Estimating Mortality Risk Using Predictive Modeling

Tom Edwalds, FSA, ACAS, MAAA
Steve Craighead, ASA, MAAA
Philip Adams, ASA, MAAA

Copyright 2008 by the Society of Actuaries.
All rights reserved by the Society of Actuaries. Permission is granted to make brief excerpts for a published review. Permission is also granted to make limited numbers of copies of items in this monograph for personal, internal, classroom or other instructional use, on condition that the foregoing copyright notice is used so as to give reasonable notice of the Society's copyright. This consent for free limited copying without prior consent of the Society does not extend to making copies for general distribution, for advertising or promotional purposes, for inclusion in new collective works or for resale.

Abstract

As a life reinsurer, understanding mortality risk is mission critical. Superior models of mortality risk lead to more competitive pricing in profitable market segments and more efficient use of capital. Surprisingly, modern statistical tools such as predictive modeling have not previously been used to model mortality data to gain this edge. In this paper, we present the use of a predictive modeling technique called projection pursuit regression (PPR) to model mortality data. We first describe the process of creating a PPR model using inter-company mortality data collected by the Individual Life Experience Committee of the Society of Actuaries, and present the resulting model. We then discuss the process of using this model as a starting point for creating a more detailed PPR model of internal company mortality data.

1. Introduction

When implementing enterprise risk management within a company, one can take a bottom-up or top-down perspective. From a top-down view, we need to understand our different types of capital (such as target or economic) as well as their management and utilization. We also need to understand the integrated impact of various risks such as market, credit or operational risk on the determination and possible consumption of our existing capital. In ERM, once a top-down approach has been implemented, we begin to look for the major risks that the enterprise is exposed to outside of the general list of market, credit, interest, operational, etc. From this list, then, one either improves the underlying risk models or adds new ones.

Taking a bottom-up view, we would first determine all of the specific major risks, either by interviewing or through the use of the Delphi method within each major business function. Once these are determined, further analysis will lead one to either set up risk management strategies or, hopefully, if there is sufficient data, actual risk models, which allow the analyst the ability to associate financial value to these risks. If these financial models are available, then these would become components of the aggregate ERM models to determine the capital impact of these risks.

In this paper, we will take a bottom-up approach, by placing emphasis on our major risk, which is mortality. As a life reinsurer, our need to understand mortality risk is mission critical. If we are able to develop superior models of mortality risk, this will allow us to competitively price in profitable market segments and allow us to more effectively use capital. Surprisingly, modern statistical tools such as predictive modeling have not previously been used to model mortality data to gain this edge.

Predictive modeling is in general a process by which a model is created or chosen such that its characteristics are the best at predicting the probability of an outcome. There have been multiple approaches used within the casualty insurance industry such as Generalized Linear Models (GLM) or Generalized Additive Models (GAM) to improve various pricing models.

In this paper, we present the use of a predictive modeling technique called projection pursuit regression (PPR) to model mortality data. PPR is a form of GAM models, but PPR has extra flexibility allowing one to model interactions between various predictors without requiring additional effort. We have found that this technique is extremely effective because of its ability to extract complex relationships within large datasets.

To illustrate the PPR technique, we begin by constructing a model of mortality for the U.S. life insurance industry. The underlying data used in developing this model is the Society of Actuaries Individual Life Experience Committee's 2002–2004 Intercompany Study (SOA 2002–2004 Study). This study includes $7.4 trillion in exposure by amount, over 75 million life-years of exposure by policy count, and nearly 700,000 death claims from 35 contributing companies. It includes over 200,000 deaths in the select period and almost 500,000 deaths in the ultimate period. The new technology of predictive modeling is designed to yield meaningful information from large, complex datasets such as this. The attributes of this data are discussed further in Section 2.

From this data, we developed several PPR models to model first the ultimate mortality experience and, based on these results, we then modeled the select data. These models are discussed in Section 3.

In Section 4, we then discuss the process of using these models as a starting point for creating more detailed PPR or other predictive models of our internal company mortality data and how we can use these models in business management as well as in ERM.

In Section 5, we discuss the strengths and weaknesses of using PPR, our conclusions, and future research and development.

In Appendix A, we give an outline of the projection pursuit regression algorithm.

In Appendix B, we describe the use of the ppr algorithm as implemented in the statistical software package R.

In Appendix C, we give a derivation of the Poisson weights, which are the weights used within the PPR models we construct.

2. Data Collection

The Individual Life Experience Committee (ILEC) of the Society of Actuaries (SOA) regularly collects data on the mortality experience of those U.S. life insurers who volunteer to submit their data. The data collection effort for the 2002–2004 experience period was particularly successful, with 35 companies contributing over 75 million life-years of exposure by policy count totaling $7.4 trillion in exposure by amount. This dataset includes nearly 700,000 death claims: over 200,000 deaths in the select period and almost 500,000 deaths in the ultimate period. Due to the confidential nature of the contributed data, the full detail of this dataset can only be accessed by the SOA Research Department staff and by the staff at MIB Solutions, the data collection agent for the SOA. The data from which the model in this paper is built was prepared from the SOA 2002–2004 Study by MIB Solutions for the ILEC as two separate datasets: an ultimate durations dataset and a select durations dataset.

The ultimate durations dataset included all of the experience at durations 26 and higher for all attained ages, plus all experience at durations 3 and higher for attained ages 90 and higher. Note that there was no experience in the study at durations 1 and 2 for attained ages 90 and higher. The data was grouped into cells by individual attained ages, gender and face amount band (five bands: under $10,000; $10,000 to $24,999; $25,000 to $49,999; $50,000 to $99,999; and $100,000 and up). The values provided were: policies exposed, amounts exposed, actual deaths, actual claim amounts, calculated qx by policy, calculated qx by amount, expected deaths on the 2001 VBT basis, expected claim amounts on the 2001 VBT basis, calculated actual-to-expected ratio (A/E) by policy on the 2001 VBT basis and calculated A/E by amount on the 2001 VBT basis.

The select durations dataset included all of the experience at durations 1–25 for attained ages 0–89. The data was grouped into cells by individual attained ages and durations, gender,

smoker status and face amount band (nine bands: under $10,000; $10,000 to $24,999; $25,000 to $49,999; $50,000 to $99,999; $100,000 to $249,999; $250,000 to $499,999; $500,000 to $999,999; $1,000,000 to $2,499,999; and $2,500,000 and up). The values provided were the same as for the ultimate durations dataset: policies exposed, amounts exposed, actual deaths, actual claim amounts, calculated qx by policy, calculated qx by amount, expected deaths on the 2001 VBT basis, expected claim amounts on the 2001 VBT basis, calculated A/E by policy on the 2001 VBT basis and calculated A/E by amount on the 2001 VBT basis.

In Table 1, observe the fields contained within the separate select and ultimate data sets. Note: The ultimate data set does not contain Smoker Status since the smoker status of duration 26+ policies is unknown. In addition, the ultimate data does not contain the duration field either.

TABLE 1
Relevant Fields of Mortality Dataset

Field Name	Data Type	Descriptor
Attained Age	Numeric	Attained age of policy holder
Face Amount Band	Categorical	Face amount bands
Gender	Categorical	Male or Female
Smoker Status	Categorical	Non-Smoker, Smoker, Unknown
Duration	Numeric	Duration of policy
Policies Exposed	Numeric	Number of policies exposed
Actual Deaths	Numeric	Number of deaths
AEpol	Numeric	Per policy A/E ratio
WAEpol	Numeric	Associated weight of AEpol

As previously mentioned, the data is separated into select and ultimate. First, we examine the underlying ultimate data, since we use the output of the ultimate model to construct the select model.

2.1 Ultimate Data

The data to create the PPR models in Section 3.1 is restricted to attained ages (AA) ranging from 29 to 89. The original data was wider, but we decided to restrict to this range due to the structural break in the data at age 90 and inadequate volume of data below age 29. In the 29 to 89 age range all of the ultimate durations data is at durations 26+. The face amount bands (FAB) are: "$1–$9,999," "$10,000–$24,999" and "$25,000+." All of these policies have an Unknown Smoker Status.

Table 2 contains the basic statistics on the policies within the ultimate data. The A/E Ratio is based on the 2001 VBT table.

TABLE 2
Ultimate Policy Statistics

	Attained Age	Policies Exposed	Actual Deaths	A/E Ratio
Minimum	29	54.6	0.0	0.0%
1st Quartile	44	403.5	2.2	66.3%
Median	59	2,710.5	31.0	86.7%
Mean	59	32,1354	695.7	88.1%
3rd Quartile	74	40,8894	175.8	105.9%
Maximum	89	196,441.4	12,007.0	919.1%

In Table 3, observe the statistics of the exposed policies and total actual deaths within the data split by gender and face amount band. The ultimate dataset contains 610 records representing 19,602,611 policies and 424,384 deaths.

TABLE 3
Ultimate Policy Statistics by Sex and Face Amount Band

Gender	Face Amount	Exposure	Actual Deaths
Females	$1–$9,999	7,052,843	141,945
Females	$10,000–$24,999	760,630	5,497
Females	$25,000 and Over	85,788	1,180
	Females Total	**7,899,261**	**148,622**
Males	$1–$9,999	7,969,243	215,864
Males	$10,000–$24,999	3,153,884	49,995
Males	$25,000 and Over	580,223	9,903
	Males Total	**11,703,350**	**275,762**
	Grand Total	**19,602,611**	**424,384**

In the next section, we examine the select data.

2.2 Select Data

In the select data, the attained ages range from 0 to 89 and the durations range from 1 to 25. The FAB are: "$1–$9,999," "$10,000–$24,999," "$25,000–$49,999," "$50,000–$99,999," "$100,000–$249,999" and "$250,000+."

In Table 4, observe the summary statistics of the policies.

TABLE 4
Select Policy Statistics

	Attained Age	Duration	Policies Exposed	Actual Deaths	A/E Ratio
Minimum	0	1	0.00027	0.0	0.0%
1st Quartile	30	6	8.0	0.0	0.0%
Median	50	11	80.3	0.0	0.0%
Mean	49	12	644.1	2.4	99.2%
3rd Quartile	68	18	522.8	2.0	86.3%
Maximum	89	25	20,000.0	183.0	250,000.0%

In Table 5, observe the statistics of the exposed policies and actual deaths within the data split by gender, smoker status and face amount band.

TABLE 5
Select Policy Statistics by Smoker Status, Sex and Face Amount Band

Smoker Status	Gender	Face Amount Band	Exposure	Actual Deaths
Smoker	Male	$1–$9,999	191,512	4,505
Smoker	Male	$10,000–$24,999	682,983	8,098
Smoker	Male	$25,000–$49,999	721,312	6,293
Smoker	Male	$50,000–$99,999	1,012,244	6,280
Smoker	Male	$100,000–$249,999	1,114,460	4,269
Smoker	Male	$250,000 and Over	372,254	1,048
Smoker	Male	Total	4,094,765	30,493
Smoker	Female	$1–$9,999	344,480	8,473
Smoker	Female	$10,000–$24,999	751,849	6,469
Smoker	Female	$25,000–$49,999	582,868	2,837
Smoker	Female	$50,000–$99,999	708,757	2,363
Smoker	Female	$100,000–$249,999	589,685	1,365
Smoker	Female	$250,000 and Over	126,779	338
Smoker	Female	Total	3,104,418	21,845
		Smoker Total	**7,199,183**	**52,338**
Non-Smoker	Male	$1–$9,999	439,804	6,903
Non-Smoker	Male	$10,000–$24,999	2,265,277	12,245
Non-Smoker	Male	$25,000–$49,999	2,775,073	12,667
Non-Smoker	Male	$50,000–$99,999	4,505,110	15,638
Non-Smoker	Male	$100,000–$249,999	7,071,060	15,639
Non-Smoker	Male	$250,000 and Over	5,539,328	6,810
Non-Smoker	Male	Total	22,595,652	69,902
Non-Smoker	Female	$1–$9,999	875,068	11,435
Non-Smoker	Female	$10,000–$24,999	2,789,220	10,427
Non-Smoker	Female	$25,000–$49,999	2,766,953	6,332
Non-Smoker	Female	$50,000–$99,999	4,009,797	6,872
Non-Smoker	Female	$100,000–$249,999	5,376,581	5,604
Non-Smoker	Female	$250,000 and Over	2,474,775	1,960
Non-Smoker	Female	Total	18,292,394	42,630
		Non-Smoker Total	**40,888,046**	**112,532**

TABLE 5 (continued)
Select Policy Statistics by Smoker Status, Sex and Face Amount Band

Smoker	Gender	Face Amount Band	Exposure	Actual Deaths
Unknown	Male	$1–$9,999	545,335	8,186
Unknown	Male	$10,000–$24,999	1,422,408	6,214
Unknown	Male	$25,000–$49,999	727,887	1,754
Unknown	Male	$50,000–$99,999	328,545	858
Unknown	Male	$100,000–$249,999	157,110	560
Unknown	Male	$250,000 and Over	22,603	98
Unknown	Male	Total	3,203,888	17,670
Unknown	Female	$1–$9,999	953,975	14,869
Unknown	Female	$10,000–$24,999	1,339,343	3,328
Unknown	Female	$25,000–$49,999	632,298	604
Unknown	Female	$50,000–$99,999	260,837	291
Unknown	Female	$100,000–$249,999	107,398	168
Unknown	Female	$250,000 and Over	11,924	23
Unknown	Female	Total	3,305,775	19,283
		Unknown Total	**6,509,663**	**36,953**
Total	Male		29,894,305	118,065
Total	Female		24,702,587	83,758
Total		$1–$9,999	3,350,174	54,371
Total		$10,000–$24,999	9,251,080	46,781
Total		$25,000–$49,999	8,206,391	30,487
Total		$50,000–$99,999	10,825,290	32,302
Total		$100,000–$249,999	14,416,294	27,605
Total		$250,000 and Over	8,547,663	10,277
		Grand Total	**54,596,892**	**201,823**

The select dataset has 84,764 records representing 201,823 deaths and total policy exposure of 54,596,892.

We now progress to the construction of the model based on the ultimate data.

3. Model Description

3.1 Ultimate

In this section, we develop the ultimate projection pursuit regression (PPR) model based on the ultimate data described in Section 2.1. Please refer to Appendix A for an overall explanation of the various properties of PPR.

The response variable will be the A/E per policy ratio. This ratio is defined as $AEpol = \frac{Actual_Deaths}{Policies_Exposed \times q_x^{VBT}}$. Here the q_x^{VBT} mortality rate is the composite 2001 Valuation Basic Table ultimate rate.

To fit the model, we used the ppr function in the R package stats, with the following formula: AEpol ~ AttainedAge+Gender+FaceAmountBand. This formula displays only three predictors, but Gender and FAB are categorical variables. The ppr algorithm in R converts multiple levels of categorical variables into additional predictor variables. For this specific model, the predictor variables are in Table 6.

TABLE 6
Ultimate PPR Predictors

X1	Attained Age
X2	Gender: Female
X3	Gender: Male
X4	FAB: $10K–$25K
X5	FAB: $25K+

For example, for a male with attained age 40 with a face amount of $15,000, the predictor vector is coded as (X1 = 40, X2 = 0, X3 = 1, X4 = 1, X5 = 0). Note how the binary nature of the variable is indicated by 0 as absent and 1 as present. For clarity, take another example. Assume that there is a female with attained age 55, and a face amount of $5,000. The predictor vector would be coded as (X1 = 55, X2 = 1, X3 = 0, X4 = 0, X5 = 0). Note how the lowest FAB is considered the default setting of the model and so the higher bands are coded with a 0.

In the determination of the model, ppr uses weighted least squares, and the responses are weighted by $Policies_Exposed \times q_x^{VBT}$. The derivation of these weights is contained in Appendix C.

The model is described in Table 7. Observe the seven α projection vectors and the corresponding β weights. In Figure 1, you can observe the seven ridge functions for the ultimate PPR model.

This PPR model can produce negative A/E Ratios, and we floor the results at zero, but we have found this model to be extremely useful because it parameterizes and smoothes the actual data.

Though there are seven ridge functions and one could consider the formula as too flexible and hard to fully understand, we can still look at the major components and gain some insight into what influences the responses.

TABLE 7
A/E Policy Ultimate PPR Parameters

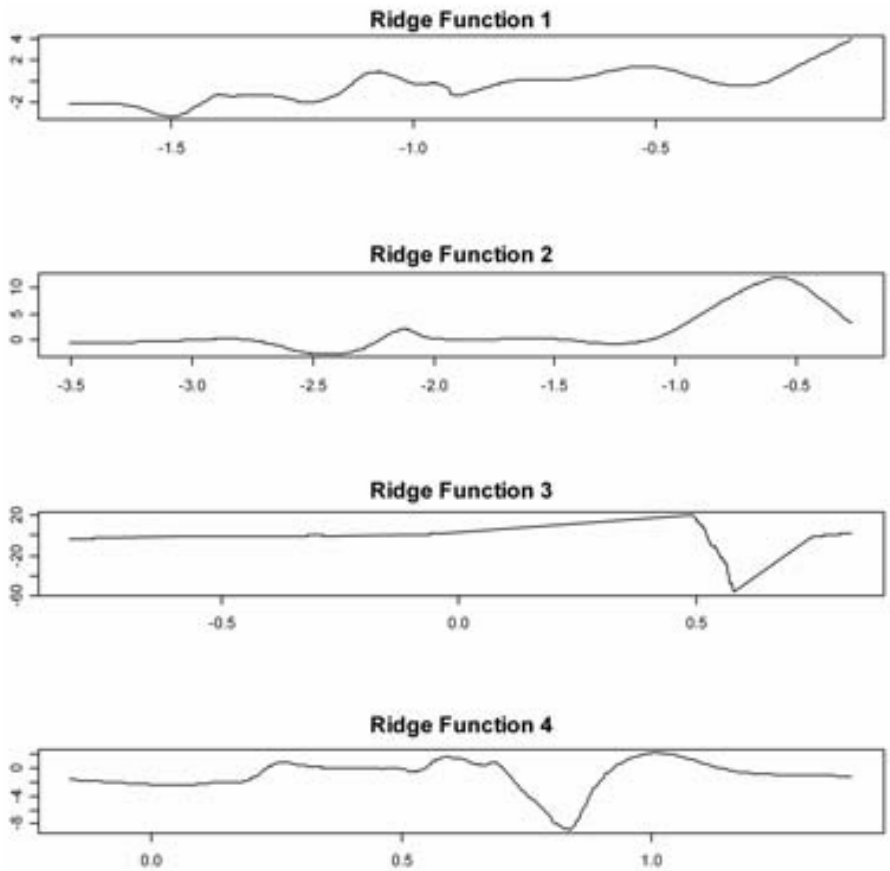
	α_1	α_2	α_3	α_4
Attained_Age	-0.008042938	-0.028385108	-0.001469583	0.009533508
GenderFemale	0.137395775	-0.805290783	-0.239808299	-0.260636678
GenderMale	-0.166948620	0.551038151	0.008455668	-0.438233960
Face_Amount$10,000-$24,999	-0.523370064	-0.131924488	-0.447968477	0.811852060
Face_Amount$25,000 and Over	-0.824178164	-0.172203156	0.861244776	0.284297140
	α_5	α_6	α_7	
Attained_Age	-0.033419098	-0.052332058	-0.003441216	
GenderFemale	0.438724899	0.077571604	-0.496024138	
GenderMale	0.066093973	0.378422791	-0.436555622	
Face_Amount$10,000-$24,999	0.839568700	-0.538174543	-0.193073652	
Face_Amount$25,000 and Over	0.311704365	-0.747267257	0.725320595	
Coefficients of ridge terms:				
β_1	0.080833163			
β_2	0.047155409			
β_3	0.011981005			
β_4	0.011425075			
β_5	0.011428112			
β_6	0.008777865			
β_7	0.003059683			
$\overline{AE}$	0.9676315			
Goodness of fit:	0.7192942			

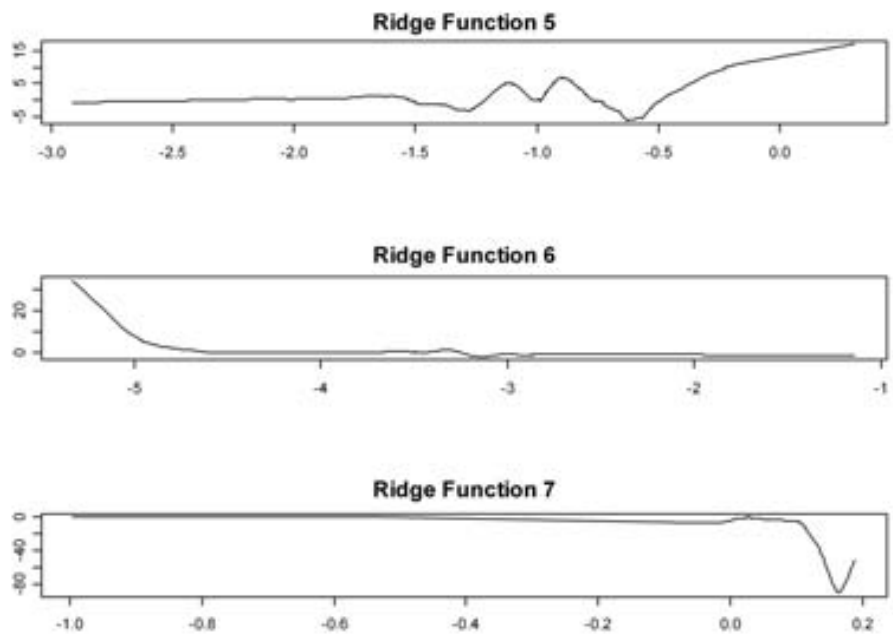
Our ultimate PPR model is:

$$AE = \overline{AE} + \sum_{i=1}^{7} \beta_i \cdot f_i\left(\sum_{k=0}^{5} \alpha_{ik} X_k\right) \qquad (1)$$

Observe that the second half of the formula is the calculation of the offset to the average of the A/E ratio, which is 0.9676315. Since the ppr algorithm is coded to assure that $E\left[f_i(\alpha_i \cdot X)\right] = 0$ and $E\left[f_i^2(\alpha_i \cdot X)\right] = 1$, the magnitude of the β_i determines which ridge function f_i contributes the most influence on the overall PPR model. Since $\beta_1 \approx 0.081$ and $\beta_2 \approx 0.047$, if we examine the terms $\beta_1 f_1(\alpha_1 \cdot X)$ and $\beta_2 f_2(\alpha_2 \cdot X)$, we should understand the major components that make up the model.

Figure 1
Ultimate Per Policy A/E Ratio Ridge Functions





Observe in Figure 1, the graph of f_1. We see that it is mostly an increasing function. It has a domain of $(-1.7,-.1)$ and a range of $(-3.3,4.0)$.

If one examines $\alpha_1 \cdot X$ and determines what characteristics within the data allow $\sum_k^5 \alpha_{1k} X_k$ to take on different values in the domain of f_1, then the understanding of the overall PPR model will improve. Note that all of the α_{1k} are negative except for the α that corresponds to the Gender:Female predictor. So, one thing that we can observe is that a male with a high FAB and a high AA will lower the value of $\alpha_1 \cdot X$. However, a female with a low AA and low FAB will make the $\alpha_1 \cdot X$ high. So $\beta_1 f_1(\alpha_1 \cdot X)$ will be negative for males with increasing FAB and AA and hence the A/E ratio is reduced from the average, whereas females with low FAB and low AA will augment the average A/E Ratio.

Since β_2 is also significant in magnitude, we need to look at the $\beta_2 f_2(\alpha_2 \cdot X)$ component as well. The domain of f_2 is $(-3.6,-0.25)$ and the range is $(-2.8,12.1)$. Note that from Figure 1 that f_2 is mostly flat with the value of 0 for $x < -1$, where when $x \in [-1,-0.4]$ f_2 rises sharply to 12. As x moves to the right of the domain the range declines to approximately 5.0. Since f_2 is mostly zero on the range, we need to see what the behavior of $\sum_{k=1} \alpha_{2k} X_k$ produces values in the subdomain $(-1,-.25)$. Since all of the α_{2k} are negative except the one that corresponds to Gender:Male, the value of $\alpha_2 \cdot X$ will be the highest when evaluating males with low FAB and lower AA. So, in this situation the average A/E ratio will be augmented.

In summary the PPR model will move the average A/E ratio down from the average for males with high FAB and high AA, where the value will be augmented if the policy is for a female with low FAB and low AA. To a lesser extent, the second component will push the PPR value above the average when examining male policy holders with low AA and low FAB.

3.2 Adjusted Ultimate

The ILEC ultimate experience contains exposures and deaths for males and females distinct by attained ages and five face amount bands. The PPR model was fitted onto the A/E ratio, where the expecteds are taken with respect to the ultimate 2001 VBT. The model was moreover restricted to attained ages 29 to 89, as the exposure volume was sufficiently high to provide acceptable fits. While the model was distinct by band, the top three bands (over $25,000) were consolidated into a single category to ensure that sufficient data volume was available for the PPR fit, while the other bands ($1–$10K, $10–$25K) appeared to have sufficient volume for most cells.

We wish to accomplish the following:

- Banding consistency: ensure that that the banding effect is consistent for every attained age, i.e., as face amount band increases, the mortality rate is non-increasing.
- Band splitting: split the top band where possible into $25–$50K, $50–$100K and greater than $100K bands, as seen in the original data.
- Juvenile extension: extend the PPR model down to juvenile ages by smoothly blending the 2003 Social Security Administration mortality rates into the PPR model.
- Older-age extension: extend the PPR model upward to attained age 120 by fitting a separate model to ages 89 and above and grading the rate at the terminal age of this model to a rate of .5 at attained age 120.

In the foregoing, we aim to modify the model results as little as possible in order to arrive at an ultimate mortality table that is consistent with both intuition and experience.

3.2.1 Ensuring Banding Consistency

Experience has indicated that for a given attained age, one should see mortality rates decline as face amount increases. The PPR model typically exhibits this pattern, except among the following groups of attained age, band and gender combinations:

- Males, ages 29 to 36, bands <$10K and $10–$25K
- Females, ages 44 to 52, bands <$10K and $10–$25K
- Males, ages 84 to 89, bands $10–$25K and $25K+
- Females, ages 81 to 88, bands $10–$25K and $25K+
- Females, age 89, bands <$10K, $10–$25K and $25K+

These groups typically had inverted banding relationships (e.g., band $10–$25K had higher mortality than band <$10K, etc.) To resolve this issue, the experience within the affected groups was blended by attained age, and the blended experience was substituted for the PPR rates.

3.2.2 Band Splitting

For males, it was discovered that credible data exists to split the top band at $50,000. To determine the split, we examined actual-to-PPR-expected ratios for bands $25–$50K and $50K and up, and applied those ratios to the top band of the male PPR model. Those ratios were 101.72 percent and 93.9 percent respectively.

For females, it was discovered that there is no credible band split above $25,000. Thus, females retain a three-band ultimate model.

3.2.3 Juvenile Extension

To extend the model to age 0, we augmented the model results with the 2003 SSA mortality rates. Starting with attained age 10, the SSA rates were graded into the PPR model. To capture the differing juvenile mortality risk profiles of males and females, we resorted to different interpolation methods to yield mortality rates consistent with intuition and experience.

For males, we computed the ratio of the age 29 PPR mortality rate to the age 29 SSA mortality rate for each band. Starting with attained age 10, we linearly interpolated the ratio to age 29 and applied it to the SSA rate. This allowed the SSA rate to smoothly blend into the PPR model.

For females, it was determined that linearly blending the SSA rates into age 29 of the PPR model created unusual and unintuitive results. Thus, the ratio was taken at age 36, and an exponential interpolation was used between ages 10 and 36.

3.2.4 Older-Age Extension

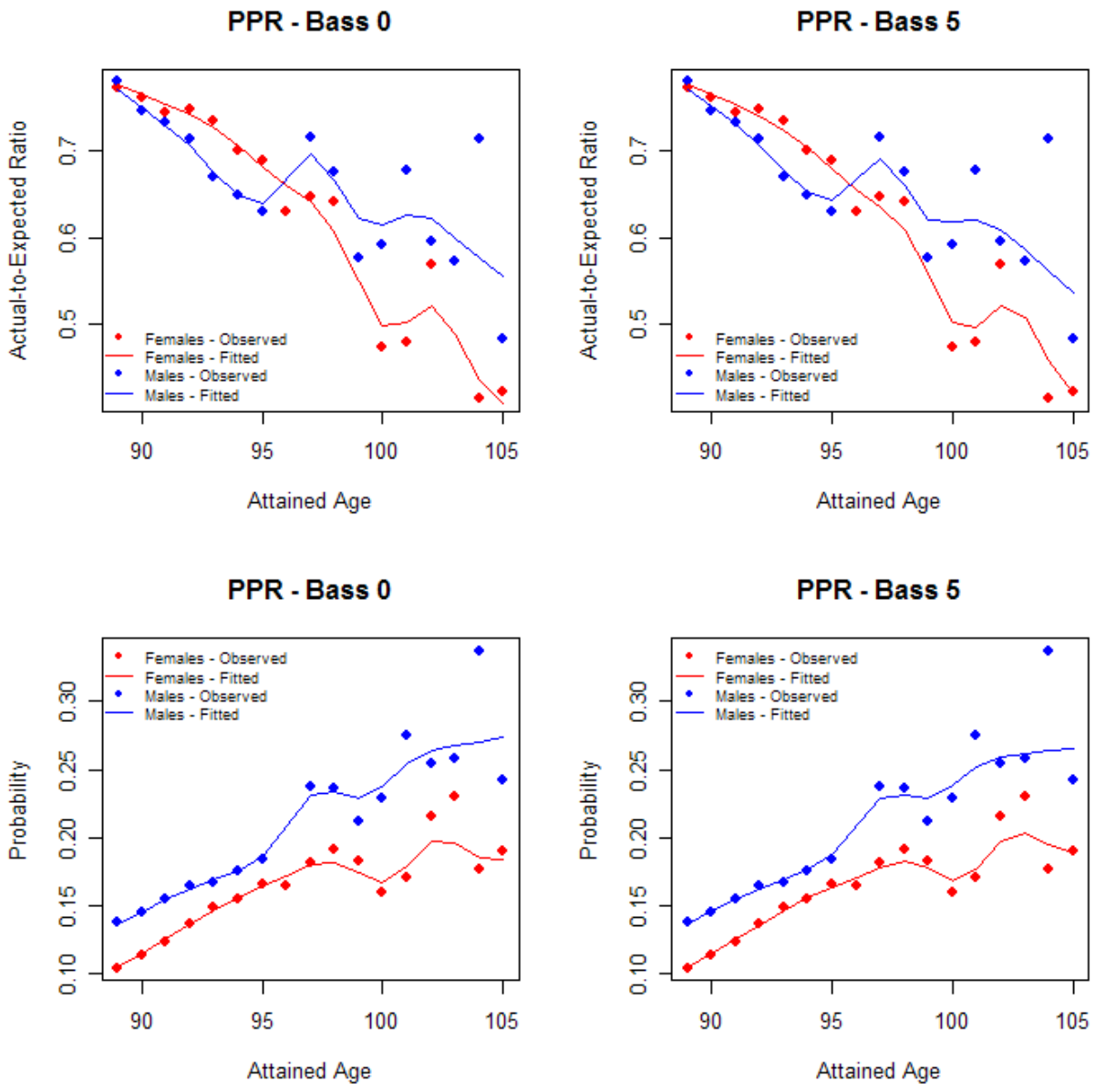
The PPR model was stopped at attained age 89 due to its being sensitive to increasing fluctuations above this age. The data volume is already diminishing by age 89 and thins considerably by attained age 100. As a guide, we expect that above age 89, mortality rates steadily increase and then level off after some late age.

We fit two models to data for attained ages 89 and higher, excluding attained age 96 for males, since the data for males attained age 96 seems to exhibit anomalously low mortality relative to ages 95 and 97. In the first model, we constructed a model analogous to the 29–89 PPR model by fitting a PPR model to the A/E ratios, where “E” is the expected mortality with respect to the 2003 SSA mortality rates. Note that band was no longer credible at these ages, as most experience was in the lowest face amount band. For attained age 89–93, the fit was excellent. Otherwise, the fitted values were prone to fluctuation, and we had to increase the model smoothness value (bass of 5). Increasing the smoothness is one of two departures from the

model parameters of the 29–89 model. The other is that on the basis of goodness-of-fit, only two ridge functions were employed in the model.

Figure 2 compares the two PPR models.

Figure 2
PPR Model Fit Graphs



Our chosen model is a logistic regression model. A priori, we are aware that males with attained age 96 appeared to have anomalously low mortality given the exposure, and hence they were excluded from any modeling steps. In the course of modeling, we found that the logistic model including both gender and attained age fits well when restricted to ages 89 through 105, and moreover, the attained age variable can be modeled as a continuous predictor at the cost of a mild reduction in fit. This model's AIC is 452.48, while that of the saturated model with categorical attained age is 402.8. The drop in AIC is significant, but it appears that the saturated model overfits the data.

Figure 3 compares the two models.

Figure 3
GLM Model Fit Graphs

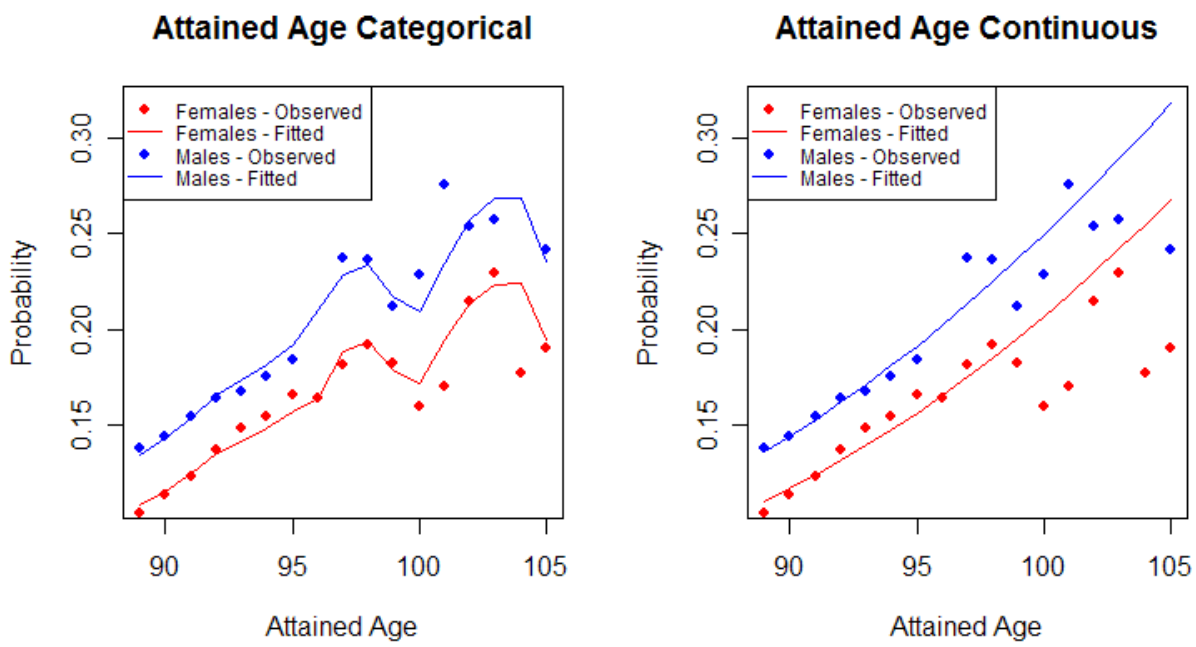


Table 8 displays the parameters of the chosen model.

TABLE 8
GLM Parameters

	Estimate	Std. Error	p-value
Intercept	-8.1353	0.1298	< 2e-16
Attained_Age	0.0679	0.0014	< 2e-16
Gender:Male	0.2437	0.0073	< 2e-16

We obtain predicted values for attained ages 89 through 114. Since the slopes differ mildly between the later PPR model and early GLM model, we use an exponential blend of age 89 in the PPR model and age 93 in the GLM to smooth the transition for ages 90 through 92. From attained age 115 to 120, we exponentially interpolate the last value of the GLM model to 0.5 at attained age 120.

In Table 9 and Table 10 observe the face amount banded male and female mortality rates for ages 0 to 89. These rates are then used to create the relative weighting used in the PPR select model described in the next section.

TABLE 9
Ultimate Per Policy Male Mortality

Attained Age	FAB < 10K	FAB 10–25K	FAB 25–50K	FAB > 50K
0	0.00717	0.00717	0.00717	0.00717
1	0.00049	0.00049	0.00049	0.00049
2	0.00033	0.00033	0.00033	0.00033
3	0.00027	0.00027	0.00027	0.00027
4	0.00021	0.00021	0.00021	0.00021
5	0.00019	0.00019	0.00019	0.00019
6	0.00018	0.00018	0.00018	0.00018
7	0.00016	0.00016	0.00016	0.00016
8	0.00015	0.00015	0.00015	0.00015
9	0.00012	0.00012	0.00012	0.00012
10	0.00011	0.00011	0.00011	0.00011
11	0.00011	0.00011	0.00011	0.00011
12	0.00016	0.00016	0.00016	0.00016
13	0.00027	0.00027	0.00026	0.00026
14	0.00042	0.00042	0.0004	0.0004
15	0.00057	0.00057	0.00055	0.00054
16	0.00072	0.00072	0.00068	0.00066
17	0.00085	0.00085	0.00079	0.00077
18	0.00095	0.00095	0.00088	0.00086
19	0.00103	0.00103	0.00094	0.00092
20	0.00112	0.00112	0.00101	0.00097
21	0.00119	0.00119	0.00106	0.00102
22	0.00123	0.00123	0.00108	0.00104
23	0.00121	0.00121	0.00106	0.00101
24	0.00117	0.00117	0.001	0.00095
25	0.00111	0.00111	0.00094	0.00089
26	0.00106	0.00106	0.00089	0.00083
27	0.00103	0.00103	0.00085	0.00079
28	0.00102	0.00102	0.00083	0.00077
29	0.00103	0.00103	0.00082	0.00076
30	0.00113	0.00113	0.00088	0.00081
31	0.00122	0.00122	0.00094	0.00087

TABLE 9 (continued)
Ultimate Per Policy Male Mortality

Attained Age	FAB < 10K	FAB 10–25K	FAB 25–50K	FAB > 50K
32	0.00128	0.00128	0.00102	0.00094
33	0.00136	0.00136	0.00106	0.00098
34	0.00145	0.00145	0.00109	0.001
35	0.00155	0.00155	0.00112	0.00104
36	0.00164	0.00164	0.00117	0.00108
37	0.00176	0.00171	0.00124	0.00114
38	0.00191	0.00183	0.00131	0.00121
39	0.0021	0.00197	0.0014	0.00129
40	0.00228	0.00209	0.0015	0.00138
41	0.00245	0.00224	0.00161	0.00148
42	0.00262	0.00236	0.00172	0.00159
43	0.00284	0.00243	0.00183	0.00169
44	0.00309	0.00255	0.00196	0.00181
45	0.00333	0.0027	0.00213	0.00197
46	0.0035	0.00286	0.00232	0.00215
47	0.00363	0.00297	0.0025	0.00231
48	0.00379	0.00306	0.00263	0.00242
49	0.00403	0.00318	0.00273	0.00252
50	0.00432	0.00339	0.00289	0.00266
51	0.0046	0.00367	0.00311	0.00287
52	0.00487	0.00397	0.00337	0.00311
53	0.00518	0.00427	0.00365	0.00337
54	0.00551	0.00462	0.00394	0.00363
55	0.00583	0.00498	0.00426	0.00393
56	0.00619	0.00535	0.00462	0.00427
57	0.00661	0.00574	0.00501	0.00462
58	0.00718	0.00617	0.00539	0.00498
59	0.00786	0.00676	0.00585	0.0054
60	0.00863	0.00745	0.00643	0.00594
61	0.00946	0.00825	0.00712	0.00658
62	0.01037	0.00908	0.00792	0.00731
63	0.01137	0.00998	0.00878	0.0081
64	0.01252	0.01099	0.00972	0.00898

TABLE 9 (continued)
Ultimate Per Policy Male Mortality

Attained Age	FAB < 10K	FAB 10–25K	FAB 25–50K	FAB > 50K
65	0.01384	0.01213	0.01076	0.00993
66	0.01533	0.01347	0.01198	0.01106
67	0.01696	0.01495	0.01337	0.01234
68	0.01871	0.01658	0.01492	0.01377
69	0.02072	0.01834	0.01662	0.01534
70	0.02293	0.02033	0.0185	0.01708
71	0.02539	0.02254	0.02063	0.01904
72	0.02805	0.02502	0.02305	0.02128
73	0.03096	0.02769	0.02581	0.02382
74	0.03417	0.03064	0.02887	0.02664
75	0.03764	0.03391	0.03231	0.02982
76	0.04145	0.03743	0.03608	0.03331
77	0.04549	0.04123	0.04003	0.03695
78	0.05013	0.04532	0.04423	0.04082
79	0.05521	0.04995	0.04854	0.0448
80	0.06082	0.05509	0.05335	0.04925
81	0.0667	0.06078	0.05884	0.05431
82	0.0734	0.06677	0.06519	0.06017
83	0.08059	0.07355	0.07247	0.0669
84	0.0887	0.08063	0.08063	0.08063
85	0.09718	0.08917	0.08917	0.08917
86	0.10619	0.09826	0.09826	0.09826
87	0.11534	0.10786	0.10786	0.10786
88	0.12451	0.11765	0.11765	0.11765
89	0.13368	0.12754	0.12754	0.12754

TABLE 10
Ultimate Per Policy Female Mortality

Attained Age	FAB < 10K	FAB 10–25K	FAB 25–50K	FAB > 50K
0	0.00596	0.00596	0.00596	0.00596
1	0.00041	0.00041	0.00041	0.00041
2	0.00027	0.00027	0.00027	0.00027
3	0.0002	0.0002	0.0002	0.0002
4	0.00015	0.00015	0.00015	0.00015
5	0.00014	0.00014	0.00014	0.00014
6	0.00014	0.00014	0.00014	0.00014
7	0.00013	0.00013	0.00013	0.00013
8	0.00013	0.00013	0.00013	0.00013
9	0.00012	0.00012	0.00012	0.00012
10	0.00011	0.00011	0.00011	0.00011
11	0.00012	0.00012	0.00012	0.00012
12	0.00013	0.00013	0.00012	0.00012
13	0.00014	0.00014	0.00013	0.00013
14	0.00015	0.00015	0.00014	0.00014
15	0.00017	0.00016	0.00015	0.00015
16	0.00018	0.00018	0.00015	0.00015
17	0.0002	0.00019	0.00016	0.00016
18	0.00021	0.00021	0.00017	0.00017
19	0.00023	0.00022	0.00018	0.00018
20	0.00025	0.00024	0.00019	0.00019
21	0.00028	0.00026	0.0002	0.0002
22	0.0003	0.00028	0.00021	0.00021
23	0.00033	0.0003	0.00023	0.00023
24	0.00035	0.00033	0.00024	0.00024
25	0.00038	0.00036	0.00025	0.00025
26	0.00042	0.00039	0.00027	0.00027
27	0.00045	0.00042	0.00028	0.00028
28	0.00049	0.00045	0.0003	0.0003
29	0.00054	0.00049	0.00032	0.00032
30	0.00058	0.00053	0.00033	0.00033
31	0.00063	0.00057	0.00035	0.00035

TABLE 10 (continued)
Ultimate Per Policy Female Mortality

Attained Age	FAB < 10K	FAB 10–25K	FAB 25–50K	FAB > 50K
32	0.00069	0.00062	0.00037	0.00037
33	0.00075	0.00067	0.0004	0.0004
34	0.00081	0.00072	0.00042	0.00042
35	0.00089	0.00078	0.00044	0.00044
36	0.00096	0.00084	0.00047	0.00047
37	0.00102	0.00089	0.00055	0.00055
38	0.00108	0.00095	0.00061	0.00061
39	0.00115	0.00103	0.00069	0.00069
40	0.00121	0.00113	0.00076	0.00076
41	0.00128	0.00123	0.00084	0.00084
42	0.00137	0.00134	0.00092	0.00092
43	0.00148	0.00145	0.00101	0.00101
44	0.00158	0.00158	0.00112	0.00112
45	0.00169	0.00169	0.00121	0.00121
46	0.00179	0.00179	0.00129	0.00129
47	0.00194	0.00194	0.00139	0.00139
48	0.00211	0.00211	0.00152	0.00152
49	0.00226	0.00226	0.00172	0.00172
50	0.00237	0.00237	0.00198	0.00198
51	0.00247	0.00247	0.00223	0.00223
52	0.00263	0.00263	0.00242	0.00242
53	0.00288	0.00283	0.00261	0.00261
54	0.0032	0.00306	0.00288	0.00288
55	0.00353	0.00335	0.0032	0.0032
56	0.0039	0.00367	0.00355	0.00355
57	0.00433	0.00402	0.00385	0.00385
58	0.00478	0.00444	0.00415	0.00415
59	0.00523	0.00488	0.00446	0.00446
60	0.00572	0.00531	0.00477	0.00477
61	0.0063	0.00577	0.00506	0.00506
62	0.00702	0.0063	0.00541	0.00541
63	0.00786	0.00697	0.00601	0.00601
64	0.00876	0.00777	0.00677	0.00677

TABLE 10 (continued)
Ultimate Per Policy Female Mortality

Attained Age	FAB < 10K	FAB 10–25K	FAB 25–50K	FAB > 50K
65	0.00969	0.00872	0.00772	0.00772
66	0.01068	0.00976	0.00876	0.00876
67	0.01176	0.01094	0.00996	0.00996
68	0.01298	0.01218	0.01122	0.01122
69	0.01435	0.01359	0.01255	0.01255
70	0.01588	0.01509	0.01403	0.01403
71	0.0175	0.01662	0.01572	0.01572
72	0.01924	0.01811	0.0173	0.0173
73	0.02124	0.01957	0.01897	0.01897
74	0.02346	0.02126	0.02061	0.02061
75	0.02594	0.02323	0.02238	0.02238
76	0.02862	0.02554	0.02435	0.02435
77	0.03157	0.02813	0.02664	0.02664
78	0.03483	0.03114	0.02945	0.02945
79	0.03834	0.03469	0.03314	0.03314
80	0.04218	0.0385	0.03773	0.03773
81	0.04643	0.0425	0.0425	0.0425
82	0.05137	0.04745	0.04745	0.04745
83	0.05678	0.05311	0.05311	0.05311
84	0.06277	0.05909	0.05909	0.05909
85	0.06905	0.06561	0.06561	0.06561
86	0.07621	0.07321	0.07321	0.07321
87	0.08381	0.0822	0.0822	0.0822
88	0.09196	0.09144	0.09144	0.09144
89	0.10038	0.10038	0.10038	0.10038

3.3 Select

In this section we develop the select PPR model based on the select data described in Section 2.2.

The response variable will be the A/E per policy ratio as the ratio of actual deaths to the expected deaths based on the ultimate PPR model results shown above.

Fitting the model, we use the following formula:

AEpol ~ Attained_Age + Duration + Smoker_Status + Gender + Face_Amount_Band.

Though this formula displays five predictors, Gender, Smoker_Status and Face_Amount_Band are categorical variables. The ppr algorithm converts the multiple levels of these categorical variables into additional predictor variables. For this specific model the predictor variables are in Table 11.

TABLE 11
Select PPR Predictors

X_1	Attained Age
X_2	Duration
X_3	Gender:Female
X_4	Gender:Male
X_5	Smoker_Status:Smoker
X_6	Smoker_Status:Unknown
X_7	FAB \$10K–\$25K
X_8	FAB \$25K–\$50K
X_9	FAB \$50K–\$100K
X_{10}	FAB \$100K–\$250K
X_{11}	FAB \$250K+

For example, for a nonsmoking male with an attained age of 40, in duration 10 and a face amount of \$15,000 the predictor vector is coded as:

$$(X_1 = 40, X_2 = 10, X_3 = 0, X_4 = 1, X_5 = 0, X_6 = 0, X_7 = 1, X_8 = 0,$$
$$X_9 = 0, X_{10} = 0, X_{11} = 0).$$

The determination of this model is based on a weighted least squares. The weights are the expected deaths based on the ultimate PPR model results.

$$W_{AEpol} = Select_Policies_Exposed \times q_x \tag{2}$$

Note that these weights vary by gender, duration, AA, FAB and smoker status. The ultimate q_x for FAB above \$50K is used for all bands above \$50K.

Observe in Figure 4 the ridge functions associated with the Select A/E Ratio model. Table 12 contains the parameters associated with the corresponding PPR model.

Figure 4
Select Per Policy A/E Ratio Ridge Functions

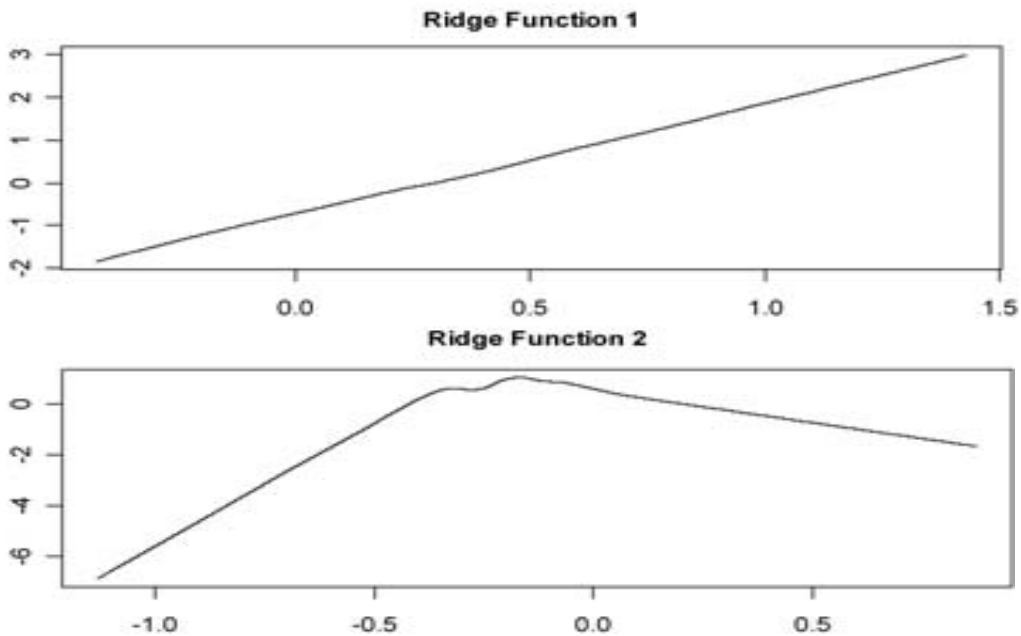


TABLE 12
A/E Policy Select PPR Parameters

	α_1	α_2
Attained_Age	-0.001320526	-0.003153862
Duration	0.018731901	0.013552710
GenderFemale	-0.004146304	-0.061148480
GenderMale	0.072364767	-0.185082074
Smoker_StatusSmoker	0.897275344	-0.522394983
Smoker_StatusUnknown	0.177385436	0.004201492
Face_Amount_Band $10,000-$24,999	-0.027797072	0.344371979
Face_Amount_Band $25,000-$49,999	0.038380290	0.665815168
Face_Amount_Band $50,000-$99,999	-0.121659701	0.278434934
Face_Amount_Band $100,000-$249,999	-0.191195703	-0.164331181
Face_Amount_Band $250,000 and Over	-0.322834759	-0.149883452
Coefficients of ridge terms:		
β_1	0.36087680	
β_2	0.02964484	
$\overline{AE}$	0.8666025	
Goodness of fit:	34856.98	

This PPR model required that the bass tone level be set to 5 instead of 0, which smoothes the model. A bass tone level of 0 (no smoothing) causes a reduction in the model A/E ratio from duration 15 to duration 25, for various ages. We also used bass tone levels of 1 through 4, but the behavior was not completely removed until we reached 5. Please refer to Appendix B for a brief discussion of the bass tone control.

Note that β_1 is approximately 0.4, and the other β_i is much smaller. In our interpretation of the select PPR we will only emphasize the first ridge function contribution. Also observe that f_1 is an increasing function (effectively linear) with a domain of $(-1.2,1.5)$ and a range of $(-6.9,3.0)$. Since f_1 is increasing, whatever increases $\alpha_1 \cdot X$ will increase the PPR model A/E ratio over the average and vice versa.

The first observation is that $\sum_{k=1}^{11} \alpha_{1k} X_k$ decreases as the FAB increases and AA increases and duration decreases. Observe that the major component to the increase of $\alpha_1 \cdot X$ is if the policyholder is a smoker, and to a much lesser degree if the policyholder is male and as duration increases.

Using the PPR Select A/E ratio model, we took the model results and multiplied them by the respective ultimate q_x rates described in Section 3.2. Note: We placed a ceiling of 0.94 on the Non-Smoker model A/E ratio.

We calculated the deaths per thousand for face amount bands \$10K–\$25K and \$100K–\$250K. Tables 13, 14, 15 and 16 are the tables for FAB \$10K–\$25K. Tables 17, 18, 19 and 20 are the tables for FAB \$100K–\$250K.

TABLE 13
Select Male Non-Smoker Band 10K–25K Deaths per Thousand

Age	Dur 1	Dur 2	Dur 3	Dur 5	Dur 10	Dur 11	Dur 15	Dur 20	Dur 25
25	0.7	0.7	0.8	0.8	0.9	0.9	1.0	1.0	1.0
35	1.0	1.0	1.0	1.1	1.2	1.2	1.3	1.4	1.5
45	1.7	1.7	1.8	1.9	2.1	2.1	2.3	2.5	2.5
55	3.1	3.2	3.2	3.4	3.8	3.9	4.2	4.5	4.7
65	7.4	7.6	7.8	8.1	9.1	9.3	10.0	10.9	11.4
75	20.4	20.9	21.5	22.5	25.2	25.7	27.8	30.1	31.9
85	52.8	54.2	55.6	58.4	65.3	66.8	72.3	78.3	83.8

TABLE 14
Select Female Non-Smoker Band 10K–25K Deaths per Thousand

Age	Dur 1	Dur 2	Dur 3	Dur 5	Dur 10	Dur 11	Dur 15	Dur 20	Dur 25
25	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.3	0.3
35	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.7	0.7
45	0.9	1.0	1.0	1.0	1.2	1.2	1.3	1.4	1.5
55	1.8	1.9	1.9	2.0	2.3	2.3	2.6	2.8	3.0
65	4.6	4.8	4.9	5.2	5.8	6.0	6.6	7.2	7.8
75	12.1	12.5	12.8	13.5	15.3	15.7	17.3	19.0	20.6
85	33.5	34.6	35.6	37.6	42.7	43.8	48.1	53.2	57.6

TABLE 15
Select Male Smoker Band 10K–25K Deaths per Thousand

Age	Dur 1	Dur 2	Dur 3	Dur 5	Dur 10	Dur 11	Dur 15	Dur 20	Dur 25
25	1.6	1.7	1.7	1.7	1.8	1.9	1.9	2.0	2.1
35	2.2	2.3	2.3	2.4	2.5	2.6	2.7	2.8	3.0
45	3.8	3.9	4.0	4.1	4.4	4.4	4.6	4.9	5.1
55	7.0	7.1	7.2	7.4	8.0	8.1	8.5	8.9	9.4
65	16.7	17.0	17.3	17.8	19.2	19.4	20.5	21.6	22.7
75	46.1	46.8	47.6	49.1	52.8	53.6	56.5	60.0	62.9
85	119.2	121.2	123.1	127.1	137.0	138.9	146.8	156.2	164.3

TABLE 16
Select Female Smoker Band 10K–25K Deaths per Thousand

Age	Dur 1	Dur 2	Dur 3	Dur 5	Dur 10	Dur 11	Dur 15	Dur 20	Dur 25
25	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.6	0.7
35	1.1	1.1	1.1	1.1	1.2	1.2	1.3	1.4	1.4
45	2.3	2.4	2.4	2.5	2.6	2.7	2.8	2.9	3.1
55	4.6	4.6	4.7	4.9	5.2	5.2	5.5	5.8	6.1
65	11.7	11.9	12.1	12.5	13.3	13.5	14.1	15.0	15.7
75	30.7	31.2	31.7	32.7	35.2	35.6	37.3	39.6	41.7
85	85.1	86.6	88.0	90.9	98.1	99.5	104.5	110.6	117.1

TABLE 17
Select Male Non-Smoker Band 100K–250K Deaths per Thousand

Age	Dur 1	Dur 2	Dur 3	Dur 5	Dur 10	Dur 11	Dur 15	Dur 20	Dur 25
25	0.4	0.5	0.5	0.5	0.6	0.6	0.7	0.7	0.8
35	0.5	0.5	0.5	0.6	0.7	0.7	0.8	0.9	0.9
45	0.9	0.9	1.0	1.0	1.2	1.3	1.4	1.6	1.8
55	1.7	1.8	1.8	2.0	2.4	2.5	2.8	3.1	3.5
65	4.0	4.2	4.5	4.9	5.9	6.1	6.9	7.8	8.6
75	11.4	12.1	12.7	14.0	17.1	17.7	20.1	23.0	25.4
85	32.3	34.3	36.2	40.0	49.3	51.1	58.4	67.5	74.8

TABLE 18
Select Female Non-Smoker Band 100K–250K Deaths per Thousand

Age	Dur 1	Dur 2	Dur 3	Dur 5	Dur 10	Dur 11	Dur 15	Dur 20	Dur 25
25	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.2	0.2
35	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.3	0.4
45	0.5	0.5	0.5	0.6	0.7	0.7	0.8	0.9	1.0
55	1.2	1.3	1.4	1.5	1.8	1.8	2.1	2.3	2.6
65	2.8	3.0	3.2	3.5	4.2	4.3	4.9	5.6	6.2
75	7.7	8.2	8.7	9.7	11.9	12.3	13.8	15.9	17.8
85	21.1	22.6	24.1	27.0	33.9	35.1	39.5	45.5	51.6

TABLE 19
Select Male Smoker Band 100K–250K Deaths per Thousand

Age	Dur 1	Dur 2	Dur 3	Dur 5	Dur 10	Dur 11	Dur 15	Dur 20	Dur 25
25	1.0	1.1	1.1	1.1	1.2	1.2	1.3	1.4	1.5
35	1.2	1.2	1.2	1.3	1.4	1.4	1.5	1.6	1.7
45	2.2	2.2	2.3	2.4	2.6	2.6	2.8	3.0	3.2
55	4.3	4.4	4.5	4.7	5.1	5.2	5.5	6.0	6.4
65	10.7	10.9	11.1	11.5	12.6	12.9	13.8	14.9	15.9
75	31.5	32.1	32.7	34.0	37.3	38.0	40.7	43.9	47.2
85	92.2	94.1	96.0	99.8	109.6	111.6	119.6	129.5	139.2

TABLE 20
Select Female Smoker 100K–250K Deaths per Thousand

Age	Dur 1	Dur 2	Dur 3	Dur 5	Dur 10	Dur 11	Dur 15	Dur 20	Dur 25
25	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.4
35	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7
45	1.3	1.3	1.4	1.4	1.5	1.6	1.7	1.8	1.9
55	3.4	3.5	3.5	3.7	4.0	4.1	4.4	4.7	5.1
65	8.0	8.2	8.4	8.7	9.5	9.7	10.4	11.3	12.1
75	22.8	23.3	23.8	24.7	27.1	27.6	29.6	32.2	34.6
85	65.6	67.0	68.3	71.1	78.2	79.6	85.4	92.8	100.0

In the next section we will discuss how this model will be first applied regarding ERM.

4. Results and ERM Management

Enterprise risk management takes on many different shades and hues according to which industry the specific enterprise belongs. The needs of ERM run the gamut from modeling or managing risks to determining how these risks impact the enterprise's capital. Within life reinsurance, we first need to understand our risk exposure. From the use of predictive models, we have already realized how much our models have improved our understanding of this risk. If the predictive models are very realistic (which is often the case with PPR models), we are then able to better manage our business as well.

In order to construct a model of mortality risk to use for ERM for our company, we use this PPR model of the SOA 2002–2004 Study as a basis for developing a predictive model of our own internal mortality experience. Our internal mortality experience database includes nearly 100,000 deaths from nearly 50 million life-years of exposure, grouped into cells by individual attained ages and durations, gender, smoker status and face amount band, and further differentiated by client, experience quarter, reinsurance type (coinsurance or YRT, excess or quota share, automatic or facultative), product type (term or permanent), underwriting type (guaranteed issue, simplified issue or fully underwritten) and risk class rank (e.g., second best preferred in a four class structure) or substandard table rating. Due to the confidential nature of the data, we may not share the details of our internal model. However, the techniques described in this paper are used to develop a model of the A/E ratios of the cells in our internal mortality experience database to the PPR model of the SOA 2002–2004 Study. From these model A/E ratios, the implied model mortality rates are calculated.

This internal mortality risk model can be used to monitor and manage risk in many ways. First, by comparing the model mortality rates to the mortality rates used when business now in force was originally priced, we can determine where there may be cross-subsidies hidden in our existing treaties. These treaties can then be monitored for shifts in distribution of business that could affect the profitability of the account. Cells for which the original pricing mortality rates are higher than the model mortality rates present an opportunity to seek new business by pricing these cells more aggressively. Similarly, cells for which the original pricing mortality rates are lower than the model mortality rates should be priced more conservatively in future deals.

We are effectively using the industrial model as a major component of both key risk and key performance indicators. Note that these approaches also work across blocks of business, so we would be able to examine our aggregated risk exposure as well.

In the next section, we will outline the advantages and disadvantages of the use of PPR, but one nice advantage that we mention here is that once the data is modeled within PPR, the use of the model in a predictive manner is very efficient. This makes PPR a natural candidate to use as a dashboard for upper management. As we continue to collect and model our internal data using PPR or some corresponding predictive model, we will be able to communicate our ongoing mortality performance on a timely basis to our management.

In the next section, not only do we discuss the strengths and weaknesses of PPR, we also discuss our conclusions and briefly outline our future work integrating this work with our ERM models.

5. Conclusions and Further Research

The following is a list of advantages of using PPR as a model for mortality:

1. The model is a continuous function. According to Venables and Ripley [9] while citing Diaconis and Shahshahani [1], given a large enough number of ridge functions, PPR can approximate arbitrary continuous functions.
2. One is able to obtain fractional age and duration estimates.
3. It is the best possible fit since every component is solved to minimize the weighted least squares.
4. Each ridge function does not extrapolate outside of its specific domain. If the specific $\alpha_i \cdot X$ is outside the domain, the relevant domain endpoint is used.
5. The model handles the interactions between the different predictors. See Appendix A for an example of how this holds.
6. PPR models categorical predictors as easily as continuous predictors.
7. PPR lessens the "curse of dimensionality" by projecting the data into a one-dimensional space, and thus reduces data sparseness that renders other methods less useful.

The disadvantages are:

1. The range of the PPR model may fall outside of the range of acceptable values. For instance, both A/E models above could create negative or zero A/E ratios. To offset this weakness, we place a lower bound on the output.
2. The implementation of ppr in R produces only point estimates on all of the parameters. The lack of the ability to create confidence intervals around each of the f_k, α_{ik} or the β_i makes it difficult to determine if a specific parameter is significant to the model. In fact one is unable to test if the actual model is significant, other than the use of the goodness of fit statistic. All research results that discuss the distributional issues of the fitted ridge functions and projection vectors are quite difficult to read, let alone implement. If the reader is interested, these topics are covered by Park and Sun [6], Diaconis and Freedman [2] and Hall [5]. Another approach to alleviate these difficulties would be to implement some type of resampling and empirically determine the confidence intervals
3. One can easily overfit or overexplain the data. See Venables and Ripley [9].
4. The model can be too flexible, which makes the model difficult to interpret. One can observe this in the discussions around the components of the ultimate PPR model above. Again, see Venables and Ripley [9].

Even with these weaknesses, we have found that the PPR algorithm is extremely valuable in our industry mortality models, and it has enabled us to not only confirm our empirical knowledge, but to actually have models calibrated to the data that corresponds to our understanding. We are strongly encouraged by the results that we have obtained, and we have gone from thinking that predictive modeling would be an interesting intellectual pursuit to finding that the use of the models makes good business sense.

Our future work is to continue the development of these models for key risk and performance indictors as well as potential dashboard models. It is our desire to be able to actually look at different blocks of business from an economic capital perspective and observe the risk adjusted return from these blocks. Once we are able to look at this business from a risk-adjusted basis, it is our desire to use this knowledge in strategic decision-making.

Acknowledgments

Steve Craighead: I would like to thank my wife Patricia and my children for their patience. I would also like to thank the SOA for providing the necessary mortality data that has enabled us to conduct this study.

Appendix A. Projection Pursuit Regression (PPR)

In linear regression, one fits a response variable Y to a collection of n predictor variables X_i in the familiar form:

$$Y = \alpha + \sum_{i=1}^{n} \beta_i X_i + \varepsilon \tag{3}$$

In additive models, the $\beta_i X_i$ are replaced with various functions $f_i(X_i)$, with this form:

$$Y = \alpha + \sum_{i=1}^{n} f_i(X_i) + \varepsilon \tag{4}$$

Projection pursuit regression (PPR) is a modification of this structure in that there are:

1. M different f_i.
2. Each f_i acts on a different linear combination of all n of the X_k.
3. A specific coefficient of these linear combinations is denoted by α_{ik}.
4. Each f_i is multiplied by a β_i.
5. The constant term is the average of the response variable.

So PPR takes on the following form:

$$Y = \bar{y} + \sum_{i=1}^{M} \beta_i f_i(\sum_{k=1}^{n} \alpha_{ik} X_k) + \varepsilon \tag{5}$$

or in vector format:

$$Y = \bar{y} + \sum_{i=1}^{M} \beta_i f_i(\alpha_i \cdot X) + \varepsilon \tag{6}$$

where $X = (X_1, X_2, ..., X_n)$ is the predictor vector, and $\alpha_i = (\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in})$.

The term "Projection" in PPR comes from the projection of X on to the directional vector α_i for each i.

"Pursuit" arises from the algorithm that is used to determine optimal direction vectors $\alpha_1, \alpha_2, \ldots, \alpha_M$.

Each f_i is called a ridge function. This is because they only have values in the α_i direction and are considered constant elsewhere. Effectively, what occurs is that the overall PPR model is a linear combination (β_i are the coefficients) of the ridge functions. These functions

only take on values that arise from the projection of the predictors against the direction vectors, and the functions are assumed to take on a constant value in any other direction. So, each ridge function is like the profile of a mountain range, and we linearly combine these functions along all different ridges (as pointed out by the α_i).

On a formal basis, Y and X are assumed to satisfy the following conditional expectation:

$$E[Y \mid X_1, X_2, \ldots, X_n] = \mu_y + \sum_{i=1}^{M} \beta_i f_i(\alpha_i \cdot X) \tag{7}$$

with $\mu_y = E[Y]$ and the f_i have been standardized to have zero mean and a unit variance. That is: $E[f_i(a_i \cdot X)] = 0$ and $E[f_i^2(a_i \cdot X)] = 1$, where i takes on values from 1 to M. We assume that the realized sample values for the random variables Y and $X = (X_1, X_2, \cdots, X_n)$ are independent and identically distributed to the distributions of Y and X, respectively.

The ppr algorithm in the R stats library [7] estimates the best $\beta_i, f_i,$ and the α_i by minimizing the following target function for the mean square error:

$$E\left[Y - \mu_y - \sum_{i=1}^{M} \beta_i f_i(\alpha_i \cdot X)\right]^2 \tag{8}$$

across all the data samples for Y and X.

A powerful trait of PPR models, since the predictor vector X is projected, is that interactions between different X_j and X_k are included within the model, whereas other model algorithms cannot do this without user intervention.

We justify this by using an algebraic demonstration based on the S-Plus Guide to Statistics [8] recast into our notation as follows:

Suppose that the actual data model is $E[Y \mid X_1, X_2] = X_1 X_2$.

Let $\overline{Y} = 0$, $M = 2$, $\beta_1 = \beta_2 = .25$ and assume that $\alpha_1 = (1,1)$ and $\alpha_2 = (1,-1)$. Furthermore assume that $f_1(t) = t^2$ and $f_2(t) = -t^2$. Let $X = (X_1, X_2)$.
Now $f_1(\alpha_1 \cdot X) = (X_1 + X_2)^2 = X_1^2 + 2X_1X_2 + X_2^2$. Similarly $f_2(\alpha_2 \cdot X) = -X_1^2 + 2X_1X_2 - X_2^2$.
Finally, $\sum_{i=1}^{M} \beta_i f_i(\alpha_i \cdot X) = X_1 X_2$.

So, we can see that if the β_i, f_i and α_i are optimally selected and the underlying model has interactions between different predictors, PPR should capture this.

Appendix B. Techniques and Diagnostics for ppr

The procedure when using the R ppr algorithm [7] is as follows:

First, one specifies that M should range between $M_{MIN} = 1$ and some positive integer M_{MAX}. The ppr algorithm then creates a PPR model for each M from M_{MAX} to M_{MIN} in a descending fashion, and at the same time produces a goodness of fit statistic for each value of M. Scanning this list of goodness of fit values should display a local minimum. If this local minimum is at M_{MAX} one should reprocess the experiment with a larger M_{MAX}. Once one determines the local minimum, say s, reset $M_{MIN} = s$ and reprocess the ppr algorithm with the same M_{MAX} as before. The resultant model arising from the backward iteration from M_{MAX} to M_{MIN} will then be the best PPR model.

Two other components that are implemented in ppr are the concepts of bass and optlevel.

bass is Friedman's [4] super smoother bass tone control that is used with automatic span selection. It is used in ppr to smooth the results. The range of values allowed with this component is from 0 to 10. To increase smoothing within the data increase this value. The default is 0 and this setting gives the best fit to the underlying data. bass is similar to the *h* smoothness parameter used within the Whitaker-Henderson graduation formula.

optlevel is an integer from 0 to 3, which determines the optimization thoroughness. The best models usually are obtained if this is set to 3. At level 0, the ridge functions are not refitted. At level 1, the projection directions are not refitted, but the ridge functions and the regression coefficients are. Levels 2 and 3 refit everything, but level 3 takes pains to re-balance each regressors' contribution at each step and so reduces the chance of converging to a saddle point in the sum of squares.

One diagnostic aid in PPR model building is to plot the ridge functions. If these ridge functions are very noisy or discontinuous, you should expect that the resultant PPR model will behave oddly.

Another effective diagnostic aid is to both plot the fitted $\hat{Y}$ against the actual Y and do a simple linear regression of Y against $\hat{Y}$, assuming no intercept. The scatterplot should display symmetry around the 45 degree line and the coefficient of the regression should be approximately 1. These two diagnostics will indicate how well the PPR model will perform as a predictive model.

Note: A PPR model does not extrapolate outside of the sample data. Therefore, the resultant fitted values from a PPR model will frequently hit a maximum value and will not grow any larger no matter how one manipulates the predictors. This is not the case for linear regression models, where there are no natural limits placed on how one sets any respective X_i. One may revise the prediction object to conduct extrapolations; however, one must first feel comfortable with the continuity of the separate ridge functions. If these functions are very noisy or appear not to be differentiable, you might want to avoid all extrapolation.

Appendix C. Poisson Weights Derivation

We would like to maximize the probability of the actual deaths given the model, assuming a Poisson distribution of deaths in each cell. For cell i,

$$P\left\{d_i \mid AE_i^M\right\} = \exp(-\lambda_i)\frac{\lambda_i^{d_i}}{d_i!} \tag{9}$$

where $\lambda_i = E_i q_i AE_i^M$, d_i is the actual number of deaths incurred in cell i, E_i is the number of policy years exposed in cell i, q_i is the reference mortality rate for cell i and AE_i^M is the model A/E ratio for cell i.

For the ultimate model, q_i is taken from the composite 2001 VBT. For the select model, q_i is taken from the PPR ultimate model.

Let $M = \{AE_i^M\}$ be the set of model A/E ratios for all cells, and let $D = \{d_i\}$ be the actual deaths for all cells. Then:

$$P\{D \mid M\} = \prod_i \exp(-\lambda_i)\frac{\lambda_i^{d_i}}{d_i!} \tag{10}$$

and:

$$\ln[P\{D \mid M\}] = \sum_i (-\lambda_i + d_i \ln \lambda_i - \ln d_i!) \tag{11}$$

We would like to maximize the loglikelihood, or equivalently, minimize its negative:

$$\Xi = \sum_i \lambda_i - \sum_i d_i \ln \lambda_i + \sum_i \ln d_i! \tag{12}$$

Using the definition of λ_i, we see that:

$$\begin{aligned}\Xi &= \sum_i E_i q_i AE_i^M - \sum_i d_i (\ln E_i + \ln q_i + \ln AE_i^M) + \sum_i \ln d_i! \\ &= \sum_i E_i q_i AE_i^M - \sum_i d_i \ln E_i - \sum_i d_i \ln q_i - \sum_i d_i \ln AE_i^M + \sum_i \ln d_i!\end{aligned}$$

Since the d_i and the E_i are given in the data and the q_i are specified, we see that:

$$\begin{aligned}\frac{\partial \Xi}{\partial AE_i^M} &= E_i q_i - \frac{d_i}{AE_i^M} \\ &= \frac{AE_i^M E_i q_i - d_i}{AE_i^M}\end{aligned} \tag{13}$$

Ξ is the function we would like to minimize, and its partial derivatives with respect to M are given by (13).

As discussed in Appendix A, the error function that will be minimized by the PPR algorithm is given in equation (8) as:

$$E\left[Y - \mu_y - \sum_{i=1}^{M} \beta_i f_i(\alpha_i \cdot X)\right]^2$$

In our model, we can write this as:

$$\Xi' = \sum_i w_i (AE_i^A - AE_i^M)^2 \tag{14}$$

where AE_i^A is the actual observed A/E ratio for cell i, w_i is the weight for cell i, and $\sum_i w_i = 1$. Note that:

$$AE_i^A = \frac{d_i}{q_i E_i}. \tag{15}$$

Substituting (15) into (14), we see that:

$$\begin{aligned}\Xi' &= \sum_i w_i \left(\frac{d_i}{q_i E_i} - AE_i^M\right)^2 \\ &= \sum_i w_i \left(\frac{d_i^2}{q_i^2 E_i^2} - 2\frac{d_i}{q_i E_i} AE_i^M + (AE_i^M)^2\right)\end{aligned}$$

The partial derivatives of Ξ' with respect to M are given by:

$$\begin{aligned}\frac{\partial \Xi'}{\partial AE_i^M} &= -2w_i \frac{d_i}{q_i E_i} + 2w_i AE_i^M \\ &= \frac{2w_i}{q_i E_i}(AE_i^M E_i q_i - d_i)\end{aligned} \tag{16}$$

If we choose w_i so that the partial derivatives of Ξ' with respect to M are the same as the partial derivatives of Ξ with respect to M, then Ξ and Ξ' will both be minimized by the same set M and the PPR algorithm will produce the model that minimizes our desired error function.

From (16) and (13), we see that these partial derivatives will be equal when

$$\frac{2w_i}{q_i E_i}(AE_i^M E_i q_i - d_i) = \frac{AE_i^M E_i q_i - d_i}{AE_i^M} \tag{17}$$

Solving for w_i, we get:

$$w_i = \frac{1}{2}\frac{q_i E_i}{AE_i^M} \tag{18}$$

Since we do not yet know AE_i^M, we assume that it is a constant, and we set $w_i = kE_i q_i$ where $k = \left(\sum_i q_i E_i\right)^{-1}$ is the normalizing constant.

References

[1 Diaconis, P., and Shahshahani, M. 1984. "On Non-Linear Functions of Linear Combinations." SIAM Journal of Scientific and Statistical Computing 5:175–191.

[2] Diaconis, P., and Freedman, D. 1984. "Asymptotics of Graphical Projection Pursuit." The Annals of Statistics 12(3): 793–815.

[3] Friedman, J.H., and Stuetzle, W. 1981 "Projection Pursuit Regression." Journal of the American Statistical Association 76: 817–823.

[4 Friedman, J.H. 1984. "A Variable Span Scatterplot Smoother." Stanford University Technical Report No. 5 Laboratory for Computational Statistics.

[5] Hall, P. 1989. "On Projection Pursuit Regression." The Annals of Statistics 17(2): 573–588.

[6] Park, M.G., and Sun, J. 1998. "Tests in Projection Pursuit Regression." J. Statist. Plann. Inference 75: 65–90.

[7] R Development Core Team. 2006. "R: A language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.r-project.org.

[8] S-Plus 6 for Windows Guide to Statistics, Vol 1. 2001. Insightful Corporation.

[9] Venables, W.N., and Ripley, B.D. 2002. *Modern Applied Statistics with S*, Fourth Edition. Springer.