivative with respect to "x" and setting it equal to zero gives

$$2(2-x)(-1) + 2(4-2x)(-2) + 2(12-3x)(-3) = 0$$

$$4 + 16 + 72 = 2x + 8x + 18x$$
$$92 = 28x$$
$$x = 92 / 28 = 3.286$$

Squared Error Minimization

Upon reflection, the illustration above seems odd to some statisticians. We are choosing a value to minimize the squared error among a series of observations. An elementary statistical theorem, which we review below, is that the average minimizes the sum of the squared errors. This seems inconsistent with the comments above.

Were we dealing with a single classification dimension, squared error minimization indeed produces the arithmetic average. The following illustration explains this statement.

Illustration: We are measuring a patient's fever with an old thermometer that is in poor working order. The thermometer is unbiased, but it is very inaccurate, and the observed readings are highly distorted by sampling error. We perform nine trials, and we observe readings of (100.1, 100.2, . . . , 100.9). The readings were not in this order, so there is no observed trend; we have simply arranged them in ascending numerical order. Using the least squared error function, we wish to determine the best estimate of the patient's temperature.

We rephrase the illustration mathematically. We have observed values of $z_1, z_2, \ldots, z_n$, and we must choose a single value for the z's – call it z^* – to minimize the squared error.

The sum of the squared errors is $\sum (z_i - z^*)^2$. The partial derivative of this sum with respect to z^* is $\sum 2(z_i - z^*)(-1)$. Setting this equal to zero gives $z^* = \sum z_i \div n$. The indicated z^* is the average of the z_i's.

In the temperature measurement illustration, the average of the nine observations is 100.5. This is the solution using the squared error bias function.

If we had chosen instead some other value, such as 100.3, we could correct this estimate by the average of the errors. The error in each observation is the observation minus 100.3. This is the series (–0.02, –0.01, 0, +0.01, . . . , +0.06). The average is +0.02. The corrected estimate is 100.3 + 0.02 = 100.5.

Multi-Dimensional Systems

This is not true for multi-dimensional systems. In a multiplicative model with two dimensions, the z_i's are the observed values. The z^* is the indicated relativity for one of the two dimensions. The other dimension has relativities of $y_1, y_2, \ldots, y_n$.

The sum of the squared errors is $\sum (z_i - y_i \times z^*)^2$. The partial derivative of this sum with respect to z^* is $\sum 2(z_i - y_i \times z^*)(-y_i)$.

Setting this equal to zero gives $z^* = \sum z_i \div \sum y_i^2$.

The indicated z^* is no longer the average of the z_i's. Rather, this result is the solution to the minimum bias procedure using the squared error bias function, as we show next.

BALANCE PRINCIPLE OPTIMIZATION

Let us contrast squared error minimization with the balance principle. When we deal with a single classification dimension, squared error minimization produces the arithmetic average. The balance principle selects the multi-dimensional equivalent to the mean.

The balance principle provides the economically optimal solution to the minimum bias problem. This is the economic corollary to Bailey's 1963 statement that the balance principle provides the only unbiased solution; see below.[27]

GENERAL SOLUTION

Throughout this study note, we solve the elementary 2 by 2 illustration before deriving the general formulas. The general formulas require readers to keep too many subscripts in mind. Although this is not difficult once one is accustomed to the minimum bias procedure, it hampers the initial grasp of the intuition.

Let us consider now a more general two dimensional classification system.

- We still assume one exposure per cell or the same number of exposures per cell. We deal with varying exposures per cell when we deal with credibility.
- The extension to more than two dimensions is straight-forward, though the additional subscripts obscure the intuition.

Suppose we have two dimensions, age of driver and territory, with "n" age classes and "m" territories. The observed loss cost in the i^{th} age class and the j^{th} territory is r_{ij}. The indicated pure premium in the i^{th} age class and the j^{th} territory is $x_i \times y_j$. This is the standard notation for the minimum bias computations.

The squared error in any cell is $(r_{ij} - x_i \times y_j)^2$. The sum of the squared errors is

---

[27] By economically optimal, we mean the bias function that maximizes the expected income of the firm in most scenarios. Clearly, there are exceptional scenarios when a different bias function may be better. In a jurisdiction that places restrictions on risk classification, the bias function may have to be changed to accommodate these restrictions. If the ratemaking data contain data errors, these errors should be corrected before the bias function is applied. If the insurer seeks to expand in certain classifications for competitive or marketing reasons, the minimum bias procedure may not accommodate the insurer's strategy. In most scenarios, however, the balance principle serves the economic interests of the firm.

$$\sum_{i=1}^{n}\sum_{j=1}^{m}(r_{ij} - x_i y_j)^2$$

We take partial derivatives with respect to each variable and set them equal to zero. We have a total of (n+m) variables, and we have a total of (n+m) equations. The constraints for least squares minimization are the same as the constraints for the balance principle. There is one totality constraint, since taking the sum of the squared errors along the rows is the same as taking the sum of the squared errors along the columns. This means that we have only (n+m−1) equations, since the (n+m) equations are not independent. In addition, we can multiply all the relativities along any dimension by a constant and divide the base rate by the same constant.

The (n+m) equations are not linear, so we must search for a solution by numerical methods. We choose starting values for one dimension – say, the y_j's. To solve for the value of x_1, we take the partial derivative with respect to x_1 and set it equal to zero:

$$\sum_{J=1}^{m} 2(r_{1j} - x_1 y_j)(-y_j) = 0$$

This gives us

$$x_1 = \sum (r_{1j} \times y_j) \div \sum y_j^2.$$

The x_1 is a variable. The summation signs in the last two equations above are over the j subscript. The y values are fixed; they are not variables once we have assigned starting values to the y values.

We repeat this procedure to solve for $x_2, x_3, \ldots, x_n$. Having solved for all the x values, we discard the starting y values and solve for new values of the y variables using the same procedure as for the x variables.

ADDITIVE MODEL

We can use an additive model with the least squares bias function. We show first the results for the elementary 2 by 2 illustration.

We repeat the observed loss costs and the indicated pure premiums for the additive model in the 2 by 2 illustration.

	Urban	Rural		$terr_1$	$terr_2$
Male	\$800	\$500	sex_1	$200 + s_1 + t_1$	$200 + s_1 + t_2$
Female	\$400	\$200	sex_2	$200 + s_2 + t_1$	$200 + s_2 + t_2$

As mentioned earlier, there are three mathematically equivalent ways of defining the additive model; the solution method is the same for each of them. The rate in cell x_iy_j is

A. Base rate + $x_i + y_j$,
B. Base rate $\times (1 + u_i + v_j)$, or
C. Base rate $\times (p_i + q_j)$

We use the first of these three equations for the intuition here, though we would use one of the other two methods in practice, thereby avoiding the need to adjust the relativities for inflation each year. A multiplicative relationship between the base rate and the relativities does not make the model multiplicative. Since the relationship among the factors is additive, the model is additive. A combined multiplicative and additive model has relationships among the relativities that are both multiplicative and additive; see below.

For the male urban cell, the squared error is $(\$800 - \$200 - s_1 - t_1)^2$. The sum of the squared errors for all four cells is

$$
\begin{aligned}
& (\$800 - \$200 - s_1 - t_1)^2 \\
+ \; & (\$500 - \$200 - s_1 - t_2)^2 \\
+ \; & (\$400 - \$200 - s_2 - t_1)^2 \\
+ \; & (\$200 - \$200 - s_2 - t_2)^2
\end{aligned}
$$

We take partial derivatives with respect to each variable and set them equal to zero. The partial derivative with respect to s_1 is

$$2(\$800 - \$200 - s_1 - t_1)(-1) + 2(\$500 - \$200 - s_1 - t_2)(-1) = 0.$$

or

$$s_1 = (\$900 - t_1 - t_2) \div 2.$$

For the additive model with the least squares bias function, the simultaneous equations are linear, and we can solve them directly. Nevertheless, it is easier to program the solution using numerical methods.

If we choose starting values of t_1 = \$250 and t_2 = \$0, we get s_1 = \$325.

General Formula

For the general formula, we let B = the base rate. The sum of the squared errors is

$$\sum_{i=1}^{n}\sum_{j=1}^{m}(r_{ij} - B - x_i - y_j)^2$$

We take the partial derivative with respect to x_1 and set it equal to zero:

$$\sum_{j=1}^{m} 2(r_{1j} - B - x_1 - y_j)(-1) = 0$$

or

$$x_1 = \sum(r_{1j} - y_{1j})/m - B$$

where the summation is over the *j* subscript.

The Bias Function

The optimal relativities depend on the choice of the bias function. The choice of bias function can be viewed from three perspectives.

1. Mathematical tractability
2. Social equity
3. Economic optimization

Mathematical tractability is of most concern when some bias functions give simple relationships and some bias functions give equations that defy simple solutions. For the minimum bias procedure, we get relatively simple equations for the bias functions discussed in this paper. We do not get simple equations if we use the average absolute error as the bias function, so we do not consider that method.[28]

Social equity is subjective, though it is vital to the success of a highly regulated industry like insurance. The balance principle sometimes results in large errors for outlying cells. The errors are particularly large in absolute value for high rated cells. If a multiplicative model is

---

[28] With modern spreadsheets, the average absolute error no longer poses tractability issues. Just like the solution for the balance principle is the mean, the solution for the average absolute error is the median. It is not uncommon for actuaries to use the median instead of the mean in practical problems.

used when an additive model is more appropriate, the errors for outlying cells are frequently overcharges. The squared error bias function reduces these large errors.[29]

Of the bias functions which we consider in this paper, the squared error bias function is the best at reducing large overcharges for individual cells. Analysts who are concerned with large overcharges might prefer the squared error bias function. Ferreira's critique of insurance industry classification systems in Massachusetts illustrates this social position.[30]

Economic optimization drives the behavior of firms in free markets. There is disagreement regarding these economic forces in complex markets, but the major attributes of these forces can be described.

We take the perspective of the firm (the insurer), not the perspective of the consumer. Firms seek to maximize profits and to minimize losses (among other firm objectives). Suppose an insurer issues 3 policies. It must choose between two rating systems.

A. Under rating system A, it loses $1 each on the first two policies and it breaks even on the third policy.
B. Under rating system B, it breaks even on the first two policies and it loses $1.50 on the third policy.

Rating system A is off by $2 using the balance principle while rating system B is off by $1.50. Using the squared error bias function, rating system A is off by $2 while rating system B is off by $2.25. The balance principle says we should choose rating system B, and the squared error bias function says we should choose rating system A.

The economic principle of profit maximization (or loss minimization) says we should choose rating system B, as the balance principle says. Using our simple assumptions, the profit maximization principle generally agrees with the balance principle.

Economic forces are not trivial. There are many economic reasons for avoiding large errors, including consumer dissatisfaction, consumer switching, and public relations. In democratic

---

[29] This is the same as saying that the least squares bias function is sensitive to outliers, since a single large outlier can significantly change the results when the least squares bias function is used.

[30] See Ferreira [1978], as well as Cummins *et al.*, *Risk Classification in Life Insurance*, chapter 4, pages —. We are not endorsing Ferreira's views, which are inconsistent with competitive insurance markets. See the continuing discussion in the text of this *Practitioner's Guide*. [Ferreira, Joseph Jr., "Identifying Equitable Insurance Premiums for Risk Classes: An Alternative to the Classical Approach," in Andrew F. Giffin, Vincent Travis, and William Owen (eds.), *Automobile Insurance Risk Classification: Equity and Accuracy* (Boston: Massachusetts Division of Insurance, 1978), pages 74-120.][Cummins, J. David, Barry D. Smith, R. Neil Vance, and Jack L. VanDerhei, *Risk Classification in Life Insurance* (Boston: Kluwer Academic Publishers, 1983).]

systems where social opinion and political pressures are strong, firms may sacrifice short-term profit maximization to achieve other ends, such as workforce diversity and environmental protection. Furthermore, manager incentives may encourage the pursuit of other goals, such as corporate growth instead of profit maximization. Nevertheless, profit maximization remains the dominant corporate goal. The pricing actuary should keep these social and economic desiderata in mind when choosing a bias function for the minimum bias procedure.

X-SQUARED

The χ-squared bias function is similar to the squared error bias function, except that each "bias" is divided by the expected value.

Let us return to the simple illustration with which we began, as reproduced below.

	Urban	Rural		$terr_1$	$terr_2$
Male	\$800	\$500	sex_1	$200 \times s_1 \times t_1$	$200 \times s_1 \times t_2$
Female	\$400	\$200	sex_2	$200 \times s_2 \times t_1$	$200 \times s_2 \times t_2$

The χ-squared value for each cell is (the square of the difference between the observed loss cost and the indicated pure premium) divided by the indicated pure premium. For urban male drivers in our basic illustration, this number is

$$(\$800 - \$200 \times s_1 \times t_1)^2 \div (\$200 \times s_1 \times t_1).$$

We sum the squared errors for the four cells to get the sum of χ-squared values:

$$\begin{aligned}
&= (\$800 - \$200 \times s_1 \times t_1)^2 \div (\$200 \times s_1 \times t_1) \text{ urban male}\\
&+ (\$500 - \$200 \times s_1 \times t_2)^2 \div (\$200 \times s_1 \times t_2) \text{ rural male}\\
&+ (\$400 - \$200 \times s_2 \times t_1)^2 \div (\$200 \times s_2 \times t_1) \text{ urban female}\\
&+ (\$200 - \$200 \times s_2 \times t_2)^2 \div (\$200 \times s_2 \times t_2) \text{ rural female}
\end{aligned}$$

To minimize the sum of the squared errors, we take partial derivatives with respect to each variable and set them to zero. We use the quotient rule:

$$\text{If } y(x) = f(x)/g(x), \text{ then } \partial y/\partial x = [g(x) \times \partial f/\partial x - f(x) \times \partial g/\partial x] / g^2(x).$$

For the "male" classification relativity ("s_1"), we have

$$\begin{aligned}
0 = \partial SE/\partial s_1 &= [(\$200 \times s_1 \times t_1) \times 2 \times (\$800 - \$200 \times s_1 \times t_1) \times (- \$200 \times t_1)\\
&- (\$800 - \$200 \times s_1 \times t_1)^2 \times (\$200 \times t_1)] / (\$200 \times s_1 \times t_1)^2\\
&+ [(\$200 \times s_1 \times t_2) \times 2 \times (\$500 - \$200 \times s_1 \times t_2) \times (- \$200 \times t_2)\\
&- (\$500 - \$200 \times s_1 \times t_2)^2 \times (\$200 \times t_2)] / (\$200 \times s_1 \times t_2)^2
\end{aligned}$$

Although the arithmetic looks cumbersome, the equation can be reduced to a simple form. To avoid needless arithmetic, we show the general solution, and we resume the illustration after deriving the appropriate recursive equation.

X-SQUARED RECURSIVE EQUATIONS

We show the general recursive equations for the χ-squared bias function. The horizontal axis is the "j" dimension, and the vertical axis is the "i" dimension. We show two dimensions with two classes in each dimension to aid visualization of the example. The equations themselves have no constraints on the number of classes in each dimension. The extension of the equations to three or more dimensions is straight-forward.

	Urban	Rural		$terr_1$	$terr_2$
Male	\$800	\$500	sex_1	$200 \times s_1 \times t_1$	$200 \times s_1 \times t_2$
Female	\$400	\$200	sex_2	$200 \times s_2 \times t_1$	$200 \times s_2 \times t_2$

We form the χ-squared bias function as a double summation covering all the cells in the array.

$$\sum \sum (n_{ij}\, r_{ij} - n_{ij}\, x_i\, y_j)^2 / n_{ij}\, x_i\, y_j$$

We factor out the number of exposures from the equation to give

$$\sum \sum n_{ij} (r_{ij} - x_i\, y_j)^2 / x_i\, y_j$$

We seek to minimize the χ-squared value.

Given starting values for either dimension we determine the intermediate values for the other dimension. Assume we have chosen starting values for the "y" relativities and we are solving for the intermediate value of x_i. Only the cells in the "i"th row have terms withe x_i in them. We take the partial derivative of this row with respect to x_i, and we set it equal to 0.

We use the quotient rule for taking derivatives: if $f(x) = g(x)/h(x)$, then $\partial f/\partial x = [h(x) \times \partial g/\partial x + g(x) \times \partial h/\partial x] / h^2(x)$.[31]

In the equation below, we take the summation over the "j" dimension. The value of "i" is fixed.

$$\sum n_{ij} [x_i\, y_j\, 2(r_{ij} - x_i\, y_j) \times (-y_j) - (r_{ij} - x_i\, y_j)^2\, y_j] / (x_i\, y_j)^2 = 0$$

The value $x_i = 0$ will not minimize this equation, so we can multiply both sides of the equation by $(x_i)^2$. We separate the left side of the equation into two fractions, and we factor out $(y_j)^2$ from the first fraction and y_j from the second fraction:

---

[31] This is the same as the product rule for taking derivatives, with $f(x) = g(x) \times (1/h(x))$.

$$\sum -2 \times n_{ij}\, x_i\, (r_{ij} - x_i\, y_j) - (n_{ij} / y_j)\, (\, r_{ij} - x_i\, y_j)^2 = 0$$

We expand the square and we combine like terms:

$$\sum -2 \times n_{ij}\, x_i\, r_{ij} + 2 \times n_{ij}\, x_i^2 y_j - (n_{ij} / y_j)\, (r_{ij}^2 - 2\, r_{ij}\, x_i\, y_j + x_i^2\, y_j^2) = 0$$

$$\sum -2 \times n_{ij}\, x_i\, r_{ij} + 2 \times n_{ij}\, x_i^2 y_j - (n_{ij} / y_j)\, (r_{ij}^2) + 2 \times n_{ij}\, x_i\, r_{ij} - n_{ij}\, x_i^2\, y_j = 0$$

$$\sum\; n_{ij}\, x_i^2 y_j - (n_{ij} / y_j)\, (r_{ij}^2) = 0$$

This gives a relatively simple expression for each x_i in terms of the y_j values:

$$x_i = [\, \sum\, (n_{ij} \times r_{ij}^2 / y_j) / \sum\, n_{ij}\, y_j\,]^{0.5}$$

For the illustration, there is one exposure in each cell. The starting values are $y_1 = 2$ and $y_2 = 1$. We use a base rate of \$200, and we divide all cells by \$200.

	Urban	Rural		$terr_1 = 2$	$terr_2 = 1$
Male	\$4	\$2.5	sex_1	$s_1 \times 2$	$s_1 \times 1$
Female	\$2	\$1	sex_2	$s_2 \times 2$	$s_2 \times 1$

Using the χ-squared bias function along the first row, we get

$$s_1 \text{ (male relativity)} = [(½ \times 4^2 + 1 \times 2.5^2) / (2 + 1)]^{0.5} = 2.179.$$

Using the χ-squared bias function along the second row, we get

$$s_2 \text{ (female relativity)} = [(½ \times 2^2 + 1 \times 1^2) / (2 + 1)]^{0.5} = 1.000.$$

The male to female relativity is 2.179 to 1.

The least squares bias function gave a relativity of 2.1 to 1. The dollar values in the urban-male cell are larger than the dollar values in the rural-male cell, so the least squares bias function gives more weight to the urban-male cell as compared to the rural-male cell than the χ-squared bias function does.

ADDITIVE MODEL WITH Χ-SQUARED

The χ-squared bias function can be used with any type of model, whether multiplicative, additive, or combined. If an additive model is used, we minimize the following expression:

$$\sum \sum n_{ij} (r_{ij} - x_i - y_j)^2 / (x_i + y_j)$$

We set the partial derivative with respect to each relativity equal to zero. It is easiest to solve the resulting set of simultaneous equations by iteration. Bailey and Simon [1960], followed by Brown [1988] give the recursive equations as

$$\Delta x_i = \frac{\sum_j n_{i,j}\left(\frac{r_{i,j}}{x_i + y_j}\right)^2 - \sum_j n_{i,j}}{2\sum_j n_{i,j}\left(\frac{r_{i,j}}{x_i + y_j}\right)^2 \left(\frac{1}{x_i + y_j}\right)}$$

The form of the recursive equation is not as simple as that for other rating models, but the work needed to implement this model in a spreadsheet is not significantly greater.[32]

[32] We have not used this model in our own applications, and we have not attempted to verify the recursive equation or to examine the appropriateness of this model for specific scenarios.

Goodness-of-Fit

There are various rating models, and there are various bias functions. For a given rating model and bias function, the minimum bias procedure optimizes the relativities. We now wish to optimize the rating system by choosing the best rating model and bias function.

The optimal procedure depends on two items.

- The choice of rating model, such as multiplicative, additive, or combined, depends on the characteristics of the observed loss costs. For some types of coverage, a multiplicative model is preferred; for other types of coverage, an additive model is preferred.
- The choice of the bias function depends on the objective.
  - The statistician seeking the best fit might use a maximum likelihood function if a tractable distribution function is appropriate for this coverage or a χ-squared function if the probability distribution function is not known or not tractable.
  - The regulator seeking to avoid large dollar mismatches between observed loss costs and indicated pure premiums might use a least squares function.
  - The insurer seeking to avoid monetary losses might use the balance principle.

Each bias function may be associated with a particular use. The preferences listed above are possibilities; other preferences are also possible. In particular, a regulator might prefer the balance principle to provide the most efficient rating system.

- The objective of avoiding large dollar overcharges and undercharges is a dubious goal. It is not always compatible with free markets, and introduces inefficiencies into the insurance system.
- A pricing actuary working with a new line of business might prefer the χ-squared function to examine whether the rating system chosen is compatible with the observed loss costs.

Empirical Tests

We can test the choice of rating model empirically.

Illustration: We are using a χ-squared bias function to optimize classification relativities. We do not know whether a multiplicative model or an additive model is more appropriate.

We perform the minimum bias procedure twice: once with the multiplicative model and a χ-squared bias function and once with an additive model and a χ-squared bias function. After optimizing the relativities for each model, we compare the final χ-squared difference between the observed loss costs and the indicated pure premiums for each model. The model with the lower χ-squared value is preferred.

Illustration: We are using the balance principle to optimize classification relativities. We do not know whether a multiplicative model or an additive model is more appropriate.

We perform the minimum bias procedure twice: once with the multiplicative model and the balance principle and once with an additive model and the balance principle. After optimizing the relativities for each model, we compare the average absolute difference between the observed loss costs and the indicated pure premiums for each model. The model with the lower average absolute difference is preferred.

We can not empirically test the suitability of the bias function. The illustration below explains why.

Illustration: We are using a multiplicative model, and we are deciding between the balance principle and the χ-squared function.

We perform the minimum bias procedure twice: once with the multiplicative model and the balance principle and once with the multiplicative model and a χ-squared bias function.

If we use a χ-squared function to measure the difference between the observed loss costs and the indicated pure premiums to test the performance of the two models, the χ-squared bias function does better. This result is tautological, since the χ-squared bias function minimized the χ-squared difference between the observed loss costs and the indicated pure premiums.

If we use the average absolute difference between the observed loss costs and the indicated pure premiums to test the performance of the two models, the balance principle does better.[33] The χ-squared bias function minimizes large percentage errors. The balance principle and the average absolute difference minimize dollar differences.

The choice of bias function is a qualitative choice, depending on the objectives of the rating system. It is not subject to a quantitative test of suitability. We examine these qualitative issues in the following section of this *Practitioner's Guide.*

Empirical tests of actual insurance rating systems may help dispel some of the rancor in public policy decisions. Some persons have criticized the insurance industry for using multiplicative models that overcharge high rated classifications. It has been suggested that an additive model might be more equitable.

Many insurers tend to view this criticism as politically motivated, intended to curry support among urban voting blocs. An empirical test of a multiplicative model against an additive model should help resolve some of the actuarial questions.

---

[33] This result is (perhaps) not always true, but exceptions are rare.

The relative merits of a multiplicative versus an additive model are unclear. In their 1960 *Proceedings* paper, Bailey and Simon concluded that an additive model was preferred to a multiplicative model for the Canadian private passenger automobile data. In his 1988 *Proceedings* paper, Rob Brown concluded that a multiplicative model was preferred to an additive model for the Canadian private passenger automobile data. In his discussion of Brown's paper, Gary Venter has suggested that a combined multiplicative and additive model might be superior to either of the models tested.

SQUARED ERROR VS X-SQUARED

The squared error bias function is similar to the χ-squared bias function. We examine the relative advantages of each.

The χ-squared test looks at percentage differences; the squared error test looks at absolute differences. For fitting distributions, statisticians often prefer the χ-squared test to a least squares test.

Illustration: We are fitting a distribution to two empirical data points.

- Point A has an observed value of \$101 and a fitted value of \$100.
- Point B has an observed value of \$1.50 and a fitted value of \$1.00.

We examine the errors for each point.

- The squared error is $(101 - 100)^2 = 1.00$ for point A and $(1.50 - 1.00)^2 = 0.25$ for point B. Point B fits better.
- The χ-squared value is $(101 - 100)^2 / 100 = 0.01$ for point A and $(1.50 - 1.00)^2 / 1.00 = 0.25$ for point B. Point A fits better.

The statistician might prefer the χ-squared test to the squared error test. The practical businessperson might argue that the insurance enterprise is not concerned with optimizing a statistical fit. It is concerned with optimizing net income. At point A, the insurer has a gain or loss of \$1.00. At point B, the gain or loss is \$0.50. The squared error test is preferred.[34]

This argument does not fully reflect the purpose of the minimum bias procedure. The argument would be correct if we fully believed the observed values – that is, if the observed values were fully credible. But if the observed values were fully credible, we would have no need to use the minimum bias procedure; we would just use the rates indicated by the observed loss costs in each cell.

---

[34] This is similar to the argument that we discussed earlier when comparing the balance principle bias function and the squared error bias function.

We are using the minimum bias procedure because the individual observed values are not fully credible, and we believe that the relationships among all the cells in the observed matrix provides useful information for choosing the true expected values.

When we say that a particular fit "X" has less of an error than another fit "Y," we do not mean that fit "X" will produce a smaller error in the future period. Rather, we mean that fit "X" is probably closer to the true values of the cells, and so it is a better pricing procedure than fit "Y" is. Our assumption is that we don't know the true expected loss costs. The χ-squared bias function does a better job showing us the true expected loss costs than the squared error bias function does.

The 1960 Bailey and Simon paper says (page 10) in defense of the χ-squared bias function:

> *The indication of each group should be given a weight inversely proportional to the standard deviation of the indication.*

This is a traditional justification for classical credibility, as Bailey and Simon continue:

> *The standard deviation of the indication is inversely proportional to the square root of the expected number of losses for the group.*[35]

BALANCE PRINCIPLE VS X-SQUARED

In the preceding sections, we compared the χ-squared bias function to the squared error bias function, and the balance principle to the squared error bias function. We now compare the χ-squared bias function with the balance principle.

We can not give an unequivocal answer. The 1960 Bailey and Simon paper prefers the χ-squared bias function, whereas the 1963 Bailey paper argues for the balance principle.[36]

- The χ-squared bias function uses proportional departures; the balance principle does not use proportional departures.

---

[35] After the writings of Hans Bühlmann, Gary Venter, Howard Mahler, and others, this statement is no longer accepted uncritically. We do not attempt to judge it more rigorously in this paper.

[36] Another bias function is the absolute proportional departure of the indicated pure premiums from the observed loss costs. The absolute proportional departure is not as mathematically tractable as the other two.

- The balance principle uses the first order departure, which is economically optimal.[37] The χ-squared bias function uses the squared departure, which is not economically optimal.
- The balance principle is unbiased; the χ-squared bias function is not unbiased.

The last statement above warrants explanation. The 1963 Bailey paper argues that the balance principle constrains the relativities so that the total indicated pure premiums along any dimension equal the total observed loss costs along that dimension.

Illustration: If the balance principle is used as the bias function, the total indicated pure premiums for all urban drivers equals the total observed loss costs for all urban drivers. Similarly, the total indicated pure premiums for all male drivers equals the total observed loss costs for all male drivers.

COMMON PRACTICE

Common practice among casualty actuaries is to use the balance principle, not the χ-squared bias function. One might argue that since more effective procedures drive out less effective procedures in a competitive market, this is an argument in favor of the balance principle.

In truth, many ratemaking procedures were selected for ease of implementation, not necessarily for their mathematical accuracy. The balance principle was easier before the widespread use of desktop computers, and it gained widespread acceptance. Few actuaries have tried the χ-squared bias function or the least squares bias function. No conclusions should be drawn from the common practice among actuaries.

CREDIBILITY

Many practitioners combine the minimum bias procedure with credibility weighting of the indicated pure premiums either with the observed loss costs or with the underlying pure premiums. We show illustrations of each method.

INDICATED AND OBSERVED

The minimum bias procedure gives the indicated pure premiums for each class in an array. The pure premiums used for the final rates is a weighted average of the indicated pure premiums and the observed loss costs for that class. The credibility for the observed loss

[37] As noted earlier, no bias function is economically optimal in all scenarios. We mean here simply that the balance principle would be the best bias function in most scenarios for a firm seeking to maximize profits.

costs is a function of the volume of business in the class. Classes with greater volume place more weight on the observed loss costs.[38]

Various credibility parameters are used; the classical credibility formulas are most common. Classes with a certain volume of claims or of exposures are given full credibility. The square root rule is used for classes with lower volume of claims or exposures.

Illustration: Suppose that classes with exposure of 10,000 or more car-years are accorded full credibility. A class with 3,600 car-years of exposure, an $800 observed loss cost, and a $700 indicated pure premium, is accorded $(3{,}600/10{,}000)^{0.5} = 60\%$ credibility. The credibility weighted pure premium is $60\% \times \$800 + (1 - 60\%) \times \$700 = \$760$.

INDICATED AND UNDERLYING

For premises and operations ratemaking, ISO uses a balance principle minimum bias procedure with observed loss ratios to determine the indicated changes to class group and type of policy relativities.[39]

- An indicated relativity change of 1.08 for type of policy 12 means that the existing relativity for type of policy 12 should be increased by 8%.
- The full credibility standard is based on the number of claims in the class during the experience period. These standards are 2,500 claims for OL&T BI, 3,000 claims for M&C BI, and 7,500 claims for M&C PD.
- Partial credibility is based on the square root rule. For example, 1,080 claims in M&C BI gives $(1{,}080/3{,}000)^{0.5} = 60\%$ credibility.
- The indicated relativity change for the class is raised to the power of the credibility. If the indicated relativity change is 1.08 and the credibility is 60%, the credibility weighted relativity change is $1.080^{0.6} = 1.047$.

These two illustration show different uses of credibility. ISO credibility weights the indicated classification *relativities* with the current classification relativities to dampen the changes from year to year. The first illustration credibility weights the observed loss costs with the indicated pure premiums to increase the accuracy of the final rates.[40]

---

[38] See Venter's [1990] discussion of Brown's 1988 paper cited earlier.

[39] See Nancy C. Graves and Richard Castillo, "Commercial General Liability Ratemaking for Premises and Operations," *Pricing* (Casualty Actuarial Society 1990 Discussion Paper Program), Volume II, pages 631-696, for a more complete discussion of the ISO procedure.

[40] Compare Gary Venter's distinction between classical credibility, which is used to minimize rate fluctuations from year to year, and Bayesian-Bühlmann credibility, which is used to increase the accuracy of the estimate.

Embedded Credibility

The minimum bias procedure has credibility embedded in the calculations, since each cell is weighted by the number of exposures in that cell. The traditional credibility weighting in classification ratemaking is embedded in the procedure; it need not be added a second time.

A comparison with the single-dimensional classification ratemaking procedure should clarify this. Suppose there are three territories in a state with the experience shown below. The exposures are car-years, and the dollar figures are in thousands.

	Exposures	Claims	Premium	Losses	Loss Ratio	Indication
Terr 01	$5,000	$500	$5,000	$3,500	70.0%	0.972
Terr 02	$10,000	$1,000	$15,000	$10,800	72.0%	1.000
Terr 03	$2,000	$200	$4,000	$2,980	74.5%	1.035
Total	$17,000	$1,700	$24,000	$17,280	72.0%	

The observed data suggest that

- Territory 01 should have a reduction of 2.8% in its base rate.
- Territory 02 should have no change in its base rate.
- Territory 03 should have an increase of 3.5% in its base rate.

The indications in the table above take no account of the number of exposures or the number of claims in each territory. Since territory 03 has only 200 claims in the experience period, the +3.5% indication may be distorted by random loss fluctuations. To adjust for the volume of business in each territory, the raw indications may be credibility weighted with the overall average of unity, where the credibility depends on the number of exposures or the number of claims.

In the minimum bias procedure, the number or exposures in each cell affects the computation. The weight accorded to the observed loss costs in the cell is proportional to the number of exposures in the cell. From this perspective, credibility weighting the observed loss costs by the number of exposures would be applying credibility twice.

Nevertheless, some justification remains for a credibility adjustment. To determine the indicated pure premium for a cell, the minimum bias procedure uses all the cells in the array and the type of rating model. The credibility embedded in the minimum bias procedure deals with random loss fluctuations. A second credibility adjustment deals with model specification risk. We explain these concepts with an illustration.

Illustration: We are setting classification relativities with a minimum bias procedure. The observed loss costs for young unmarried urban male drivers is $2,500 per car. The indicated pure premiums for these drivers is $3,000 per car. There are two explanations for the difference.

1. The rating model is correct, and random loss fluctuations account for the difference. Random loss fluctuations may have reduced the observed loss costs for this cell, or random loss fluctuations in neighboring cells may have increased the indicated pure premium for this cell. If random loss fluctuations are the cause of the difference between the observed loss costs and the indicated pure premium, the credibility embedded in the minimum bias procedure is sufficient. No additional credibility adjustment should be used.

2. There were no random loss fluctuations causing the difference, but the rating model is not correct. The minimum bias procedure may be using a multiplicative model, which produces high indicated pure premiums for high risk drivers, when an additive model is proper, which would lead to lower pure premiums for these drivers. This is model specification risk, and a second credibility adjustment is warranted.

Classical credibility procedures are not an ideal compensation for model specification risk. The ideal approach is to use several models, such as multiplicative, additive, and combined models, and to test the goodness-of-fit for each model. Time constraints preclude this ideal approach in most cases, and a credibility adjustment may be a reasonable alternative.

Rate Fluctuations

The use of credibility to temper rate fluctuations from year to year is a dubious practice. In practice, most actuaries conceive of credibility as a means to price more accurately. Although Venter correctly notes that the stated rationale for classical credibility deals with tempering rate fluctuations, even classical credibility does serve the objective of increasing the accuracy of the rate indications.[41]

When rating bureaus made advisory rates, they had more incentive to temper rate fluctuations from year to year than private insurers have. Since the objective is rarely well defined, the credibility procedures are often arbitrary. ISO's credibility procedure may not have had firm statistical justification, but it fulfilled the objective of tempering the requested rate changes.

---

[41] See Venter's chapter of "Credibility" in any of the first three editions of the CAS textbook, *Foundations of Casualty Actuarial Science.*

MAXIMUM LIKELIHOOD

Some statisticians prefer a maximum likelihood test to either a χ-squared test or a least squares test when fitting a distribution to observed data. In his 1988 *Proceedings* paper, Rob Brown illustrated the use of a maximum likelihood test to optimize classification relativities.

The use of a maximum likelihood test requires an assumption about the distribution of values in each class. The appropriate distribution for loss costs is not evident. It probably is not a simple mathematical distribution that would be amenable to the procedure discussed here, such as an exponential distribution or a Poisson distribution. If the appropriate distribution is not known, the statistical merits of a maximum likelihood test are less clear.

The maximum likelihood test is rarely used in practical work, and not all actuaries are familiar with it. We explain the use of the maximum likelihood test by a series of illustrations.

> *Illustration:* We are fitting an exponential curve to a set of insurance losses. The exponential distribution function says that the likelihood of a loss of size "x" is proportional to $e^{-\lambda x}$. We first determine the constant of proportionality.
>
> Given that a loss has occurred, the likelihood that the loss is between zero and infinity is 1. If "k" is the constant of proportionality, the integral of $ke^{-\lambda x}$ between 0 and infinity equals k/λ. For this to be unity, the constant of proportionality must be λ. The exponential distribution function is $\lambda e^{-\lambda x}$.

LIKELIHOOD AND PROBABILITY

We use the term likelihood, not the term probability. If the exponential distribution function has a λ of 0.0001, the likelihood of a loss of size \$20,000 is $0.0001 \times e^{-2}$.

If losses are spread throughout the positive numbers, the probability of a loss exactly equal to \$20,000 is zero.[42] We may conceive of the likelihood that a loss is equal to \$X as the probability that the loss is between $\$X - \epsilon$ and $\$X + \epsilon$, divided by $2 \times \epsilon$. The limit of this ratio as ϵ tends to zero is the likelihood.

[42] In practice, losses cluster at round dollar figures, so the probability of loss exactly equal to \$20,000 is greater than zero. The statement in the text assumes an ideal model, where losses can be any amount, down to fractions of a penny, with no rounding to dollar amounts.

Before showing the use of the maximum likelihood test, we examine the mean of the exponential distribution function. The mean equals

$$\int_0^\infty x\lambda e^{-\lambda x}\,dx = 1/\lambda$$

We resume the illustration. We would like to fit an exponential curve to a set of insurance losses. We seek to determine the value of λ.

We have four methods of doing this. We show the full procedure only for the maximum likelihood method.

METHOD OF MOMENTS

The mean of the exponential distribution is $1/\lambda$. We take the average of the observations, and we set λ equal to the reciprocal of this average.

LEAST SQUARES

We divide the loss sizes into ranges, such as $0 to $5,000, $5,001 to $25,000, $25,001 to $100,000, and so forth. We calculate the percentage of observed losses which fall into each range. For any given λ, we determine the percentage of theoretical losses that would fall into each range.

For each range, we calculate the squared difference between the observed percentage and the theoretical percentage. We sum the squared differences over all the ranges. The result is a function of λ. To minimize this squared difference, we set the partial derivative with respect to λ equal to zero.

χ-SQUARED

The χ-squared procedure is similar to the least squares procedure, but instead of taking the squared difference we take the χ-squared difference. For each range, we divide the squared difference by the expected value.

MAXIMUM LIKELIHOOD

We dispense with the ranges. Suppose we have observed five losses, with sizes of $3,000, $5,000, $15,000, $20,000, and $80,000. For a given value of λ, the likelihood of a loss equal to $3,000 is $\lambda e^{-\lambda \times 3{,}000}$. The likelihood of five losses for the values listed above is the product of the likelihoods of each individual loss, or

$$\lambda e^{-\lambda \times 3{,}000} \times \lambda e^{-\lambda \times 5{,}000} \times \lambda e^{-\lambda \times 15{,}000} \times \lambda e^{-\lambda \times 20{,}000} \times \lambda e^{-\lambda \times 80{,}000}$$

The sum of the five losses is $123,000. We simplify the likelihood to $\lambda^5 e^{-\lambda \times 123{,}000}$. To find the optimal λ, we must choose the value that gives the greatest likelihood. To do this, we set the partial derivative with respect to λ equal to zero.

Before taking the partial derivative, we make one simplification. Maximizing the likelihood is the same as maximizing the logarithm of the likelihood. The logarithm of the likelihood is

$$5 \ln \lambda - 123{,}000 \times \lambda.$$

Setting the partial derivative with respect to λ to zero gives $5/\lambda - 123{,}000 = 0$, or $\lambda = 5/123{,}000$.

The method of moments provides the same answer. The mean of the five losses is 123,000/5, and λ is the reciprocal of the mean. In many cases, the method of moments is not practicable, or it gives a different answer than the maximum likelihood procedure.

Maximum Likelihood and Minimum Bias Procedure

The rating model uses the classification relativities to determine the expected loss in each cell, or the mean loss in each cell. The maximum likelihood test is most practicable as a bias function when

- a single parameter distribution is used
- the mean of the distribution equals this parameter or some simple function of this parameter, such as its reciprocal
- the distribution extends over the positive real numbers
- the distribution is a reasonable reflection of some insurance process.

The exponential distribution and the Poisson distribution meet these conditions.

We illustrate a multiplicative model with the exponential distribution function. We use the same illustration as for the other models. The observed loss costs are shown below.

	Urban	Rural		$terr_1$	$terr_2$
Male	$800	$500	sex_1	$200 \times s_1 \times t_1$	$200 \times s_1 \times t_2$
Female	$400	$200	sex_2	$200 \times s_2 \times t_1$	$200 \times s_2 \times t_2$

Each class has an exponential distribution of loss costs. If the indicated pure premium is \$200, we don't expect every driver in that class to have losses of \$200 each year. Rather, we expect the observed losses to follow an exponential distribution with a mean of \$200.

The λ differs by cell. The reciprocal of λ is the indicated pure premium in that cell.

Illustration: For the urban/male cell, the loss costs have an exponential distribution with the parameter λ equal to $1/(\$200 \times s_1 \times t_1)$.

We choose starting values for $t_1 = 2.00$ and $t_2 = 1.00$. We determine the likelihood of the observed loss costs. The value of λ for the urban/male cell is $1/(200 \times s_1 \times t_1) = 1/(400 \times s_1)$. The likelihood of the \$800 loss cost in the urban/male cell is

$$1/400s_1 \times \exp(-800/400s_1) = 1/(400s_1) \times \exp(-2/s_1).$$

The value of λ for the rural/male cell is $1/(200 \times s_1 \times t_2) = 1/(200s_1)$. The likelihood of the \$500 loss cost in the rural/male cell is

$$1/200s_1 \times \exp(-500/200s_1) = 1/(200s_1) \times \exp(-2.5/s_1).$$

The likelihoods of the observed values in the female cells are determined in the same manner.

To maximize the likelihood, we maximize the logarithm of the likelihood, also knows as the loglikelihood.

- The likelihood of the set of four observed values is the *product* of the four individual likelihoods.
- The loglikelihood of the set of four observed values is the *sum* of the four individual loglikelihoods.

The partial derivative of the loglikelihood with respect to s_1 depends on the loglikelihoods in the male row only. This is the same simplification that we used for the least squares method and the χ-squared method.

The loglikelihood of the values in the male row is $-\ln(400s_1) - 2/s_1 - \ln(200s_1) - 2.5/s_1$. The partial derivative with respect to s_1 is $-1/s_1 + 2s_1^{-2} - 1/s_1 + 2.5s_1^{-2}$. We set this equal to zero.

$$-1/s_1 + 2s_1^{-2} - 1/s_1 + 2.5s_1^{-2} = 0$$
$$-s_1 + 2 - s_1 + 2.5 = 0$$
$$s_1 = 2.25.$$

The likelihood of the \$400 loss cost in the urban/female cell is

$$1/400s_2 \times \exp(-400/400s_2) = 1/(400s_2) \times \exp(-1/s_2).$$

The likelihood of the $200 loss cost in the rural/female cell is

$$1/200s_2 \times \exp(-200/200s_2) = 1/(200s_2) \times \exp(-1/s_2).$$

The loglikelihood of the values in the female row is $-\ln(400s_2) - 1/s_2 - \ln(200s_2) - 1/s_2$. The partial derivative with respect to s_2 is $-1/s_2 + 1s_2^{-2} - 1/s_2 + 1s_2^{-2}$. We set this equal to zero.

$$-1/s_2 + 1s_2^{-2} - 1/s_2 + 1s_2^{-2} = 0$$
$$-s_2 + 1 - s_2 + 1 = 0$$
$$s_2 = 1.00.$$

JUSTIFICATION

The maximum likelihood method has strong statistical justification. If the distribution of loss costs is a simple mathematical function, such as a Poisson distribution, a normal distribution, a lognormal distribution, or an exponential distribution, we can solve for simple recursive equations; see Brown [1988].

In practice, we don't know the proper distributions. The distributions that have been suggested for use in the minimum bias procedure, such as the exponential distribution, the Poisson distribution, and the normal distribution, are not assumed to be the correct distribution. They are tractable distributions that allow simple recursive functions.

If the tractable distribution is not a reasonable reflection of the true distribution, the use of the maximum likelihood method adds bias to the indicated relativities. The Poisson distribution and the normal distribution are not reasonable reflections of the true loss costs distribution. The true distribution has much fatter tails than these two mathematical functions indicate. Using these distributions with the maximum likelihood bias function may distort the solution.

RECURSIVE FUNCTIONS: PRACTICE SUMMARY

This *Practitioner's Guide* emphasizes intuition. Modern spreadsheet software has built-in iterative functions that perform the needed calculations. Mildenhall [1999] shows that the commonly used minimum bias recursive functions are equivalent to certain multiple regression equations or generalized linear models. Commercially available GLM models can be adapted to perform the minimum bias calculations.[43]

For each model discussed in this *Guide*, there are simple iterative functions. The task of the pricing actuary is to determine the type of rating function – multiplicative, additive, or combined – and the type of bias function (balance principle, least squares, χ-squared, or maximum likelihood). If the maximum likelihood bias function is used, the actuary must also select a probability distribution function for the loss costs (or other values) in each cell.

The type of data in each cell will generally be either loss costs or loss ratios. If the pricing actuary is using all the dimensions of the classification system in the minimum bias analysis, it is easiest to use loss costs. If there are significant classification dimensions that are not included, and if there may be an uneven distribution of exposures along these other classification dimensions, the pricing actuary may prefer to use loss ratios.

[43] As Mildenhall points out, the minimum bias procedures discussed here are a subset of the potential generalized linear models that can be used for classification ratemaking. Actuaries without a strong statistical background may find it difficult to understand the intuition for generalized linear models. We hope to complete the companion *Practitioner's Guide* on generalized linear models in the near future.

The table below shows the models which have been proposed for insurance use, followed by the recursive equations.

Rating Model	*Bias Function*	*Distribution Function*
multiplicative	balance principle	N/A
additive	balance principle	N/A
multiplicative	least squares	N/A
additive	least squares	N/A
multiplicative	χ-squared	N/A
additive	χ-squared	N/A
multiplicative	maximum likelihood	normal
additive	maximum likelihood	normal
multiplicative	maximum likelihood	exponential
multiplicative	maximum likelihood	Poisson

Multiplicative model, balance principle:

$$x_i = \frac{\sum_j n_{ij} r_{ij}}{\sum_j n_{ij} y_j}$$

Multiplicative model, balance principle, three dimensions:

$$x_i = \frac{\sum_{j,k} n_{ijk} r_{ijk}}{\sum_{j,k} n_{ijk} y_j z_k}$$

Additive model, balance principle:

$$x_i = \frac{\sum_j n_{ij}(r_{ij} - y_j)}{\sum_j n_{ij}}$$

Multiplicative model, least squares:

$$x_i = \sum (n_{ij} \times r_{ij} \times y_j) \div \sum (n_{ij} \times y_j^2).$$

Additive model, least squares:

$$x_1 = \sum(r_{1j} - y_{1j})/m - B$$

Multiplicative model, χ-squared:

$$x_i = [\sum (n_{ij} \times r_{ij}^2 / y_j) / \sum n_{ij} y_j]^{0.5}$$

Additive model, χ-squared:

$$\Delta x_i = \frac{\sum_j n_{i,j}\left(\frac{r_{i,j}}{x_i + y_j}\right)^2 - \sum_j n_{i,j}}{2\sum_j n_{i,j}\left(\frac{r_{i,j}}{x_i + y_j}\right)^2\left(\frac{1}{x_i + y_j}\right)}$$

Multiplicative model, maximum likelihood, normal density function:

$$x_i = \frac{\sum_j n_{ij}^2 r_{ij} y_j}{\sum_j n_{ij}^2 y_j^2}$$

Additive model, maximum likelihood, normal density function:

$$x_i = \frac{\sum_j n_{ij}^2 (r_{ij} - y_j)}{\sum_j n_{ij}^2}$$

Multiplicative model, maximum likelihood, exponential density function:

$$x_i = \frac{\sum_j \frac{r_{ij}}{y_j}}{k}$$

where "k" is the number of classes in the "j" dimension.

The recursive functions for a multiplicative model, maximum likelihood, Poisson distribution function are the same as those for the multiplicative model, balance principle.

Derivations of the formulas for the maximum likelihood models may be found in Brown [1988].

Conclusion

Accurate classification systems are the bedrock of insurance pricing. Accurate and unbiased rating systems enable insurers to attain competitive advantages over their peer companies. Inaccurate rating systems lead to unsatisfactory profits and losses of market share.

As competition in the insurance industry increases, and as companies are forced to rely on their own pricing prowess instead of bureau rates, the need for more accurate ratemaking procedures increases. The minimum bias procedure can be used to optimize a variety of rating models.

In the past, the iterative computational methods and the lack of clear documentation hindered many practicing actuaries from using minimum bias methods. The availability of built-in functions to perform iterative calculations in popular spreadsheets and programming languages has removed the major obstacle to effective use of minimum bias methods. This *Practitioner's Guide* provides clear documentation for actuaries desiring to implement minimum bias methods.