Relaxing Heteroscedasticity Assumptions in Area-Yield Crop Insurance Rating

Ardian Harri, Keith H. Coble, Alan P. Ker, and Barry J. Goodwin

This article focuses on the effect of differing heteroscedasticity assumptions on derived premium rates of area-yield crop insurance. Tests of the proportional and absolute heteroscedasticity assumptions are conducted using both in-sample and out-of-sample measures. Our results suggest that arbitrarily imposing a specific form of heteroscedasticity or homoscedasticity in insurance rate calculations limits actuarial soundness. Our results have practical implications for the federal crop insurance programs, as we reject the traditional rating assumptions for many cotton regions and lower-yielding/higher-risk corn and soybean counties but not in the heart of the Cornbelt.

Key words: area-yield insurance, heteroscedasticity, out-of-sample forecasting.

JEL code: Q18.

Area-yield insurance has attracted significant attention as a means to avoid the problems of moral hazard, inaccurate rates, and transactions cost associated with individual coverage crop insurance policies (Deng, Barnett, and Vedenov 2007; Glauber 2004; Miranda 1991). With an area design, coverage and payments are based on the observed yield for an area that encompasses several producers. By doing so, moral hazard is mitigated because a producer usually directly controls inputs for a relatively small proportion of aggregate production. Area-yield insurance does not conform to the traditional model of adverse selection in that producers in a county are exposed to the same payout, so there can be no separating equilibrium between low- and high-risk insureds. However, producers in a county with superior information about the common county yield risk could select against a poorly rated program.

Federally subsidized area-yield insurance is offered for several major crops in the United States. Introduced in 1993, the Group Risk Plan (GRP) provides area-yield insurance based on county-yield data of the National Agricultural Statistics Service (NASS) of the USDA. GRP covered $25 million of liability in its first year. In recent years a revenue insurance variant, Group Risk Income Protection (GRIP) has been added. By 2008 the liability for these two programs neared $8.5 billion and acreage insured exceeded 34 million acres. A number of other area-based policies based on weather and vegetation indexes are also available.

While area yield designs are often expected to reduce moral hazard and improve actuarial performance relative to individual coverage, this conclusion is predicated upon producers not having superior knowledge of local yield trends and yield variability. Thus, recognizably erroneous rating and modeling assumptions can lead to a program that violates the legislative mandate that rates must be actuarially sound—neither overcharging nor undercharging producers.[1]

Several studies have either explicitly or implicitly proposed systems to set rates for area-yield designs. Halcrow (1949) proposed an early area-wide plan, but the current U.S. design is largely based on the work of Skees, Black, and Barnett (1997). Alternative procedures have been considered by Goodwin and Ker (1998), Ker and Coble (2003),

Ardian Harri and Keith H. Coble are an assistant professor and professor at Mississippi State University, Alan Ker is a professor at University of Guelph, and Barry Goodwin is a professor at North Carolina State University. This study was supported by the Mississippi Agricultural and Forestry Experiment Station at Mississippi State University. Approved for publication as Journal Article No. J-11950 of the Mississippi Agricultural and Forestry Experiment Station, Mississippi State University.

[1] The actuarial soundness target loss ratio was revised to 1.0 in the 2008 Farm Bill. This target is applied prior to legislatively mandated rate subsidies.

Amer. J. Agr. Econ. 93(3): 707–717; doi: 10.1093/ajae/aar009
Received April 2010; accepted January 2011; published online March 7, 2011

© The Author (2011). Published by Oxford University Press on behalf of the Agricultural and Applied Economics Association. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

Downloaded from http://ajae.oxfordjournals.org/ at Pennsylvania State University on May 18, 2016

and Ker and Goodwin (2000). The typical approach is to regress historical yield against a deterministic trend and then use the estimated residuals—along with assumptions about the existence and form of heteroscedasticity—to estimate expected insurance indemnities. Skees, Black, and Barnett (1997) fitted a two-knot, linear spline to historical yield data and then estimated rates using the regression residuals. They assumed that the standard deviation of the residuals increased proportionally with increases in yields. Their approach encompasses technological change, adjustments in input use, and any other deterministic yield effects and treats all remaining variation as random shocks arising primarily from weather, insects, diseases, and other assumed uncontrollable events.[2]

Two primary heteroscedasticity assumptions have been maintained in the area-yield insurance literature. The assumption of a constant coefficient of variation where changes in the yield standard deviation are proportional to the changes in the yield mean has been extensively used (Deng, Barnett, and Vedenov 2007; Ker and Coble 2003; Miranda and Glauber 1997; Skees, Black, and Barnett 1997). We refer to the insurance rates calculated using this heteroscedasticity assumption as "proportional rates." The assumption of homoscedasticity has also been used extensively (Coble, Heifner, and Zuniga 2000; Mahul 1999; Miranda 1991). We identify insurance rates calculated using the homoscedasticity assumption as "absolute rates." The literature on heteroscedasticity remains ambiguous for two reasons: (*a*) assumptions have often been explicitly or implicitly made without testing; and (*b*) empirical evidence has been provided to support specific applications of both the homoscedastic model and the constant coefficient of variation (CV) assumptions, but neither holds universally across location or crop type. Such assumptions about heteroscedasticity are not always fully examined; yet there is a significant empirical literature (starting with the seminal work by Just and Pope [1978]) suggesting that heteroscedasticity may be present in crop yields and that, if present, it varies spatially across regions.

[2] A better in-sample fit might be achieved by modeling inputs more explicitly but since those inputs are not observable at the time of rating, the model cannot be used to forecast yields (unless those inputs are first forecasted). Further, crop insurance plans typically require that participating producers follow "good farming practices," though verification of such practices may present a significant challenge for insurers.

In this article, we focus on the effect of the different heteroscedasticity assumptions on area-yield insurance rates. When the estimated residuals are used for insurance rating, heteroscedasticity assumptions have significant effects on the derived rates. Woodard, Sherrick, and Schnitkey (2009) have recently raised the issue that heteroscedasticity assumptions are also relevant for the individual coverage products offered by the Risk Management Agency (RMA). They suggest that lost-cost rating procedures, as used by RMA, implicitly impose proportional heteroscedasticity. We assume that the yield variance is a function of predicted yield. However, we estimate the relationship between the variance and predicted yield rather than a priori impose either the proportional or the absolute form of heteroscedasticity. We refer to the insurance rates calculated using this approach as "empirical rates" and compare them to conventional proportional and absolute rates.

Our results demonstrate that no single heteroscedasticity assumption is appropriate in every case. Being correct on average or even in a majority of cases may still lead to significant problems in the design, rating, and performance of crop insurance contracts. Further, the federal crop insurance program, like any other government program offering benefits, is subject to a wide range of political considerations. Inequities in the provision of program benefits often trigger political debate and realignment of policies. Therefore, inaccuracies and inequities across regions and crops are important issues, even when the distortions occur at the margin. We demonstrate that standard heteroscedasticity assumptions often break down when one moves outside of principal growing areas. In particular, we show that in excess of $1.3 billion in total premium existed in crops and regions in 2009 for which the conventional rating assumptions were rejected.

In order to provide a measure of the economic differences between the alternative heteroscedasticity assumptions, we estimate county-level models for corn, soybeans, and cotton in several states. We consider the differences between the proportional, absolute, and empirical rates for each county and multiply these differences by total liability to obtain the respective differences in total premiums for each county. We summarize (results are reported in table 1) these county total premiums for several major corn, soybean, and cotton producing states—respectively, Illinois, Indiana, Iowa, Minnesota, Missouri, Ohio, and

Downloaded from http://ajae.oxfordjournals.org/ at Pennsylvania State University on May 18, 2016

Table 1. Summary of the Overcharged or Underpaid Premium Amounts Based on the Difference Between the Proportional and Empirical Rates for Several Major Corn Producing States (in US$)

State	Average overcharge/ underpayment	Std Dev of overcharge/ underpayment	Maximum underpayment to the program[a]	Maximum overcharge to farmers[a]	Total underpayment to the program	Total overcharge to farmers
			Corn			
IL	−82,671	1,081,640	−6,429,329	2,255,108	−34,278,506	26,342,114
IN	−31,623	354,660	−1,185,032	918,680	−10,315,362	7,690,691
IA	18,583	255,213	−915,362	707,298	−6,848,936	8,670,087
MN	20,387	114,208	−285,213	634,442	−790,959	1,769,524
MO	45,928	119,733	−115,943	588,272	−345,000	2,595,473
OH	27,138	94,657	−255,477	241,508	−1,406,242	3,251,649
WI	−36,088	393,574	−1,835,075	1,202,875	−5,054,275	3,069,418
			Soybeans			
IL	8,743	51,105	−274,364	112,104	−1,453,729	2,319,302
IN	30,189	57,813	−127,378	250,010	−531,108	3,036,793
IA	3,157	14,278	−77,605	40,504	−316,818	623,056
MN	3,164	6,813	−6,532	34,956	−17,728	125,317
MO	7,916	26,944	−81,411	112,420	−146,483	550,200
OH	22,744	31,207	−59,640	92,734	−213,963	1,669,607
WI	58,023	9,858	−6,387	51,774	−6,523	216,145
			Cotton			
AL	36,452	55,313	13,563	155,571	–	255,164
AR	1,281,594	1,183,008	361,009	2,615,901	–	3,847,824
GA	133,642	138,411	2,166	365,164	–	1,202,773
LA	26,984	13,923	−21,777	16,355	−23,419	23,554
MS	65,479	168,862	−30,117	408,293	−45,201	438,077
TN	165,044	149,392	35,453	324,159	–	660,175

Note: [a]These are calculated at the county level.

Wisconsin for corn and soybeans and Alabama, Arkansas, Georgia, Louisiana, Mississippi, and Tennessee for cotton.

Given these differences, we investigate the effect of differing heteroscedasticity assumptions on derived premium rates at varying coverage levels. We conduct a repeated game of insurance program selection. Specifically, using the RMA rating methodology and heteroscedasticity assumptions for area-yield insurance, we derive actuarially fair rates. We assume the role of a private insurance company and re-derive the rates using the RMA methodology but instead of using the RMA's heteroscedasticity assumption we estimate the form of the heteroscedasticity. Note that both procedures are estimates of actuarially fair rates and differ only in the treatment of heteroscedasticity. This allows us to consider the effect of the heteroscedasticity assumption on the actuarial performance of the program.

Our results suggest that arbitrarily imposing a specific form of heteroscedasticity or homoscedasticity in insurance rate calculations is unwise. The magnitude of the rate effects in many cases is quite large and could potentially introduce significant rate inaccuracies into an area-yield program. Based on earlier versions of this work, the USDA/RMA has now adopted the approach suggested in this article and applied it to GRP and GRIP programs nationally. While the standard RMA assumption is often not rejected in major corn and soybean producing regions, the new approach appears to provide more actuarially sound results in many cotton counties and in many lower-yielding, higher-risk corn and soybean producing counties.

Data and Methods

County-yield data for corn, soybeans, and cotton were obtained from NASS. The time span for the data is 1955–2006. We include only those counties with a complete yield history. This filtering of counties was consistent with the program design procedures followed by RMA. In addition, we attempt to avoid thin data problems that could potentially distort our

Downloaded from http://ajae.oxfordjournals.org/ at Pennsylvania State University on May 18, 2016

results and thus eliminated counties with less than 5,000 total acres of the relevant crop.

Trend Estimation

Researchers have considered numerous functional forms and specifications for yield trend models. Several previous studies explored the relative merits of alternative functional forms for trend estimation (e.g., Goodwin and Ker 1998; Harri et al. 2009; Ker and Coble 2003). Determining the correct functional form or model to describe technological change is important for crop insurance rate estimation. Since crop yields tend to trend upward, it is necessary to remove the trend effect of technology. Approaches include deterministic and stochastic trend models, although Harri et al. (2009) found limited support for stochastic trends in crop yields.

In evaluating alternative ways of modeling trend, we considered a number of different functional forms and estimation methods. These include a simple linear trend and more sophisticated functional forms like the one- or two-knot linear spline model currently used in the design of the GRP and GRIP programs and various nonparametric models. We also considered L_1, L_2, M-estimation and Bayesian estimation procedures.[3] Alternative models and estimation methods resulted in no measurable change in our results and conclusions regarding heteroscedasticity assumptions.[4]

In this article we report the results from a two-knot linear spline model with robust M-estimation, which is the approach RMA recently adopted on the basis of work related to this study. M-estimation techniques are a hybrid of L_2 and L_1 methods in that they are less sensitive to outliers than L_2 but tend to be more efficient than the L_1 methods. M-estimators can be used when outliers in the dependent variable (yields) are possible. In our estimation, we iterate using Huber weights until convergence and then use bisquare weights for two iterations.[5] Specifically, we minimize $\sum_{t=1}^{T} f(e_t, c) = \sum_{t=1}^{T} f(y_t - \beta' x_t)$ using first the Huber function and then the bi-square function for two iterations. They are defined as follows:

$$\text{Huber function} =$$

$$f(e,c) = \begin{cases} 1 & \textit{if } |e| < c \\ \dfrac{c}{|e|} & \text{otherwise} \end{cases}$$

$$\text{Bisquare function:}$$

$$f(e,c) = \begin{cases} \left(1 - \left(\dfrac{e}{c}\right)^2\right)^2 & \textit{if } |e| < c \\ 0 & \text{otherwise} \end{cases}$$

where c is a constant (we use the default value, 1.345, for the Huber function, and 4.685 for the bi-square function).

To increase the stability of trend estimators over time and across space, we impose temporal and spatial priors[6] on the knots for the spline models—an approach that RMA has also recently adopted. The two-knot linear spline model is defined as:

$$(1) \qquad y_t = \gamma_1 + \gamma_2 t + \gamma_3 d_1 (t - knot_1) + \gamma_4 d_2 (t - knot_2) + e_t$$
$$d_1 = 1 \text{ if } t \geq knot_1,\ 0 \text{ otherwise},$$
$$d_2 = 1 \text{ if } t \geq knot_2,\ 0 \text{ otherwise}$$

where y_t is the yield at time t, $t = 1, T$, and, γ_1, γ_2, γ_3, γ_4, $knot_1$, and $knot_2$ are parameters to be estimated. We make no specific assumptions about the distribution of e_t.

Heteroscedasticity Assumptions and Derived Rates

We calculate the proportional, absolute, and empirical GRP rates by modeling yield heteroscedasticity as a function of the predicted yield. Traditionally, the RMA procedures have implicitly assumed a special case of this heteroscedasticity function. Modeling yield heteroscedasticity as a function of predicted yield will allow testing the various assumptions regarding heteroscedasticity that have been used in previous research. These different assumptions on the form of

[3] L_1 and L_2 regressions are respectively, median and SSE-minimizing regression, and M estimation is a hybrid approach between L_1 and L_2.

[4] Results with alternative models and estimation methods not presented here are available from the authors.

[5] This is the default in S-Plus.

[6] We imposed a Bayesian uniform prior on the knot point by imposing the temporal restriction that the knot cannot change by more than three years in either direction from the previous year. We also imposed a Bayesian uniform prior on the knot point by imposing the spatial restriction that the knot cannot be more than three years in either direction from the average of the knots for all counties within the crop reporting district.

Downloaded from http://ajae.oxfordjournals.org/ at Pennsylvania State University on May 18, 2016

Table 2. Differences of Absolute (A) and Data Determined (E) Rates and Proportional (P) and Data Determined (E) Rates for Restricted Spline Trend

Crop	Het form	Mean	SD	Min	Max
			Coverage Level 65%		
Corn	A-E[a]	−0.001	0.014	−0.316	0.202
	P-E[b]	0.016	0.041	−0.187	0.431
Soybeans	A-E[a]	−0.003	0.009	−0.088	0.039
	P-E[b]	0.001	0.009	−0.082	0.042
Cotton	A-E[a]	−0.003	0.016	−0.161	0.007
	P-E[b]	0.138	0.149	−0.168	0.362
			Coverage Level 90%		
Corn	A-E[a]	−0.003	0.014	−0.272	0.135
	P-E[b]	0.028	0.041	−0.157	0.447
Soybeans	A-E[a]	−0.006	0.014	−0.115	0.036
	P-E[b]	0.007	0.015	−0.089	0.057
Cotton	A-E[a]	−0.006	0.017	−0.101	0.032
	P-E[b]	0.139	0.142	−0.105	0.358

Note: Table 2 presents the differences between rates for the three different heteroscedasticity assumptions for each crop and two coverage levels. The mean, standard deviation, maximum, and minimum of the absolute differences by crop and coverage level are reported.
[a] Difference between absolute and empirical heteroscedasticity.
[b] Difference between proportional and empirical heteroscedasticity.

heteroscedasticity can be represented by the following relationship:

$$(2) \qquad var(e_t) = \sigma^2[E(y_t)]^\beta = \sigma^2\hat{y}_t^\beta$$

where $var(e_t)$ is the variance of the error term in (1). The case of $\beta = 0$ indicates homoscedastic errors. When $\beta = 1$ the variance of yields moves in direct proportion to the predicted (trending) yield. Finally, the case of $\beta = 2$ suggests that the standard deviation of yields moves in proportion to the predicted (trending) yield (i.e., that the coefficient of variation is constant).

We estimate equation (2) using the following:

$$(3) \qquad \ln(\hat{e}_t^2) = \alpha + \beta\ln(\hat{y}_t) + \varepsilon_t$$

where $\hat{e}_t$ is the error term from equation (1), $\hat{y}$ is the predicted value from a trend line estimation. To calculate rates, the residuals from equation (1) are adjusted as follows:[7]

$$(4) \qquad e\_adj_t = \hat{e}_t \frac{\hat{y}_{T+2}^{\hat{\beta}}}{\hat{y}_t^{\hat{\beta}}}$$

where T is the length of the yield series and $\hat{y}_{T+2}$ is the two-year-ahead out-of-sample yield forecast. We use $\hat{\beta} = 2$ to obtain the adjusted residuals for calculating the proportional rate, $\hat{\beta} = 0$ to obtain the adjusted residuals for calculating the absolute rate, and the estimate of $\hat{\beta}$ from equation (3) to obtain the adjusted residuals for calculating the empirical rate. We calculate the indemnity received by the insured, *indem*, for each of these three scenarios using the following formula:

$$(5)\ indem_t = \max\left[\left(\frac{y_c - y_t^*}{y_c}\right)(\hat{y}_{T+2})(scale), 0\right]$$

where

$$y_c = \hat{y}_{T+2}{}^*cov$$
$$y_t^* = \hat{y}_{T+2} + e\_adj_t$$

cov is the coverage level, and $scale = 1$. The GRP rate is then calculated using:

$$(6) \qquad r = E(indem_t)/\hat{y}_{T+2}.$$

In table 2 we present the rates for the three different heteroscedasticity assumptions for each crop and two coverage levels. Because aggregation will hide important differences, we present the mean, standard deviation, maximum, and minimum of the absolute differences by crop and coverage level. The results indicate that while on average the absolute rates are

[7] When heteroscedasticity is present this would lead to inefficient parameters for the mean equation and invalid standard errors. Correcting for heteroscedasticity would certainly address both issues. However, obtaining efficient parameter estimates for the mean equation and corrected standard errors for these parameters would have no effect on the main results of the study, the effect of heteroscedasticity assumptions on insurance premium rates.

Downloaded from http://ajae.oxfordjournals.org/ at Pennsylvania State University on May 18, 2016

lower than the empirical rates, the difference between the two rates can be large. For example, for corn the differences range from as low as 32% to as high as 20%. Also, while on average the proportional rates are higher than the empirical rates, the difference between the two rates can be as low as 19% and as high as 45%. Similar differences are found for cotton, while smaller differences are revealed for soybeans.

To provide a measure of the insurance premium differences between the alternative heteroscedasticity assumptions, we conducted a county-level analysis for corn, soybeans, and cotton in several states. We first obtained the total actual 2007 area plan insurance liability for all acres covered under an area product policy and for each county. Because the RMA's experience has traditionally been based on the proportional heteroscedasticity assumption, we examine the difference between the rates calculated under this assumption and the empirical rates for each county. We multiply this difference by the total liability in each county to obtain the difference in the total premium if the empirical rates were used instead of the proportional rates. We calculate total premiums for major corn, soybean, and cotton producing states, including Illinois, Indiana, Iowa, Minnesota, Missouri, Ohio, and Wisconsin for corn and soybeans, and Alabama, Arkansas, Georgia, Louisiana, Mississippi, and Tennessee for cotton.

These results, reported in table 1, suggest that while the mean total premium differences for each state are not large, there is considerable variation in total premium differences, even within a certain state. As an example, for Logan County, Illinois, the proportional rate is 3.9% and the empirical rate is 10.7% or 2.75 times higher. This rate difference results in a difference in total premiums of $6.4 million. In another example, in Macoupin County, Illinois, the proportional rate is 4.5% and the empirical rate is 2.3%, or about half of the proportional rate, resulting in a difference in total premium of $2.3 million. These large county-level differences in the premiums obviously may lead to important selection and actuarial performance issues.

Heteroscedasticity Tests

The results noted above demonstrate that heteroscedasticity assumptions can have significant impacts on premium rates and thus on total premium. We find that neither homoscedastic error terms, nor the constant proportional heteroscedasticity assumption are strongly supported. Previous work usually tests different assumptions regarding the form of heteroscedasticity using equation (3). For example, if the test[8] fails to reject $H_0: \beta = 0$, then one can assume homoscedasticity when calculating insurance rates. Alternatively, if a test of $H_0: \beta = 2$ is not rejected, one may impose the proportional heteroscedasticity correction in rate calculations.

We tested different forms of heteroscedasticity using the RMA temporal model, a two-knot linear spline model with robust M-estimation and temporal and spatial priors. Our results were robust across the various temporal models and estimation methodologies previously discussed. Using the estimated residuals resulting from the two-knot linear spline model, we tested the assumptions of homoscedasticity and a constant coefficient of variation using equation (3) and L_2 estimation. The results of these tests are reported in table 3. We also tested for time-varying heteroscedasticity using a spline model similar to equation (1), and the results, not surprisingly, are very similar to those presented with respect to predicted yields, since predicted yields follow a time trend.

The results in table 3 indicate that neither assumption is overwhelmingly supported by the data and both assumptions are sometimes rejected. For example, for corn the proportional heteroscedasticity assumption is rejected for 467 counties, or 41% of the total. In addition, in 95 out of 668 counties where we fail to reject the proportional heteroscedasticity assumption, we also fail to reject the alternative assumption of homoscedasticity. This means that for these 95 counties the current RMA rates that are based on the proportional heteroscedasticity assumption may also be incorrect. In 185 of the 467 corn counties (or 40%) where we reject proportional heteroscedasticity the parameter $\hat{\beta}$ in equation (3) was greater than two, and in the remaining 282 counties this parameter was less than two.

Geographically, counties for which the $\hat{\beta}$ parameters are less than two are concentrated in the riskier portion of the Great Plains, while those with $\hat{\beta}$ parameters estimated to be greater than two are scattered in the Corn Belt and in the more productive portions of the eastern United States. Similarly, of the 230 soybeans counties where we reject proportional

[8] We use a simple *t*-test to test these hypotheses.

Downloaded from http://ajae.oxfordjournals.org/ at Pennsylvania State University on May 18, 2016

Table 3. Fail to Reject Cases of Different Heteroscedasticity Assumptions for Corn, Soybeans and Cotton Counties for the Restricted Spline Trend Estimator

Crop	Heteroscedasticity assumption	Fail to reject one assumption	Fail to reject both assumptions
Corn	Proportional Heteroscedasticity	668 (59)	95 (14)
	Absolute Heteroscedasticity	276 (24)	95 (14)
Soybeans	Proportional Heteroscedasticity	653 (74)	140 (16)
	Absolute Heteroscedasticity	357 (40)	140 (16)
Cotton	Proportional Heteroscedasticity	52 (36)	37 (26)
	Absolute Heteroscedasticity	66 (46)	37 (26)

Note: Numbers in parentheses are percentage values.

heteroscedasticity, in 130 counties (or 57%) the parameter $\hat{\beta}$ in equation (3) was greater than two and in the remaining 100 counties this parameter was less than two. The geographical pattern for the soybeans estimates of $\hat{\beta}$ is quite similar to that of corn. For cotton, in 6 of the 91 counties where we reject proportional heteroscedasticity (or 7%) the $\hat{\beta}$ parameter in equation (3) was greater than two, while in the remaining 85 counties this parameter was less than two. In contrast to corn and soybeans, the counties found to have $\hat{\beta}$ less than two tend to be in higher-yield regions and those areas with irrigation.

Table 3 also shows that we reject homoscedasticity with corn in 859 counties or 76 percent of the total. Results are similar for the other two crops, soybeans and cotton.

To examine whether spatial patterns exist in our test of proportional heteroscedasticity, we mapped the county-by-county results for each of the three crops. A spatial pattern of rejecting the assumption suggests differences are in part due to the nature of mean yield and yield risk. Figure 1 maps the county-by-county tests of proportional heteroscedasticity forcorn. For the most part, we fail to reject proportional heteroscedasticity in the major Corn Belt counties. Specifically, most of southern Minnesota, Iowa, Central Illinois, and western Indiana fail to reject proportional heteroscedasticity. Conversely, we reject proportional heteroscedasticity in most counties in the Great Plains regions

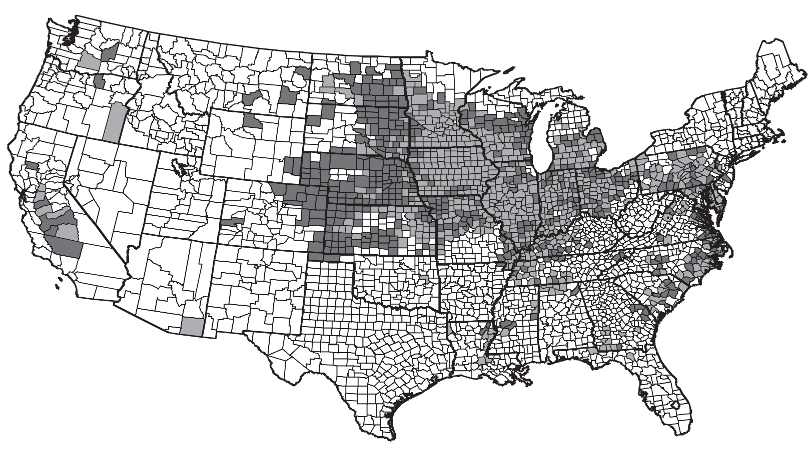

Figure 1. Test of the proportionality assumption for corn

Downloaded from http://ajae.oxfordjournals.org/ at Pennsylvania State University on May 18, 2016

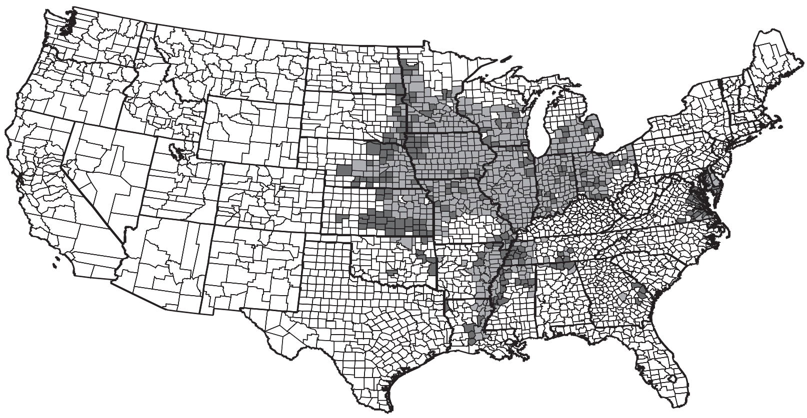

Figure 2. Test of the proportionality assumption for soybeans

of Kansas, Nebraska, South Dakota and North Dakota that have mostly dryland production. Other areas which tend to be more marginal corn producing regions, such as the Eastern Seaboard and the Southeast, are shown to have mixed results.

Figure 2 illustrates the results for soybeans. Again rejection of proportional heteroscedasticity occurs more often in lower-yielding, higher-risk soybean producing counties. In contrast, in very few counties in Illinois, Indiana, Iowa, Missouri, Ohio, and Wisconsin is proportional heteroscedasticity rejected. Overall, proportional heteroscedasticity is rejected less often for soybeans than corn.

The proportional heteroscedasticity tests for cotton are shown in figure 3. Cotton has some very distinct spatial production patterns. Proportional heteroscedasticity is rejected in the Texas high plains, Georgia, and North Carolina. Conversely, results for Arizona, California, Louisiana, and most of Mississippi are consistent with proportional heteroscedasticity.

In general, the results for these three crops show that proportional heteroscedasticity is most often rejected in lower-yielding, riskier regions, which in many cases are also regions with a less pronounced yield trend. An interesting implication is that we fail to reject the RMA maintained assumption in the regions where most production and insured acreage occur. Conversely, this also suggests that RMA assumptions have resulted in inaccurate rates and poor program performance in riskier, more marginal regions.

In order to provide a picture of the distribution of the insurance policies in areas where we reject or fail to reject the proportional heteroscedasticity assumption, we calculate the percentages of total area-based insurance premium located in counties where we reject the assumption. These are around 18%, 20%, and 34% for corn, soybeans, and cotton, respectively. In light of Woodard, Sherrick, and Schnitkey's (2009) arguments that a constant yield CV is inappropriate for the individual coverage products offered by RMA, we also calculate the percentage of total premium for the Actual Production History (APH) program located in counties where we reject the proportional heteroscedasticity assumption. The APH program has around 32%, 30%, and 40% for corn, soybeans and cotton, respectively.

Out-of-Sample Comparisons

In this section we ascertain whether the heteroscedasticity assumptions which drive differences in premium rates are economically significant. We do so by assuming the role of an insurance company that can choose to retain or cede policies, such as is the case under the Standard Reinsurance Agreement (SRA) currently in place in the federal crop insurance program (see Coble, Dismukes, and Glauber (2007) for details of the SRA). As the insurance company, we calculate rates the same as RMA with one exception—we estimate the degree of heteroscedasticity (as reflected in the beta parameter in the above regression model) and correct the residuals accordingly, whereas the RMA assumes a constant CV. If our rates are higher than the RMA, we cede the policies as we believe them to be underpriced and

Downloaded from http://ajae.oxfordjournals.org/ at Pennsylvania State University on May 18, 2016

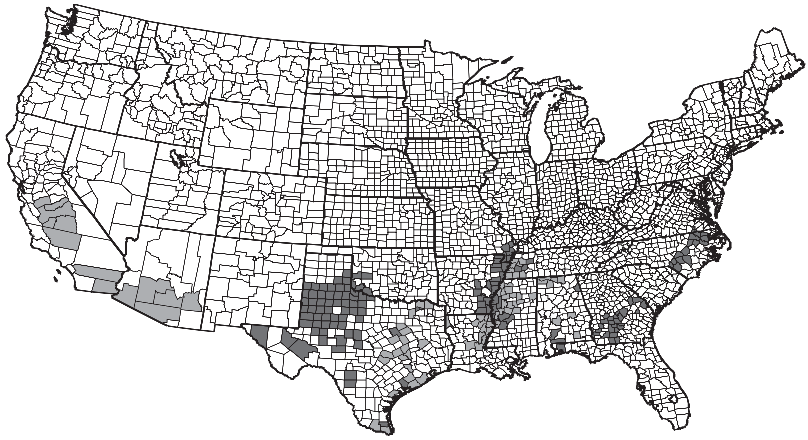

Figure 3. Test of the proportionality assumption for cotton

thus expect a loss. Conversely, we retain policies where our rates are lower than the RMA's rates as we believe they are overpriced and thus we expect to earn a profit. Given the realized yields, we can compare the loss ratio of the policies we retained to the loss ratio of the policies we ceded to RMA to derive a measure of the economic rents that could be lost by RMA in the event of incorrect heteroscedasticity assumptions. We do this at the 90% coverage level for each crop.

To operationalize this approach, we use the data from 1955–1992 and then calculate the rates for RMA and the insurance company (under the two different heteroscedasticity assumptions) for 1993. The underwriting gains and losses are then calculated using the actual 1993 yields. We then repeat this using the data from 1955–1993 and recover the underwriting gains and losses for each contract for 1994. We continue until we use the data from 1955–2006 and calculate the underwriting gains and losses for 2007. We then calculate the loss ratio of the policies retained and ceded over the fifteen years. Using bootstrapping techniques,[9] we calculate the statistical significance of differences under the null hypothesis that distinguishing the degree of heteroscedasticity does not give rise to economic rents. Under this null hypothesis, the ratio of loss ratios from ceded and retained policies should equal one, meaning the insurance company is indifferent in selecting the policies.

In the case described above, the RMA assumes a constant CV for all crop/county combinations. For interest and completeness we consider two additional cases and repeat the above out-of-sample evaluation of the gain in rating accuracy. In the first case, RMA assumes homoscedasticity for all crop/county combinations. In the second case, RMA tests both assumptions and uses the assumption that is not rejected. When both or neither assumption are rejected, a constant CV is assumed.

The results when RMA assumes a constant CV are presented in table 4. Similar results[10] were found for the other two cases when RMA assumes homoscedasticity and when RMA selects between these two forms of heteroscedasticity, respectively. For each crop/state combination in table 4 the following measures are reported. The "program loss ratio" is the loss ratio based on the base or current RMA methodology and the respective heteroscedasticity assumption. The "loss ratio of retained policies" is the loss ratio of the policies retained in the repeated comparison between the relevant alternative methodology

[9] We start with the pool of insurance rates for each state calculated using the RMA's methodology and the alternative methodology. The pool of these rates for each state contains n times 15 observations, where n is the number of counties in the state and 15 is the number of years that we repeat the process. From this pool, we select 10,000 bootstrapped samples of size 15. For each of these samples, we compare the rates from the two methodologies and group the policies into ceded and retained. Finally, we calculate the loss ratio of the policies retained and ceded. This process gives us the distribution of the loss ratio under the null hypothesis that the rates from the two methods are the same. By comparing the actual loss ratio with these 10,000 simulated values of the loss ratio under the null, we obtain the bootstrapped p-value.

[10] These results, not presented here, are available from the authors.

Downloaded from http://ajae.oxfordjournals.org/ at Pennsylvania State University on May 18, 2016

Table 4. Out-of-Sample Comparison of Data Determined (E) and Proportional (P) Rates for Restricted Spline Trend for Coverage Level 90 Percent

Crop	State	Program loss ratio	Loss ratio of ceded policies	Loss ratio of retained policies	Percent of policies retained	Ceded to retained ratio	Bootstrapped *P* value
Corn	IL	0.693	0.928	0.670	0.914	1.386	0.093
	IN	0.983	1.172	0.955	0.872	1.226	0.126
	IA	1.355	2.499	1.197	0.879	2.089	0.047
	MN	1.496	2.458	1.399	0.908	1.758	0.016
	MO	1.004	1.463	0.984	0.959	1.490	0.011
	OH	1.518	1.658	1.065	0.854	1.557	0.078
	WI	1.545	1.892	1.318	0.604	1.436	0.043
Soybeans	IL	1.427	1.978	1.120	0.641	1.764	0.092
	IN	1.338	1.482	1.275	0.698	1.162	0.145
	IA	2.206	2.385	2.142	0.738	1.113	0.106
	MN	3.173	3.522	3.078	0.785	1.145	0.055
	MO	1.059	1.157	1.011	0.673	1.142	0.142
	OH	2.471	2.904	2.325	0.750	1.249	0.114
	WI	2.189	2.892	2.099	0.886	1.378	0.034
Cotton	AL	7.971	15.598	5.197	0.733	2.996	0.061
	LA	6.400	8.251	5.432	0.656	1.516	0.051
	MS	2.332	3.505	1.917	0.739	1.818	0.051
	TN	2.091	2.995	1.457	0.588	2.058	0.070

Note: Table 4 presents the results of an out-of-sample test of the hypothesis that significant economic rents can be lost by RMA as a result of incorrect assumptions about heteroscedasticity. The results support the hypothesis.

and the RMA methodology. The "loss ratio of ceded policies" is the loss ratio of the policies ceded. The "ceded/retained ratio" is obtained by dividing the loss ratios of ceded and retained policies. The "bootstrapped *p*-value" provides the probability that the alternative methodology, when the degree of heteroscedasticity is estimated, is less accurate than the RMA methodology. The percentage of retained policies is the percentage of policies retained under the alternative methodology.

The results in table 4 support the hypothesis that significant economic rents can be lost by RMA as a result of incorrect assumptions about heteroscedasticity. The "ceded/retained ratios" for all crop/state combinations are greater than one and the bootstrapped *p*-values show statistical significance in the majority of crop/state combinations.

Conclusions

Based on the results of this analysis, we conclude that one should not arbitrarily impose proportional heteroscedasticity or homoscedasticity in crop risk simulations that may be used for insurance rate calculations. Our results show that researchers may fail to reject either of the commonly used heteroscedasticity assumptions.

We also consider allowing the data to determine the form of heteroscedasticity. Admittedly, these estimates are limited by the same small samples as the tests of standard assumptions. This led us to also consider an out-of-sample simulation that showed the empirically estimated heteroscedasticity model is superior in the majority of cases.

Our results have important practical implications for the RMA of the USDA in determining rates for area-based yield and revenue insurance programs. The agency must rate thousands of county-crop programs each year. Our work shows that assuming proportional heteroscedasticity or homoscedasticity could lead to selection problems across counties and crops and therefore give rise to poor program performance. However, the standard RMA assumption is most often rejected in marginal production regions where less liability exposure occurs. Ultimately, our results suggest that rates can be quite sensitive to maintained heteroscedasticity assumptions. We therefore recommend allowing the data to select the most appropriate specification of heteroscedasticity. USDA/RMA adopted these recommendations for rating

Downloaded from http://ajae.oxfordjournals.org/ at Pennsylvania State University on May 18, 2016

area products beginning in the 2009 crop year.

References

Coble, K. H., R. Dismukes, and J. W. Glauber, 2007. Private Crop Insurers and the Reinsurance Fund Allocation Decision. *American Journal of Agricultural Economics* 89(August): 582–595.

Coble, K. H., R. Heifner, and M. Zuniga. 2000. Implications for Crop Yield and Revenue Insurance for Produces Hedging. *Journal of Agricultural and Resource Economics* 25: 539–551.

Deng, X. H., B. J. Barnett, and D. V. Vedenov. 2007. Is There a Viable Market for Area-Based Crop Insurance? *American Journal of Agricultural Economics* 89 (2): 508–519.

Fackler, P., D. L. Young, and G. A. Carlson. 1993. Estimates of Trend and Variability Patterns in U.S. Crop Yield Series. Unpublished manuscript, Department of Agricultural and Resource Economics, North Carolina State University.

Glauber, J. W. Crop Insurance Reconsidered. 2004. *American Journal of Agricultural Economics* 86 (5): 1179–1195.

Goodwin, B. K., and A. P. Ker. 1998. Nonparametric Estimation of Crop Yield Distributions: Implications for Rating Group-Risk Crop Insurance Contracts. *American Journal of Agricultural Economics* 80(February): 139–153.

Halcrow, H. G. 1949. Actuarial Structures for Crop Insurance. *Journal of Farm Economics*, 31(August): 418–443.

Harri, A., K. H. Coble, C. Erdem, and T. O. Knight. 2009. Crop Yield Normality: A Reconciliation of Previous Research. *Review of Agricultural Economics* 31(1): 163–182.

Just, R. E. and R. D. Pope. 1978. Stochastic Specification of Production Functions and Economic Implications, *Journal of Econometrics* 7: 67–86.

Just, R. E., and Q. Weninger. 1999. Are Crop Yields Normally Distributed? *American Journal of Agricultural Economics* 81(May): 287–304.

Ker, A. P., and K. H. Coble. 2003. Modeling Conditional Yield Distributions. *American Journal of Agricultural Economics* 85: 291–304.

Ker, A, P. and B. K Goodwin. 2000. Nonparametric Estimation of Crop Insurance Rates Revisited. *American Journal of Agricultural Economics* 82 (2): 463–478.

Miranda, M. J. 1991. Area-Yield Crop Insurance Reconsidered. *American Journal of Agricultural Economics* 73(May): 233–242.

Miranda, M. J., and J. W. Glauber. 1997. Systemic Risk, Reinsurance and the Failure of Crop Insurance Markets. *American Journal of Agricultural Economics* 79: 206–215.

Moss, C. B., and J. S. Shonkwiler. 1993. Estimating Yield Distribution with a Stochastic Trend and Nonnormal Errors. *American Journal of Agricultural Economics* 75(November): 1056–1062.

Norwood, B., M. C. Roberts, and J. L. Lusk. 2004. Ranking Crop Yield Models Using Out-of-Sample Likelihood Functions. *American Journal of Agricultural Economics* 86(November): 1032–1043.

Ozaki, V., B. K. Goodwin, and R. Shirota. 2008. Parametric and Nonparametric Statistical Modeling of Crop Yield: Implications for Pricing Crop Insurance Contracts. *Applied Economics* 40: 1151–1164.

Ramirez, O. A. 1997. Estimation and Use of a Multivariate Parametric Model for Simulating, Heteroskedastic, Correlated, Nonnormal Random Variables: The Case of Corn Belt Corn, Soybean and Wheat Yields. *American Journal of Agricultural Economics* 79(February): 191–205.

Ramirez, O. A., S. Misra, and J. Field. 2003. Crop-Yield Distributions Revisited. *American Journal of Agricultural Economics* 85(February): 108–120.

Ramirez, O. A., T. McDonald. Ranking Crop Yield Models: 2006. A Comment. *American Journal of Agricultural Economics* 88(November): 1105–1110.

Skees, I. R., I. R. Black, and B. I. Barnett. 1997. Designing and Rating an Area Yield Crop Insurance Contract. *American Journal of Agricultural Economics* 79(May): 430–438.

Ozaki, V. A., S. K. Ghosh, B. K. Goodwin, R. Shirota. 2008. Spatio-Temporal Modeling of Agricultural Yield Data with an Application to Pricing Crop Insurance Contracts. *American Journal of Agricultural Economics* 90(4): 951–961.

Woodard, J. D., B. J. Sherrick, and G. D. Schnitkey. 2009. Crop Insurance Ratemaking under Trending Liabilities. Unpublished working paper, University of Illinois at Urbana–Champaign.

Downloaded from http://ajae.oxfordjournals.org/ at Pennsylvania State University on May 18, 2016