

Index Based Crop Insurance Product Design and Ratemaking

The Case of Modified NAIS in India

Daniel J. Clarke

Olivier Mahul

Niraj Verma

The World Bank
Financial and Private Sector Development Vice Presidency
Non-Banking Financial Institutions Unit
&
South Asia Region
Finance and Private Sector Development Unit
March 2012



Abstract

Designing and rating insurance products requires both science and judgment. In developing and emerging economies, actuarial procedures must be robust and implementable, as well as offering a sufficient degree of transparency and flexibility so as to allow expert judgment to be incorporated. This paper outlines an approach to designing and rating a portfolio of index

insurance products that uses both temporal and spatial aspects of the data to increase the efficiency of statistical estimates. The approach has formed the basis for the design and ratemaking methodology implemented by the Agriculture Insurance Company of India for the modified National Agricultural Insurance Scheme, which was initiated by the Government of India in late 2010.

This paper is a product of the Non-Banking Financial Institutions Unit of the Capital Market Practice, Finance and Private Sector Development Vice Presidency; and Finance and Private Sector Development Unit, South Asia Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at omahul@worldbank.org and nverma@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Index Based Crop Insurance Product Design and Ratemaking: The Case of Modified NAIS in India¹

Daniel J. Clarke, *Department of Statistics, University of Oxford*

Olivier Mahul, *Program Coordinator, Disaster Risk Financing and Insurance Program, Non-Banking Financial Institutions Unit and Global Facility for Disaster Reduction and Recovery, World Bank*

Niraj Verma, *Senior Financial Sector Specialist, South Asia Finance and Private-Sector Development Unit, World Bank*

JEL Classification Numbers:

G22, O13, O16.

Keywords:

Agriculture, area yield crop insurance, developing countries, India, product design, pricing.

Sector Board:

Capital Market Practice, Finance and Private-Sector Development

¹This paper was motivated by a larger program of technical work by the World Bank, requested by the Government of India and the Agricultural Insurance Company of India. This technical assistance was funded by the Swiss Development and Cooperation Agency (Phase 1, 2005), the FIRST Initiative (Phase 2, 2005–2007) and the Global Facility for Disaster Reduction and Recovery (Phase 3, 2008–2011). The authors thank M. Parshad, Kolli Rao and M. K. Poddar of AICI, as well as other government officials with whom we interacted, for useful discussions and access to data, and to Ivan Rossignol, Loic Chiquier, Stefan Dercon and Avery Cook for useful comments. Views expressed in this paper are the authors' and should not be attributed to the World Bank, University of Oxford or Government of India. Email: omahul@worldbank.org, nverma@worldbank.org, clarke@stats.ox.ac.uk.

1. Introduction

There is a vast body of actuarial literature on design and ratemaking for crop insurance; for example Josephson et al. (2000) and Skees et al. (1997) document the methodologies implemented for the US Multiple Peril Crop Insurance (MPCI) program and Group Risk Plan (GRP) respectively, and Stohs and LaFrance (2004) propose a more sophisticated methodology for the MPCI program. However, despite the recent growth of crop insurance programs in low and middle-income countries, there is little literature on how to apply actuarial techniques outside high-income countries.

Any methodology for design and ratemaking should be robust and statistically efficient. Moreover in low-income environments where actuarial expertise is often lacking, it should be transparent enough for senior management to scrutinize and flexible enough to allow the inclusion of expert judgment.

Currently, many of the design and ratemaking procedures being used in developing countries are product-based approaches, where calculations are conducted separately for each product (Syroka 2007), rather than portfolio-based approaches, where calculations are conducted jointly for all products. By utilizing only the time dimension of historical data, these approaches are statistically inefficient (Efron and Morris 1975). There are a number of ways of incorporating the spatial dimension of data into statistical procedures, such as Hierarchical Bayes methods or seemingly unrelated regressions (SUR). However, these approaches may be challenging to implement and scrutinize when there is limited actuarial capacity. Empirical Bayes Credibility Theory is suggested as one way to increase the statistical efficiency of design and ratemaking while retaining transparency in intermediate calculations and the flexibility to apply expert judgment (Bühlmann 1967).

In this paper we document the design and rating methodology implemented by the Agriculture Insurance Company of India for the modified National Agricultural Insurance Scheme (mNAIS), which was piloted during the spring (Rabi) season 2010-11. The approach uses long term data series, available in India, to calculate premium rates that are stable yet reflective of regional differences and responsive to changes in risk over time. This enhanced risk classification increases the equity of the mNAIS as compared to its predecessor, the NAIS, and helps to protect the crop insurance portfolio from adverse selection.

In the Indian context the development of this design and ratemaking methodology had much wider implications than just improving the current crop insurance scheme; it has allowed the Government of India to transition to a market-based crop insurance program with involvement from private sector insurers and reinsurers, improved fiscal management for government by enabling ex-ante budgeting for government subsidies towards premium, and increased equity and faster claim settlement for farmers. While the current paper is targeted at a technical actuarial audience, we refer readers interested in the wider policy agenda to Mahul et al. (2011).

While the product details are specific to India, the principles can be applied much more generally to indexed insurance products in low and middle income countries. However, while we hope that others will find this paper useful, it does not substitute for the expert advice of an appointed actuary.

The Modified NAIS

At the time of writing the modified NAIS (mNAIS) is in a pilot phase and may be amended in the future. In this paper the following structure for mNAIS is assumed:

- The primary benefit to farmers from mNAIS is calculated on an area-yield-based approach: if the observed seasonal area-yield per hectare of the insured crop for the defined geographical area (the *insurance unit*) falls below a specific threshold yield, all insured

farmers growing that crop in the defined area will receive the same claim payment per unit of sum insured. For a given threshold yield (TY) and actual yield, the claim payment per unit of sum insured is given by:

$$\text{Area yield based claim payment rate} = \max(0, TY - \text{Actual Yield})/TY$$

- The Threshold Yield is then calculated as

$$\text{Threshold Yield} = \text{Indemnity level} \times \text{Probable Yield}$$

where indemnity levels of 70%, 80% and 90% are set at the crop/district level and probable yields are set at the crop/insurance unit level as the moving average yield of the last 7 years.²

- Commercial premium rates are to be set at the crop/district level, calculated to be actuarially sound estimates of the expected value of all future costs associated with insurance products sold for the respective crop in the respective district.
- The mNAIS scheme operates under an actuarial regime whereby the insurer receives premiums (premiums paid by farmers plus premium subsidies from government) and is responsible for the management of the mNAIS portfolio through risk transfer to private reinsurance markets and risk retention through reserves and possibly contingent credit. Government's liability is therefore predominantly in the form of upfront premium subsidies, paid to the insurer and funded ex-ante;
- Under the previous scheme, the NAIS, the insurance unit was typically at the sub-district level. Under mNAIS the insurance unit is reduced to the village Panchayat level for major crops;
- There may be a partial on-account settlement of claims due to mid-season adversity during the cropping season, whereby an early part-claim settlement based on informal, local information and, potentially, weather data may be made in some insurance units;
- Additional benefits may be offered for prevention of sowing, replanting or post-harvest losses, or for localized risk, such as hail losses or landslides.

Actuarial Concepts

Under mNAIS the commercial premium for an individual product should be an actuarially sound estimate of the expected value of all future costs associated with the product. Different insurance units have different characteristics and so this objective could only be met if the insurance product or premium could vary between insurance units. Such *risk classification* would promote equity and protect the insurer from the threat of adverse selection, whereby farmers would be able to adversely select against the insurance program by participating when the cost of coverage is low relative to the benefits, and not participating when the cost of coverage is high relative to the benefits (American Academy of Actuaries 1980).

Risk classification should be based, whenever possible, on statistical analysis, modified by informed judgment. American Academy of Actuaries (1980) classifies the statistical considerations when designing a risk classification system into homogeneity, credibility and predictive stability: the principle of homogeneity requires that "The expected costs for each of the individual risks in a class should be reasonably similar. In a given class, there should be no clearly identifiable subclasses with significantly different potential for losses;" credibility requires that "each of the classes in a risk

²Under mNAIS state governments may notify up to two years in the last seven as *calamity years*, which are excluded from the Probable Yield calculation, but included in the rating calculation. Notifying years with low yields as calamity years therefore leads to a higher level of cover being offered to farmers, at a correspondingly higher premium.

classification system be large enough to allow credible statistical predictions about that class;” and predictive stability requires that “elements of a risk classification system be useful for predictive purposes. The predictive capability must be responsive to changes in the nature of insurance losses, yet stable in avoiding unwarranted abrupt changes in resulting prices.”

Many of the statistical steps suggested in this paper are motivated by the observation that designing or pricing a product based on no more than ten years history for that individual product is not ideal from the point of view of credibility or predictive stability. Credibility and predictive stability can be increased by grouping similar products into classes and jointly designing or pricing all products within the same class. The procedure outlined in this paper involves smoothing products at the crop/district level and smoothing premium rates at the crop/state level. Such smoothing can increase statistical efficiency if there is some degree of agronomic homogeneity within the district and the state.

While improved risk classification increases protection against adverse selection, the move to upfront state premium subsidies leaves the insurer vulnerable to moral hazard. Section 0 offers a discussion of potential responses to the threat of moral hazard.

Overview

This paper provides a description of each key step of the implemented design and ratemaking calculation along with illustrative examples and a discussion of the actuarial justification. For each of the steps in the process, there may be alternative approaches that could be used and which could produce reasonable results. The paper does not intend to try to identify all possible alternatives to the suggested approach. However, in many cases alternatives that are believed to be appropriate are discussed.

The rest of this paper is organized as follows. The actuarial calculation framework for mNAIS comprises four main steps, each of which is discussed in detail in a section of this paper: cleaning and de-trending data; calculating the Threshold Yield; calculating the Pure Premium Rate; and calculating the Commercial Premium Rate. The penultimate section offers a proposal for how the suggested procedure could be amended when the insurance unit has changed in the last ten years, and the final section concludes.

First, data must be checked for any obvious errors, and trends or structural breaks in the data caused by changes to agronomic productivity or the recording of data may be removed (Section 2). Data errors, such as missing yield values being erroneously entered as zero yield values, can have a significant effect on calculations. Data should be checked for any such errors.

Statistical analysis and informed judgment can be used to determine if a trend has occurred in historical data. For example, widespread adoption of high yielding crop varieties may lead to a sustained increase in yields for the affected crop. Such trends should be removed from the data to ensure equity and appropriate risk classification. If such trends were not removed, farmers would be offered poor value products, with low threshold yield and high premium rates.

Second, the Threshold Yield, the crop/insurance unit level trigger for area yield based claim payments, is calculated, determining the product to be sold (Section 3). The Threshold Yield is calculated to be the product of the Probable Yield and the Indemnity Level.

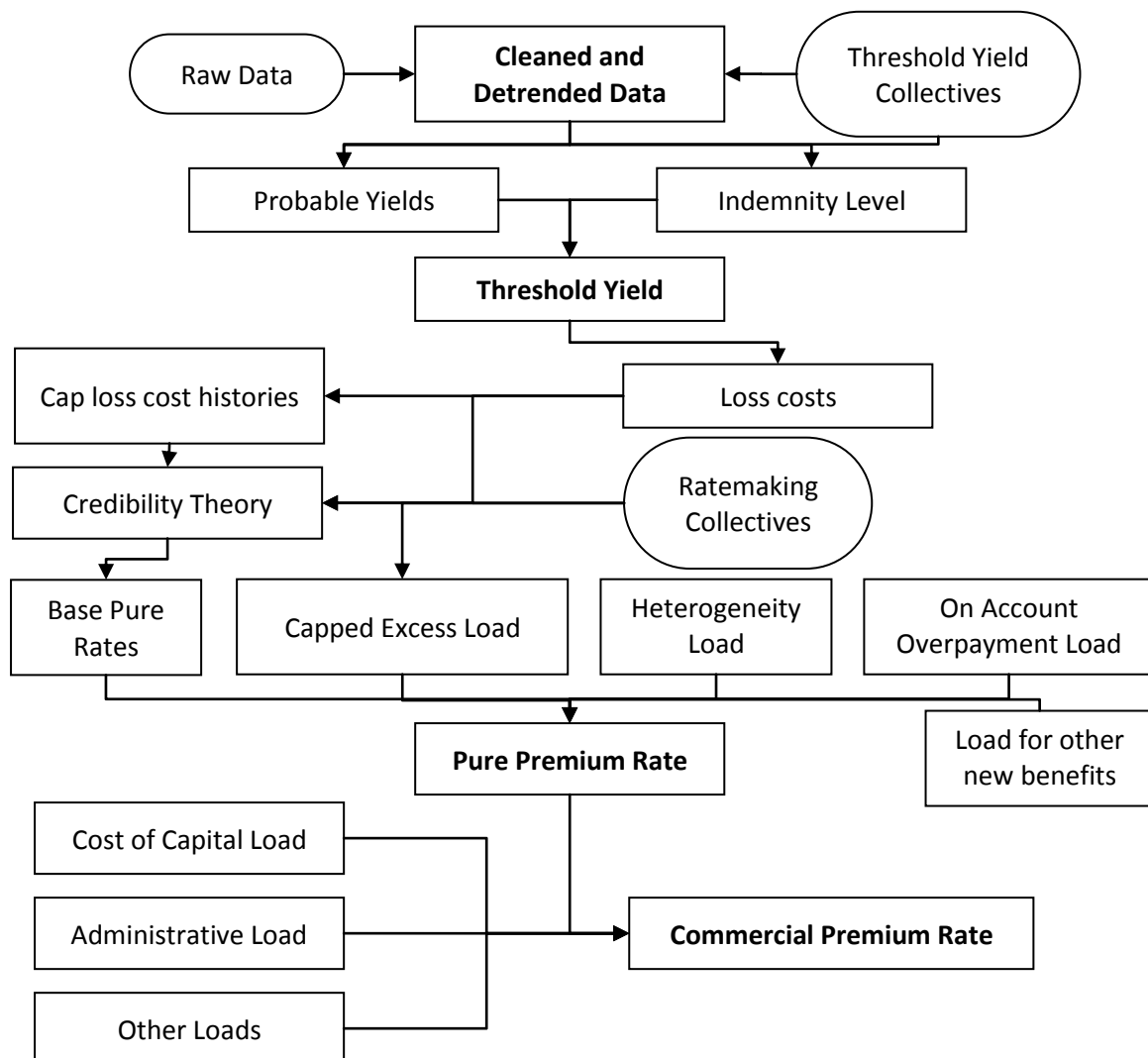
The Probable Yield is calculated based on the seven year average yield for each insurance unit, with two potential adjustments. First, up to two years ‘calamity’ years may be removed from this calculation. This has no statistical justification but, by increasing threshold yield and associated commercial premium rate and government subsidy rate for those insurance units in which calamities have been officially reported, it could increase equity between farmers who are only partially insured. Second, Probable Yields may be smoothed within a *Threshold Yield Collective*, which in this

paper is taken to be a crop/district, according to the size of the Credibility Factor. Credibility smoothing does not change the average Probable Yield within a Threshold Yield Collective, but reduces the variation in Probable Yields when some of that variation is not statistically credible. This can substantially increase risk classification, thereby protecting the mNAIS portfolio from adverse selection.

The Indemnity Level rule targets the Pure Premium Rate, with high risk crop/districts offered an Indemnity Level of 70%, low risk crop/districts offered 90%, and medium risk districts offered 80%. Crop/districts are classified as high, medium, or low risk according to the average historical loss cost at 70% and 90% indemnity levels.

Third the Pure Premium Rate is calculated as a statistical estimate of the crop/district-level expected loss cost, with allowance made for all benefits offered under mNAIS (Sections 4 and 0). Pure Premium Rates are calculated jointly for each crop, with credibility smoothing of premium rates at the level of the crop/agronomic region.

Figure 1: Overview of framework for mNAIS actuarial calculations



Where ten years of historical data is not available at the current insurance unit level, Pure Premium Rates must be set with reference to aggregated data from the old insurance unit. If the yield distribution varies across the old insurance unit, use of data from the old insurance unit may lead to an underestimate of the expected cost of providing the product in some current insurance units. To protect against adverse selection the Pure Premium Rate therefore includes a Heterogeneity Load whenever a sufficient history of appropriate data is not available at the current insurance unit level.

Fourth the Commercial Premium Rate is calculated, giving the premium and premium subsidy rates payable by farmers and government (Section 0). The Commercial Premium Rate is higher than the Pure Premium Rate both to account for costs such as administrative and reinsurance costs, and to allow a margin against future contingencies.

2. Data Cleaning and De-trending

An actuarial approach to design and ratemaking requires stability in the process for collecting data. Before the actuarial calculations described in future sections, data must be checked for any obvious errors, and trends or structural breaks caused by changes to farming practices or the recording of data. The de-trending step, in particular, requires actuarial judgment in addition to statistical analysis.

This section focuses on yield data, but the principles equally apply to other datasets. A related procedure for cleaning and de-trending weather data was suggested in Syroka (2007).

Sections 2, 3 and 4 of this paper illustrate the suggested procedure using worked examples of groundnut products to be sold across India in 2010 and cotton products to be sold in Gujarat in 2008. These examples use real yield data but district and subdistrict names are not disclosed for reasons of confidentiality.

Data Cleaning

Description of step

All historical yield data to be used for design and ratemaking should be collated into a standard format for use in actuarial calculations. Ten years of historical yield data is used in the procedure suggested in this paper.

Missing data should be recorded as distinct from a zero yield. At this stage missing yield data should *not* be 'filled in', but recorded as missing.

Where the insurance unit has changed in the last ten years two historical yield series may be recorded for each crop. First, the series for data from the new insurance unit only; and second the series for data from the old insurance unit only. The choice of data format to be used when the insurance unit has changed in the last ten years will depend on the ultimate use of the data. Section 0 provides an overview of one possible framework for using this data in actuarial calculations.

Actuarial justification

Data should be cleaned and collated into a standard format to enable efficient checking and scrutiny of actuarial calculations. International best practice for robust risk management requires that each actuarial calculation should be performed by one individual, checked by another and scrutinized by a suitably qualified actuary.

Threshold Yield Collectives and Ratemaking Collectives

Description of step

A Threshold Yield Collective is a group of products which will receive the same premium rate. Throughout this paper Threshold Yield Collectives are assumed to be set at the crop/district level, containing all products to be sold for the same crop within a given district. However, if agronomic characteristics are known to vary within a district there may be multiple Threshold Yield Collectives for the same crop/district, each containing a conceptually contiguous group of insurance units. For example, for districts containing insurance units classified by the insurer as 'hilly', there may be two Threshold Yield Collectives for each crop in the district: one containing the hilly insurance units and one containing the non-hilly insurance units. De-trending and Threshold Yield calculations will be performed at the level of the Threshold Yield Collective. The name Threshold Yield Collective refers to the credibility smoothing of Threshold Yields to be performed within these collectives.

A Ratemaking Collective is a group of one or more Threshold Yield Collectives which share similar agronomic fundamentals. Throughout this paper Ratemaking Collectives are assumed to be set at the level of the crop/state/indemnity level. However, it may be more appropriate to set Ratemaking Collectives at the level of the crop/agronomic region/indemnity level. The name Ratemaking Collective refers to the credibility smoothing of premium rates to be performed within these collectives.

Actuarial justification

Expected yields should be reasonably similar for products in the same Threshold Yield Collective. The shape of the yield distribution should be believed to be reasonably similar for products in the same Ratemaking Collective. Although the procedure described in this paper is robust to these two conditions not holding, the application of credibility theory is more useful if the conditions do hold.

Products in a Collective should be chosen based on a sound spatial, agronomic or practical rationale and should not be overly influenced by historical yield data for individual Insurance Units. Although different Collective specifications could be considered with the intention of lowering the resulting credibility factor, any specifications should contain a conceptually contiguous group of products. Products should not be included or excluded from a Collective based on their historical yields. Districts and States have been suggested as the main geographic determinant of Threshold Yield Collectives for practical reasons.

Data De-trending

Description of step

All data should be analyzed for a linear trend. It is suggested that calculations are performed at the Threshold Yield Collective level, but that judgment is applied at the crop level. The suggested procedure for such analysis is as follows.

First, at both the Threshold Yield Collective level and the crop level, calculate the weighted average yield for each of the last ten years. It is suggested that the weight used is the best estimate of the coming season's area insured, or zero if there is no yield data for that IU for that year:

$$\begin{aligned} & \text{Weighted average yield for year } n \\ &= \frac{\sum_{IUs \text{ with data for year } n} IU \text{ yield for year } n \times \text{estimate of area to be insured in IU}}{\sum_{IUs \text{ with data for year } n} \text{estimate of area to be insured in IU}} \end{aligned}$$

These weights could be based on area insured or area sown in one or more previous years. For example Table 1 lists the average historical yield for groundnut in six mNAIS districts. Data for area

sown or area insured was not available so all districts have been weighted equally, and within each district each insurance unit has been weighted equally.

Table 1. Groundnut: weighted average yield by district

State	District	Weighted Average Yield (kg/ha) for year									
		1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
AP	AP1	1,428	1,994	2,062	1,679	2,618	2,130	2,389	2,669	2,967	-
AP	AP2	1,164	1,449	1,098	1,735	1,770	2,471	2,719	2,121	2,771	-
AP	AP3	1,465	1,288	1,150	1,394	1,339	1,773	1,627	1,732	1,627	-
GA	GA1	1,957	1,535	1,959	1,796	1,436	1,949	1,483	1,617	2,277	3,126
GJ	GJ6	1,094	942	1,399	636	1,424	1,556	1,826	1,703	1,706	1,668
GJ	GJ13	1,243	1,124	1,361	1,280	1,488	2,007	1,123	1,444	1,220	1,367
All districts		1,363	1,548	1,425	1,598	1,857	2,121	2,215	2,126	2,426	2,347

Second, at both the Threshold Yield Collective level and the crop level, calculate the best estimate linear trend in weighted average yield and the p-value of a two tailed t-test with null hypothesis that there is no trend.³ Table 2 lists best estimate linear trends and p-value of trends for groundnut. As can be seen, the best estimate trends for groundnut vary from 14 kg/ha/year to 207 kg/ha/year for the six districts considered.

Table 2. Groundnut: Linear trend analysis by district

State	District	Best estimate linear trend (kg/ha/year)	P-value of trend	Percentage of statistically significant best estimate trend to remove for this crop	Amount of trend to be removed
[1]	[2]	[3]	[4]	[5]	[6]
AP	AP1	155	0.2%	75%	116
AP	AP2	207	0.1%	75%	155
AP	AP3	55	2.9%	75%	41
GA	GA1	82	14.4%	75%	-
GJ	GJ6	95	1.3%	75%	72
GJ	GJ13	14	65.7%	75%	-
All districts		125	0.0%		

Judgment should be applied to determine where, if anywhere, de-trending should be applied and the degree of de-trending, where de-trending is to be applied. The decision to de-trend data should be made with proper consideration of the impact any such decision will have on design and ratemaking. Removal of a statistically significant trend is likely to increase the threshold yield and reduce the commercial premium rate, resulting in a better value product for the farmer.

One potential approach would be to remove the same percentage of any statistically significant best estimate trend for each crop. This percentage could be set to zero for crops which do not appear to

³ These may be calculated using MS Excel functions as follows. Suppose cells A1:A10 contained ten yield data years (e.g. the numbers 1998, 1999, ... , 2007) and cells B1:B10 contained weighted average yields from year 1 to year 10. The linear trend, assumed to be calculated in cell C1, is given by the formula $SLOPE(B1:B10, A1:A10)$. The two sided t-value, assumed to be calculated in cell D1, is then given by $INDEX(LINEST(B1:B10, A1:A10, TRUE), 2)/C1$. Finally, the p-value is given by $TDIST(ABS(D1), COUNT(A1:A10) - 2, 2)$.

display justifiable trends across India and, for example, 75% for crops which do appear to display justifiable trends across India. For the case of groundnut, yield histories for four out of the six districts and the aggregate yield history for all six districts all show a strongly statistically significant linear trend, with p-value below 5%. Statistical analysis therefore supports the view that the trends have not occurred “by accident”. If qualitative analysis also supported this view, a trend removal percentage of 75% could be chosen for groundnut. Trends were not significant for districts GA1 or GJ13 and therefore no trend would be removed for these districts. However, 75% of the best estimate trend would be removed for the other four districts (column [6] of Table 2). Table 3 shows the raw and de-trended groundnut yield histories for an insurance unit in district GJ6, Gujarat before and after the removal of a trend of 72 kg/ha/year.

Table 3. Groundnut in a selected insurance unit in district GJ6 of Gujarat: illustration of yield de-trending

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Yield for insurance unit (kg/ha)	1,073	867	1,399	597	1,463	1,929	2,061	2,176	2,270	1,997
Addition to insurance unit-level yields to remove trend of 72kg/ha/year	+720	+648	+576	+504	+432	+360	+288	+216	+144	+72
Detrended yield for insurance unit (kg/ha)	1,793	1,515	1,975	1,101	1,895	2,289	2,349	2,392	2,414	2,069

The above procedure is likely to be sufficient for many technological or agronomic trends. However, if there are grounds for concern that historical data may feature any other forms of artificial trend or structural break, due to technological changes or the recording of yield data, further analysis would be necessary. For example, if the trend is believed to be constant across an agronomic region, the degree of de-trending could be based on based on crop/agronomic region-level analysis.

Actuarial justification

The statistical procedures suggested in this paper rely on the assumption that past experience is, at least in a probabilistic sense, a good guide to the future. If this assumption is not valid the data must be adjusted before application of the statistical procedures. For example, increased application of certain agricultural technologies or inputs could mean that the best estimate expected yield for next year is higher than the best estimate expected yield from ten years ago. In agriculture, such trends in yields are common and are usually corrected for by de-trending.

However, any statistical procedure designed to determine whether there is a trend in the data is exposed to the problem of a type-I error. Specifically, when testing at a statistical significance level of $5\% = \frac{1}{20}$, around 1 in 20 weighted average yield histories with no underlying trend in agronomic fundamentals would be expected to display statistically significant trends ‘by accident’. For example, if the most recent three years were exceptionally good for farmers due to unusually good weather, this may appear a statistically significant trend, even though it is actually just the historical realization of uncertainty.

When unsure about whether a pattern in yields is pure uncertainty or a structural trend, actuarial prudence suggests that an insurer should assume whichever gives the higher premium. The removal of a positive trend in yields from data will typically increase the threshold yield and decrease the premium rate, resulting in a better product being offered to farmers at a lower price. For a positive trend, removing a smaller trend is likely to be more prudent. Any trends that are considered to

have occurred by accident should therefore be left in the data; that is, no de-trending should be performed for this data. De-trending should aim to remove technological trends, such as the adoption of a high yielding crop variety, without removing natural variation in experience.

Expert judgment must be applied, not only in deciding whether to de-trend data or not, but also in choosing the actual trend to be removed. Statistical tests for de-trending typically test whether the data exhibits a statistically significant trend, not what the true trend is. For example, the low p-value for groundnut across the mNAIS portfolio suggests that there has been a statistically significant trend in yields, but it does not tell you that the true trend is equal to the best estimate trend of 125kg/ha/year. The true trend could be higher or lower, and could vary across districts. Removing a smaller trend is likely to be more prudent.

A 5% level of significance is suggested. A trend with p-value greater than 5% is not treated as being statistically significant but a trend with a p-value less than or equal to 5% is. A low p-value means that it is unlikely that the trend could have occurred 'by chance' and therefore it is more plausible that there has in fact been a trend in yields. Choice of significance level is a matter of judgment: a lower (higher) level leads to less (more) de-trending.

Other approaches to de-trending are possible. Aggregate linear de-trending could be conducted by crop/state if adoption of technology is considered to be similar across the state. Aggregate stepwise de-trending might be appropriate if a change is believed to have occurred over a very small period of time. Piecewise linear de-trending might be appropriate if more than ten years of data was used for ratemaking. It would also be possible to conduct de-trending individually for each insurance unit. However, in addition to removing any trends, it is likely that such a procedure would remove much of the natural variation in the data. For example, for an insurance unit that suffered poor weather for the first few years but very good weather for the final few years, de-trending would remove this natural variation even if it was not caused by a trend that is expected to continue over time.

If both the mean and variance of yields are increasing, a smaller trend should be removed than that suggested by the above approach. One implicit assumption of the suggested method is that the variance of yields does not change over time. For a more sophisticated approach to de-trending designed for the U.S. Federal Crop Insurance program, see Zhu et al. (2011).

3. Product Design: Calculation of Threshold Yields

A core principle of actuarial design and ratemaking is that the commercial premium received by an insurer for a given policy should reflect the actuarial cost to the insurer of that policy. It is understood that under modified NAIS, this *risk classification* will be conducted through a combination of crop/district-level commercial premium rates and indemnity levels and crop/insurance unit level probable yields.

The level of the threshold yield affects the actuarial cost of an insurance policy and therefore the commercial premium rate. Where commercial premium rates are determined on an actuarial basis, a decrease in the threshold yield, the deductible for yield-based payments under mNAIS, would decrease the commercial premium rate, and an increase in the threshold yield would correspondingly increase the commercial premium rate. The threshold yield is the multiple of the indemnity level and the probable yield, and so an increase in either the indemnity level or probable yield would increase the threshold yield.

One objective for the design of mNAIS insurance products is that the value of insurance products for the same crop in the same district should not vary significantly between insurance units.⁴ If the value of products does vary considerably between insurance units in the same district the insurer would be exposed to the threat of adverse selection, and insurance purchase in the district may become dominated by farmers in the insurance units for which the products were particularly good value.

One technically feasible solution would be for farmer and government premium rates to be fixed by government at a crop or crop/state level, and threshold yields to be calculated on an actuarial basis according to the principles of risk classification. Such an approach would have no actuarial disadvantages compared to risk classification through premium rates, but could yield significant political economy advantages by allowing fixed premium rates, and therefore fixed premium subsidies, across a state in addition to a low variation in threshold yields for a given crop across the state.

However, if rates under mNAIS are to be set at the crop/district level and the threshold yield in each insurance unit must be based on a seven year moving average yield, at the very least these moving average yields should be smoothed within the district to improve predictive stability. This second best solution is described and motivated in the suggested calculation of probable yields.

A rule for calculating indemnity limits is also suggested. The calculation rule for threshold yields affects the equity of mNAIS; if the variation in commercial premium rates is high within a state then state and central government subsidies may be unevenly spread between farmers, with farmers from insurance units with high commercial premium rates capturing a larger share of subsidies than other farmers. The indemnity limit rule is intended to reduce the variation in commercial premium rates within a state.

Probable Yields

Description of step

First, the yield data to be used in calculating the probable yield must be determined according to the rules of mNAIS. For the worked example of cotton in Gujarat, it is assumed that the probable yield is to be calculated using the seven years of yield data between 2004 and 2007 inclusive, after detrending by 50 kg/ha/year (columns [2] to [8] of Table 4).

Second, calculate the average detrended yield for each insurance unit (column [9] of Table 4).

⁴ There are much broader questions of political economy as to the degree to which the value of a mNAIS insurance product should be the same for different farmers growing different crops in different states. These broader questions are not addressed by this objective. Instead, the focus is on the technical issue of how products could be designed and priced so that, for example, the mNAIS cotton product sold in one insurance unit in one district offers the same expected claim payment to farmers as the cotton product sold in another insurance unit in the same district.

Table 4. Cotton in district GJ8 of Gujarat: calculation of seven year average yields

Insurance Unit	Yield, after removal of trend of 50 kg/ha/year (kg/ha)							Seven year average detrended yield for insurance unit (kg/ha)	Sample variance of detrended yield for insurance unit (kg ² /ha ²)
	2001	2002	2003	2004	2005	2006	2007		
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
GJ8_1	648	2,178	4,971	3,458	4,408	2,842	2,979	3,069	2,046,099
GJ8_2	1,054	324	4,541	3,716	3,985	3,357	2,451	2,775	2,480,213
GJ8_3	715	460	1,647	2,757	1,354	798	1,064	1,256	598,068
GJ8_4	656	379	2,805	2,651	2,995	2,428	1,912	1,975	1,113,738
GJ8_5	1,954	1,878	3,713	4,210	3,591	2,954	2,555	2,979	811,981
GJ8_6	3,077	1,654	4,285	3,272	3,377	2,719	2,133	2,931	747,051
GJ8_7	1,000	745	3,010	2,847	2,815	2,509	1,416	2,049	926,698
GJ8_8	1,000	745	3,010	2,847	2,815	2,509	1,522	2,064	905,942
GJ8_9	1,158	685	3,060	2,811	3,265	2,590	3,375	2,421	1,136,716
Average:								2,391	1,196,278
Sample variance:								366,300	

Third, calculate the Empirical Bayes Credibility Factor for probable yields Z_{PY} at the Threshold Yield Collective level using the yield data. For cotton in Gujarat the credibility factors for each district are given in Table 5.

Table 5. Cotton in Gujarat: Empirical Bayes credibility factors for probable yields

District	Number of insurance units in district	Empirical Bayes Credibility Factor for probable yields, Z_{PY}
GJ1	10	93%
GJ2	11	18%
GJ3	4	93%
GJ4	13	79%
GJ5	11	91%
GJ6	12	50%
GJ7	2	0%
GJ8	9	53%
GJ9	8	72%
GJ10	8	94%
GJ11	14	91%
GJ12	14	24%
GJ13	11	0%
GJ14	10	87%

A detailed breakdown of the credibility factor for district GJ8 is as follows:

Let X_{ij} denote the detrended yield for insurance unit i in year j . These are tabulated in columns [2] to [8] of Table 4;

$N = 9$, since there are nine insurance units in the district;

$n_i = \bar{n} = 7$, since all products have seven years of yield data to be used;

$\bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}$, is given in column [9] of Table 4;

- $\bar{X} = \frac{1}{N} \sum_{i=1}^N \bar{X}_i = 2,391$ kg/ha (see the bottom of column [9] of Table 4);
- $VAR = \frac{1}{N-1} \sum_{i=1}^N (\bar{X}_i - \bar{X})^2 = 366,300$ (see the bottom of column [9] of Table 4);
- $VAR_i = \frac{1}{n_i-1} \sum_{j=1}^{n_i} (LC_{ij} - \bar{LC}_{ij})^2$ the sample variance of yields for each insurance unit i is given in column [10] of Table 4;
- $\mathbb{E}[s^2(\theta)] = \frac{1}{N} \sum_{i=1}^N VAR_j(i) = 1,196,279$ (see the bottom of column [10] of Table 4);
- $VAR[m(\theta)] = \max\left(VAR - \frac{\mathbb{E}[s^2(\theta)]}{\bar{n}}, 0\right) = 195,403$;
- $K = \frac{\mathbb{E}[s^2(\theta)]}{VAR[m(\theta)]} = 6.122$;
- $Z_{PY} = \frac{n_i}{n_i + K} = 53\%$.

Fourth, calculate the weighted average de-trended yield at the Threshold Yield Collective level, where the weight is the same portfolio weight used in the de-trending step. For cotton in Gujarat, the de-trending step used area sown in 2007 as the portfolio weight. This gives a weighted average de-trended yield of 2,465 kg/ha for cotton in district GJ8 of Gujarat (column [4] of Table 6). Note that this differs from the unweighted average of 2,391 calculated in Table 4.

Fifth, calculate the smoothed probable yield for each insurance unit as the weighted average of the insurance unit and Threshold Yield Collective de-trended yields, where the Empirical Bayes Credibility Factor is used as the weight. For cotton in Gujarat the smoothed probable yields are given in column [6] of Table 6.

Table 6. Cotton in district GJ8 of Gujarat: probable yield calculations

Insurance Unit	Area Sown 2007 (ha)	Average yield for IU (kg/ha)	Weighted average yield for district (kg/ha)	Probable Yield Credibility Factor, Z_{PY}	Smoothed Probable Yield (kg/ha) [3] × [5] + [4] × (1 - [5])
[1]	[2]	[3]	[4]	[5]	[6]
GJ8_1	29,000	3,069	2,465	53%	2,787
GJ8_2	20,300	2,775	2,465	53%	2,631
GJ8_3	21,894	1,256	2,465	53%	1,820
GJ8_4	26,440	1,975	2,465	53%	2,204
GJ8_5	38,000	2,979	2,465	53%	2,739
GJ8_6	16,500	2,931	2,465	53%	2,714
GJ8_7	10,000	2,049	2,465	53%	2,243
GJ8_8	18,000	2,064	2,465	53%	2,251
GJ8_9	4,810	2,421	2,465	53%	2,441

Actuarial justification

If it has been judged that the yield data should be de-trended it is appropriate and consistent to use de-trended data for calculation of both probable yields and premium rates. If, for example, there is judged to have been a large positive trend in yields over the last ten years but this trend is not removed for the probable yield calculations, the resulting threshold yields will be unreasonably low. Farmers would be offered products with low commercial premium rates that offered very little protection. Calculating probable yields based on de-trended threshold yields would result in products that offered a more typical degree of protection for a commensurately higher commercial premium rate.

The removal of (up to two) official calamity years in calculating the probable yield acts to increase the threshold yield and associated commercial premium rate and government subsidy rate for those insurance units in which droughts have been officially reported. The effect is therefore to target subsidies to districts in which droughts have been reported in the last seven years.

Appropriate smoothing of probable yields within districts decreases the standard error of average yield estimates, increasing equity and decreasing the potential threat from adverse selection. It is understood that probable yields are to be based on the most recent seven year average of non-drought year yields. Using seven years of data is likely to increase the stability of probable yields, relative to using three or five years of data as was required under the NAIS. The probable yield for a given insurance unit could be calculated as a weighted average of historical yields for that insurance unit only, as done for NAIS. However, if this was the case, a large part of the resulting variation in threshold yields within a district could be caused by natural statistical variation, rather than a variation in underlying agronomic characteristics; seven data points is not enough to obtain an accurate estimate of the mean yield unless the risk faced is very low. If premium rates were then set at the crop/district level farmers would be able to adversely select against the insurance program by participating when coverage is high or not insuring when coverage is low.

It is suggested that probable yields within a Threshold Yield Collective are smoothed through the use of within- Threshold Yield Collective credibility weighting. The formula for the probable yield is then:

$$\text{Probable yield (PY)} = \text{IU average yield} \times Z_{PY} + \text{district average yield} \times (1 - Z_{PY})$$

where Credibility Factor Z_{PY} is calculated objectively and can range from 0% (no credibility assigned to yield history for insurance unit) to 100% (full credibility assigned to yield history for insurance unit). Z_{PY} should be 0% when the crop/insurance unit average yield history is statistically uninformative, compared with the history for the full district. Correspondingly, Z_{PY} should be 100% when the crop/district level average yields are not statistically useful for estimating the average yield.

Of particular interest is the credibility factor approach derived by Bühlmann (1967), which offers the best linear approximation to unconstrained Bayesian estimates.⁵ Denoting the historical loss from product i in year j as X_{ij} this *Empirical Bayes Credibility Factor* Z is given by

$$Z = \frac{n}{n + \mathbb{E}[\widehat{s^2(\theta)}] / \text{Var}[\widehat{m(\theta)}]} \quad (1)$$

where n and N denote the number of years of data and the number of products in the Threshold Yield Collective respectively, and

$$\bar{X}_i = \frac{1}{n} \sum_{j=1}^n X_{ij}$$

⁵ Textbook treatments include Herzog (1999) and Bühlmann and Gisler (2005).

$$\bar{X} = \frac{1}{nN} \sum_{i=1}^N \sum_{j=1}^n X_{ij}$$

$$\mathbb{E}[s^2(\theta)] = \sum_{i=1}^N \sum_{j=1}^n \frac{(X_{ij} - \bar{X}_i)^2}{N(n-1)}$$

$$Var[m(\theta)] = \left[\sum_{i=1}^N \frac{(\bar{X}_i - \bar{X})^2}{N-1} - \frac{\mathbb{E}[s^2(\theta)]}{N} \right]^+$$

This Credibility Factor satisfies intuitive properties and is robust in a range of scenarios. Z increases if there is more data for the insurance unit itself, the variation of average yields for each insurance unit decreases, or the variation of average yields between insurance units in the same district increases. Smoothing probable yields within a district using the Empirical Bayes Credibility Factor formula is most likely to reduce the variation in the actuarial value of products within the district, protecting the mNAIS portfolio against adverse selection.⁶ However, other objective formulae for a credibility factor are possible.

For the current example, the Empirical Bayes Credibility Factor varies from 0% to 94% for districts in Gujarat. A high credibility factor is suggestive of agronomic heterogeneity within the district and a low credibility factor is correspondingly suggestive of agronomic homogeneity. For district GJ8 of Gujarat, only 53% of the variation in average yields appears to be credible, with the remaining 47% arising from natural statistical variation.

Indemnity Levels

Description of step

It is understood that indemnity levels of 70%, 80% or 90% are to be set at the crop/district level under the modified NAIS. The indemnity level could be calculated as follows.

First, for each insurance unit calculate the ten year average loss costs at both 70% and 90% indemnity levels ($ALC_{70\%}$ and $ALC_{90\%}$) where the average loss cost at an indemnity level of IL is given by:

$$ALC_{IL} = \sum_{\text{last ten years}} \frac{\max(0, PY \times IL - \text{Actual detrended yield})}{PY \times IL}$$

Average historical loss costs for insurance units in district GJ8 of Gujarat are given in columns [4] and [5] of Table 7.

Second, calculate the ten year weighted average loss cost at the crop/district level at both 70% and 90% indemnity levels. The weighted average is taken over all insurance units in the district, where the weight is the same portfolio weight used in the de-trending and probable yield steps (bottom of columns [4] and [5] of Table 7).

⁶ Since seven years of data is used for probable yield calculations but ten years is used for ratemaking there is a possibility that this probable yield credibility factor step could reduce the variation in probable yields but increase the variation of commercial premium rates. However, based on data already analysed this seems to be unlikely.

Table 7. Cotton in district GJ8 of Gujarat: average and weighted average loss costs at 70% and 90% indemnity levels

Insurance Unit	Area Sown 2007 (ha)	Probable Yield (kg/ha)	Average historical loss cost	
			70% indemnity level ($ALC_{70\%}$)	90% indemnity level ($ALC_{90\%}$)
[1]	[2]	[3]	[4]	[5]
GJ8_1	29,000	2,787	17%	21%
GJ8_2	20,300	2,631	14%	19%
GJ8_3	21,894	1,820	26%	37%
GJ8_4	26,440	2,204	19%	24%
GJ8_5	38,000	2,739	10%	17%
GJ8_6	16,500	2,714	8%	14%
GJ8_7	10,000	2,243	23%	31%
GJ8_8	18,000	2,251	19%	26%
GJ8_9	4,810	2,441	18%	23%
Weighted average loss cost			16%	23%

The indemnity level for each district may then be chosen as a function of the weighted average loss costs at 70% and 90% indemnity levels (see Table 8). Cutoffs x and y in Table 8 should be chosen to target commercial premium rates. An increase in x or y would increase the average indemnity level across the mNAIS portfolio, and commensurately increase the average commercial premium rate and government subsidy rate.

Table 8. Indemnity level loss cost cutoffs

	Rule	Indemnity limit
	District Loss Cost at 90% indemnity limit less than x	90%
	District Loss Cost at 70% indemnity limit greater than y	70%
	Otherwise	80%

Note that selecting x and y to be between 2.5% and 3.0% is broadly equivalent to the CV rule under the current NAIS, based on the Normal Theory Method. Using this indemnity level cutoff rule it is possible to calculate the indemnity levels for cotton for all districts in Gujarat (column [4] of Table 9). Under the NAIS NTM method, if the coefficient of variation (CV) of yields for the last ten years was less than 16, corresponding to an NTM pure premium rate of 2.9%, an indemnity level of 90% was chosen. If the CV was greater than 31, corresponding to an NTM pure premium rate of 2.4%, an indemnity level of 60% was chosen. Otherwise an indemnity level of 80% was chosen.

Table 9. Cotton in Gujarat: indemnity levels

District	Weighted average loss cost at 70% indemnity level	Weighted average loss cost at 90% indemnity level	Indemnity level
[1]	[2]	[3]	[4]
GJ1	0.6%	5.9%	80%
GJ2	8.6%	14.8%	70%
GJ3	4.8%	11.4%	70%
GJ4	2.3%	6.1%	80%
GJ5	0.4%	3.7%	80%
GJ6	10.6%	17.8%	70%
GJ7	0.4%	3.1%	80%
GJ8	16.3%	22.6%	70%
GJ9	8.2%	15.1%	70%
GJ10	4.9%	10.5%	70%
GJ11	0.6%	5.1%	80%
GJ12	22.1%	27.2%	70%
GJ13	6.7%	14.9%	70%
GJ14	3.3%	9.2%	70%

Actuarial justification

For a particular crop the average yield may vary across a district, but the shape of the yield distribution may be fairly constant across the district. In so far as this is the case it is appropriate to set the indemnity level to be fixed across a district.

The specific rule suggested for the calculation of indemnity level corresponds to the existing rule for NAIS, based on the Normal Theory Method.

The large variation in weighted average loss costs at both 70% and 90% indemnity levels (column [2] of Table 9) indicates that cotton yield risk varies considerably between districts in Gujarat, with low risk in districts like GJ7 and GJ5 and very high risk in districts like GJ12 and GJ8. The actuarial commercial premium rates of mNAIS would allow different rates to be charged in different insurance units within the same state. Moreover, the suggested rule for calculating indemnity levels would allow farmers in high risk districts to purchase a lower level of cover at a lower premium.

Threshold Yields

Description of step

The threshold yield for each insurance unit may then be calculated as:

$$\text{Threshold Yield (TY)} = \text{Probable Yield (PY)} \times \text{Indemnity level (IL)}$$

Actuarial justification

This approach is appropriate when the shape of the yield distribution is fairly constant across the district. However, if the shape of the yield distribution varies across the district, with differences in risk in addition to differences in average yields, this will result in a variation in premium rates across a district. In such circumstances a flexible experience based approach to setting threshold yields, with fixed premium rates would be more appropriate.

4. Pure Premium Rates

Under mNAIS premium rates are set at the crop/district level and so the pure premium rate is taken to be an actuarially sound estimate of the expected weighted average loss cost across the crop/district, that is, the expected claim payment divided by the sum insured, for the coming crop season. At this stage the reader may find it helpful to review Figure 1, which offers an overview of how the following steps fit together.

Loss Costs

Description of step

First calculate the historical loss costs for each insurance unit with the threshold yields calculated in Section 3. This is the same as the calculation of loss cost histories for the calculation of the indemnity level (see column [4] of Table 7), except that the indemnity limit is no longer necessarily 90%. For example, the indemnity level for cotton in district GJ8 of Gujarat was calculated to be 70% (see column [3] of Table 9). Threshold yields with indemnity level of 70% are reported in column [2] of Table 10 and historical loss costs, based on these threshold yields and using de-trended data, are given in columns [3] to [12] of Table 10.

Table 10. Cotton in district GJ8 of Gujarat: historical loss costs

Insurance Unit	Threshold yield (kg/ha)	Historical loss cost									
		1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
GJ8_1	1,951	0%	21%	78%	67%	0%	0%	0%	0%	0%	0%
GJ8_2	1,842	0%	0%	16%	43%	82%	0%	0%	0%	0%	0%
GJ8_3	1,274	0%	46%	54%	44%	64%	0%	0%	0%	37%	17%
GJ8_4	1,543	0%	0%	56%	57%	75%	0%	0%	0%	0%	0%
GJ8_5	1,918	0%	42%	58%	0%	2%	0%	0%	0%	0%	0%
GJ8_6	1,900	0%	18%	48%	0%	13%	0%	0%	0%	0%	0%
GJ8_7	1,570	33%	29%	66%	36%	53%	0%	0%	0%	0%	10%
GJ8_8	1,576	0%	29%	66%	37%	53%	0%	0%	0%	0%	3%
GJ8_9	1,709	0%	35%	53%	32%	60%	0%	0%	0%	0%	0%

Second calculate the weighted average historical loss costs for each district for each of the last ten years. This is the average loss cost that would have been payable in the district that year, if the portfolio had been as per the assumed portfolio weights. Table 11 gives these for each district in Gujarat.

Table 11. Cotton in Gujarat: weighted average historical loss costs for each district in each year

District	District Area Sown 2007 (ha)	Weighted average historical loss cost for district									
		1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
GJ1	190,999	0%	1%	5%	3%	8%	1%	1%	1%	5%	2%
GJ2	198,051	0%	9%	34%	30%	11%	0%	2%	0%	0%	0%
GJ3	37,195	2%	9%	9%	13%	11%	0%	0%	0%	3%	0%
GJ4	186,303	0%	2%	9%	25%	1%	0%	0%	0%	4%	1%
GJ5	160,974	0%	0%	2%	2%	0%	3%	0%	0%	1%	7%
GJ6	343,553	1%	14%	39%	36%	9%	5%	2%	0%	0%	1%
GJ7	21,024	0%	0%	7%	6%	0%	0%	0%	0%	0%	0%
GJ8	184,944	2%	24%	56%	35%	39%	0%	0%	0%	4%	3%
GJ9	55,857	0%	17%	8%	16%	31%	5%	1%	1%	2%	2%
GJ10	56,054	4%	12%	12%	4%	14%	0%	4%	0%	0%	0%
GJ11	94,726	0%	2%	3%	2%	4%	3%	2%	0%	2%	6%
GJ12	281,750	2%	27%	62%	49%	68%	2%	3%	2%	4%	2%
GJ13	98,831	0%	4%	5%	3%	34%	0%	0%	0%	19%	2%
GJ14	448,126	0%	0%	10%	4%	14%	4%	1%	0%	0%	0%

Actuarial justification

Historical loss costs are the fundamental building block for experience-based approaches to ratemaking. This differs from the Normal Theory Method used for NAIS, under which the average historical yield and coefficient of variation are the fundamental building blocks. Since premium rates will be calculated at the crop/district level, it is appropriate for the weighted average crop/district level loss costs to be the fundamental building block for rates.

Historical loss costs are statistically informative if historical yields are, at least in a probabilistic sense, a good guide to the future. If the actuary has judged that the yield data should be de-trended it is appropriate and consistent to use de-trended data for calculation of both probable yields and premium rates. This increases the predictive quality of historical loss costs. If, for example, there is judged to have been a large positive trend in yields over the last ten years but this trend is not removed for ratemaking calculations, the resulting premium rate will be unreasonably high.

Capping of Loss Costs

Description of step

Judgment should be applied to choose a loss cost cap percentile (LCCP) at the crop/state level. This LCCP is applied to the data as follows.

For each district, calculate the LCCPth percentile historical loss cost. This is called the loss cost cap (LCC). For example, if the LCCP was the 90th percentile, the LCC for cotton in district GJ8 would be 40%. This is calculated by linearly interpolating between the 100th percentile historical loss cost of 56.0% and the $100 \times \frac{8}{9}$ th percentile of 38.5% to give $38.5\% + \frac{56.0\% - 38.5\%}{90} = 40\%$.⁷

⁷This can be implemented in MS Excel using the PERCENTILE function as follows. If E11:N11 is the range for the most recent ten loss costs and LCCP=90%, LCC= PERCENTILE(E11:N11,90%).

The historical capped loss costs for each insurance unit are then given by $\min(\text{loss cost}, LCC)$. For cotton in district GJ8, only the 2000 loss cost of 56% exceeds the LCC of 40% and so only that loss cost is reduced to 40% in this step. Table 12 lists the LCC based on an LCCP of 90, and capped loss cost histories for all insurance units in the Gujarat.

Table 12. Cotton in Gujarat: capped weighted average historical loss costs

District	Loss cost cap (LCC) at 90%	Capped Weighted average historical loss cost for district									
		1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
GJ1	5%	0%	1%	5%	3%	5%	1%	1%	1%	5%	2%
GJ2	30%	0%	9%	30%	30%	11%	0%	2%	0%	0%	0%
GJ3	11%	2%	9%	9%	11%	11%	0%	0%	0%	3%	0%
GJ4	10%	0%	2%	9%	10%	1%	0%	0%	0%	4%	1%
GJ5	3%	0%	0%	2%	2%	0%	3%	0%	0%	1%	3%
GJ6	36%	1%	14%	36%	36%	9%	5%	2%	0%	0%	1%
GJ7	6%	0%	0%	6%	6%	0%	0%	0%	0%	0%	0%
GJ8	40%	2%	24%	40%	35%	39%	0%	0%	0%	4%	3%
GJ9	18%	0%	17%	8%	16%	18%	5%	1%	1%	2%	2%
GJ10	12%	4%	12%	12%	4%	12%	0%	4%	0%	0%	0%
GJ11	4%	0%	2%	3%	2%	4%	3%	2%	0%	2%	4%
GJ12	62%	2%	27%	62%	49%	62%	2%	3%	2%	4%	2%
GJ13	21%	0%	4%	5%	3%	21%	0%	0%	0%	19%	2%
GJ14	11%	0%	0%	10%	4%	11%	4%	1%	0%	0%	0%

Actuarial justification

The intention of the cap and Capped Excess Load is to spread statistically insignificant large losses over a large base so that pricing is robust to outliers. For example, one very bad crop year in a particularly district may be just as likely to have happened in other districts within the state. It may therefore be appropriate to spread the effect on rates of such a year over a broader base than the district.

Bühlmann's Empirical Bayes Credibility Theory, which will be applied in the next step, is a linear process, and thus extreme outliers can present difficulties requiring special attention. A technique commonly used in conjunction with credibility weighting is to cap large losses before the application of Credibility Theory. Capping large losses then adding back the probability mass may be actuarially sound as infrequent events lack statistically credibility. A properly chosen cap may not only add stability, but may even make the methodology more accurate by eliminating extremes.

Actuarial judgment is required in the choice of cap percentile, as this may substantially affect the degree of smoothing of rates within the state. If the LCCP is set to be too high, the presence of outliers may lead to a low credibility factor Z . If the LCCP is set to be too low, the LCC could be 0% for a large number of districts. Based on the data analysed in this note a LCCP between 85% and 100% is suggested, although this may depend on the specific historical data and historical claim payment frequency. The LCCP could be chosen at the crop/state level.

Since extreme losses are removed from each insurance unit's experience in this step, it is appropriate to add them back later at a broader level, so that rates are not underestimated. These excess amounts will be added back in the *Capped excess load* subsection below.

Credibility Theory

Description of step

This step is similar to the credibility step in calculating probable yields. However, instead of smoothing average yields within a district, weighted average historical loss costs are smoothed within the state.

First, the Product Base Rate (PBR) is defined as the average capped loss cost for that district. For district GJ8, using the capped loss cost history of Table 12, this is calculated as

$$\frac{(2\%+24\%+40\%+35\%+39\%+0\%+0\%+0\%+4\%+3\%)}{10} = 14.7\%$$
. Product base rates are displayed for all insurance units in the Gujarat in column [3] of Table 13.

Table 13. Cotton in Gujarat: calculation of base pure rate

District	District Area Sown 2007 (ha)	Product Base Rate (PBR)	Sample variance of historical capped loss costs	Weighted average PBR across state	Ratemaking Credibility Factor, Z_R	Base Pure Rate [3] × [6] + [5] × (1 - [6])
[1]	[2]	[3]	[4]	[5]	[6]	[7]
GJ1	190,999	2.5%	0.0%	7.6%	67%	4.2%
GJ2	198,051	8.2%	1.5%	7.6%	67%	8.0%
GJ3	37,195	4.5%	0.2%	7.6%	67%	5.6%
GJ4	186,303	2.6%	0.1%	7.6%	67%	4.3%
GJ5	160,974	1.2%	0.0%	7.6%	67%	3.3%
GJ6	343,553	10.3%	2.0%	7.6%	67%	9.4%
GJ7	21,024	1.2%	0.1%	7.6%	67%	3.3%
GJ8	184,944	14.7%	3.1%	7.6%	67%	12.4%
GJ9	55,857	6.9%	0.5%	7.6%	67%	7.2%
GJ10	56,054	4.8%	0.3%	7.6%	67%	5.7%
GJ11	94,726	2.1%	0.0%	7.6%	67%	3.9%
GJ12	281,750	21.6%	6.9%	7.6%	67%	16.9%
GJ13	98,831	5.4%	0.6%	7.6%	67%	6.1%
GJ14	448,126	3.0%	0.2%	7.6%	67%	4.6%
Unweighted average:		6.3%	1.12%			
Sample variance:		0.338%				

Second, calculate the Empirical Bayes Credibility Factor for ratemaking (Z_R) at the crop/state level using the capped loss cost histories of Table 12. For cotton in Gujarat the credibility factor of 67% is calculated as follows:

Let X_{ij} denote the weighted average capped loss cost for district i in year j . These are tabulated in columns [3] to [12] of Table 12;

$N = 14$, since there are fourteen districts in the state;

$n_i = \bar{n} = 10$, since all products have ten years of loss cost history to be used;

$\bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}$, the product base rate is given in column [3] of Table 13;

- $\bar{X} = \frac{1}{N} \sum_{i=1}^N \bar{X}_i = 6.3\%$ (see the bottom of column [3] of Table 13);
- $VAR = \frac{1}{N-1} \sum_{i=1}^N (\bar{X}_i - \bar{X})^2 = 0.338\%$ (see the bottom of column [3] of Table 13);

- $VAR_i = \frac{1}{n_i-1} \sum_{j=1}^{n_i} (LC_{ij} - \overline{LC}_{ij})^2$ the sample variance of yields for each insurance unit i is given in column [4] of Table 13;
- $\mathbb{E}[s^2(\theta)] = \frac{1}{N} \sum_{i=1}^N VAR_j(i) = 1.12\%$ (see the bottom of column [4] of Table 13);
- $VAR[m(\theta)] = \max\left(VAR - \frac{\mathbb{E}[s^2(\theta)]}{\bar{n}}, 0\right) = 0.226\%$;
- $K = \frac{\mathbb{E}[s^2(\theta)]}{VAR[m(\theta)]} = 4.95$;
- $Z_{PY} = \frac{n_i}{n_i+K} = 67\%$.

Third, calculate the weighted product base rate at the crop/state level, where the weight is the same portfolio weight used in the de-trending step. For cotton in Gujarat, the de-trending step used area sown in 2007 as the portfolio weight. This gives a weighted base rate of 7.6% for cotton in Gujarat (column [5] of Table 13). Note that this differs from the unweighted average of 6.3%, reported at the bottom of column [3] of Table 13.

Fourth, calculate the base pure rate for each insurance unit as the weighted average of the product base rate and weighted base rate, where the Empirical Bayes Credibility Factor for ratemaking (Z_R) is used as the weight. For cotton in Gujarat the base pure rates are given in column [7] of Table 13.

Actuarial justification

Credibility refers to the degree of belief in a particular source of data. Credibility is a relative concept and is greater the more relevant the data source, and the greater the number of observations in the data source.

The justification for smoothing product base rates within each state is similar to the earlier justification for smoothing average yields within each district. In both cases, some of the within class variation may be attributed to natural statistically variation, rather than a within-class variation in the fundamental data generating process. However, while the credibility smoothing of probable yields reduced the variation in threshold yields within the district, the credibility smoothing of base rates reduces the variation in premium rates across the state.

The credibility theory step does not change the weighted average base rate within the state; instead it re-spreads so that rates in the state only differ in as much as the difference is statistically significant. If historical capped loss costs in different districts are not statistically different, the premium will be the same for all these products. The greater is the statistical significance of any difference, the greater is the difference in base rates after the credibility step.

As for the credibility factor for probable yields, credibility factor Z_R can range from 100% (full credibility assigned to capped loss cost history for insurance unit) to 0% (no credibility assigned to capped loss cost history for insurance unit). Z_R should be 100% when the crop/state level average capped loss cost is not statistically useful for determining the district-level average capped loss cost. Correspondingly, Z_R should be 0% when the crop/district level average capped loss cost history is statistically uninformative, compared with the history for the full state.

It is suggested that credibility factor Z_{PY} be calculated using Bühlmann's Empirical Bayes Credibility Factor formula. This Credibility Factor satisfies intuitive properties and is robust in a range of scenarios. Z increases if there is more data for the district itself, the variation of loss costs for each district decreases, or the variation of average loss costs between districts in the same state increases.

The Empirical Bayes Credibility Factor is 67% for Gujarat, suggesting that two thirds of the variation in premium rates within Gujarat is due to differences in agronomic fundamentals, but one third is due to natural statistical variation. A high credibility factor is suggestive of heterogeneity in the

value of products to be sold within the district and a low credibility factor is correspondingly suggestive of homogeneity in the value of products to be sold in the district.

Capped Excess Load

Description of step

First calculate the weighted average *uncapped* loss cost within the Ratemaking Collective. This is similar to the calculation of the weighted base rate (7.6% for cotton in Gujarat, reported in column [5] of Table 13) but where the uncapped loss costs of Table 11 are used instead of the capped loss costs of Table 12. For cotton in Gujarat, this gives a weighted average uncapped loss cost of 8.23%.

The capped excess load is then this weighted average uncapped loss cost minus the weighted base rate. For cotton in Gujarat this is 0.59%.

Table 14. Cotton in Gujarat: capped excess load

State	Weighted average uncapped loss cost	Weighted average capped loss cost (=weighted base rate)	Capped Excess Load = [3] – [2]
[1]	[2]	[3]	[4]
Gujarat	8.23%	7.65%	0.59%

Actuarial justification

The capped excess load is calculated so that, when added to base pure rates, the weighted average pure premium rate across the Ratemaking Collective is equal to the historical weighted average loss cost across the Ratemaking Collective. In the section *Capping of Loss Costs*, each Threshold Yield Collective loss cost history was capped at the Loss Cost Cap Percentile. The excess amounts must now be added back or the rates will be underestimated. The capped excess load could be added in proportion to pure premium rates or sum insured. Here the capped excess load has been added back in proportion to the sum insured.

Other Additional Benefits

Description of step

When there is the possibility that an early on account payment could be higher than the eventual area yield-based payment, and any such on account overpayment could not be retrieved by the insurer, this must be allowed for in rates. If it is judged that such an eventuality occurring very unlikely, so as not to materially affect the average future loss cost, then no allowance is necessary in rates. Otherwise, an additional load may need to be added to the pure premium rate, based on the expected state-wide on account overpayment rate.

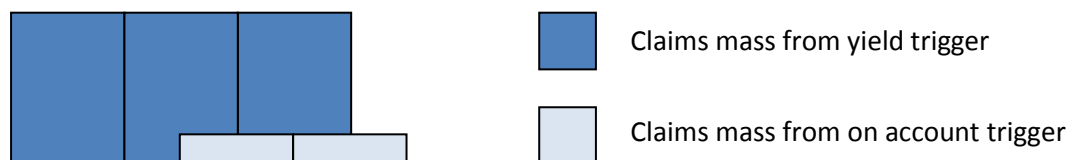
If other additional benefits are to be offered to farmers, for example due to extended coverage for planting and post-harvest risks, the cost of any such benefits must also be allowed for in the pure premium rate.

Actuarial justification

With on account payments limited to 25% of the likely claim, and payable only when yields are believed to be less than 50% of the probable yield, the effect on rates of on account overpayment is likely to be small. However, the insurer should ensure that any on account overpayment is included in the pure premium rate. As a purely hypothetical example, suppose that in a particular district, the yield triggers a claim to farmers every 3 in 10 years. When a claim occurs, the average claim payment to farmers is 25% of total sum insured, and so the average loss cost is 7.5%. Now suppose

that there is an on account payment to farmers every 2 in 10 years with average on account payment of 5% of total sum insured. Finally, suppose that in one quarter of those years there is no yield-triggered claim payable.

Figure 2: Motivation for on account load



In this hypothetical example, the on account payment would increase the pure premium rate from 7.50% to 7.75% of total sum insured.

The actuary may judge that additional benefits from the on account payment do not merit an adjustment to be made to rates. Any such judgment should be noted and clearly documented.

Pure Premium Rate

Description of step

The pure premium rate at the crop/district level is then taken to be:

$$\text{Pure Premium Rate} = \text{Base Pure Rate} + \text{State Excess Load} + \text{Rate for other benefits}$$

For cotton in Gujarat, assuming that the *Rate for other benefits* is zero, the pure premium rate calculation for each district are given in Table 15.

Table 15. Cotton in Gujarat: pure premium rates

District	Base Pure Rate	Capped Excess Load	Assumed rate for other benefits	Pure Premium Rate = [2] + [3] + [4]
[1]	[2]	[3]	[4]	[5]
GJ1	4.2%	0.6%	0.0%	4.8%
GJ2	8.0%	0.6%	0.0%	8.6%
GJ3	5.6%	0.6%	0.0%	6.1%
GJ4	4.3%	0.6%	0.0%	4.9%
GJ5	3.3%	0.6%	0.0%	3.9%
GJ6	9.4%	0.6%	0.0%	10.0%
GJ7	3.3%	0.6%	0.0%	3.9%
GJ8	12.4%	0.6%	0.0%	12.9%
GJ9	7.2%	0.6%	0.0%	7.8%
GJ10	5.7%	0.6%	0.0%	6.3%
GJ11	3.9%	0.6%	0.0%	4.5%
GJ12	16.9%	0.6%	0.0%	17.5%
GJ13	6.1%	0.6%	0.0%	6.7%
GJ14	4.6%	0.6%	0.0%	5.1%

5. Commercial Premium Rates

The commercial premium rate is an actuarially sound estimate of the expected total cost associated with products sold across the crop/district for the coming crop season, divided by the expected sum insured. Typically, the commercial premium rate is based on the pure premium rate, but with additional loads to account for administrative expenses, the cost of risk capital (including reinsurance costs) and a profit margin.

The insurer may find it prudent to add loads to allow for the uncertainty associated with ratemaking with inadequate data and the potential threat of moral hazard. In the normal course of events such margins would be expected to lead to growth of the insurer's reserves. Under an actuarial regime it is in the long terms interests of farmers and government for insurers to develop and maintain healthy reserves, so that it is able to sustainably retain an appropriate amount of risk thereby reducing reinsurance costs.

Loads are typically calculated at the level of the portfolio and then split between products through use of an additive or multiplicative load. The basic formula for the commercial premium is:

$$\text{Commercial Premium Rate} = (\text{Pure Premium Rate} + \text{Additive Load}) \times \text{Multiplicative Load}$$

An additive load splits the total cost between policies in proportion to the total sum insured, whereas a multiplicative load splits the total cost in proportion to the commercial premium rate. The cost of capital load is typically added additively and the administrative load is typically added multiplicatively.

Although in the remainder of this section a number of separate loads are discussed, it is important to consider the loading on an aggregate basis, to ensure that the aggregate load is prudent, but not excessively high. Constituent and aggregate loads should be determined by management on the advice of the actuary. Discussion of loading when ten years of data is not available at the insurance level is deferred to Section 0.

Cost of Capital

Description of step

Designing a risk financing strategy is a complex topic that is not covered here. This paper is instead focused on design and ratemaking for mNAIS. Broadly, an insurer will wish to acquire risk capital in the form of reserves, reinsurance and contingent credit to ensure that it is able to pay all claims as they fall due, with sufficiently likelihood. International experience suggests that an insurer would acquire enough risk capital to remain solvent over the coming year with probability greater or equal to 99.5%.

It is suggested that the cost of capital is calculated using an aggregate portfolio approach, based on an assumed portfolio and the insurer's choice of risk financing strategy. This means that the cost of capital is calculated so that, over the entire expected portfolio, the total additional premium income from the cost of capital element should equal the total cost to the insurer of acquiring risk capital, where:

The (expected) cost of reinsurance capital is equal to

$$\text{reinsurance premium} - \text{expected income from reinsurance claims};$$

The cost of reserves is equal to

$$\text{reserves} \times \text{required return on capital over the season};$$

The cost of contingent credit is equal to

$$\text{maximum loan} \times \max(0, \text{contingent credit multiple} - 1).^8$$

The cost of capital is expressed as a rate, to be added to the pure premium rate, and is therefore equal to this aggregate cost to the insurer of acquiring risk capital for a given portfolio divided by the total sum insured for that portfolio.

Where the insurer must quote premiums for products before the entire portfolio is finalized, judgment must be applied in assessing the level of diversification in the forthcoming season's mNAIS portfolio. The more diversified the portfolio, the less risk capital will be required, and the lower the cost of capital will be. If there is some uncertainty surrounding the level of diversification in the forthcoming portfolio the insurer may wish to estimate the cost of capital on a prudent, or cautious, basis.

The risk management team can support management decisions about the risk financing strategy by estimating the portfolio Probable Maximum Loss (PML) using historical data. See Chapters 4 and 5 and Annex G of World Bank (2011) for examples of such calculations for the WBCIS and NAIS portfolios.

Actuarial justification

The cost of acquiring capital to allow for the risk of random variation from expected claims should be included in the commercial premium rate. The insurer must acquire capital in the form of reserves, reinsurance and contingent credit to ensure that it is able to pay all claims as they fall due. Most regulatory regimes require private insurance companies to demonstrate that they hold enough capital to be able to survive specified extreme events; for example the Europe Union's forthcoming Solvency II will require each (re)insurer to demonstrate that it is be able to meet its obligations over the next 12 months with a probability of at least 99.5%.

Administrative Load

Description of step

The administrative load should be sufficient to cover all costs to the insurer of administering the mNAIS and is a decision for the insurer's management team.

Actuarial justification

All costs to the insurer of providing benefits, including administrative costs, should be included in commercial premium rates.

⁸ A formula for the contingent credit multiple is given in Clarke and Mahul (2011). The contingent credit multiple is less than unity, and therefore the contribution of contingent credit to the cost of capital is zero, when the opportunity cost of capital is greater than the rate of interest that would be paid on drawn down credit. If this does not hold then the contingent credit multiple will be greater than unity. For example, if the opportunity cost of capital is 2%, the rate of interest payable on drawn down credit is 5%, and the term of any drawn down credit is 30 years with a five year grace period the contingent credit multiple would be approximately 1.4 (Clarke and Mahul 2011).

6. Moving to Implementation

Moral Hazard Load

Description of step

Two complementary approaches to loading for moral hazard could be considered. First, management could commit to a moral hazard load schedule to account for the perceived threat of moral hazard. The load would be decreasing in the quality and reliability of crop cutting experiments (CCEs). This load would be zero for states with high quality, reliable CCEs and high for states with low quality, unreliable CCEs.

Second, the insurer could add a load to account for the estimated effect of moral hazard on reported claims payable across the state. The degree of systematic misreporting of CCEs within a state could be estimated using a statistical sample of co-observed CCEs, free from moral hazard, and the CCEs reported across the state.

Actuarial justification

Under NAIS, states were partly responsible for settling claims and therefore had financial incentives to protect the NAIS portfolio from moral hazard. However, under modified NAIS, states pay subsidies in the form of upfront premiums and the insurer is responsible for settling all claims, weakening the financial incentive for states to ensure that CCEs are reliable. Indeed, in theory, the upfront premium subsidy structure offers short term incentives for states to encourage the underreporting of yields, leading to large claim payments to citizens at no additional short term cost to the state (Mahul et al. 2011).

Claim payments to farmers are based on the yield estimates from local Crop Cutting Experiments (CCEs) but the insurer cannot efficiently monitor all CCEs conducted within a state and has limited freedom to train, monitor and incentivize the individuals tasked with the conduct of CCEs, since they are typically state government employees. However the insurer can offer incentives to state governments to protect against this moral hazard from the top down, through career incentives, punishments and implementation of robust procedures. One option for the insurer would be to charge a high moral hazard load for those states with unreliable CCEs, along with a commitment that a lower load would be applied if CCE reliability across the state was to improve. Such a structure could spur states to improve the reliability of CCEs, thereby reducing the threat of moral hazard. However, the insurer may find it difficult to judge the reliability of CCEs across a state, and to determine the appropriate financial incentive to offer.

Under the suggested design and ratemaking methodology, any overstatement of yields in good years would not affect the robustness of the pure premium rate, since this depends only on the yields in bad years. This differs from ratemaking under the Normal Theory Method, whereby overstatement of yields in good years would increase the coefficient of variation, thereby increasing the pure premium rate. Overstatement of yields in good years would, however, increase the Threshold Yield, and therefore the state and central government subsidy rates.

Design and ratemaking for without sufficiently granular historical data

A reduction in the size of the insurance unit (IU) to the level of the Gram Panchayat (GP) could introduce technical challenges to actuarial ratemaking if ten years of yield data is not available at the GP level. Both a sound technical framework and expert actuarial judgment will be required to respond to these challenges. This section outlines a framework that could be used as the basis for design and ratemaking with limited data. This framework will require expert interpretation and adjustment on application to actual yield data.

Description of step

It is suggested that when the level of the IU has changed in the last ten years, with limited data at the new IU-level, the calculation procedure outlined in sections 2, 3, and 4 is directly applied to data at the level of the new IU, where any missing new-IU level data for the last ten years is filled in with the corresponding data for the old IU. For example, for groundnut in Gram Panchayat GP1 of Andhra Pradesh, yield data was collected at the mandal level from 2000 to 2006, with GP-level yield data recorded from 2007 to 2009 (see Table 16). If products are to be offered at the level of the GP, the data to be used in de-trending, calculation of the threshold yield and pure premium rate is therefore the concatenation of mandal-level data from 2000 to 2006 and GP-level data from 2007 to 2009.

Table 16. Groundnut in one gram panchayat in Andhra pradesh: data for threshold yield and pure premium rate calculations

District	Historical yield data, before de-trending (kg/ha)									
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
New insurance unit: village cluster (gram panchayat)	-	-	-	-	-	-	-	612	1,243	978
Old insurance unit: subdistrict (mandal)	167	916	1,061	840	1,682	738	1,143	-	-	-
Data to use in threshold yield and pure premium rate calculations	167	916	1,061	840	1,682	738	1,143	612	1,243	978

Having calculated the pure premium rate, as described in the previous section, the commercial premium rate should be calculated as suggested in Section 0, but with an additional multiplicative load at the crop/district-level to allow for the degree of missing data. One simple option would be to use a multiplicative load of the following form:

$$\text{Heterogeneity Multiple} = 1 + (\alpha + \beta Z_{PY}) \times (7 - n)$$

where:

- n denotes the number of years of data at the new insurance unit level, up to a maximum of 7
- Z_{PY} is the Credibility Factor in the probable yields step
- α and β are constants, to be chosen based on expert judgment.

Actuarial justification

The suggested procedure involves two key steps. First, probable yields and pure premium rates are calculated as described above. These calculations are unbiased estimates of the average non-drought year yield and average loss cost under the following assumptions:

1. Yields are homogenous across the district. Specifically, for any possible yield threshold, the 'true' probability that yields will fall below that yield threshold does not vary significantly across the district; and
2. The number of crop cutting experiments conducted per insurance unit has remained constant, even as the size of the insurance unit has changed.

Second, rates within each district will be universally increased to account for the possibility that the above two assumptions do not hold.

Where data is insufficient to enable adequate risk classification, it may be necessary to add a load to rates to protect the mNAIS portfolio from the threat of adverse selection. When there is limited data available at the level of the new insurance unit, it may not be possible to ascertain which insurance units have a higher average yield or are subject to lower yield risk, and so risk classification is only partially possible. Risk classification within the district is not necessary if the district is agronomically homogenous. However, if the district is agronomically heterogenous then the inability to classify risks could expose the mNAIS portfolio to inequity and adverse selection. In such circumstances it may therefore be appropriate to add an additional load to protect the insurer from adverse selection, the risk that insurance purchase is unusually high in those insurance units for which rates have been underestimated due to the lack of data.

A decrease in the number of crop cutting experiments conducted per insurance unit would increase the variation of yield estimates, therefore increasing the expected claim payment to farmers. The minimum number of CCEs conducted for insurance purposes within a given IU varies with the size of the IU. If an IU is a district the minimum number of CCEs conducted is 24; if it is a Taluka / Tehsil / Block the minimum number is 16; if it is a Block / Phirka / Revenue Circle / Hobli, the minimum number is 10; and if it is a Village Panchayat the minimum number is 8 (AICI, 1999). If the number of CCEs conducted per insurance unit has decreased as the size of the insurance unit has decreased then the standard error of the calculated average yields would be expected to increase in the future. Other things being equal this would motivate an increase in premium rates.

With appropriate data and portfolio stability it would be possible to calculate the Moral Hazard Load and Heterogeneity Multiple in a reasonably precise manner. However, in the absence of such data, the load formula must be determined based on expert judgment. The above Heterogeneity Multiple is increasing both in the Credibility Factor, which may be considered to be a proxy for the degree of statistical heterogeneity, and in the number of years of missing data at the new IU level.

7. Conclusion

This paper has proposed a framework for actuarial calculations that has been operationalized in a large developing country. This is one of the few cases worldwide where the ratemaking methodology of such a large agricultural insurance scheme has been completed revisited, moving from the Normal Theory Method to an experience based approach.

The method suggested in this paper has been designed to be appropriate to the political economy constraints of the mNAIS. The rules of the mNAIS were such that the threshold yield, the trigger below which farmers would receive claim payments, would be determined at the level of the product but that the premium would be determined at the crop/district level. Moreover, political economy constraints demanded that the formula for calculating threshold yields be simple. The suggested procedure uses Empirical Bayes credibility weighting of average historical yields for insurance units within the same district, and credibility weighting of premium rates within the same state. Such smoothing has been implemented.

While the methodology outlined in this paper was designed to be implemented in India for area yield indexed insurance, we hope that this paper may prove useful as a reference for other index insurance schemes around the world.

Bibliography

- AICI**, 1999. "National Agricultural Insurance Scheme: Scheme & Guidelines" New Delhi: Agriculture Insurance Company of India
- AICI**, 2010. "A Discussion note for Actuarial Design and Ratemaking" New Delhi: Agriculture Insurance Company of India
- American Academy of Actuaries**, 2008. *Credibility Practice Note*.
- Bühlmann, H.**, 1967. "Experience rating and credibility," ASTIN Bulletin, 4, 199–207.
- Bühlmann, H. and A. Gisler**, 2005. *A course in credibility theory and its applications*, Springer Verlag.
- American Academy of Actuaries**, 1980, *Risk Classification Statement of Principles*, Committee on Risk Classification.
- Clarke, D. and O. Mahul**, 2011. "Catastrophe risk financing and contingent credit: a dynamic analysis," World Bank, mimeo.
- Efron, B. and C. Morris**, 1975. "Data Analysis Using Stein's Estimator and its Generalizations," Journal of the American Statistical Association, 70 (350), 311–319.
- Herzog, T.N.**, 1999. *Introduction to Credibility Theory*, Actex Publications.
- Josephson, G.R., R.B. Lord, and C.W. Mitchell**, 2000. *Actuarial documentation of multiple peril crop insurance ratemaking procedures*, Report prepared for the US Department of Agriculture, Risk Management Agency, by Milliman and Robertson, Inc.
- Mahul, O., N. Verma, and D.J. Clarke**, 2011. "Improving farmers' access to agricultural insurance in India," World Bank, mimeo.
- Skees, J.R., J.R. Black and B.J. Barnett**, 1997. *Designing and rating an area yield crop insurance contract*, American Journal of Agricultural Economics 79:430-438.
- Stohs, S.M. and J.T. LaFrance**, 2004. "A learning rule for inferring local distributions over space and time," in J.P. LeSage and R.K. Pace, eds., *Spatial and Spatiotemporal Econometrics*, (Advances in econometrics, Vol. 18), chapter 11.
- Syroka, J.**, 2007. *Index-Based Weather Insurance for Indian Farmers: Recommendations for Ratemaking*, Washington DC: World Bank.
- World Bank**. 2007. "India National Agricultural Insurance Scheme: Market-based Solutions for Better Risk Sharing" Report No. 39353. Washington DC: World Bank.
- World Bank**. 2011. "Enhancing Crop Insurance in India" Washington DC: World Bank.
- Zhu, Y., B. Goodwin, and S. Ghosh**, 2011. "Time-varying Yield Distributions and the US Crop Insurance Program," presented at Agricultural and Applied Economics Association 2011 Annual Meeting, July 24-26, 2011, Pittsburgh, Pennsylvania.