

Using a Farmer's Beta for Improved Estimation of Actual Production History (APH) Yields

Miguel Carriquiry, Bruce A. Babcock, and Chad E. Hart

Working Paper 05-WP 387

March 2005

**Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011-1070
www.card.iastate.edu**

Miguel Carriquiry is a postdoctoral research associate in the Center for Agricultural and Rural Development (CARD) at Iowa State University; Bruce Babcock is a professor of economics and director of CARD; and Chad Hart is a scientist and U.S. agricultural policy and insurance analyst at CARD.

This paper is available online on the CARD Web site: www.card.iastate.edu. Permission is granted to reproduce this information with appropriate attribution to the authors.

For questions or comments about the contents of this paper, please contact Bruce Babcock, 578F Heady Hall, Iowa State University, Ames, IA 50011-1070; Ph: 515-294-6785; Fax: 515-294-6336; E-mail: babcock@iastate.edu.

Iowa State University does not discriminate on the basis of race, color, age, religion, national origin, sexual orientation, sex, marital status, disability, or status as a U.S. Vietnam Era Veteran. Any persons having inquiries concerning this may contact the Director of Equal Opportunity and Diversity, 1350 Beardshear Hall, 515-294-7612.

Abstract

The effect of sampling error in estimation of farmers' mean yields for crop insurance purposes is explored using farm-level corn yield data in Iowa from 1990 to 2000 and Monte Carlo simulations. We find that sampling error combined with nonlinearities in the insurance indemnity function will result in empirically estimated crop insurance rates that exceed actuarially fair values by between 2 and 16 percent, depending on the coverage level and the number of observations used to estimate mean yields. Accounting for the adverse selection caused by sampling error results in crop insurance rates that will exceed fair values by between 42 and 127 percent. We propose a new estimator for mean yields based on a common decomposition of farm yields into systemic and idiosyncratic components. The proposed estimator reduces sampling variance by approximately 45 percent relative to the current estimator.

Keywords: actual production history (APH), crop insurance, mean yields estimation, sampling error.

USING A FARMER'S BETA FOR IMPROVED ESTIMATION OF ACTUAL PRODUCTION HISTORY (APH) YIELDS

The basis for the amount of crop insurance that a U.S. farmer can purchase and the cost of the insurance is a producer's Actual Production History (APH) yield. A producer's APH yield is used to estimate her expected yield. Producers with a high APH yield can buy more insurance while paying a lower insurance rate. (An insurance rate equals the total premium for the insurance policy divided by the total liability for the policy.) This type of relationship between the insurance rate and expected yield is consistent with yield variability—measured by the coefficient of variation—that decreases as mean yield increases.

The exact current rules for calculating APH yield are complicated.¹ In their simplest form, APH yields equal the simple average of past yields. A minimum of 4 years and a maximum of 10 years of yield data go into the calculation of the average yield. A common complaint about the rules is that they unduly penalize farmers in a region that experiences abnormally poor growing weather. For example, when drought hits, most farmers' yields in the region will fall, thus decreasing their APH yields. The weight given to the drought yield is equal to the weight given to every other yield in calculating APH yields regardless of the likelihood of a future drought. Similarly, a bumper crop will increase a farmer's APH yields and will be weighed equally with other yield observations, even if the conditions that led to the bumper crop are extremely unlikely to repeat themselves. This suggests that farmers who reside in regions that have had good (poor) recent growing conditions will tend to have APH yields that exceed (are below) their true expected yields.

APH yields can also provide a poor estimate of a farmer's mean yield because of individual farmer luck. Individual yield luck can take many forms. A farmer can suffer significant hail damage while the neighbor across the street suffers none. A farmer can choose to plant Bt corn in a year when a significant corn borer infestation occurs. A

farmer can obtain significant rainfall at a critical stage of crop growth while a neighboring farmer misses out. A farmer may miss the only suitable planting period because of a broken planter.

Noisy estimates of farmers' expected yields imply that farmers with positive average noise (APH yield greater than expected yield) will be undercharged for their insurance coverage whereas farmers with negative average noise will be overcharged. If both groups of farmers participate in the program, then one can hope that the overcharge will tend to compensate for the undercharge so that the financial soundness of the program is unaffected. However, for most yield distributions, actuarially fair insurance rates are convex with respect to coverage level (Babcock, Hart, and Hayes 2004). Thus, the average rate charged will tend to be less than the average of the actuarially fair rates, resulting in a net undercharge. A potentially larger problem is the adverse selection that will result when the producers who are overcharged for insurance drop out of the program and those who are undercharged increase their participation (Just, Calvin, and Quiggin 1999).

Of course, all estimators of expected yield will result in this problem, because a farmer's true expected yield can never be observed. However, an estimator with a lower variance than the current estimator should lead to a reduction in adverse selection and more accurate crop insurance rates.

The objective of this paper is to estimate the magnitude of the problem caused by noisy estimation of farmers' mean yields and to propose a new estimator for use in the crop insurance program. The proposed estimator is based on a commonly used decomposition of farm yield into systemic and idiosyncratic components. The importance of systemic yields depends on a farmer's "beta," which measures the sensitivity of farm yields to changes in area yields. The conditions under which the proposed estimator will reduce variance are determined and a combination of bootstrapping and Monte Carlo simulations is used to show the likely magnitude of the reduction for corn producers in Iowa. The potential economic impact on the Iowa corn insurance program of moving to the new estimator is examined. We find that sizeable improvements on the actuarial performance of the program can be achieved by simply switching estimators.

Conceptual Model

This section adapts the conceptual model developed by Just, Calvin, and Quiggin (hereafter abbreviated JCQ) to the problem at hand. As previously stated, the amount of insurance producers can buy and the premium they have to pay for it is a function of their APH yield. However, the APH yield is just an estimate of farmers' expected mean yield. As an estimator, the APH yield of any given producer is itself a random variable. APH insurance (the government's main form of yield insurance) makes an indemnity payment if yields are below a yield guarantee that is the product of a producer-selected coverage level and the producer's APH yield. The offered coverage levels range from 50 to 85 percent in 5 percent increments. Producers can also choose a price guarantee level (p_g). The insurance premium ($\omega_{g\alpha}$) is contingent upon these choices and the yield distribution $G(y)$ on which the program is based, which has a mean equal to the farmer's APH yield (μ_a). The yield distribution captures the yield variability assigned by the Risk Management Agency (RMA) to a farm with an observed APH yield of μ_a . Adapting JCQ's model to include insurance premium subsidies explicitly (Coble and Knight 2001), the revenue under insurance for a farmer for a given output price and yield realization is given by

$$\phi(y) = \begin{cases} py - (1 - s_\alpha)\omega_{g\alpha} & \text{if } y \geq \alpha\mu_a \\ py + p_g(\alpha\mu_a - y) - (1 - s_\alpha)\omega_{g\alpha} & \text{if } y < \alpha\mu_a \end{cases}$$

where α is the selected coverage level, y denotes yield, and s_α is the subsidy rate, which is contingent on the coverage level. For any realization of the APH yield estimator and the actual farm level yield density $f(y)$, expected returns to crop insurance are

$$E(\phi(y) - py) = I(\mu_a) - C(\mu_a) = \int_0^{\alpha\mu_a} p_g(\alpha\mu_a - y)f(y)dy - (1 - s_\alpha)\omega_{g\alpha}, \quad (1)$$

where $I(\mu_a)$ denotes the indemnity expected by a farmer with an APH yield μ_a , and

$C(\mu_a)$ represents that farmer's premiums. Note that $\frac{\partial^2 I}{\partial \mu_a^2} = p_g \alpha^2 f(\alpha\mu_a) \geq 0$. For any

yield distribution, expected indemnities are a convex function of the APH yield. Hence, Jensen's inequality indicates that in an insurance program in which premiums are actuarially fair (in expectations), farmers in aggregate expect to collect indemnities in excess of the premiums in the long run, which works against the insurance program's actuarial soundness. This happens even when farmers choose to participate independently of their realized APH yield.

Adverse selection aggravates this problem. To look at the role that APH yield estimation plays in adverse selection we first need to estimate how expected net returns from the program vary with the realization of a farmer's APH yield. We examine the current rate structure for all Iowa counties. Premiums for corn are first decreasing and then increasing in a farmer's APH yield. The APH yield as a proportion of the county expected yield at which the minimum premium is obtained (and its slope thereafter) varies by county. There is only one county (Webster) in which the minimum premium is obtained for an APH yield below the county mean yield (at 97 percent of the county expected yield). At the other extreme, the minimum premium would be paid by a farmer who has an APH yield that doubles the county mean yield (199.5 percent of the expected yield) in Clarke County. The slope of the premium schedule after the minimum is attained is quite modest for all counties. The average slope of the 75 percent coverage level premium schedule when the APH yield increases above the county mean is $0.0121p_g$ \$/acre, with a minimum of $0.0001p_g$ \$/acre in Clarke County and a maximum of $0.0158p_g$ \$/acre in Cherokee County. In short, given that expected indemnities are increasing in APH yields

$$\left(\frac{\partial I}{\partial \mu_a} = p_g \alpha F(\alpha \mu_a) > 0 \right),$$

expected returns to insurance are likely to increase with the APH yield assigned.²

Figure 1 illustrates this point for a representative farmer who has a yield distribution with mean μ . Details of the methods used to generate the expected indemnities and premiums portrayed in the figures are presented next.

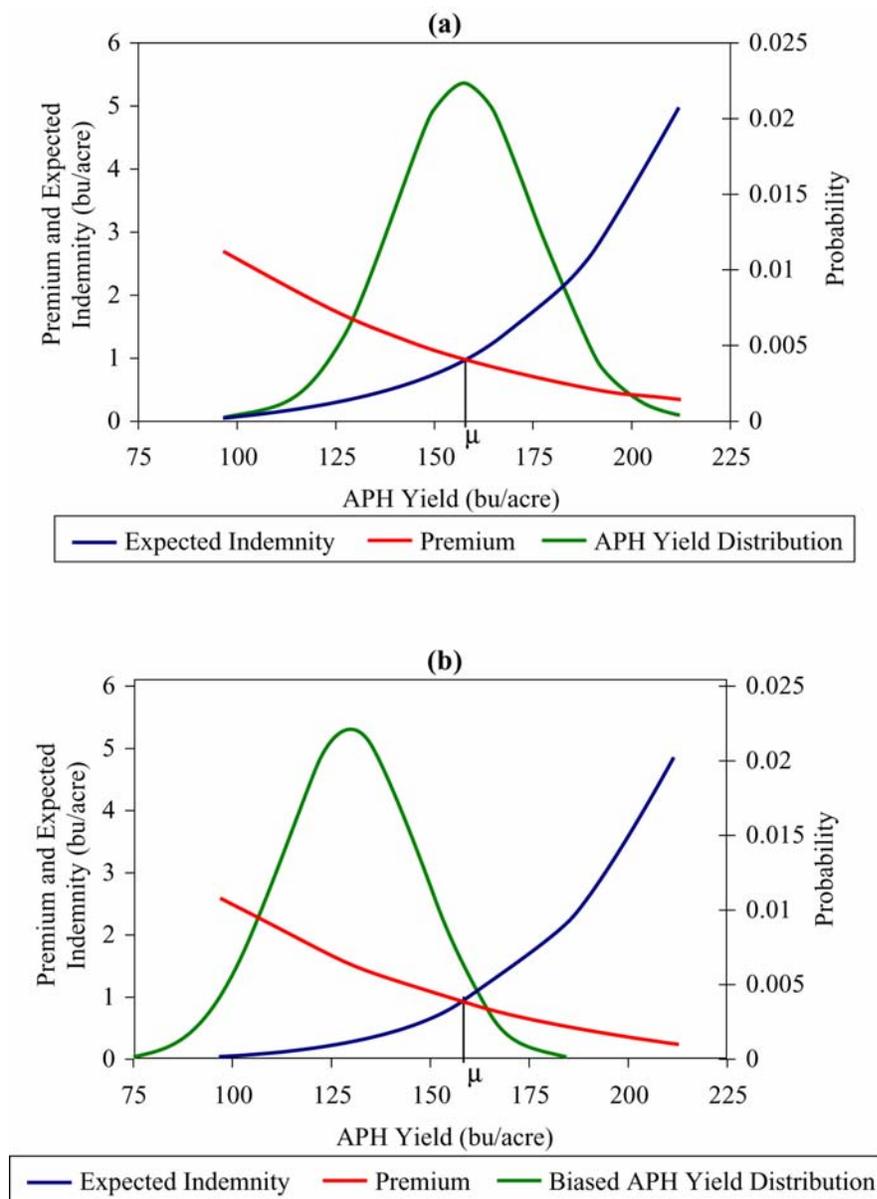


FIGURE 1. Expected indemnities and premiums for a representative farmer with mean yield μ and a 65 percent coverage level

When the realization of the APH rule is below (above) the farmer's true mean yield, the farmer will find insurance overpriced (underpriced) and will expect negative (positive) returns from participation. The farmer will expect a zero return on crop insurance participation only when her APH coincides with her true mean (assuming that the premium is actuarially fair on average). This is exactly the asymmetric information effect quantified by JCQ, where farmers with APH yields above their true mean yield will be

offered “fair premiums” (from the RMA’s standpoint) that are lower than their expected indemnities. Figure 1a shows a possible distribution of APH yields, in the case in which the APH rule is an unbiased estimator of the mean yield. With an unbiased rule, it can easily be seen that taking expectations over all farmers’ APH yields will result in expected indemnities exceeding expected premiums, even with no adverse selection.

Because time trends in yields are not accounted for in the RMA’s APH yield determination, as JCQ and other authors (e.g., Skees and Reed 1986) correctly point out, a farmer’s APH yield will, on average, be lower than the expected mean yield, and the region of negative returns to insurance is expanded or has a higher probability of occurrence (see Figure 1b). JCQ found that on average, both insuring and non-insuring farmers expected negative returns to unsubsidized crop insurance.

Adverse selection occurs when program participation rates are higher for farmers who expect a positive return on insurance. A positive return occurs when the assigned APH yield is higher than farmers’ true mean yield (see Figure 1). The magnitude of the expected return to crop insurance depends on the difference between the APH yield and the subjective mean yield. When a farmer’s APH yield is larger than the true mean yield, the effective coverage level provided is higher than the coverage level contracted. A farmer with mean yield μ who chooses coverage level α and is assigned an APH yield of μ_a receives an effective coverage level of $\alpha(\mu_a/\mu)$.³ In this situation, the RMA ends up with an adversely selected risk pool and incurs excess losses. Increases in rates to mitigate losses will only result in a more adversely selected risk pool, since only producers whose APH yield exceeds their mean yield by a larger amount will stay in the program. As shown in Figure 1b, a biased APH yield reduces the proportion of farmers who find that insurance increases expected returns. This lowers participation in the crop insurance program.

One way to improve the actuarial soundness of the program and mitigate the incidence of adverse selection is to subsidize premiums. The effects of this policy are illustrated in Figure 2.

Subsidizing the premium increases the probability of participation because it makes participation profitable for a portion of farmers with APH yields that are below their true mean yields. Our representative farmer will now participate as long as her realized APH

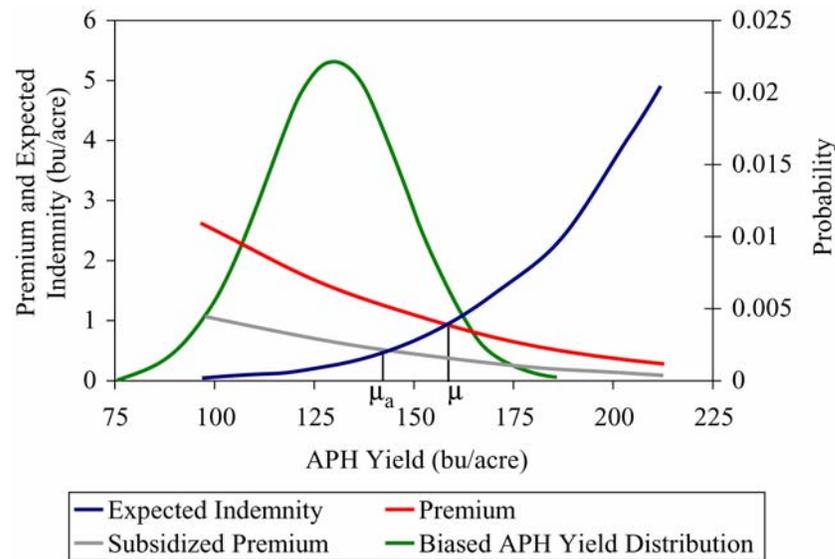


FIGURE 2. Expected indemnities, premiums, and subsidized premiums for a representative farmer

yield is above μ_a in Figure 2. Because $\mu_a < \mu$, the premium subsidy reduces adverse selection. The subsidy effect improves the prospects of the crop insurance program for all producers but may not be enough to overcome the asymmetric information effect for some producers.

This result agrees with JCQ's analysis. For the 65 percent coverage level, JCQ did not find differences between the mean subsidy effects for insuring and non-insuring producers of corn and soybeans.⁴ On average, insuring producers had a less negative asymmetric information effect. In our framework, that would correspond with a representative farmer who has an APH between μ_a and μ .

The current analysis makes clear that subsidies only partially mitigate the adverse selection problem. From the perspective of expected revenue incentives, the risk pool will be adversely selected whenever there is noise in the estimation of a farmer's expected yield and the insurance is not fully subsidized.

These problems are likely to be compounded at higher coverage levels, since both the slope and the curvature of the expected indemnity–APH yields schedule increase with the coverage level.⁵ Figure 3 depicts the expected indemnities and subsidized premium schedules for our representative farmer for two coverage levels.

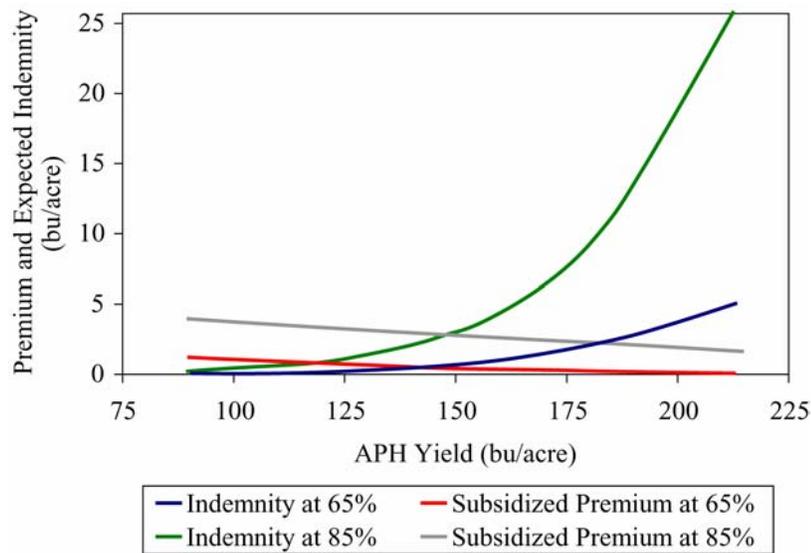


FIGURE 3. Expected indemnities and subsidized premiums for a representative farmer

Figure 3 indicates again that the producer will not buy the offered insurance (from an expected returns standpoint) if the realized APH yield is low enough. Since the break-even APH yield for the 65 percent coverage level is lower than for the higher coverage levels, the producer will choose first to participate at the 65 percent coverage level as APH yields increase. At some point after the realized APH yield surpasses the next break-even point, the producer will choose to switch to the higher coverage level. The exact APH yield that will trigger the switch will depend on the slopes of the expected indemnities and premiums under both coverage levels. Figure 3 indicates that the insurer can expect that the losses from the adversely selected pool will increase with both the APH yield and the coverage level. That is, Figure 3 suggests that losses in excess of premiums will be greater at the 85 percent coverage level than at the 65 or 75 percent coverage levels. This finding is consistent with the empirical findings of Coble et al. (2002) that resulted in premium surcharges for 75 and 85 percent coverage levels.

In JCQ, corn farmers who insured at the 75 percent coverage level (the highest available at that time) expected larger returns to insurance than did farmers who insured at the 65 percent coverage level. Our framework explains that result. If the insurance parameter, namely the APH yield, is high enough, producers can increase returns by switching to higher levels of coverage.

The foregoing discussion reveals that in order to improve the actuarial soundness of the crop insurance program, it is crucial for the RMA to obtain better estimates of expected farm-level yields. Before presenting an alternative estimator and the conditions upon which the new estimator performs better than the APH rule, we explore the economic magnitude of the problem for the case of Iowa corn insurance.

Impact of Uncertainty in the Estimation of Expected Yields

A Model of Yields

Farm yields can be decomposed into systemic and idiosyncratic components (Miranda 1991; Mahul 1999; Vercaemmen 2000):

$$y_{it} = \mu_c + \delta_i + \beta_i (y_{ct} - \mu_c) + \varepsilon_{it} = (1 - \beta_i) \mu_c + \delta_i + \beta_i y_{ct} + \varepsilon_{it}, \quad (2)$$

where μ_c is the area mean yield, δ_i is the difference between the area mean yield and farm mean yield, y_{ct} is the area yield in year t , and ε_{it} is the farm-yield deviation in year t . Also,

$$E\varepsilon_{it} = 0 \quad \text{Var}(\varepsilon_{it}) = \sigma_{\varepsilon_i}^2 \quad \text{Cov}(\varepsilon_{it}, y_{ct}) = 0 \quad \text{Var}(y_{ct}) = \sigma_c^2$$

$$Ey_{it} = \mu_i = \mu_c + \delta_i \quad \text{Var}(y_{it}) = \beta_i^2 \sigma_c^2 + \sigma_{\varepsilon_i}^2 \quad Ey_{ct} = \mu_c.$$

This decomposition is appealing because of its simplicity and its similarity with the capital asset pricing model of finance, which provides clear interpretations for its main parameter β_i . In this model, β_i measures the sensitivity of farm yields to the systemic factors affecting area yield. Recently, Ramaswami and Roe (2004) derived this linear form from the aggregation of microproduction functions. They showed that linearity only attains if the systemic and non-systemic individual risk components of the production functions are additive. Additionally, to estimate the beta parameter using ordinary least squares (OLS) requires that the decomposition should consist of two uncorrelated risks, which requires the aggregation to be large (eliminating all individual risk from the area aggregate).

A variant of this model was used to rate Income Protection (Atwood, Baquet, and Watts 1996). This rating method uses equation (2) with a slight modification in order to

obtain estimates of the residuals. These residuals are then used to construct simulated farm yields.⁶ While recognizing that premiums are affected by the precision involved in the estimation of a farmer's expected yields, Atwood, Baquet, and Watts do not elaborate further on this point and do not investigate the implications of using this widely employed model for obtaining better estimators than the simple averages of farm-level yield history. Other authors (Miller, Kahl, and Rathwell 2000a,b), rating insurance products for peaches in Georgia and South Carolina, have restricted beta to be one. Miranda (1991) showed that the area-weighted average of the betas of all farmers in the relevant region must equal one. However, there is no reason why beta should equal one for all farmers, and empirical estimates seem to indicate that farmers' beta values in an area are symmetrically distributed around one.

How Severe Is the Problem?

We use the model of farm-level yields in equation (2) to quantify the magnitude of the problem resulting from the current rating practices for corn insurance for the nine crop reporting districts (CRD) in Iowa. Hence, the relevant area for a given farmer is her CRD. We obtained farm-level yield data for farmers in each CRD of Iowa from 1990 to 2000. The farm-level yield data were made available to us by the RMA. For each CRD, we obtained 49 years (1956-2004) of corn yield data from the National Agricultural Statistics Service. All yields, whether at the farm or CRD level, are multiplicatively detrended to a base year (2004).

Equation (2) was estimated using OLS. We assume that the variance of the idiosyncratic shock is equal for all farmers within a CRD. The regression residuals were used to estimate this variance as a (degrees of freedom) weighted average of the error variance estimated for each farmer. That is, the variance of the idiosyncratic shock in district k is

$$\hat{\sigma}_{\varepsilon,k}^2 = \frac{\sum_{j=1}^J \sum_{t=1}^{T_j} \left(y_{it,k} - (1 - \hat{\beta}_i) \mu_c - \hat{\delta}_i - \hat{\beta}_i y_{ct,k} \right)^2}{\sum_{j=1}^J (T_j - 2)},$$

where T_j denotes the number of yield observations available for farmer j , and J is the total number of farmers in the CRD. As in Atwood, Baquet, and Watts, only data from farmers with at least eight yield observations were used in the analysis. Note that this

variance would increase if the model were estimated using field-level rather than farm-level data.

Expected indemnities and pure premiums were estimated based on yield simulations (based on equation (2)) for the average farmer in the pool. The average farmer in the pool has $\beta = 1$ (see Miranda 1991) and $\delta_i = 0$. For each CRD, 10,000 yield samples of size 10 were constructed by bootstrapping from the set of detrended yields for that CRD. Sets of farm-level yields were obtained by adding an error term independently drawn from a normal distribution with zero mean and variance $\hat{\sigma}_{\epsilon,k}^2$ to each area yield sampled. We constructed APH yields by averaging over 4 through 10 observations on each sample, obtaining 7 different APH yield distributions by CRD. By construction, APH distributions based on fewer observations are mean-preserving spreads of the distributions that average more years of data; hence, we can compare the benefits of obtaining more precise estimators of a farmer's expected yield.

For each distribution of APH yields in each CRD, a lattice is constructed by obtaining the empirical quantiles in 1 percent increments. Each point in the lattice is a possible realization of the farmer's APH yield and is what the insurer (RMA) observes and uses in the inference about the farmer's levels of risk. From the insurer's perspective, a farmer's expected yield deviates from the area yield by $\delta_a = \mu_a - \mu_c$. The farmer is assumed to know her yield distribution.⁷ We construct 200,000 farm-level yields for each sampled point in each APH yield distribution from both the farmer's perspective, through equation (2) by bootstrapping area yields and adding a draw from the error distribution, and the insurer's perspective, by replacing $\delta_i = 0$ with $\delta_a = \mu_a - \mu_c$ for each realization of the APH yield. Hence, we can estimate the expected indemnity for the farmer and the actuarially fair premium from the insurer's perspective for each realized μ_a of each distribution. The exercise is repeated for three coverage levels, namely 65 percent, 75 percent, and 85 percent, and the nine CRDs in Iowa. What we are trying to accomplish with these simulations is the measurement of the impacts of random APH yields on the ratio of indemnities that will be paid out under random APH yields to the indemnities that one expects to be paid out if the farmer's true mean yield were known. Results, for a price guarantee p_g normalized to one, are presented in Tables 1-5.

TABLE 1. Ratio of the average value of expected indemnities across all APH yields to expected indemnities calculated at the true mean yield by Iowa crop reporting district

Coverage	APH Distribution Based On					
	4 Observations			10 Observations		
	65%	75%	85%	65%	75%	85%
CRD	Participate independently of realized APH yield					
10	1.198	1.210	1.205	1.079	1.087	1.088
20	1.111	1.203	1.221	1.036	1.073	1.083
30	1.135	1.177	1.189	1.052	1.074	1.077
40	1.175	1.173	1.158	1.074	1.080	1.070
50	1.119	1.140	1.171	1.051	1.063	1.075
60	1.134	1.165	1.193	1.052	1.057	1.073
70	1.163	1.169	1.166	1.067	1.068	1.071
80	1.126	1.151	1.158	1.050	1.064	1.067
90	1.114	1.145	1.163	1.046	1.059	1.070
	Participate only when expected indemnity > premium					
10	1.724	1.737	1.729	1.399	1.406	1.409
20	1.491	1.707	1.793	1.265	1.371	1.432
30	1.568	1.665	1.709	1.315	1.366	1.394
40	1.733	1.713	1.680	1.415	1.411	1.391
50	1.559	1.576	1.647	1.329	1.333	1.368
60	1.512	1.603	1.680	1.279	1.320	1.368
70	1.670	1.684	1.680	1.381	1.386	1.390
80	1.559	1.620	1.644	1.310	1.346	1.363
90	1.499	1.573	1.626	1.289	1.327	1.363

Table 1 presents the ratio of the expected value of indemnities under random APH yields to the expected indemnities that would result if the true mean yield ($\delta_i = 0$) was known. The top half of Table 1 assumes that all farmers participate in the program regardless of the level of their APH yield. The bottom half of Table 1 assumes that farmers only participate if their APH yield is greater than their mean yield. The average value of the expected indemnities are taken with respect to the APH yield distributions that would result from 4 or 10 farm-level yield observations in a district. As a consequence of the convexity of the expected indemnities curve with respect to assigned APH yields (analyzed in the previous section), all ratios in the table exceed one, with the magnitude contingent on the underlying yield distribution. This has a discouraging implication. It indicates that the farmer expects to obtain positive returns (hence the insurer expects losses) from insurance that is fairly priced when the farmer's APH yield

equals her true mean yield. The results also show that as the level of uncertainty in the estimation of the APH yield decreases (as more yield observations are used in its calculation), so does the ratio, indicating an expected indemnity closer to that when the true mean yield is known. For example, a representative farmer in CRD 50 would expect to collect indemnities 14 percent higher (if she participates regardless of her APH yield) than the fair premium calculated at the true mean yield for a coverage level of 75 percent and an APH yield based on four observations. If farmers have 10 observations, they then expect indemnities to exceed fair premiums by 6.3 percent.

This problem is aggravated if adverse selection is present (bottom half of the table). If the farmer decides to buy insurance (again at 75 percent coverage) only when she knows that the realized APH yield exceeds her expected yield, she will expect to receive 57.6 percent more in indemnities for an APH yield rule based on four observations than what she would have expected if her true mean yield was known by the insurer. The expected excess payout drops to 33 percent when 10 yield observations are used to construct APH yields.

Table 1 provides some insight into why crop insurance rate making is so difficult. Rates are typically based on loss experience so that all farmers with an APH yield equal to, say, 150 bu/ac are grouped together, and their historical losses are averaged to come up with a premium rate. The results in Table 1 suggest that the resulting premiums will be significantly greater than the expected indemnities that a farmer who has a true mean yield of 150 bu/ac will expect to receive from the program. This then discourages farmers from joining the program. Of course, the rate-making problem is magnified under adverse selection because it is more likely that those farmers who have true mean yields less than 150 bu/ac will be in the program.

Table 1 does not account for changes in premiums that will be charged as the APH yield changes. That is, we assume that all farmers are charged the actuarially fair premium for the true mean yield. Of course, the true mean yield is unobservable, so it is useful to calculate the premium for each APH yield draw that is viewed as being actuarially fair from the perspective of the insurer, who believes that the farmer's true mean yield equals the farmer's APH yield. Table 2 reports the ratio of the average of expected indemnities to the expected premium charged, with the average taken with respect to the

TABLE 2. Ratio of the average value of expected indemnities to expected actuarially fair premiums by Iowa crop reporting district

Coverage	APH Distribution Based On					
	4 Observations			10 Observations		
	65%	75%	85%	65%	75%	85%
CRD	Participate independently of realized APH yield					
10	1.133	1.170	1.184	1.051	1.063	1.071
20	1.075	1.173	1.215	1.023	1.063	1.083
30	1.085	1.146	1.177	1.035	1.060	1.073
40	1.109	1.140	1.150	1.058	1.067	1.070
50	1.078	1.113	1.160	1.036	1.049	1.070
60	1.089	1.150	1.185	1.026	1.051	1.066
70	1.109	1.149	1.162	1.053	1.066	1.072
80	1.102	1.137	1.154	1.039	1.053	1.062
90	1.080	1.126	1.157	1.030	1.049	1.065
	Participate only when expected indemnity > premium					
10	2.219	2.010	1.862	1.633	1.534	1.468
20	1.785	1.968	1.973	1.419	1.505	1.527
30	1.932	1.910	1.852	1.502	1.489	1.469
40	2.276	2.006	1.832	1.696	1.562	1.474
50	1.956	1.789	1.770	1.542	1.444	1.434
60	1.788	1.833	1.817	1.417	1.436	1.435
70	2.151	1.978	1.835	1.631	1.545	1.475
80	1.919	1.848	1.771	1.493	1.459	1.425
90	1.812	1.784	1.753	1.459	1.440	1.429

distribution of APH yields. Whether this ratio is greater or less than one depends crucially on the relative curvature of the expected indemnities and premium schedules with respect to APH yields.

The entries in the upper half of Table 2 are greater than one, indicating that the expected indemnities schedule is more convex than the premium schedule. Comparing the upper halves of Tables 1 and 2, it is easy to see that the ratios are larger in Table 1, which indicates that the premium schedule is also convex. This is in agreement with the curvature of premium schedules (with respect to APH yields) for corn currently in use in all Iowa counties. The small magnitude of the difference (which averages 3.7 percent and 2.2 percent for 4 and 10 observations respectively) indicates that the premium schedule with respect to the APH yield departs slightly from linearity.⁸

Comparing the results in the bottom halves of Tables 1 and 2 confirms that the expected fair premium for APH yields that are greater than the true mean yield are less than the premium at the true mean yield. Hence, if our average farmer in CRD 50 buys insurance only when expected indemnities exceed premiums, then expected returns are 78.9 percent and 44.4 percent greater than expected premiums for APH distributions based on 4 and 10 observations, respectively, and a 75 percent coverage level.

Table 3 presents the expected returns to unsubsidized crop insurance (i.e., equation (1) with $s_\alpha = 0$ for all coverage levels) for a price guarantee normalized to one. The table includes both the situation in which farmers buy insurance regardless of the realized APH yields and in which farmers participate only for APH yields that result in expected indemnities that exceed the actuarially fair premium (from the insurer's perspective) based on that APH yield. As expected after the analysis of Table 2, all entries in Table 3 are positive, indicating again that, on average, farmers will obtain positive returns to insurance for all contracts considered, even if the decision to purchase does not depend on the realized APH yield.⁹ This differs from the results of JCQ, who estimated negative expected returns to unsubsidized crop insurance. But this difference in results should be expected because JCQ, in accordance with the rules used to compute APH yields, did not detrend the farm-level yield series.

We observe again that the expected returns to crop insurance decrease when the precision with which expected yields are estimated increases. This can be seen by comparing expected returns by district and coverage level for the estimators based on 4 and 10 observations. In the same line, the correlations (not shown) between the variance of yields of the average farmer (in each CRD) and expected returns under the "buy always" rule range from 0.93 (at 65 percent coverage and an APH yield based on 10 observations) to 0.98 (65 and 75 percent coverage levels and an APH yield based on 4 observations). The correlations are even higher for the other participation rule. This is expected since larger farm level-variances will result *ceteris paribus* in a less precise estimator under the current APH rule. The current discussion highlights the importance, at least from an actuarial sustainability standpoint, of obtaining better estimators of the farmer's expected yield.

TABLE 3. Expected returns (bu/ac) from unsubsidized corn insurance to the average farmer by Iowa crop reporting district

Coverage		APH Distribution Based On					
		4 Observations			10 Observations		
		65%	75%	85%	65%	75%	85%
CRD	Yield Variance	Participate independently of realized APH yield					
10	907.9	0.111	0.338	0.786	0.041	0.124	0.302
20	783.8	0.051	0.250	0.711	0.015	0.089	0.272
30	939.8	0.090	0.321	0.788	0.037	0.129	0.324
40	964.0	0.115	0.336	0.734	0.058	0.157	0.339
50	898.8	0.076	0.239	0.674	0.035	0.103	0.295
60	1044.9	0.123	0.391	0.908	0.036	0.132	0.325
70	1035.3	0.149	0.415	0.864	0.069	0.182	0.382
80	1550.7	0.339	0.727	1.274	0.126	0.283	0.510
90	1236.2	0.174	0.473	0.998	0.064	0.183	0.409
		Participate only when expected indemnity > premium					
10		0.748	1.674	3.364	0.428	0.939	1.886
20		0.434	1.180	2.910	0.247	0.647	1.626
30		0.772	1.698	3.465	0.449	0.960	1.960
40		0.967	2.001	3.699	0.578	1.181	2.168
50		0.720	1.437	2.988	0.441	0.847	1.727
60		0.891	1.872	3.684	0.503	1.029	2.023
70		1.158	2.284	4.070	0.692	1.341	2.384
80		2.421	3.896	5.907	1.402	2.219	3.354
90		1.421	2.552	4.406	0.858	1.499	2.579

It is also clear from Table 3 that expected returns can be increased in the presence of uncertainty in the estimate of APH yields by the average farmer in each district by simply increasing the coverage level. This confirms the analysis of the previous section.

Table 4 presents the expected returns to insurance (equation (1)) for the average farmer for the case in which premiums are subsidized at the current program rates (59, 55, and 38 percent for coverage levels 65, 75, and 85 percent, respectively). Obviously, each entry in Table 4 is larger than its analogous entry in Table 3, and the patterns with regard to yield variability are maintained. Our results are consistent with those reported by JCQ. These authors reported expected returns of \$0.95/acre and \$3.93/acre for corn farmers choosing 65 and 75 percent coverage levels, respectively, for a price guarantee of \$2.00/bu. The averages of the expected returns for the 75 percent coverage level across districts when farmers will only purchase if they expect positive returns are 2.32 bu/acre and 1.76 bu/acre for the APH rules based on 4 and 10 observations, respectively. Given

TABLE 4. Expected returns from subsidized corn insurance to the average farmer by Iowa crop reporting district

Coverage	APH Distribution Based On					
	4 Observations			10 Observations		
	65%	75%	85%	65%	75%	85%
CRD	Participate independently of realized APH yield					
10	0.603	1.428	2.411	0.520	1.202	1.925
20	0.455	1.043	1.968	0.411	0.870	1.523
30	0.721	1.531	2.479	0.649	1.322	2.004
40	0.738	1.654	2.590	0.655	1.454	2.182
50	0.655	1.404	2.274	0.600	1.255	1.888
60	0.945	1.821	2.775	0.844	1.552	2.190
70	0.950	1.951	2.895	0.845	1.694	2.403
80	2.292	3.646	4.419	2.061	3.195	3.661
90	1.461	2.539	3.407	1.331	2.234	2.818
	Participate only when expected indemnity > subsidized premium					
10	0.789	1.816	3.579	0.577	1.317	2.390
20	0.519	1.247	2.917	0.428	0.927	1.952
30	0.884	1.888	3.681	0.692	1.411	2.484
40	0.989	2.176	4.013	0.738	1.612	2.714
50	0.820	1.688	3.246	0.651	1.340	2.265
60	1.089	2.149	3.964	0.878	1.634	2.673
70	1.228	2.492	4.405	0.928	1.857	2.994
80	2.779	4.470	6.447	2.171	3.402	4.405
90	1.731	3.036	4.881	1.404	2.357	3.387

that the price is normalized to one in our analysis, the bu/acre figures also represent \$/acre figures. The same statistics are 1.89 bu/acre and 1.64 bu/acre if farmers buy the insurance regardless of the realized APH yield.

Table 5 presents an alternative way of visualizing the importance of obtaining more precise estimators of farm-level expected yields. Table 5 reports the elasticity of expected returns to corn insurance in Iowa with respect to the variance with which the expected farm-level mean yield is estimated for the average farmer. The table provides information about the relative changes in expected returns and hence the impacts on the sustainability or cost reductions of the program that can be achieved through better estimators of a farmer's expected yield. The impacts are highest when the farmer's participation decision is independent of the APH yield assigned and the insurance is unsubsidized. In that situation, a 1 percent reduction in the variance of the estimator results in about the same

TABLE 5. Elasticities of expected returns to insurance with respect to the variance of the estimator of the expected yield by Iowa crop reporting district

Coverage	Participate Independently of APH			Participate Only If E(Indemnity)>Premium		
	65%	75%	85%	65%	75%	85%
CRD	Unsubsidized insurance					
10	1.07	1.08	1.04	0.63	0.66	0.66
20	1.27	1.10	1.04	0.64	0.68	0.66
30	0.99	0.99	0.97	0.62	0.65	0.65
40	0.77	0.85	0.86	0.59	0.60	0.61
50	0.88	0.93	0.91	0.56	0.60	0.62
60	1.28	1.16	1.10	0.65	0.68	0.68
70	0.85	0.91	0.90	0.59	0.61	0.61
80	1.06	1.03	1.00	0.62	0.64	0.64
90	1.09	1.03	0.98	0.58	0.61	0.61
Average	1.03	1.01	0.98	0.61	0.64	0.64
	Subsidized insurance					
10	0.17	0.20	0.26	0.36	0.37	0.46
20	0.12	0.21	0.30	0.22	0.34	0.46
30	0.12	0.17	0.25	0.28	0.34	0.45
40	0.14	0.15	0.20	0.34	0.35	0.45
50	0.10	0.13	0.22	0.27	0.27	0.42
60	0.13	0.19	0.27	0.25	0.32	0.45
70	0.14	0.16	0.22	0.32	0.34	0.45
80	0.12	0.15	0.22	0.29	0.32	0.44
90	0.11	0.15	0.22	0.24	0.29	0.42
Average	0.13	0.17	0.24	0.29	0.33	0.45

Note: Elasticities are calculated through the midpoint formula.

percent reduction in expected returns to insurance. The other extreme is the situation in which participation is still independent of the APH yield but the crop insurance is subsidized. In this case, a 1 percent reduction in the variance of the estimator decreases expected returns to insurance from 0.1 to 0.3 percent. However, the absolute reduction is similar for both cases. When a farmer's participation is contingent on the APH yield realized, the same elasticities are between 0.56 and 0.68 for the unsubsidized insurance and between 0.22 and 0.46 if insurance is subsidized.

Increasing the number of observations that enter the calculation of APH yields from 4 to 10 leads to a 60 percent reduction in the variance with which the expected yields are estimated¹⁰ and therefore to a reduction in expected returns (losses) for the average

farmer (insurer) that ranges from 8.4 to 69.6 percent depending on the CRD, coverage level, and participation strategy followed. If one believes that farmers will only purchase insurance when expected returns are non-negative, the reduction averages 42 and 26 percent for the unsubsidized and subsidized insurance, respectively.

An Alternative Estimator

Equation (2) gives a justification for current APH rules. In any given year, a farmer's APH yield is simply the average of yields in the preceding T years:

$$\bar{y}_{it} = T^{-1} \sum_{k=t-T}^{t-1} (\mu_c + \delta_i + \beta_i (y_{ck} - \mu_c) + \varepsilon_{ik}) = \mu_c + \delta_i + \beta_i (\bar{y}_{ct} - \mu_c) + \bar{\varepsilon}_{it}.^{11}$$

If farm luck ($\bar{\varepsilon}_{it}$) and area luck ($\bar{y}_{ct} - \mu_c$) are both zero, then the average of past yields exactly equals the mean farm yield. Justification for assuming that both sources of luck are zero is that the expected value of luck is zero, and the standard deviation of luck decreases with the square root of the number of observations.

However, the probability that luck is exactly zero is zero. Furthermore, the maximum number of observations that can go into calculating APH yields is 10, so the standard errors of area and farm-yield deviations are reduced by no more than a factor of about three. In addition, we have not yet accounted for possible yield trends. If we assume that area yields may be changing over time, and if the goal is to estimate the current expected yield for insurance purposes for a farmer, the average of past yields is a downward-biased estimator. The magnitude of the bias increases with the number of periods considered.

Equation (2) allows us to identify the factors (other than the time trend) that make a farmer's APH yield different from her expected yield. Based on T observations of farm- and area-level yields, the difference between a farmer's APH yield and expected yield is $\bar{y}_{it} - E(y_{it}) = \beta_i (\bar{y}_{ct} - \mu_{ct}) + \bar{\varepsilon}_{it}$. Thus, if an area has had good (bad) luck and β_i is positive, then the farmer's APH yield will tend to be greater (less) than her expected yield. If the farmer has a β_i of zero, then area luck will have no influence on her APH yield. The last term shows the effect of "on-farm luck" in the difference between the APH yield and expected yield.

The decomposition presented in equation (2) suggests an alternative to the current APH rule for calculating a farmer's expected yield, taking advantage of the information embedded in the area yield data. Equation (2) can be rewritten as

$$y_{it} = \mu_c(1 - \beta_i) + \delta_i + \beta_i y_{ct} + \varepsilon_{it} = \alpha_i + \beta_i y_{ct} + \varepsilon_{it}; \quad (3)$$

hence, $E(y_i) = \alpha_i + \beta_i \mu_c$. Assuming that the aggregation in the area yield is large enough, and after detrending farm and area level yields to a common base year, the regression parameters (α_i, β_i) can be estimated applying OLS to equation (3).¹² By the law of iterated expectations, and assuming that $\hat{\mu}_c$ is unbiased for the area expected yield, an unbiased estimator of farmer i 's expected yield is given by $\hat{y}_i = \hat{\alpha}_i + \hat{\beta}_i \hat{\mu}_c$.¹³

To compare yield estimators, we use the mean square error (MSE) criterion. The MSE of the APH yield estimator is given by

$$MSE_{APH} = E(\bar{y}_{it} - \mu_i)^2 = Var(\bar{y}_{it}) + (E(\bar{y}_{it}) - \mu_i)^2.$$

The APH yield estimator is biased downwards whenever yields are increasing over time. This problem with the APH rule was already reported by Skees and Reed (1986) and by JCQ.¹⁴

Assuming that farm-level yields follow a linear trend (the form fitted by Skees and Reed, and by Sherrick et al. 2004) of the form $y_{it} = \gamma_0 + \gamma_1 t$, the bias is given by $-\gamma_1(T+1)/2$.¹⁵ Thus, the severity of the bias depends, as expected, on the rate at which yields grow over time and on the number of periods of history considered. Since the current APH rules make no attempt to correct for the time trends, it is not clear whether farmers with longer histories will have an APH yield more in line with their expected yields than farmers with fewer (but more recent) yield records.

Since, as previously argued, the proposed estimator is unbiased, its MSE equals its variance, given by $MSE_{NEW} = Var(\hat{y}_{it})$, and the proposed estimator is an improvement over the APH rule whenever $MSE_{NEW} < MSE_{APH}$, or, equivalently, whenever

$$Var(\hat{y}_{it}) < Var(\bar{y}_{it}) + (E(\bar{y}_{it}) - \mu_i)^2, \quad (4)$$

where T highlights the number of observations used.

The variance of the APH rule can be derived from equation (3) as

$$Var(\bar{y}_{iT_i}) = T^{-1}Var(y_{it}) = T^{-1}Var(\alpha_i + \beta_i y_{ct} + \varepsilon_{it}) = T^{-1}(\beta_i^2 \sigma_c^2 + \sigma_\varepsilon^2)$$

and its MSE is

$$MSE_{APH} = T^{-1}(\beta_i^2 \sigma_c^2 + \sigma_\varepsilon^2) + \gamma_i^2 ((T+1)/2)^2.$$

We are now in a position to derive the variance of the proposed estimator. We first derive the variance under the assumption that we know the expected value of the area yield. The effects of relaxing this assumption are discussed in the Appendix.

When μ_c is known, equation (2) is rewritten as $y_{it}^* = \delta_i + \beta_i y_{ct}^* + \varepsilon_{it}$, where $y_{it}^* = y_{it} - \mu_c$ and $y_{ct}^* = y_{ct} - \mu_c$, and the parameters δ_i and β_i are estimated by OLS. Let Y_{cN}^* denote the (detrended) sample of area yields, where N denotes the number of observations in the sample. The regression parameters are estimated using the T_i observations for which paired data (y_{it}^*, y_{ct}^*) for farmer i are available. For any arbitrary choice of the area yield y_c^* the estimate of farmer i 's expected yield is $\hat{y}_i | y_c^*, Y_{cT_i}^* = \mu_c + \hat{\delta}_i + \hat{\beta}_i y_c^*$, and the variance of the estimator is given by

$$Var(\hat{y}_i | y_c^*, Y_{cT_i}^*) = Var(\mu_c + \hat{\delta}_i + \hat{\beta}_i y_c^* | y_c^*, Y_{cT_i}^*) = \sigma_\varepsilon^2 \left(\frac{1}{T_i} + \frac{(y_c^* - \bar{y}_{cT_i})^2}{\sum_{j=1}^{T_i} (y_{cj} - \bar{y}_{cT_i})^2} \right),$$

which is just the variance of the regression line at $y_c = y_c^*$. Now we need to acknowledge that the regressors are stochastic. As is well known, the OLS estimates are still the minimum variance linear unbiased estimators for δ_i and β_i (Greene 1999, p. 246). By the conditional variance identity (see Casella and Berger 2001, p. 167),

$$Var(\hat{y}_i | y_c^*) = \sigma_\varepsilon^2 \left(\frac{1}{T_i} + E_{y_c} \left(\frac{(y_c^* - \bar{y}_{cT_i})^2}{\sum_{j=1}^{T_i} (y_{cj} - \bar{y}_{cT_i})^2} \right) \right).$$

Hence, the estimator for farmer i 's expected yield is $\hat{\mu}_i = \mu_c + \hat{\delta}_i$ (i.e., the regression equation evaluated at $y_c^{*'} = 0$), which has an associated variance of

$$\text{Var}\left(\hat{y}_i \mid y_c^{*'} = 0\right) = \sigma_\varepsilon^2 \left(\frac{1}{T_i} + E_{y_c} \left(\frac{(\mu_c - \bar{y}_{cT_i})^2}{\sum_{j=1}^{T_i} (y_{cj} - \bar{y}_{cT_i})^2} \right) \right). \quad (5)$$

We are now ready to compare the MSE of both estimators for farmer i 's expected yield. In this case, the proposed estimator results in a lower MSE whenever

$$\sigma_\varepsilon^2 \left(\frac{1}{T_i} + E_{y_c} \left(\frac{(\mu_c - \bar{y}_{cT_i})^2}{\sum_{j=1}^{T_i} (y_{cj} - \bar{y}_{cT_i})^2} \right) \right) < T_i^{-1} (\beta_i^2 \sigma_c^2 + \sigma_\varepsilon^2) + \gamma_1^2 ((T_i + 1)/2)^2, \quad (6)$$

or, equivalently, whenever

$$\sigma_\varepsilon^2 E_{y_c} \left(\frac{(\mu_c - \bar{y}_{cT_i})^2}{\sum_{j=1}^{T_i} (y_{cj} - \bar{y}_{cT_i})^2} \right) < T_i^{-1} (\beta_i^2 \sigma_c^2) + \gamma_1^2 ((T_i + 1)/2)^2. \quad (7)$$

Equations (6) and (7) allow us to identify conditions under which the proposed estimator will perform better than the APH rule in the sense of MSE. From equation (7), increasing the number of periods used to compute the APH rule without detrending the yield observations will likely increase MSE_{APH} and decrease MSE_{NEW} .¹⁶ Hence, the proposed rule will likely perform better than the current APH rule, and even more so when the idiosyncratic risk (σ_ε^2) is small. Equation (7) also indicates that the APH rule has a greater chance of outperforming the proposed estimator when β_i is close to zero, and/or it cannot be estimated precisely. This makes sense because we are likely to introduce noise in trying to estimate β_i 's (and control for a systemic component) that have little effect on the farm-level yield. The introduced sampling variability may outweigh the risk we are trying to remove. Also, the proposed estimator will perform better than the current APH rule when the variance of the area yield increases for a given mean area yield. To see this, note that the left-hand side of equation (6) can be rewritten as

$$MSE_{NEW} = \frac{\sigma_\varepsilon^2}{T_i} E_{y_c} \left(\frac{\sum_{j=1}^{T_i} (y_{cj}^{*2})}{\sum_{j=1}^{T_i} (y_{cj}^* - \bar{y}_c^*)^2} \right) = \sigma_\varepsilon^2 \left(\frac{1}{T_i} + E_{y_c} \frac{(\bar{y}_c^*)^2}{(T_i - 1) s_{y_c}^{*2}} \right),$$

which is non-increasing in the variance of the area yield, whereas the right-hand side of equation (6) increases in the same variable.

From this point on, we will compare the proposed estimator against an “improved” APH rule applied on detrended farm-level yields. After correcting for the bias of the APH rule, equation (7) becomes

$$\sigma_\varepsilon^2 E_{y_c} \left(\frac{(\mu_c - \bar{y}_{cT_i})^2}{\sum_{j=1}^{T_i} (y_{cj} - \bar{y}_{cT_i})^2} \right) < \frac{\beta_i^2 \sigma_c^2}{T_i}, \quad (8)$$

or

$$\sigma_\varepsilon^2 E_{y_c} \left(\frac{\mu_c - \bar{y}_{cT_i}}{s_{y_c}} \right)^2 < \beta_i^2 \sigma_c^2 \frac{(T_i - 1)}{T_i},$$

or

$$\sigma_\varepsilon^2 E_{y_c} \left(\frac{\mu_c - \bar{y}_{cT_i}}{s_{y_c} / \sqrt{T_i}} \right)^2 < \beta_i^2 \sigma_c^2 (T_i - 1), \quad (9)$$

which makes the previous discussion more transparent.¹⁷

Figures 4-6 help to illustrate when we can expect the proposed estimator to outperform the “improved” APH rule. Since both estimators are now unbiased, only their variances are compared. To save space, only figures for Iowa CRD 50 are presented. Other districts yielded the same general patterns.

The farm-level yields used to compute the variance of the estimators were constructed using the same procedures detailed in the previous section. For every case, δ was fixed at zero. Farm-level yields and pairs of farm- and area-level yields were simulated and used to compute the bootstrapped variances for both estimators.

Figure 4 provides a comparison of the variance of the two estimators as the number of observations available per farm changes. Increasing the number of observations increases the precision for both estimators. When $\beta_i = 0.5$ (Panel a), the proposed estimator performs worse than the “improved” APH rule when few farm yield observations are available; however, this difference disappears (or reverses marginally) as the amount of available information increases. The relatively poor performance of the proposed estimator in this situation was expected. Equation (8) indicates that the proposed estimator will perform worse than the “improved” APH rule when β_i is low and there are few observations (for any given σ_ε^2 and σ_c^2). It may well happen that the short series of area yields used in the estimation of β_i has low variability (in the sense of having several similar observations), preventing a precise estimation of the regression parameters. This problem will be mitigated when a longer time series of area level yield data becomes available.

Panels (b) and (c) of Figure 4 show that the proposed estimator has a uniformly lower variance than the “improved” APH rule when β_i increases, except perhaps when $T_i = 4$, when, as just mentioned, the proposed estimator may exhibit erratic behavior. Note that an increase in β_i increases the systemic component of variance; hence, the total farm-yield variance increases also because we are holding constant the idiosyncratic component of variance. The change in the relative performance is because the “improved” APH rule performs worse and worse as the variance of the farm-level yield increases.

Figure 5 shows the relationship between a farmer’s beta and the variance of the estimators of the expected farm-level yield for different sample sizes. It makes clear that the total farm-level variability does not affect the variance of the proposed estimator. We believe this is an important advantage of the proposed estimator, since the “improved” APH rule will be likely to miss the target by more for producers with larger risk. The vertical distance between the two curves in any given panel of Figure 5 measures the difference in the variance of the two estimators considered. For example, switching to the proposed estimator will result in a reduction of 65.5, 58.6, and 39.4 (bu/acre)² in the variance with which the expected yield for the average farmer (those with $\beta_i = 1$) based on 5, 8, and 10 observations, respectively, is estimated. It is worth noting that for low

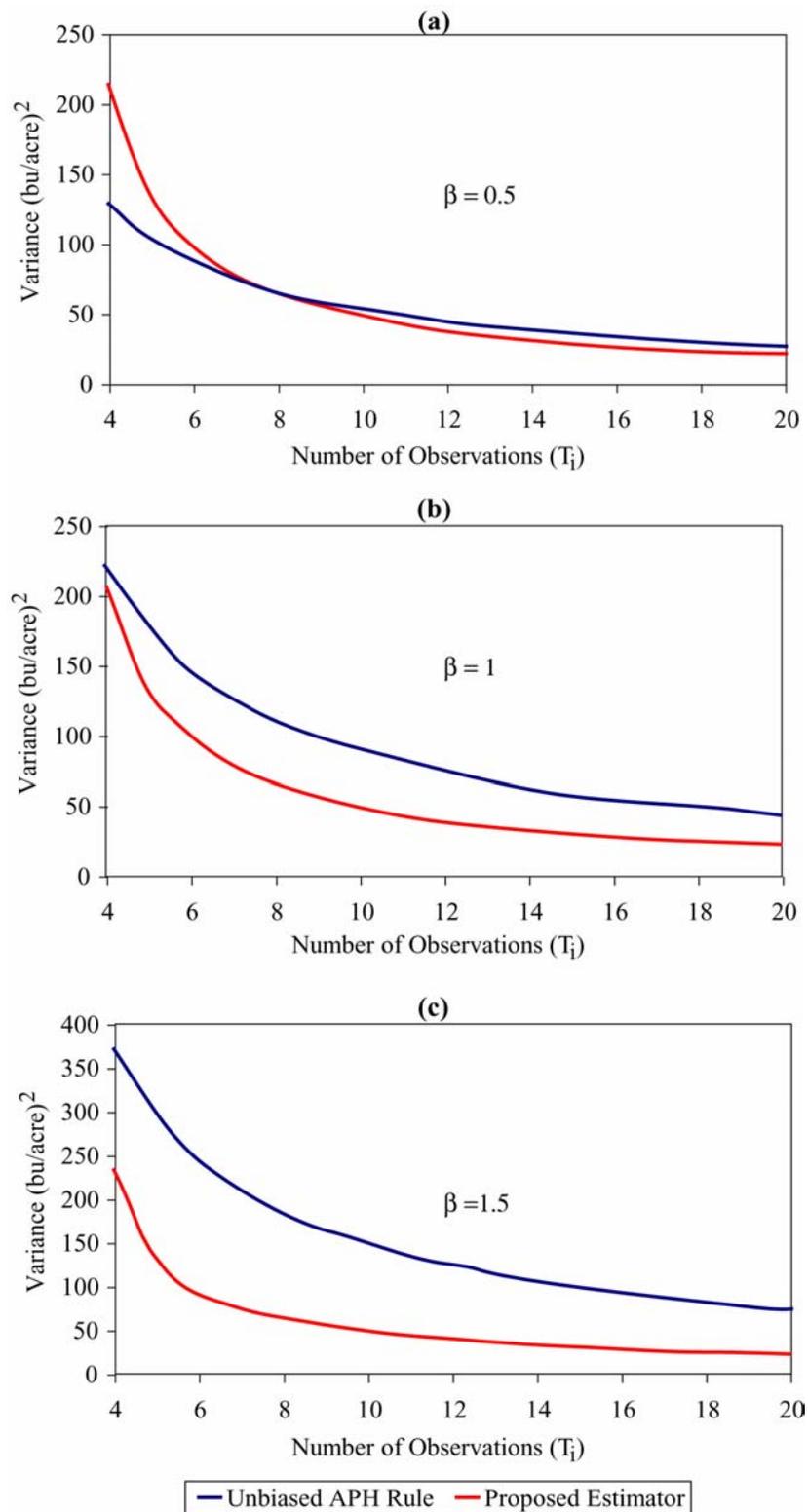


FIGURE 4. Relationship between the variance of an unbiased APH rule and the proposed estimator

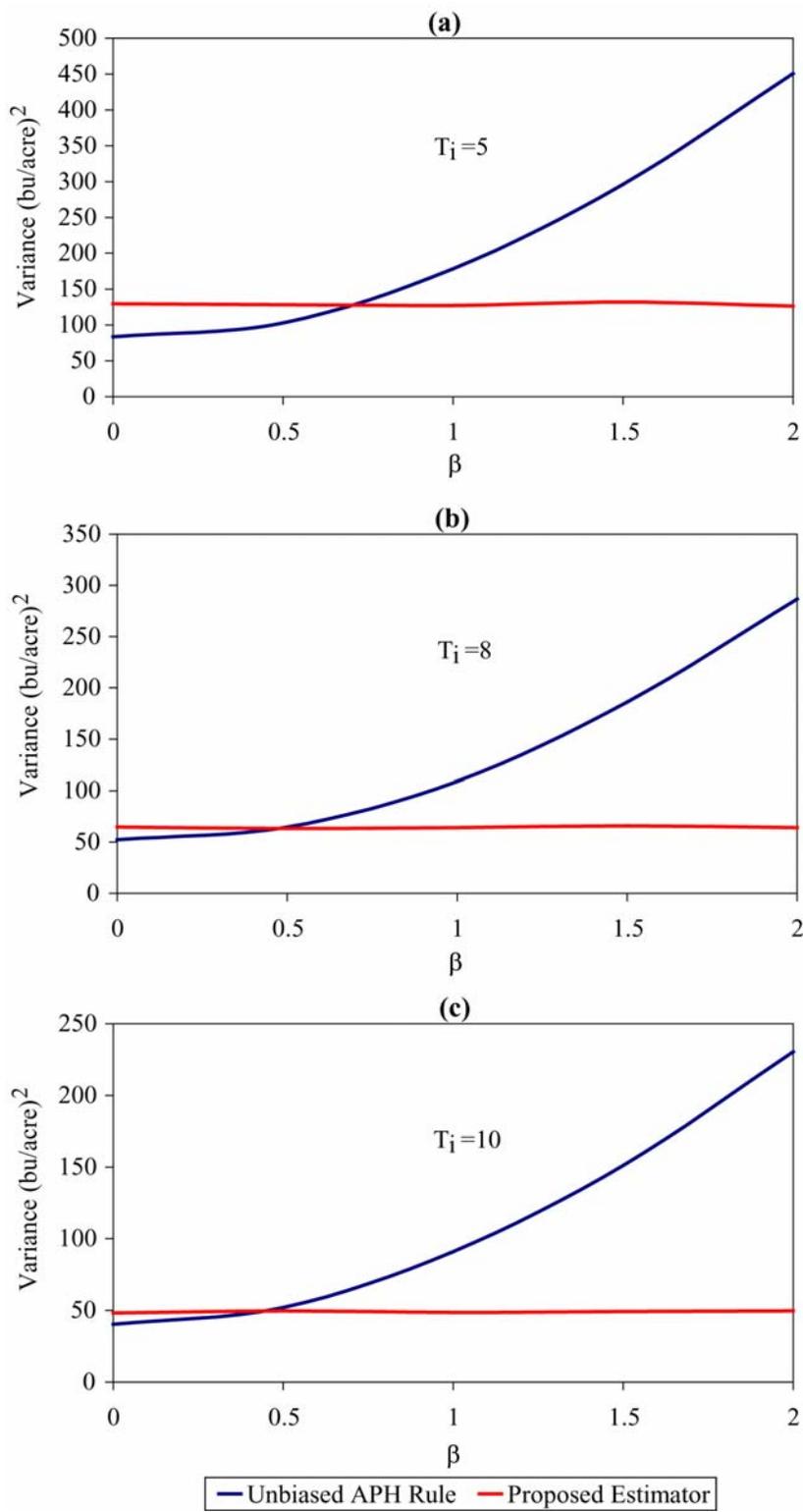


FIGURE 5. Variance of the unbiased APH rule and proposed estimator for different values of beta and number of observations

values of beta the “improved” APH rule performs marginally better than the proposed estimator, whereas the latter greatly outperforms the former for larger betas.

The analysis presented up to this point indicates that the proposed estimator will perform better than an improved APH rule when the contribution of the systemic risk to the farm-level yield risk is large and when the parameters of the hypothesized relationship can be estimated precisely. This reveals that in any given area, the variance with which expected farm-level yields are estimated will decrease with the proposed estimator for some farmers, whereas it is likely to increase for other farmers. Therefore, the next question is, Will we obtain aggregate variance reductions in a region by switching to the proposed estimator? The answer hinges on the distribution of betas within the region.

To obtain insight into the magnitude of the variance change that one could expect in aggregate by switching to the proposed estimator, we estimated the expected variance change for all Iowa CRDs when the estimators are based on 4 to 20 yield observations. The results are presented in Table 6 and Figure 6. Figure 6 presents the expected variances for both estimators as a function of the number of observations available for CRD 50.

As just mentioned, a distribution for the betas in each region is required for this task. Betas for individual farmers in a region were estimated by applying OLS to equation (2) for each farm in our data set for which we had at least eight yield observations. The average beta for each CRD was close to one, as expected, and the standard deviations were approximately 0.4 for all districts. As in Miranda (1991), the distribution of betas was centered at the mean, possessing no discernable skewness. Tests for normality failed to reject the null hypothesis (at a 5 percent confidence level) for all CRDs, and hence that distribution was used to obtain the distribution of betas by region.¹⁸

Figure 6 shows that the expected variance of both estimators decreases as the number of observations increases. More interestingly, Figure 6 shows that the expected variance of the proposed estimator is lower than the variance of the APH rule, except when there are just four observations available. The performance of both estimators improves at a decreasing rate as the number of observations used increases. Table 6, which presents the percent changes that would result from switching to the new estimator, indicates that this pattern is shared by all CRDs. Table 6 also reveals the erratic behavior, previously discussed, of the proposed estimator when only four observations are available for the current data set.

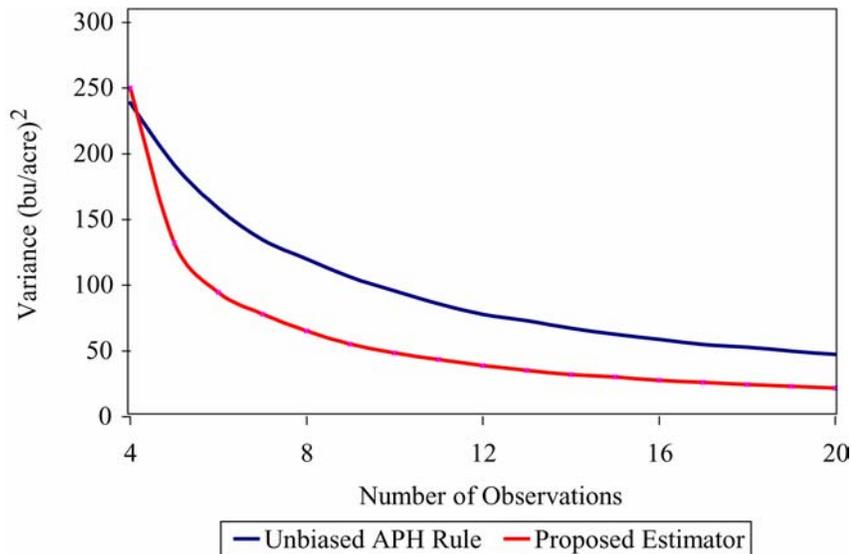


FIGURE 6. Expected variance of the estimators considered as a function of the number of observations available for Iowa crop reporting district 50

The main conclusion from Table 6 is that significant reductions in expected variances can be obtained by switching to the proposed estimator. This is especially true for farmers for which a large number of observations are available. This points out that gains in variance reduction beyond those resulting from increasing the number of observations can be achieved by switching estimators.

TABLE 6. Percent change in the expected variance as a result of moving from the current APH rule to the proposed estimator of expected yields for all CRDs in Iowa

CRD	Number of Observations in the Estimator							
	4	5	6	7	8	9	10	20
	(percent change in variance)							
10	289.0	-23.1	-34.9	-40.5	-43.5	-45.4	-47.0	-52.3
20	125.8	-15.1	-32.3	-37.3	-41.4	-42.2	-44.6	-48.8
30	20.5	-17.3	-30.4	-35.3	-40.2	-42.6	-44.0	-49.1
40	2.2	-26.9	-39.6	-43.9	-47.7	-50.0	-51.7	-56.5
50	4.7	-31.0	-40.3	-42.0	-45.6	-48.0	-49.2	-53.8
60	17.7	-18.2	-28.5	-33.1	-38.0	-39.6	-41.9	-47.9
70	143.2	-27.0	-36.9	-40.9	-44.2	-46.2	-47.3	-51.5
80	149.0	2.6	-21.0	-28.7	-33.1	-35.3	-37.1	-43.1
90	4.9	-39.9	-47.4	-51.0	-54.0	-55.2	-56.8	-60.5

Our analysis suggests that the best use of the estimator may be as a complement to the current APH yield estimator. In that line, it may be best to combine both estimators by using the current rule when there are few observations (four in our analysis), and switch to the new estimator after that. Of course, the current APH rule should be revised to account for yield trends.

The estimates of the elasticities of expected returns to crop insurance to the variance of the estimator of expected yields presented in the previous section indicate that sizeable improvements to the actuarial soundness of the Iowa corn insurance program are feasible by using the proposed estimator for expected yields for farmers with more than four years of validated yield registries.

An estimate of the economic importance of the reduction in variance shown in Table 6 can be obtained by multiplying the estimated elasticities of net profit from insurance with respect to variance by this change in variance. Table 7 reports the resulting estimates of the percent change in expected net profit from adoption of the proposed estimator.

Table 7 shows that a reduction in the variance of the estimator of expected yields will significantly reduce the expected net profit from insurance. This reduction would lead to improved rate making because premium rates would be better matched with expected yields and there would be a reduction in the problems caused by adverse selection because of a reduction in the deviations between true mean yields and APH yields.

Conclusions

This paper studies the impact of uncertainty in the estimation of farm-level expected yields on the actuarial soundness of the U.S crop insurance program. Reducing uncertainty about farm-level expected yields has the potential to significantly improve the actuarial performance of U.S. crop insurance and reduce problems associated with asymmetric information, such as adverse selection.

A conceptual model of insurance is presented and used to obtain insights about the magnitude of the impact and the potential for improvement. The framework is put to work using corn yield data from all CRDs in Iowa. The analysis suggests that significant gains from an actuarial perspective can be achieved by obtaining more precise estimators

TABLE 7. Expected percent change in producer returns to crop insurance at the 75 percent coverage level from moving to the proposed estimator

	Participate for All APH Yields				Participate Only When Expected Indemnity > Premium			
	Unsubsidized		Subsidized		Unsubsidized		Subsidized	
Number of observations ^a	5	10	5	10	5	10	5	10
CRD								
10	-24.9	-50.7	-4.6	-9.4	-15.2	-30.8	-8.6	-17.5
20	-16.7	-49.1	-3.2	-9.4	-10.3	-30.3	-5.2	-15.3
30	-17.2	-43.7	-3.0	-7.5	-11.2	-28.5	-5.8	-14.9
40	-22.8	-43.8	-4.0	-7.8	-16.2	-31.1	-9.4	-18.0
50	-28.7	-45.5	-4.0	-6.4	-18.7	-29.6	-8.3	-13.2
60	-21.1	-48.5	-3.4	-7.8	-12.3	-28.4	-5.8	-13.3
70	-24.5	-43.0	-4.4	-7.8	-16.4	-28.7	-9.2	-16.1
80	2.7	-38.1	0.4	-5.7	1.7	-23.7	0.8	-11.7
90	-41.2	-58.6	-6.0	-8.5	-24.2	-34.5	-11.7	-16.7
Average	-21.6	-46.8	-3.6	-7.8	-13.6	-29.5	-7.0	-15.2

^a Number of farm-level yield observations in the estimator.

of a farmer's expected yield. Additionally, better estimates of the farm-level yield would allow insurers to rate the products more accurately.

We propose a new estimator to replace or complement the current APH yield estimator of farm-level expected yields in the U.S. crop insurance program. The conditions upon which the proposed estimator is an improvement over the current APH rule are identified. Our findings suggest that significant improvements are feasible by moving to the new estimator for producers who have several years of yield records. Region-specific analyses are warranted to determine the number of observations needed to obtain the touted improvement for other crops and regions of the country.

Endnotes

1. See http://www.rma.usda.gov/FTP/Publications/directives/18000/pdf/05_18010.pdf for a complete description of APH rules.
2. A sufficient condition for returns to insurance to be non-decreasing in APH is $\frac{\partial \omega_{g\alpha}}{\partial \mu_a} \leq 0$. If we let $\omega_{g\alpha} = p_g \alpha \mu_a r_{\alpha \mu_a}$, where, as in JCQ, $r_{\alpha \mu_a}$ is the pure premium rate, the previous condition translates into $\varepsilon_{r, \mu_a} \leq -1$. The elasticity of the premium rate with respect to the APH yield is ε_{r, μ_a} . The necessary condition is $F(\alpha \mu_a) \geq r_{\alpha \mu_a} (1 + \varepsilon_{r, \mu_a})$ for all μ_a . From the analysis of Iowa premium schedules for corn just presented, the last inequality is very likely to hold for most farmers in every county of Iowa.
3. Clearly, the intended and actual coverage levels will coincide only when the APH yield and true mean yield coincide, which of course has a zero probability of occurrence.
4. They did find differences at the 75 percent level.
5. Note that $\frac{\partial^2 I}{\partial \mu_a \partial \alpha} = p_g (F(\alpha \mu_a) + \alpha \mu_a f(\alpha \mu_a)) \geq 0$, and

$$\frac{\partial^3 I}{\partial^2 \mu_a \partial \alpha} = p_g \alpha \left(2f(\alpha \mu_a) + \alpha \mu_a \frac{\partial f(\alpha \mu_a)}{\partial y} \right) \geq 0.$$

The latter inequality will be violated only by highly unlikely combinations of α and μ_a for which the second term is sufficiently negative (which implies a negative slope for the probability density function). Considering that the maximum coverage level is 85 percent and there is widespread agreement that yield distributions are negatively skewed (assuming unimodality), the APH yield would have to be at least 18 percent ($1/0.85$) higher than the true mean for the second term to be negative.

6. The formula used in rating Income Protection simulates farm yield j for farmer f in year q as $y_{q,j}^f = (\tilde{y}^f + \hat{u}_q^f + \tilde{e}_{q,j}^f) \hat{C}_q \tilde{C}^f$. The variables with a “tilde” are proportions of the county trend yield. The term within parentheses combines an estimate of the farm-level average proportion (\tilde{y}^f) yield, a deviation explained by the county pro-

portional variation ($\hat{u}_q^f = \hat{b}_1^f + \hat{b}_2^f (\tilde{C}_q^f - \underline{C}^f)$), and a residual term ($\tilde{\epsilon}_{q,j}^f$) obtained from a regression of farm yield on county yield and represents uncorrelated variability. The terms outside the parentheses convert the resulting proportions into actual yields.

7. Since the realized APH yield (μ_a) is an estimate of the farmer's expected yield ($\mu_c + \delta_i$), after observing the APH yield the insurer updates his beliefs about the farmer's expected yield by setting $\delta_a = \mu_a - \mu_c$.
8. Note that the percentage difference between elements in the upper halves of Tables 1 and 2 is just the ratio of expected premiums to premiums at the true mean yield.
9. Note additionally that the values reported in Tables 3 and 4 are expected returns for a given contract. In reality, farmers can change from one coverage level to another in successive periods, a choice that makes them at least as well off as when the coverage level is fixed.
10. Note that this is constant across districts since

$$\% \Delta \text{Variance} = \frac{\text{Var}(y)/10 - \text{Var}(y)/4}{\text{Var}(y)/4} * 100 = -60\%.$$

11. We keep the time subscript on the deviations because they will vary by year.
12. We detrend to take care of the time-varying expected yields.
13. $E(\hat{y}_i) = E_{\mu_c} E(\hat{\alpha}_i + \hat{\beta}_i \hat{\mu}_c | \hat{\mu}_c) = E_{\mu_c} (\alpha_i + \beta_i \hat{\mu}_c) = \alpha_i + \beta_i \mu_c$.
14. These authors identified this bias as a reason for participation problems.
15. $\text{Bias} = \left(E \left(T^{-1} \sum_{k=1}^T y_{i,t-k} \right) - E(y_{i,t}) \right)$
 $= \left(T^{-1} \sum_{k=1}^T (\gamma_0 + \gamma_1(t-k)) - \gamma_0 - \gamma_1 t \right) = T^{-1} \left(-\gamma_1 \sum_{k=1}^T k \right)$.
16. Note that the current APH rule considers only values of T_i greater than 3. For an area yield variance of 500 (bu/acre)² (the variance of corn yields for the nine CRDs of Iowa was estimated to be between 360 and 780), a β_i of 1, and a trend slope of 2 bu/acre/year, the minimum MSE_{APH} is attained when $T_i \approx 6$.
17. As the sample size increases, the central limit and Slutsky's theorems indicate that equation (9) reduces to $\sigma_\epsilon^2 < \beta_i^2 \sigma_c^2 (T_i - 1)$, since

$$\frac{\sqrt{T_i}(\mu_c - \bar{y}_{cT_i})}{\sigma_{y_c}} \frac{\sigma_{y_c}}{s_{y_c}} \rightarrow N(0,1).$$

Also, if one can argue that the average of farm-level yields (based on few observations) is approximately normally distributed, equation (9) reduces to $\sigma_\varepsilon^2 \leq \beta^2 \sigma_c^2 (T - 3)$. These formulas point to the same conclusion: the proposed estimator will be an improvement over the current rule when the idiosyncratic risk is small compared with the systemic risk, when the correlation between farm- and area-level yields is strong, and when the number of observations is large.

18. Analysis based on the empirical distributions of beta for different CRDs yielded similar results.

Appendix

Expected Area Yield Is Unknown

If we recognize that even with the amount of information available about area-level yields our estimate of its mean is still imperfect, we need to use the model described by equation (3). Note that this problem becomes less and less relevant as time goes by and more information is available. At some point, the expected area yield can be estimated with negligible variance.

In this case, the unconditional variance (applying again the conditional variance identity and assuming that our estimator for the expected area yield is independent of the observations used to estimate the parameters of the model) is given by

$$Var(\hat{y}_i) = \sigma_\varepsilon^2 \left(\frac{1}{T_i} + E_{\hat{\mu}_c} E_{y_c} \left(\frac{(\hat{\mu}_c - \bar{y}_{cT_i})^2}{\sum_{j=1}^{T_i} (y_{cj} - \bar{y}_{cT_i})^2} \right) \right) + \beta_i^2 Var(\hat{\mu}_c). \quad (A.10)$$

Note that if we can estimate the expected area yield with enough precision, equations (5) and (A.10) will yield similar results.

We are now ready to compare the MSE of both estimators for farmer i 's expected yield. In its full generality, the proposed estimator results in a lower MSE whenever

$$\sigma_\varepsilon^2 \left(\frac{1}{T^{-1}} + E_{\hat{\mu}_c} E_{y_c} \left(\frac{(\hat{\mu}_c - \bar{y}_{cT_i})^2}{\sum_{j=1}^{T_i} (y_{cj} - \bar{y}_{cT_i})^2} \right) \right) + \beta_i^2 Var(\hat{\mu}_c) < T^{-1} (\beta_i^2 \sigma_c^2 + \sigma_\varepsilon^2) + \gamma_1^2 (T(T+1)/2)^2$$

or, after replacing for the “improved” APH rule,

$$\sigma_\varepsilon^2 \left(E_{\hat{\mu}_c} E_{y_c} \left(\frac{(\hat{\mu}_c - \bar{y}_{cT_i})^2}{\sum_{j=1}^{T_i} (y_{cj} - \bar{y}_{cT_i})^2} \right) \right) + \beta_i^2 Var(\hat{\mu}_c) < \frac{\beta_i^2 \sigma_c^2}{T_i}. \quad (A.11)$$

Rearranging equation (A.11) further, and if N observations in the area yield are used to estimate its expected yield through a sample average, we obtain

$$\sigma_{\varepsilon}^2 \left(E_{\hat{\mu}_c} E_{y_c} \left(\frac{(\hat{\mu}_c - \bar{y}_{cT_i})^2}{\sum_{j=1}^{T_i} (y_{cj} - \bar{y}_{cT_i})^2} \right) \right) < \beta_i^2 \left(\frac{\sigma_c^2}{T_i} - \text{Var}(\hat{\mu}_c) \right) = \beta_i^2 \sigma_c^2 \left(\frac{1}{T_i} - \frac{1}{N} \right). \quad (\text{A.12})$$

Again, as the variance of the estimate of the area expected yield becomes smaller (for example as N goes to infinity) equation (A.12) becomes

$$\sigma_{\varepsilon}^2 E_{y_c} \left(\frac{(\hat{\mu}_c - \bar{y}_{cT_i})^2}{\sum_{j=1}^{T_i} (y_{cj} - \bar{y}_{cT_i})^2} \right) < \frac{\beta_i^2 \sigma_c^2}{T_i}.$$

References

- Atwood, J. A., A. E. Baquet, and M. J. Watts. 1996. "Income Protection." Technical Report. Montana State University, submitted to the U.S. Department of Agriculture, Economic Research Service, Commercial Agricultural Division. February 16. <http://www.rma.usda.gov/pilots/miscproj/techsum.pdf> (accessed November 2004).
- Babcock, B. A., C. E. Hart, and D. J. Hayes. 2004. "Actuarial Fairness of Crop Insurance with Constant Rate Relativities." *American Journal of Agricultural Economics* 86(3, August): 563-75.
- Casella, G., and R.L. Berger. 2001. *Statistical Inference*, 2nd ed. Duxbury Advanced Series. Toronto: Wadsworth/Thomson Learning.
- Coble, K., B. Goodwin, A. Ker, and T. Knight. 2002. "Rate Review Analysis—Report for External Review." A report for U.S. Department of Agriculture, Risk Management Agency. December 15.
- Coble, K. H., and T. O. Knight. 2001. "Crop Insurance as a Tool for Price and Yield Risk Management." Chap. 20 in *A Comprehensive Assessment of the Role of Risk in U.S. Agriculture*, edited by R. E. Just and R. D. Pope. Boston: Kluwer Academic Publishers.
- Greene, W. H. 1999. *Econometric Analysis*, 4th ed. Upper Saddle River, NJ: Prentice Hall.
- Just, R. E., L. Calvin, and J. Quiggin. 1999. "Adverse Selection in Crop Insurance." *American Journal of Agricultural Economics* 81(4, November): 835-49.
- Mahul, O. 1999. "Optimum Area Yield Crop Insurance." *American Journal of Agricultural Economics* 81(1, February): 75-82.
- Miller, S. E., K. H. Kahl, and P. J. Rathwell. 2000a. "Evaluation of Crop Insurance Premium Rates for Georgia and South Carolina Peaches." *Journal of Agribusiness* 18(3, Fall): 303-17.
- . 2000b. "Revenue Insurance for Georgia and South Carolina Peaches." *Journal of Agricultural and Applied Economics* 32(1, April): 123-32.
- Miranda, M. J. 1991. "Area-Yield Crop Insurance Reconsidered." *American Journal of Agricultural Economics* 73(2, May): 233-42.
- Ramaswami, B., and T. L. Roe. 2004. "Aggregation in Area-Yield Crop Insurance: The Linear Additive Model." *American Journal of Agricultural Economics* 86(2, May): 420-31.
- Sherrick, B. J., F. C. Zanini, G. D. Schnitkey, and S. H. Irwin. 2004. "Crop Insurance Valuation under Alternative Yield Distributions." *American Journal of Agricultural Economics* 86(2, May): 406-19.
- Skees, J., and M. Reed. 1986. "Rate Making for Farm-Level Crop Insurance: Implications for Adverse Selection." *American Journal of Agricultural Economics* 68(3, August): 653-59.
- Vercammen, J. A. 2000. "Constrained Efficient Contracts for Area Yield Crop Insurance." *American Journal of Agricultural Economics* 82(4, November): 856-64.