Rating Area-yield Crop Insurance Contracts Using Bayesian Model Averaging and Mixture Models

Alan P. Ker, University of Guelph, aker@uoguelph.ca
Tor Tolhurst, University of Guelph, ttolhurs@uoguelph.ca
Yong Liu, University of Guelph, yliu17@uoguelph.ca

Selected Paper prepared for presentation at the 2015 Agricultural & Applied Economics Association and Western Agricultural Economics Association Annual Meeting, San Francisco, CA, July 26-28

Abstract

The Agricultural Act of 2014 solidified insurance as the cornerstone of U.S. agricultural policy. The Congressional Budget Office (2014) estimates this Act will increase spending on agricultural insurance programs by $5.7 billion to a total of $89.8 billion over the next decade. In light of the sizable resources directed toward these programs, accurate rating of insurance contracts is of utmost importance to producers, private insurance companies, and the federal government. Unlike most forms of insurance -- where sufficient information exists to accurately estimate the probability and magnitude of losses (i.e. the underlying density) -- agricultural insurance is plagued by a paucity of spatially correlated data. A novel interpretation of Bayesian Model Averaging is used to estimate a set of possibly similar densities that offers greater efficiency if the set of densities are similar while seemingly not losing any if the set of densities are dissimilar. Simulations indicate finite sample performance -- in particular small sample performance -- is quite promising. The proposed approach does not require knowledge of the form or extent of any possible similarities, is relatively easy to implement, admits correlated data, and can be used with either parametric or nonparametric estimators. We use the proposed approach to estimate U.S. crop insurance premium rates for area-type programs and develop a test to evaluate its efficacy. An out-of-sample game between private insurance companies and the federal government highlights the policy implications for a variety of crop-state combinations. We repeat the empirical analyses under reduced sample sizes given: (i) new programs will dramatically expand area-type insurance to crops and states that have significantly less historical data; and (ii) changes in technology could render some historical loss data no longer representative. Consistent with the simulation results, the performance of the proposed approach with respect to rating area-type insurance -- in particular small sample performance -- remains quite promising.

Some key words: rating crop insurance contracts, Bayesian model averaging, multiple density estimation, spatial correlation, small sample estimation

Federally regulated crop insurance programs have been part of U.S. agricultural policy for upwards of 80 years. Over the past twenty years, crop insurance has spread from traditional crops like corn, soybean, and wheat to specialty and biofuel crops. In 2014, total liabilities held by the federal government exceeded $108.5 billion. The Agricultural Act of 2014, commonly referred to as the farm bill, solidified publicly subsidized insurance as the cornerstone of U.S. agricultural policy and the primary venue to funnel monies to the agricultural sector. The Congressional Budget Office (2014) estimates this Act will increase spending on agricultural insurance programs by $5.7 billion to a total of $89.8 billion for the next decade.

The federal crop insurance program is administered by the United States Department of Agriculture's Risk Management Agency (RMA). RMA operates a number of programs at the unit (sub-farm), farm and area level with additional programs being developed as a result of the latest farm bill. Of particular note is the introduction of the Supplemental Coverage Option (SCO) which allows producers to "top-up" their farm or unit level coverage with area-type coverage. RMA is responsible for a number of aspects of operating the program but most notably setting the premium rates. Interestingly, RMA is not the delivery agent to the producers; that is done by private insurance companies through a reinsurance agreement, which also stipulates how policies can be allocated and underwriting gains/losses shared.

The actuarially fair premium rate, π, is defined as the expected loss divided by total liability. That is, defining the random variable crop yield as Y, the actuarially fair premium rate for insurance coverage below a yield guarantee, denoted y_G, is:

$$\pi = \frac{1}{y_G} \int_0^{y_G} (y_G - y) f_Y(y|I) dy \tag{1}$$

where the conditional yield density $f_Y(y|I)$ needs to be estimated and I is the set of information at time of rating. Unlike most forms of insurance in which sufficient information exists to accurately estimate the probability and magnitude of losses (i.e. the underlying density), agricultural insurance is plagued by a paucity of spatially correlated data. The standard approach in the literature is to take a series of historical yields for the area of interest and estimate the temporal process, if necessary adjust the residuals for heteroscedasticity, and then estimate the conditional yield density of interest using the prediction and (adjusted) residuals. Rates are derived using the estimated density and equation (1).

Deterministic and stochastic approaches have been considered in estimating the temporal process (i.e. technological change) of yields. Deterministic approaches have dominated the literature and include a simple linear trend, two-knot linear spline (Skees and Reed, 1986), and polynomial trend (Just and Weninger, 1999). Stochastic approaches include the Kalman filter (Kaylen and Koroma, 1991) and ARIMA(p, d, q) (Goodwin and Ker, 1998). More recently, Tolhurst and Ker (2015) used mixtures to allow for heterogeneous rates of technological change across subpopulations of the yield distribution by estimating unique temporal processes for each component of the mixture. With the exception of Just and Weninger (1999) and Harri et al. (2011), heteroskedasticity has received little attention in the literature. A wide variety of density estimation approaches have been proposed. Parametric specifications include the Normal (Botts and Boles, 1958), Gamma (Gallagher, 1987), Beta (Nelson and Preckel, 1989), Logistic (Atwood, Shaik, and Watts, 2003) and Weibull distributions (Sherrick et al., 2004). Other approaches include the inverse sine transformation method (Moss and Shonkwiler, 1993; Ramirez, 1997), maximum entropy (Stochs and LaFrance, 2004; Wu and Zhang, 2012; Tack, Harri, and Coble, 2012) and Normal mixtures (Ker, 1996; Goodwin, Roberts, and Coble, 2000; Woodard and Sherrick, 2011; Tolhurst and Ker, 2015). Goodwin and Ker (1998) proposed nonparametric kernel density methods while a semi-parametric approach was forwarded by Ker and Coble (2003).

The above approaches do not make use of abundant -- albeit spatially correlated -- extraneous yield data in estimating the density for an area of interest. That is, while there exists historical corn yield data for all counties in Iowa, when estimating the density for say Adams county Iowa, only the yield data from Adams county is used.[1] The conditional yield densities of neighbouring counties are likely similar in shape given yields are highly influenced by climate, production practices, and soil type (all of which should be relatively homogeneous across the 99 Iowa counties). The purpose of this manuscript is to propose an approach to estimate a set of *possibly* similar yield densities that offers greater efficiency if the set of densities are similar while seemingly not losing any if the set of densities are dissimilar. The methodology must not require knowledge as to the form or extent of any similarities, accommodate correlated data, and maintain its performance in small samples (especially important given the pending widespread introduction of the SCO). Bayesian Model Averaging (BMA) -- simply a weighted average over a set of candidate models -- meets these criteria.

Typical applications of BMA use the *same* data to estimate alternative models (i.e. parametric forms) to account for model uncertainty. However, this is not necessarily the only interpretation of BMA: we could exploit possible similarities by forming our set of candidate models using estimates from *different* data sources (i.e. yields from different counties). To this end, we could first estimate using just its own yield data each county-level conditional yield density. The BMA estimate for county i would necessarily be a weighted average of these county-specific estimates where the weights would be derived by evaluating the likelihood of each estimate using only county i yield data. The candidate models would be the same for each county, but the weights would differ. This interpretation of BMA is both unique and statistically valid. Note: (i) yield realizations from not only county i (as is typical) but all other counties are used to recover the BMA estimate; (ii) assumptions about the extent or form of similarities are not needed; (iii) the weighting scheme is data-driven; (iv) estimated county densities that are more probable to produce county i yield data get more weight; (v) the estimated density from county i necessarily receives the highest weight as the parameter estimates for that county are the maximum likelihood estimates; and (vi) we do not need

[1]It is worth noting that after the initial rates are recovered, RMA does undertake credibility weighting which smooths the rates across the various counties. The justification for this comes from Stein's Paradox. Because the RMA's loss function is not based on a single county but rather all the counties, premium rates are shrunk toward the mean. The approach proposed here is to improve the initial density estimate of each county and thereby improve the accuracy of the premium rates before the credibility weighting takes place. As a result, credibility weighting does not negate the use of an improved density estimate nor does an improved density estimate negate the use of credibility weighting.

to explicitly account for spatial correlation. This last point is especially noteworthy. When data is pooled prior to estimation, spatial correlation can have (depending on the degree of correlation) noticeable effects on the variance of the pooled data. As a result, premium rates derived from density estimates based on spatially correlated pooled data can be biased. The proposed BMA approach mitigates spatial correlation as data is not pooled prior to estimation.

Most approaches in the literature do not attempt to use extraneous yield data in estimating the density for an area of interest, although there are some exceptions. Goodwin and Ker (1998) pool contiguous counties to increase the number of observations but the pooling is subject to spatial correlation issues and the weighting is ad hoc. In a more sophisticated pooling scheme, Annan et al. (2014) employed distributional tests and correlation estimates in an attempt to pool uncorrelated data from similar yield densities. While Ker (2002) derived an estimator of possibly similar densities, this approach is only valid with nonparametric estimators. Ozaki and Silva (2009) use a skewed-normal multivariate conditional yield distribution but assume zero spatial correlation beyond contiguous counties. In contrast to the above, the proposed BMA approach does not require use of distributional tests which tend to have low power, does not require exclusion of yield data that may be spatially correlated, and recovers estimates of appropriate weights for each county. Interestingly, the estimated (data-driven) weights reflect the relative similarity of yield densities from different counties to the county of interest where, as shown in the empirical results, there tends to be a surprising amount of weight placed on counties outside their own crop-reporting district or even state.

Bayesian Model Averaging

Model averaging, including both frequentist model averaging (Hjort and Claeskens, 2003) and Bayesian model averaging, is a popular technique to take model uncertainty into consideration by combining estimates across different models. Early work includes Roberts (1965) who proposed a distribution model which combines the opinions of two experts, as well as Leamer and Leamer (1978), who contributed the basic paradigm for BMA by suggesting a Bayesian approach to calculate the posterior probabilities of all competing models. In a review article, Kass and Raftery (1995) discussed the costs of ignoring model uncertainty and provided several important techniques for computing Bayes factors (a core ingredient of BMA). Madigan and Raftery (1994) presented a BMA approach for model selection and uncertainty in graphical models. BMA has demonstrated improved predictive performance in a wide variety statistical models including linear regression (Raftery, Madigan, and Hoeting, 1997), generalized linear models (Raftery, 1996), and survival analysis (Volinsky et al., 1997). In the agricultural economics literature, BMA has been used to estimate price changes by averaging a set of qualitative forecasts (Dorfman, 1998) and optimal hedge ratios by averaging across a set of hedge ratios conditional on different models (Dorfman and Sanders, 2006).

Draper (1995) accommodates model uncertainty based on the idea of expanding models: starting with a single "best" model indicated by the data, expanding the model space to accommodate other models, and finally averaging over the expanded model space. Our approach is similar in that we consider the crop yield model estimated following the standard approach (i.e. using data exclusively from the area of interest) to be the "best" model, while the models estimated using extraneous county data form the set of expanded models. In fact, our motivation is not explicitly model uncertainty; instead it is to exploit possible but unknown similarities amongst a set of densities to increase estimation efficiency.

Hoeting et al. (1999) provides a general guidance for the implementation of BMA, including several typical methods and related software. The general framework of BMA is as follows. Suppose there are models M_1, M_2, …, M_J to be considered and Δ is the quantity of interest. The posterior distribution of Δ given data D is

$$\mathrm{pr}(\Delta \mid D) = \sum_{i=1}^{J} \mathrm{pr}(\Delta \mid M_i, D)\mathrm{pr}(M_i \mid D), \tag{2}$$

where $\mathrm{pr}(\Delta \mid M_i, D)$ is the posterior distribution of Δ under model M_i, and $\mathrm{pr}(M_i \mid D)$ is the posterior model probability of model M_i. The posterior distribution, $\mathrm{pr}(\Delta \mid D)$, is an average of the posterior distributions under each of the models considered weighted by their posterior model probability. The posterior probability of model M_i is given as

$$\mathrm{pr}(M_i \mid D) = \frac{\mathrm{pr}(D \mid M_i)\mathrm{pr}(M_i)}{\sum_{i=1}^{J} \mathrm{pr}(D \mid M_i)\mathrm{pr}(M_i)}, \tag{3}$$

and

$$\mathrm{pr}(D \mid M_i) = \int \mathrm{pr}(D \mid \theta_i, M_i)\mathrm{pr}(\theta_i \mid M_i)\; d\theta_i, \tag{4}$$

where θ_i is the vector of parameters of model M_i, $\mathrm{pr}(\theta_i \mid M_i)$ is the prior for θ_i under model M_i, and $\mathrm{pr}(M_i)$ is the prior probability of model M_i.

Although BMA can give better predictive performance than using any single model (e.g. Madigan and Raftery, 1994), there are two major difficulties in the implementation of BMA. First, the number of models in the summation of equation (2) can be so large as to make an exhaustive summation practically impossible. Second, because the computation generally involves high dimensional integrals the posterior model probabilities $\mathrm{pr}(M_i \mid D)$ are difficult to compute. In this manuscript we restrict the number of models to the number of counties in the data thus alleviating the first difficulty. The second difficulty is handled using normal mixtures as in Tolhurst and Ker (2015) to model crop yields.

In the context of a normal mixture model the integral of the equation (4) can be approximated by the BIC (Dasgupta and Raftery, 1998). A Laplace approximation to the integral gives:

$$\ln \mathrm{pr}(D \mid M_i) = \ln \mathrm{pr}(D \mid \theta_i, M_i) - k \cdot \ln n. \tag{5}$$

Combined with the definition of BIC, $\mathrm{pr}(D \mid M_i)$ can be computed as $\mathrm{pr}(D \mid M_i) = \exp\{-\frac{1}{2}\mathrm{BIC}_i\}$ where BIC_i is the BIC value of model M_i. Assuming each model has equal prior model probability, the weights for each model considered, $\mathrm{pr}(M_i \mid D)$, are given by:

$$\mathrm{pr}(M_i \mid D) = \frac{\exp\Big\{-\frac{1}{2}\mathrm{BIC}_i\Big\}}{\sum_{i=1}^{J}\exp\Big\{-\frac{1}{2}\mathrm{BIC}_i\Big\}}, \tag{6}$$

and then, where $\hat{f}_i$ is the estimate of each model, the final estimate $\tilde{f}_{\mathrm{BMA}}$, is:

$$\tilde{f}_{\mathrm{BMA}} = \sum_{i=1}^{J}\mathrm{pr}(M_i \mid y)\hat{f}_i. \tag{7}$$

In our setting consider a set of Q crop yield densities with sample realizations $\{y_{11}, \ldots, y_{1n_1}, \ldots, y_{Q1}, \ldots, y_{Qn_Q}\}$. The standard density estimates, $\hat{f}_1, \ldots, \hat{f}_Q$, comprise the model space and use sample data exclusively from their own county and are estimated with the Expectations-Maximization (EM) algorithm as is commonly done for mixture models.[2] Then the BMA density estimate for county i is:

$$\tilde{f}_i = \sum_{j=1}^{Q}\omega_j^i\hat{f}_j \quad \text{where} \quad \omega_j^i = \frac{\exp\Big\{-\frac{1}{2}\mathrm{BIC}_j^i\Big\}}{\sum_{q=1}^{Q}\exp\Big\{-\frac{1}{2}\mathrm{BIC}_q^i\Big\}} \tag{8}$$

BIC_j^i is evaluated using the yield realizations from county i at $\hat{f}_j$, the standard density estimate from county j. The weights necessarily sum to one. To provide some intuition, consider two extreme scenarios: very dissimilar and identical densities. In the dissimilar case, BIC_j^i will be high unless $j = i$ and thus $\omega_j^i \approx 0$ for $i \neq j$. That is, the BMA estimate for county i will be close to, or even equal to, the standard density estimate for county i. Conversely, in the identical case, the BIC_j^i and ω_j^i will be relatively constant $\forall j \neq i$. Note, ω_i^i receives the maximum weight in $\tilde{f}_i$ because the likelihood is necessarily maximized (and BIC is minimized) at the parameters associated with the standard density estimate.[3]

Finite Sample Simulations

While BMA is certainly not new, our application is unique: we use BMA not to explicitly account for model uncertainty, but to exploit possible but unknown similarities amongst a set of densities to increase estimation efficiency. To that end, our set of candidate models or model space is not alternative parametric forms estimated using the same data but rather the same model estimated using alternative sources of

---

[2] While the EM algorithm exhibits monotonic convergence and admissible estimates when the initial values are within the admissible range, the likelihood function is unbounded because it goes to infinity when one of its component variances goes to zero (Karlis and Xekalaki, 2003; Chen and Li, 2009). Therefore, to ensure the admissibility of its estimates, we initialize the algorithm over a number (between 10-25) of different sets of starting values and use the penalized likelihood of Chen and Li (2009) to choose the "best" estimate over the different initializations. For a M component mixture, their penalized likelihood is defined as $pl_n(\lambda, \mu, \sigma) = l_n(\lambda, \mu, \sigma) + p_n(\sigma_1) + \cdots + p_n(\sigma_M) + p_n(\lambda_1) + \cdots + p_n(\lambda_{M-1})$ where the penalty on the component variance is $p_n(\sigma) = -\{s_n^2/\sigma^2 + \log(\sigma^2/s_n^2)\}$ where s_n^2 is the sample variance (i.e. $s_n^2 = 1/n\sum_{i=1}^{n}(X_i - \bar{X})^2$ with $\bar{X} = 1/n\sum_{i=1}^{n}X_i$) and the penalty on the $M-1$ mixture weights are given by $p_n(\lambda) = \log(1 - |1 - 2\lambda|)$. Intuitively $p_n(\sigma)$ is maximized at $p_n(\sigma){=}s_n^2$, which prevents under estimation of σ, while $p_n(\lambda)$ is maximized at the boundaries when either $\lambda = 0$ or $\lambda = 1$. These penalties were chosen to depend on the data and satisfy a number of conditions presented in Chen and Li (2009) for $p_n(\sigma)$ and Li, Chen, and Marriott (2009) for $p_n(\lambda)$.

[3] It is worth noting that the BMA approach has a flavor similar to that of shrinkage estimation. Strictly speaking, shrinkage estimation moves each of the individual estimates towards the overall mean and thus they are bounded between the individual estimate and overall mean. This is not necessarily the case with the proposed BMA approach in that the BMA estimate for a county is not (and often it isn't) bounded between the overall mean and the individual estimate.

data. To get a sense of the performance of BMA to exploit possible but unknown similarities, we evaluate its performance for estimating a set of generic densities in a best and worst case scenario. The best case scenario assumes the densities are identical while the worst case scenario uses the first nine Marron and Wand (1992) test densities.[4] These densities, which represent a large variety of more or less realistic density shapes, are commonly employed to assess finite sample performance of density estimators (see Hjort and Glad, 1995; Jones, Linton, and Nielsen, 1995; Jones and Signorini, 1997, among others). The logic is that one can be reasonably assured the estimator will perform well in an empirical setting if the proposed estimator performs well across these dissimilar density shapes. For each test density, 500 samples of size $n = \{25, 50, 100, 500\}$ were taken. We use a mixture of normals where the number of mixtures is chosen by minimizing BIC (note the number of mixtures can vary across each sample and density).

The first simulation represents a worst case scenario: a situation where the set of densities are very dissimilar and the proposed BMA approach would not readily come to mind.[5] The Mean Integrated Squared Error (MISE) between the true and estimated densities and the average weight put on the standard density estimate are located in table 1 for each of the nine test densities. There are a number of interesting results: (i) MISE of the BMA is essentially equal to the MISE of the standard density estimate; (ii) MISE decreases as the sample size increases; (iii) average weight on the standard density estimate increases as sample size increases; and (iv) density 5 which is very different has the most weight on itself. Most importantly, the results in table 1 suggest the estimation error of the BMA approach is roughly equivalent to the standard density estimate when the underlying densities are quite dissimilar.

The second simulation, summarized in table 2, considers the ideal situation in which all densities are identical $N(0, 1)$. We consider various sizes of the set of densities to be estimated, $Q = \{2, 5, 10, 25\}$. We also include simulations with moderate (0.25) and strong (0.75) correlation between the samples across the Q densities. The results are as expected in a number of ways: (i) the MISE of the BMA is bounded above by the standard density estimate; (ii) the MISE of the BMA realizes efficiency gains as Q increases; (iii) correlation decreases the efficiency gains of the BMA approach as there is less information in the sample; (iv) the weight on the standard density estimate decreases as the sample size increases but quite slowly; (v) the weight on the standard density estimate decreases as Q increases; and (vi) as correlation increases the weight on the standard density estimate marginally decreases.

Overall, these simulations suggest the BMA approach is roughly equivalent to the standard density estimate when the underlying densities are quite different and can realize significant efficiency gains when they are similar. Also, the weight on the standard density estimate versus the other candidate models is as expected: approaches one when the underlying densities are very dissimilar and decreases dramatically when the densities are similar. Correlation is seen to have two relatively marginal effects. First, there is less

[4]The densities are: (i) standard normal; (ii) skewed unimodal; (iii) strongly skewed unimodal; (iv) kurtotic unimodal; (v) outlier; (vi) bimodal; (vii) separated bimodal; (viii) asymmetric bimodal; (ix) trimodal.

[5]Given the flexibility of the proposed approach, one can transform the samples to have equal mean and variance prior to estimation thereby using BMA and extraneous data to influence the estimated shape (moments ≥ 3) only. In this case, only the individual data would be used to estimate the mean and variance.

Table 1. Worst Case Scenario Finite Sample Simulation Results

Sample Size		MISE x1000		Weight (%) on Standard
		Standard Approach	BMA Approach	Density Estimate
$n = 25$	Density 1	12.15	10.32	58.81
	Density 2	26.54	27.81	90.42
	Density 3	253.32	253.32	100.00
	Density 4	232.32	235.39	72.33
	Density 5	563.71	564.08	97.62
	Density 6	31.38	29.88	51.65
	Density 7	142.77	148.41	79.07
	Density 8	33.84	33.34	59.40
	Density 9	39.25	38.36	54.13
	Average	148.36	148.99	73.71
$n = 50$	Density 1	5.91	5.55	73.57
	Density 2	19.10	19.73	95.97
	Density 3	207.89	207.89	100.00
	Density 4	220.81	224.59	84.28
	Density 5	101.90	101.90	100.00
	Density 6	25.96	25.20	72.13
	Density 7	19.60	19.45	99.30
	Density 8	28.15	28.21	80.37
	Density 9	30.53	30.74	76.22
	Average	73.31	73.70	86.87
$n = 100$	Density 1	3.07	2.82	90.92
	Density 2	9.56	9.87	99.07
	Density 3	152.69	152.69	100.00
	Density 4	210.91	212.30	92.15
	Density 5	30.89	30.89	100.00
	Density 6	10.54	11.45	87.48
	Density 7	8.49	8.49	100.00
	Density 8	16.81	17.52	93.60
	Density 9	14.13	14.67	91.19
	Average	50.79	51.19	94.93
$n = 500$	Density 1	0.76	0.76	100.00
	Density 2	1.63	1.63	100.00
	Density 3	66.88	66.88	100.00
	Density 4	20.83	20.83	100.00
	Density 5	5.56	5.56	100.00
	Density 6	1.60	1.59	98.68
	Density 7	1.70	1.70	100.00
	Density 8	3.23	3.23	100.00
	Density 9	5.59	5.59	98.60
	Average	11.98	11.97	99.70

Table 2. Best Case Scenario Finite Sample Simulation Results

		MISE x1000				Weight (%) on Standard Density Estimate			
		$n = 25$	$n = 50$	$n = 100$	$n = 500$	$n = 25$	$n = 50$	$n = 100$	$n = 500$
No Correlation									
Standard		11.74	6.32	3.34	0.78	–	–	–	–
Proposed	$Q = 2$	9.92	5.05	2.47	0.48	78.49	78.97	76.34	72.32
	$Q = 5$	7.84	3.76	1.70	0.32	49.61	49.46	46.03	43.80
	$Q = 10$	6.28	2.84	1.30	0.25	30.14	30.18	28.64	27.26
	$Q = 25$	4.69	2.10	0.96	0.19	14.36	14.09	13.69	12.69
Pooled	$Q = 2$	5.06	2.51	1.21	0.24	–	–	–	–
	$Q = 5$	1.99	0.92	0.48	0.10	–	–	–	–
	$Q = 10$	0.94	0.47	0.25	0.05	–	–	–	–
	$Q = 25$	0.38	0.18	0.09	0.02	–	–	–	–
Correlation = 0.25									
Standard		11.46	6.01	3.35	0.77	–	–	–	–
Proposed	$Q = 2$	9.85	4.92	2.24	0.48	76.43	76.98	75.65	72.18
	$Q = 5$	7.57	3.50	1.71	0.34	45.96	44.61	44.90	43.77
	$Q = 10$	6.18	2.81	1.39	0.28	27.42	26.86	26.82	26.51
	$Q = 25$	4.96	2.28	1.15	0.23	12.61	12.42	12.35	12.04
Pooled	$Q = 2$	6.25	2.81	1.40	0.29	–	–	–	–
	$Q = 5$	3.45	1.69	0.81	0.16	–	–	–	–
	$Q = 10$	2.59	1.29	0.59	0.12	–	–	–	–
	$Q = 25$	1.97	1.06	0.47	0.09	–	–	–	–
Correlation = 0.75									
Standard		11.50	6.01	3.50	0.75	–	–	–	–
Proposed	$Q = 2$	11.32	4.91	2.67	0.48	67.13	64.96	64.55	62.50
	$Q = 5$	9.74	4.09	2.21	0.40	33.48	32.19	31.69	30.92
	$Q = 10$	8.94	3.88	2.08	0.37	18.26	17.54	16.95	16.92
	$Q = 25$	8.27	3.71	2.00	0.35	7.86	7.41	7.14	7.08
Pooled	$Q = 2$	9.41	4.14	2.23	0.42	–	–	–	–
	$Q = 5$	8.04	3.65	2.00	0.38	–	–	–	–
	$Q = 10$	7.53	3.60	1.94	0.35	–	–	–	–
	$Q = 25$	7.27	3.56	1.90	0.34	–	–	–	–

information in correlated data relative to uncorrelated data which necessarily reduces the efficiency of any estimator. Second, specific to the BMA approach, positive correlation reduces the weight on the standard density estimate; though the net effect is negligible because the candidate models are more similar. This is in contrast to when the data are pooled prior to estimation where correlation could significantly reduce the variance of the estimated density, which in turn would have non-trivial implications for estimated tail probabilities and derived premium rates.

Whereas the previous two simulations evaluate the BMA approach over the entire distribution, crop insurance is exclusively concerned lower tail probabilities. We designed a third simulation to evaluate the efficacy of the proposed BMA approach using Mean Squared Error (MSE) from the "true" rate. For this simulation we use the heteroskedasticity adjusted residuals from the RMA's detrending methodology (detailed in the next section) to estimate the conditional yield density for each county-crop combination using nonparametric kernel methods. We then assume those densities are the "true" densities and sample, both correlated and uncorrelated, 500 samples of sizes 15, 20, 25, and 50 for each county-crop combination.[6] For each sample, rates are recovered for the RMA (empirical rate), standard, and BMA approaches. The standard and BMA approaches use the same normal mixture model as in the previous two simulations.

The data used in this simulation are the yield data used throughout the rest of the manuscript. We choose corn, soybean, cotton, and winter wheat as our crops. For corn and soybean, we use county level yield data from Illinois, Indiana, Iowa, Minnesota, Missouri, Ohio, and Wisconsin. For cotton, we choose Arkansas, Georgia, Louisiana, Mississippi, and Tennessee. Finally, for winter wheat we choose Illinois, Indiana, Kansas, Maryland, Michigan, Missouri, Ohio, Oklahoma, and Tennessee. These crop-state combinations were chosen on the basis of having (any) participation in area-yield programs in the 2014 insurance year and historical NASS yields available from 1955 (counties with incomplete yield histories were excluded). In total, we have data from 472, 469, 58, and 225 counties for corn, soybean, cotton and winter wheat respectively. Given each county represents a separate model, the model space for the BMA approach consists of 472, 469, 58, and 225 models for corn, soybean, cotton and winter wheat respectively. Specifically, this simulation is based on rates for the 2013 insurance year using data from 1955-2012.[7]

Table 3 reports the MSE for each of the RMA, standard, and BMA approaches for the correlated and uncorrelated samples. A number of results are worth noting: (i) results are consistent across crops and sample sizes; (ii) the RMA and standard rates have roughly identical MSEs (speaks to the flexibility of the normal mixture model); (iii) BMA leads to significant MSE decreases for all crops with the greatest being corn and cotton, which suggests their densities are more homogeneous across space than soybean and winter wheat; (iv) the RMA and standard approach are as expected roughly identical in the correlated and uncorrelated cases; and (v) correlation has a marginal effect on the efficiency of the BMA as less information is in the sample. While we have not broken down MSE into bias and variance in Table 3, as expected, the presence of spatial correlation does not increase bias as there is no pooling prior to estimation. Importantly, this simulation allowed us to compare the estimated rates from the three methodologies (absent trend estimation) to the true rate while isolating the effect of spatial correlation. We note that applying BMA to the trend estimation should increase any efficiency gains. In the next section we apply BMA to the

---

[6]Drawing a random sample from a kernel density estimate is relatively simple. Draw, with replacement, a realization from the original data and then perturb that by adding a draw from the Kernel (in our case the normal density) with mean 0 and standard deviation h where h is the smoothing parameter. To retrieve uncorrelated samples across the counties, we sample the initial realizations independently and perturb them. To get correlated samples across the counties, we sample a year and take the realization for all counties for that year and perturb them. Given h is relatively small compared to the standard deviation of the data, this maintains the majority of the correlation structure from the initial yield realizations in the samples.

[7]All data and code will be made available by the authors.

Table 3. Mean Squared Error of Estimated Premium Rates

	MSE x1000							
	No Correlation				*Spatial Correlation*			
	$n = 15$	$n = 20$	$n = 25$	$n = 50$	$n = 15$	$n = 20$	$n = 25$	$n = 50$
Corn								
Empirical	1.9370	1.4110	1.1245	0.5501	1.9496	1.4184	1.1614	0.5260
Standard	1.8876	1.3764	1.0996	0.5463	1.9031	1.3843	1.1329	0.5231
BMA	1.1812	0.8761	0.7137	0.3753	1.3271	0.9645	0.7947	0.3910
Soybean								
Empirical	1.0186	0.7515	0.5905	0.2900	1.0302	0.7226	0.5727	0.2762
Standard	0.9777	0.7219	0.5691	0.2841	0.9877	0.6958	0.5521	0.2705
BMA	0.7747	0.5765	0.4647	0.2421	0.8235	0.5842	0.4704	0.2397
Cotton								
Empirical	1.5961	1.1979	0.9477	0.4625	1.5389	1.2292	0.9307	0.4303
Standard	1.5438	1.1553	0.9162	0.4556	1.4848	1.1880	0.9002	0.4247
BMA	1.1043	0.8619	0.6784	0.3474	1.1216	0.9041	0.6883	0.3345
Winter Wheat								
Empirical	1.4043	1.0464	0.8145	0.4014	1.4071	1.0306	0.8339	0.3910
Standard	1.3556	1.0116	0.7888	0.3934	1.3614	0.9946	0.8061	0.3834
BMA	1.0523	0.7931	0.6286	0.3240	1.0950	0.8240	0.6739	0.3279

trend estimation as well by undertaking an out-of-sample rating game and developing a test to evaluate the relative efficacy of the BMA approach where the true densities and rates are unknown.

Rating Area-Type Insurance Contracts

The motivation for the proposed BMA approach is to improve the accuracy of the estimated premium rates for area-type contracts in the U.S. crop insurance program. Ultimately, the viability of any insurance program is conditional upon the actuarial methods used to develop premium rates.[8] Several articles have proposed systems to rate area yield designs (for example see Skees, Black, and Barnett, 1997; Ker and Goodwin, 2000). While these area programs have been relatively less popular in the past, this is expected change with the introduction of the SCO in the new farm bill. The SCO is expected to account for nearly 30% of the estimated increase in government crop insurance spending (Congressional Budget Office, 2014) and thus their actuarial methods, such as those proposed here, are of the utmost policy relevance.

An interesting feature of the U.S. crop insurance program is that the government uses intermediaries or private insurance companies to deliver the program to producers. While the government sets the premium rates for each insurance policy, insurance companies sell the contracts, conduct claim adjustments, and

[8]Area yield and revenue insurance has attracted significant attention in the agricultural economics literature as a means to avoid the problems of moral hazard, adverse selection, and the transaction costs associated with individual coverage crop insurance policies. A number of articles discuss the strengths and weaknesses of area-type programs in detail (e.g. Halcrow, 1949; Miranda, 1991; Bourgeon and Chambers, 2003; Glauber, 2013).

participate in the underwriting gains and losses of those contracts. An underwriting gain is realized if indemnity payments (claims) are less than premiums paid. Conversely, an underwriting loss is realized if indemnity payments are greater than premiums paid. The insurance companies share, asymmetrically, the underwriting gains and losses of the contracts sold with the government.

While the actual agreement between the private insurance companies and government is somewhat complex (and outlined in the Standard Reinsurance Agreement), it is structured such that insurance companies can either cede or retain the majority of the underwriting gains/losses of a policy it sells. Obviously, insurance companies attempt to predict which policies will return an underwriting gain and which will return an underwriting loss. That is, which policies are overpriced and which are underpriced. To do so, insurance companies could, and likely do, re-estimate the premium rates and compare their rates to those of the government. For those policies with rates larger than the government rates they cede back to the government as they believe them to be underpriced. Conversely, for those policies with rates lower than the government rates they retain them as they believe them to be overpriced. We can use this mechanism and act as an insurance company -- by adversely selecting against the government -- to evaluate the proposed BMA approach with respect to the current RMA methodology.

Out-of-Sample Rating Game

We conduct a repeated game of out-of-sample rating accuracy as developed by Ker and McGowan (2000) and used by Ker and Coble (2003), Racine and Ker (2006), Harri et al. (2011), Annan et al. (2014) and Tolhurst and Ker (2015). To this end, we estimate the 1994 premium rates using both the proposed methodology and the RMA methodology for each crop-county combination using data from 1955-1993 only. The decision rule is to cede those counties where the RMA estimated premium rate is lower than the estimated premium rate using the proposed BMA approach. We repeat this for $1994, \ldots, 2013$ using data only from $1955 - 1993, \ldots, 1955 - 2012$. Using actual realized yields we calculate the loss ratio (defined as total indemnities divided by total premiums) for the county-year combinations retained as well as the county-year combinations ceded. The loss ratio for a set of policies, denoted Ω, is defined as:

$$LossRatio_{\Omega} = \frac{\sum_{k \in \Omega} \max(0, \hat{y}_{G_k} - y_k)}{\sum_{k \in \Omega} \hat{\pi}_{RMA}} \tag{9}$$

where $\hat{\pi}_{RMA}$ is the RMA estimated premium rate (in bushels per acre), $\hat{y}_{G_k}$ is the yield guarantee (expected yield multiplied by the coverage level), and y_k is the actual realization. We chose the 90% coverage level because roughly 95% of the recent policies sold are at this coverage level. In addition, we calculate the percent of total county-year combinations retained using our methodology. To ascertain the statistically significance of our results, we use a randomization test. That is, we randomly select the same percent of contracts as those retained under the decision rule and calculate the loss ratio for the corresponding set. We repeat this 5000 times. The 5000 loss ratios represent the distribution of the loss ratio from the policies retained using our decision rule under the null that the decision rule is ineffective in identifying more profitable policies.

While our out-of-sample rating game uses all available historical data (1955-2013), this may be of limited practical use to RMA for two reasons. First, there is significantly less historical data in many places where area-yield rates are now required under the 2014 Agricultural Act. Second, it can be argued that given technological change -- specifically advances in seed technology that have changed the plants response to climate conditions -- yield losses from more than 20 years ago may not be representative despite corrections to the first two sample moments. Therefore, we evaluate the BMA approach for reduced sample sizes as well. That is, we repeat the out-of-sample rating game by assuming we only have 25, 20, and 15 years of historical data. For example, to recover the estimate for 2002 and using only 15 years of historical data, the rates are based on data from 1987-2001 only.

While BMA can be applied to both nonparametric (using the empirical likelihood) and parametric models, we follow Tolhurst and Ker (2015) by modeling crop yields using normal mixtures with embedded trend functions to account for potentially different rates of technological change in different components of the yield distribution.[9] For a vector of yields y indexed by time t:

$$y_t \sim \sum_{m=1}^{M} \lambda_m N(h_m(t), \sigma_m^2) \tag{10}$$

where the unknown parameter vectors λ_m, σ_m^2 and functions $h_m(t)$ are estimated with a maximum likelihood approach using the EM algorithm for the M components of the mixture. Tolhurst and Ker (2015) indicate the mixture model offers many advantages: it can approximate many of the distributional structures associated with conditional yield densities; embedding possibly unique trend functions within each mixture does not restrict the effect of technological developments to the first two moments of the yield distribution; and a mixture model exhibiting different rates of technological change in different components will lead to a non-constant variance with respect to time (heteroskedasticity is common in yield data). In the out-of-sample rating we assume a mixture of two normals for all crop-county combinations given BIC is minimized with two components for 91%, 97%, 95%, and 98% of the counties for corn, soybean, cotton, and winter wheat, respectively. We isolate the effect of the BMA by comparing it to both the RMA rating methodology and the standard approach (i.e. the normal mixture model with embedded trend functions using observed data exclusively from its own area of interest). Note, whereas BMA was applied to the parameters of the conditional density in the previous section's simulations, here BMA is applied to the those parameters as well as the parameters depicting the temporal process. RMA models the temporal process of yields with a two-knot robust linear spline with spatial and temporal priors on the knots. The estimated residuals are then adjusted for heteroskedasticity following Harri et al. (2011). The RMA rate is then empirical rate from the heteroskedasticity adjusted yields centered at the predicted yield, whereas the BMA approach rate is the BMA density rate (equation (1)) at the RMA's predicted yield.

[9]We would have preferred to apply BMA directly to RMA rating procedure but given the RMA approach (robust estimation and no assumption of distribution), it can not be applied given there is no likelihood.

Empirical Results

The simulated game results between the BMA approach and the RMA methodology are presented in tables 4, 5, and 6. Table 4 illustrates the number of counties, the percent of county-year combinations that experienced losses, the percent of contracts retained, the loss ratio of the contracts ceded (loss ratio government), the loss ratio of the contracts retained (loss ratio private) and finally the p-value from the randomization test. For the 28 state-crop combinations, economically significant rents can be generated by insurance companies in 27 of the 28 cases of which 20 are statistically significant. For brevity, table 4 presents only the p-values for the analysis repeated with reduced sample sizes (detailed tables are included in the online supplement).[10] With a reduced sample size of 25, we find economically significant rents can be generated by insurance companies in 26 of the 28 cases of which 23 are statistically significant. For reduced sample size of 20 and 15, we find economically significant rents can be generated by insurance companies in 26 and 28 of the 28 cases, respectively, all but one of which are statistically significant. Results for the BMA approach versus the standard approach are similar and presented in Appendix A.[11] Overall, these results are quite strong and suggest that insurance companies could make significant economic rents using the BMA methodology.

In table 6, we present the average weight on the standard density estimate over the 20 year out-of-sample rating game by state-crop combination for the BMA approach. Specifically, the cumulative weight for the first 1, 2, 3, 5, 20, 25, and 100 weights in the BMA weighting scheme. That is, the average cumulative weight of the first three BMA weights for Illinois corn is 57.8% and represents the total of the first three highest weights averaged over all 20 years of the rating game. Note, the BMA approach is re-estimated every year and so the counties that comprise the first three highest weights could -- but do not necessarily -- remain constant over the 20 years. We find a number of results worth noting: (i) necessarily, the standard density estimate has the maximum BMA weight; (ii) there is significant weight given to other estimates, particularly in small samples; (iii) counties outside the crop-reporting-district (CRD) receive significant weight; (iv) there is surprisingly high weight given to counties outside the state; (v) there is noteworthy variation in the weights across states within crops; (v) cotton with a relatively small number of total counties places more weight on the standard density estimate; and (v) as the number of years decreases the weight on other counties increases significantly.

In appendix B, we graphically show the spatial distribution of the average BMA weights for a randomly selected soybean county (Winneshiek, Iowa) for different sample sizes. Winneshiek county is in green, receives the highest weight, whereas all other counties are shades of blue with darker blue representing greater BMA weight. Two features are of worth noting: BMA weights are not clustered around the county of interest; and, as the sample size is reduced the weight on the extraneous counties is less concentrated. These features are also evident in the online supplement which contains similar graphs for randomly selected counties for corn, cotton, and winter wheat. We also summarize graphically the average weight on the number of counties by

[10] In the full sample for Tennessee cotton all contracts were either ceded or retained not allowing for a comparison.

[11] Detailed results for all out-of-sample games are provided in the online supplement.

Table 4. Out-of-Sample Rating Game Results (BMA versus RMA)

Crop-State	Number of Counties	Payouts (%)	Retained by Private (%)	Loss Ratio Government	Loss Ratio Private	p-value
Corn						
Illinois	82	24.5	87.1	1.014	0.649	0.0332**
Indiana	69	25.6	83.1	1.080	0.758	0.0538*
Iowa	99	16.1	73.8	0.614	0.347	0.0038***
Minnesota	57	16.1	59.6	0.181	0.145	0.1858
Missouri	48	36.8	82.7	0.861	0.765	0.2694
Ohio	63	21.5	83.7	1.075	0.676	0.0262**
Wisconsin	54	20.6	64.5	0.891	0.476	0.0034***
Soybean						
Illinois	91	20.4	44.0	1.068	0.594	0.0000***
Indiana	66	27.0	51.5	1.475	0.613	0.0000***
Iowa	98	21.6	40.2	0.955	0.786	0.0882*
Minnesota	63	21.7	51.0	0.721	0.415	0.0024***
Missouri	57	33.1	39.6	1.432	0.894	0.0012***
Ohio	52	23.3	67.3	1.173	0.719	0.0050***
Wisconsin	42	34.4	74.4	1.485	0.731	0.0000***
Cotton						
Arkansas	7	25.0	9.3	0.486	0.000	0.0000***
Georgia	20	37.8	43.0	0.809	0.436	0.0012***
Louisiana	6	38.3	26.7	2.303	1.194	0.0934*
Mississippi	15	29.0	8.7	1.435	0.552	0.0588*
Tennessee	10	24.5	0.0	0.758	–	–
Winter Wheat						
Illinois	10	26.0	31.5	1.201	0.232	0.0000***
Indiana	28	27.9	45.5	1.101	0.725	0.0296**
Kansas	71	33.9	55.4	1.607	1.194	0.0048***
Maryland	7	32.9	5.7	2.725	0.336	0.1470
Michigan	32	15.2	15.5	0.466	0.505	0.6000
Missouri	23	25.4	48.3	0.893	0.728	0.2474
Ohio	19	21.8	41.8	1.198	0.910	0.2510
Oklahoma	22	37.0	80.2	3.377	1.531	0.0002***
Tennessee	13	19.6	67.7	0.734	0.367	0.0590*

Note: Counties with incomplete yield histories are excluded. Statistical significance of lower private (government) loss ratios indicated by *-10% (†-10%), **-5% (††-5%), and*** -1% (†††-1%), respectively.

crop-sample size combination in appendix B. The figures are interesting in that they are relatively similar across the four crops.

Efficacy of Proposed BMA Approach

While the out-of-sample rating game replicates the environment between RMA and private insurance companies, the results -- surprisingly -- are not indicative of the relative efficiency for the methodologies being compared. A simple example can illuminate this counter-intuitive point. Let RMA use a rating methodology that is distributed uniformly around the true rate, denoted π, with bounds equal to $\pi \pm 0.005$ (and is therefore unbiased). Let the insurance company use a rating methodology that is distributed uniformly

Table 5. Out-of-Sample Rating Game Result Summary: p-values

	BMA versus the RMA			
	Full Sample	25 Years	20 Years	15 Years
Corn				
Illinois	0.0332**	0.0002***	0.0000***	0.0000***
Indiana	0.0538*	0.1512	0.0000***	0.0000***
Iowa	0.0038***	0.7146	0.7602	0.0000***
Minnesota	0.1858	0.0530*	0.0000***	0.0000***
Missouri	0.2694	0.0006***	0.0000***	0.0000***
Ohio	0.0262**	0.0000***	0.0000***	0.0000***
Wisconsin	0.0034***	0.0000***	0.0000***	0.0000***
Soybean				
Illinois	0.0000***	0.0000***	0.0000***	0.0000***
Indiana	0.0000***	0.0000***	0.0000***	0.0000***
Iowa	0.0882*	0.0082***	0.0022***	0.2498
Minnesota	0.0024***	0.0000***	0.0000***	0.0000***
Missouri	0.0012***	0.0000***	0.0000***	0.0000***
Ohio	0.0050***	0.0000***	0.0000***	0.0000***
Wisconsin	0.0000***	0.0000***	0.0000***	0.0000***
Cotton				
Arkansas	0.0000***	0.0000***	0.0016***	0.0154**
Georgia	0.0012***	0.0342**	0.0008***	0.0000***
Louisiana	0.0934*	0.0252**	0.0010***	0.0002***
Mississippi	0.0588*	0.0000***	0.0000***	0.0000***
Tennessee	n/a	0.0000***	0.0000***	0.0002***
Winter Wheat				
Illinois	0.0000***	0.0060***	0.0000***	0.0000***
Indiana	0.0296**	0.0002***	0.0000***	0.0000***
Kansas	0.0048***	0.0000***	0.0000***	0.0000***
Maryland	0.1470	0.3640	0.0084***	0.0034***
Michigan	0.6000	0.0422**	0.0000***	0.0000***
Missouri	0.2474	0.0012***	0.0000***	0.0062***
Ohio	0.2510	0.4244	0.0020***	0.0004***
Oklahoma	0.0002***	0.0000***	0.0000***	0.0016***
Tennessee	0.0590*	0.8652	0.5030	0.0008***

Note: Statistical significance of lower private (government) loss ratios indicated by $^{*}(^{\dagger})$-10%, $^{**}(^{\dagger\dagger})$-5%, and $^{***}(^{\dagger\dagger\dagger})$-1%.

around π with bounds equal to $\pi \pm 0.015$ (also unbiased but clearly less efficient). Assume the RMA and insurance company rate draws are independent. Now assume the RMA reveals its rate and it happens to be $\pi + .0025$. The likelihood that the insurance company has a rate below this, and *retains* the contract, is 62.5%. Conversely, if the RMA reveals its rate and it happens to be $\pi - .0025$ the likelihood that the insurance company has a rate above this, and *cedes* the contract, is 62.5%. In fact, the probability that the insurance company retains an underpriced contract or cedes an overpriced contract is easily shown to be bounded above by 50%. The fact that the insurance company can react to the RMA price is a distinct advantage and represents an implicit subsidy to the insurance company.

Table 6. BMA Weights (%) During Out-of-Sample Game

	Cumulative Weight on n Counties								Weight within	
	1	2	3	5	10	25	50	100	CRD	State
Corn (472 counties)										
Illinois	43.0	51.9	57.8	66.0	77.0	90.5	97.2	99.7	51.1	66.4
Indiana	38.5	46.7	52.2	60.1	71.9	87.1	95.7	99.5	46.8	59.2
Iowa	33.0	42.6	49.0	57.9	70.9	87.2	95.8	99.5	47.4	71.0
Minnesota	49.0	58.5	63.8	70.4	79.5	90.8	96.8	99.6	57.7	68.1
Missouri	56.2	64.9	70.0	77.1	86.9	96.3	99.3	99.9	62.6	80.9
Ohio	40.9	49.6	55.3	63.2	75.0	89.5	96.6	99.6	46.2	63.5
Wisconsin	49.5	58.7	64.7	72.2	82.5	93.5	98.1	99.8	59.1	69.5
Soybean (469 counties)										
Illinois	52.9	62.6	68.5	76.5	86.4	96.1	99.3	100.0	58.9	73.1
Indiana	42.0	50.3	56.0	63.7	75.7	91.1	98.0	99.9	45.4	60.3
Iowa	44.3	53.9	60.4	69.2	80.7	93.3	98.5	99.9	54.5	71.6
Minnesota	59.0	68.8	74.7	81.8	89.7	96.8	99.3	100.0	70.1	79.7
Missouri	57.3	66.5	72.0	79.1	88.5	97.2	99.7	100.0	65.8	84.3
Ohio	54.9	61.7	66.3	72.6	81.8	93.0	98.2	99.9	58.2	66.8
Wisconsin	60.4	71.9	77.8	85.7	94.5	99.4	99.9	100.0	77.1	85.9
Cotton (58 counties)										
Arkansas	50.8	67.2	75.0	85.0	96.3	99.9	100.0	100.0	69.8	77.7
Georgia	76.5	84.6	88.4	93.6	98.6	100.0	100.0	100.0	81.0	91.9
Louisiana	69.7	78.5	84.3	93.8	99.3	100.0	100.0	100.0	72.3	75.2
Mississippi	64.3	76.2	83.7	92.7	98.7	100.0	100.0	100.0	74.5	86.9
Tennessee	58.5	70.5	78.5	88.3	97.7	100.0	100.0	100.0	77.5	93.2
Winter Wheat (225 counties)										
Illinois	52.4	66.3	72.6	80.4	90.3	98.9	100.0	100.0	60.3	70.6
Indiana	60.7	72.4	79.1	86.8	94.3	99.4	100.0	100.0	65.7	75.2
Kansas	45.2	56.7	64.4	74.7	88.3	98.4	99.9	100.0	62.5	94.9
Maryland	61.5	72.9	79.0	85.7	94.0	99.4	100.0	100.0	65.4	68.9
Michigan	54.2	65.3	71.4	79.3	90.0	98.8	100.0	100.0	62.8	74.0
Missouri	69.8	77.2	81.8	87.6	93.9	98.8	99.9	100.0	76.6	82.1
Ohio	54.9	65.9	73.2	81.3	92.0	99.3	100.0	100.0	62.8	71.5
Oklahoma	48.9	62.3	70.3	79.5	90.8	98.7	99.9	100.0	64.7	86.7
Tennessee	66.0	78.2	84.7	91.8	96.5	99.5	100.0	100.0	78.3	89.8

We can account for this implicit subsidy by considering the gains the insurance company makes relative to the gains the RMA would make if their roles were reversed. To do so, we calculate the following metric

$$D = \frac{LR_{IC}/LR_{RMA}}{LR^*_{RMA}/LR^*_{IC}} \tag{11}$$

where LR_{IC} and LR_{RMA} are the loss ratios where the insurance company adversely selects against the RMA, while LR^*_{IC} and LR^*_{RMA} are the loss ratios where RMA adversely selects against the insurance company. It is necessary to restrict the insurance company and government to retain half the contracts. To reiterate, D measures the adverse selection gains of the insurance company relative to the adverse selection gains of the

Table 7. Efficacy Test Results: p-values

	BMA versus RMA			
	Full Sample	25 Years	20 Years	15 Years
Corn				
Illinois	0.2517	0.2517	0.0059**	0.0059**
Indiana	0.1316	0.0207**	0.0013***	0.0000***
Iowa	0.1316	0.0207**	0.0059**	0.0059**
Minnesota	0.0577	0.1316	0.0002***	0.0013***
Missouri	0.4119	0.0059**	0.0013***	0.0000***
Ohio	0.0059**	0.0577	0.0059**	0.0000***
Wisconsin	0.0059**	0.0059**	0.0013***	0.0059**
Soybean				
Illinois	0.0013***	0.0059**	0.0207**	0.4119
Indiana	0.0013***	0.0013***	0.0013***	0.0059**
Iowa	0.0207**	0.0000***	0.0059**	0.0013***
Minnesota	0.1316	0.0013***	0.0013***	0.0000***
Missouri	0.0577	0.0059**	0.0059**	0.1316
Ohio	0.1316	0.0002***	0.0207**	0.0207**
Wisconsin	0.0577	0.1316	0.0207**	0.0207**
Cotton				
Arkansas	0.0059**	0.0000***	0.0002***	0.0002***
Georgia	0.1316	0.0059**	0.0000***	0.0000***
Louisiana	0.4119	0.0013***	0.0002***	0.0000***
Mississippi	0.1316	0.0013***	0.0000***	0.0002***
Tennessee	n/a	0.0000***	0.0059**	0.0013***
Winter Wheat				
Illinois	0.0013***	0.0013***	0.0013***	0.0002***
Indiana	0.2517	0.0207**	0.0059**	0.0059**
Kansas	0.5881	0.1316	0.0207**	0.0059**
Maryland	0.1316	0.0059**	0.0577	0.0002***
Michigan	0.1316	0.0059**	0.0059**	0.0013***
Missouri	0.7483	0.0207**	0.0059**	0.0059**
Ohio	0.0002***	0.0002***	0.0002***	0.0013***

Note: Statistical significance of lower private (government) loss ratios indicated by *(†)-10%, **(††)-5%, and ***(†††)-1%.

RMA.[12] Under the null that the rates are equally efficient in terms of variance, D will be distributed with median equal one. We calculate D for each of the crop-state combination for each of the 20 years. Denote C^* as the random variable depicting the number of values of $D > 1$ in the 20 years for each crop-state combination or crop overall. Note, C^* is binomial with probability 0.5 and $n = 20$ under the null that the methodologies are equally accurate.[13]

[12]In the online supplement, we show that if the insurance company rate is more (less) efficient (as measure by variance around the true rate) than the RMA rate, D will have median greater (less) than one.

[13]This assumes D is independent across time. We conducted two tests on D: (i) a two-sided Durbin-Watson test for auto-correlation and (ii) a t-test on an added autoregressive term. In every case -- including the RMA approach, county-specific approach, and over all different sample sizes -- the assumption of independence in D_t is never rejected at the 5% significance level in either test. This is not surprising because the year-to-year variation in loss ratios is dominated by year-to-year variation in losses (and thus yields) which are independent over time.

The p-values for test-statistic C^* are presented in table 7. A low p-value corresponds to BMA rates being more efficient whereas a high p-value corresponds to the RMA rates being more efficient (two-sided test). Not surprisingly, when the implicit subsidy is accounted for, the superiority of the BMA approach, relative to the RMA methodology, is reduced. In eight (as opposed to 20) of the 28 full sample cases can we conclude that the BMA approach produces more efficient rates. When the sample is reduced to 25, 20 and 15 years that number, not surprisingly, increases to 21, 27, and 26 cases respectfully. It is worth noting that the BMA approach does perform worse ($p > .5$) than the RMA methodology in two of the 112 cases (28 crop-state combinations times four sample sizes) but neither are statistically significant. Appendix A contains the results for the BMA approach versus the standard approach: results are similarly dominant for the BMA approach. Overall, the results are strongly favorable for the BMA approach and coincide with the finite sample simulations in section 3 which also yielded strong support for the BMA approach.

Conclusions

The 2014 farm bill solidified insurance as the dominant domestic farm policy and thus accurate pricing of insurance contracts is of utmost importance. A novel interpretation of Bayesian model averaging is used to estimate a set of possibly similar densities that offers greater efficiency if the set of densities are similar while seemingly not losing any if the set of densities are dissimilar. This BMA approach is relatively easy to implement, does not require knowledge of the form or extent of any possible similarities, admits correlated data, and can be used with either parametric or nonparametric estimators.

A number of simulations were performed to evaluate the efficacy of the proposed BMA approach. First, standard simulations were undertaken using common test densities for both worst and best case scenarios. These simulations suggest that the BMA does no worse than the standard density estimate when the underlying densities are quite different and can realize significant efficiency gains when the set of densities are identical. The weight on the standard density estimate versus the other candidate models approaches one when the underlying densities are very dissimilar and decreases dramatically when the densities are identical. In this simulation, correlation is shown to only effect the BMA through the weights (and does so marginally). Second, simulations were performed for hypothetical yield densities for corn, soybean, cotton, and winter wheat. The mean squared error of the rates from the proposed versus competing methodologies were recovered under both spatially correlated and uncorrelated scenarios. The proposed BMA approach significantly outperformed the competing estimators for all crops and sample size combinations with only minor slippage in efficiency due to spatial correlation.

Third, an out-of-sample game between private insurance companies and the federal government was undertaken and highlights the policy implications for a number of crops and states. With the RMA using its current methodology and the insurance companies using the BMA approach, economically significant rents can be generated by insurance companies in 27 of the 28 cases of which 20 are statistically significant. With a reduced sample size of 25 years, we find economically significant rents can be generated by insurance companies in 26 of the 28 cases of which 23 are statistically significant. For reduced sample size of 20 and 15, we find economically significant rents can be generated by insurance companies in 26 and 28 of the 28 cases, respectively, all but one of which are statistically significant. Results for the BMA versus the standard approach are qualitatively similar. While the rating game replicates the environment between RMA and private insurance companies and as such are most relevant from a policy perspective, the results do not speak to the relative efficiency of the methodologies being compared because of the inherent implicit subsidy. When the implicit subsidy is accounted for, in only eight of the 28 full sample cases can we conclude that the BMA approach leads to significantly more efficient rates. When the sample is reduced to 25, 20 and 15 years that number increases to 21, 27, and 26 cases respectfully. The BMA approach is never statistically significantly worse than the RMA or standard approach. Consistently, each simulation undertaken demonstrated the BMA approach may offer significant efficiency gains, particularly where historical yield data is limited or no longer representative.

References

Annan, F., J. Tack, A. Harri, and K. Coble. 2014. "Spatial pattern of yield distributions: Implications for crop insurance." *American Journal of Agricultural Economics* 96(1):253–268.

Atwood, J., S. Shaik, and M. Watts. 2003. "Are crop yields normally distributed? A reexamination." *American Journal of Agricultural Economics* 85(4):888–901.

Botts, R.R., and J.N. Boles. 1958. "Use of normal-curve theory in crop insurance ratemaking." *American Journal of Agricultural Economics* 40(3):733–740.

Bourgeon, J.M., and R.G. Chambers. 2003. "Optimal area-yield crop insurance reconsidered." *American Journal of Agricultural Economics* 85(3):590–604.

Chen, J., and P. Li. 2009. "Hypothesis test for normal mixture models: The EM approach." *The Annals of Statistics* 37(5A):2523–2542.

Congressional Budget Office. 2014. *H.R. 2642, Agricultural Act of 2014: Cost Estimate*. Washington, DC.

Dasgupta, A., and A.E. Raftery. 1998. "Detecting features in spatial point processes with clutter via model-based clustering." *Journal of the American Statistical Association* 93(441):294–302.

Dorfman, J.H. 1998. "Bayesian composite qualitative forecasting: Hog prices again." *American Journal of Agricultural Economics* 80(3):543–551.

Dorfman, J.H., and D.R. Sanders. 2006. "Generalized hedge ratio estimation with an unknown model." *Review of Futures Markets* 15(2):167–188.

Draper, D. 1995. "Assessment and propagation of model uncertainty." *Journal of the Royal Statistical Society. Series B (Methodological)* 57(1):45–97.

Gallagher, P. 1987. "U.S. soybean yields: Estimation and forecasting with nonsymmetric disturbances." *American Journal of Agricultural Economics* 69(4):796–803.

Glauber, J.W. 2013. "The growth of the federal crop insurance program, 1990-2011." *American Journal of Agricultural Economics* 95(2):482–488.

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(1):139–153.

Goodwin, B.K., M.C. Roberts, and K.H. Coble. 2000. "Measurement of price risk in revenue insurance: Implications of distributional assumptions." *Journal of Agricultural and Resource Economics* 25(1):195–214.

Halcrow, H.G. 1949. "Actuarial structures for crop insurance." *Journal of Farm Economics* 31(3):418–443.

Harri, A., K. Coble, A.P. Ker, and B.J. Goodwin. 2011. "Relaxing heteroscedasticity assumptions in area-yield crop insurance rating." *American Journal of Agricultural Economics* 93(3):707–717.

Hjort, N.L., and G. Claeskens. 2003. "Frequentist model average estimators." *Journal of the American Statistical Association* 98:879–899.

Hjort, N.L., and I.K. Glad. 1995. "Nonparametric density estimation with a parametric start." *The Annals of Statistics* 23(3):882–904.

Hoeting, J.A., D. Madigan, A.E. Raftery, and C.T. Volinsky. 1999. "Bayesian model averaging: A tutorial." *Statistical Science* 14(4):382–401.

Jones, M., O. Linton, and J. Nielsen. 1995. "A simple bias reduction method for density estimation." *Biometrika* 82(2):327–338.

Jones, M., and D. Signorini. 1997. "A comparison of higher-order bias kernel density estimators." *Journal of the American Statistical Association* 92(439):1063–1073.

Just, R.E., and Q. Weninger. 1999. "Are crop yields normally distributed?" *American Journal of Agricultural Economics* 81(2):287–304.

Karlis, D., and E. Xekalaki. 2003. "Choosing initial values for the EM algorithm for finite mixtures." *Computational Statistics & Data Analysis* 41(3):577–590.

Kass, R.E., and A.E. Raftery. 1995. "Bayes factors." *Journal of the American Statistical Association* 90:773–795.

Kaylen, M.S., and S.S. Koroma. 1991. "Trend, weather variables, and the distribution of U.S. corn yields." *Review of Agricultural Economics* 13(2):249–258.

Ker, A.P. 2002. "Nonparametric estimation of possibly similar densities." Unpublished, Working paper, Department of Agricultural and Resource Economics, University of Arizona.

—. 1996. "Using SNP maximum likelihood techniques to recover conditional densities: A new approach to recovering premium rates." Unpublished, Working paper, Department of Agricultural and Resource Economics, University of Arizona.

Ker, A.P., and K. Coble. 2003. "Modeling conditional yield densities." *American Journal of Agricultural Economics* 85(2):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.

Ker, A.P., and P. McGowan. 2000. “Weather-based adverse selection and the U.S. crop insurance program: The private insurance company perspective.” *Journal of Agricultural and Resource Economics* 25(2):386–410.

Leamer, E.E., and E.E. Leamer. 1978. *Specification Searches: Ad hoc Inference with Nonexperimental data*. Wiley New York.

Li, P., J. Chen, and P. Marriott. 2009. “Non-finite Fisher information and homogeneity: An EM approach.” *Biometrika* 96(2):411–426.

Madigan, D., and A.E. Raftery. 1994. “Model selection and accounting for model uncertainty in graphical models using Occam’s window.” *Journal of the American Statistical Association* 89:1535–1546.

Marron, J.S., and M.P. Wand. 1992. “Exact mean integrated squared error.” *The Annals of Statistics* 20(2):712–736.

Miranda, M.J. 1991. “Area-yield crop insurance reconsidered.” *American Journal of Agricultural Economics* 73(2):233–242.

Moss, C.B., and J.S. Shonkwiler. 1993. “Estimating yield distributions with a stochastic trend and non-normal errors.” *American Journal of Agricultural Economics* 75(5):1056–1062.

Nelson, C.H., and P.V. Preckel. 1989. “The conditional beta distribution as a stochastic production function.” *American Journal of Agricultural Economics* 71:370–378.

Ozaki, V.A., and R.S. Silva. 2009. “Bayesian ratemaking procedure of crop insurance contracts with skewed distribution.” *Journal of Applied Statistics* 36(4):443–452.

Racine, J., and A.P. Ker. 2006. “Rating crop insurance policies with efficient nonparametric estimators that admit mixed data types.” *Journal of Agricultural and Resource Economics* 31(1):27–39.

Raftery, A.E. 1996. “Approximate Bayes factors and accounting for model uncertainty in generalised linear models.” *Biometrika* 83(2):251–266.

Raftery, A.E., D. Madigan, and J.A. Hoeting. 1997. “Bayesian model averaging for linear regression models.” *Journal of the American Statistical Association* 92(437):179–191.

Ramirez, O.A. 1997. “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(1):191–205.

Roberts, H.V. 1965. “Probabilistic prediction.” *Journal of the American Statistical Association* 60:50–62.

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):406–419.

Skees, J.R., J.R. Black, and B.J. Barnett. 1997. “Designing and rating an area yield crop insurance contract.” *American Journal of Agricultural Economics* 79(2):430–438.

Skees, J.R., and M.R. Reed. 1986. "Rate making for farm-level crop insurance: Implications for adverse selection." *American Journal of Agricultural Economics* 68(3):653–659.

Stochs, S.M., and J. LaFrance. 2004. "A learning rule for inferring local distributions over space and time." In T. B. Fomby and R. C. Hill, eds. *Advances in Econometrics, Volume 18: Spatial and Spatiotemporal Econometrics*. New York: Elsevier Science, pp. 295–331.

Tack, J., A. Harri, and K. Coble. 2012. "More than mean effects: Modeling the effect of climate on the higher order moments of crop yields." *American Journal of Agricultural Economics* 94(5):1037–1054.

Tolhurst, T.N., and A.P. Ker. 2015. "On technological change in crop yields." *American Journal of Agricultural Economics* 92(1):137–158.

Volinsky, C.T., D. Madigan, A.E. Raftery, and R.A. Kronmal. 1997. "Bayesian model averaging in proportional hazard models: Assessing the risk of a stroke." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 46(4):433–448.

Woodard, J.D., and B.J. Sherrick. 2011. "Estimation of mixture models using cross-validation optimization." *American Journal of Agricultural Economics* 93(4):968–982.

Wu, X., and Y.Y. Zhang. 2012. "Nonparametric estimation of crop yield distributions: A panel data approach." AAEA 2012 Annual Meeting, August 12-14, Seattle, WA., No. 124630.

Appendix A. Summary of Rating Game Results: BMA versus Standard Approach

Table 8. Out-of-Sample Rating Game Result Summary: p-values

	BMA versus Standard Density Estimate			
	Full Sample	25 Years	20 Years	15 Years
Corn				
Illinois	0.0000***	0.0020***	0.0008***	0.0000***
Indiana	0.0002***	0.0996*	0.0000***	0.0016***
Iowa	0.0002***	0.9962†††	0.8144	0.1064
Minnesota	0.0224**	0.0002***	0.0000***	0.0000***
Missouri	0.0306**	0.0000***	0.0000***	0.0000***
Ohio	0.0000***	0.0000***	0.0000***	0.0000***
Wisconsin	0.0008***	0.0000***	0.0000***	0.0000***
Soybean				
Illinois	0.0036***	0.0172**	0.0020***	0.0024***
Indiana	0.0168**	0.0000***	0.0000***	0.0000***
Iowa	0.0000***	0.0098***	0.0028***	0.4546
Minnesota	0.0000***	0.0000***	0.0000***	0.0000***
Missouri	0.0054***	0.0000***	0.0000***	0.0000***
Ohio	0.0000***	0.0000***	0.0000***	0.0000***
Wisconsin	0.0406**	0.0000***	0.0000***	0.0000***
Cotton				
Arkansas	0.1662	0.0160**	0.3894	0.5562
Georgia	0.0558*	0.0398**	0.0056***	0.0258**
Louisiana	0.0340**	0.0008***	0.0002***	0.0004***
Mississippi	0.0038***	0.0004***	0.0000***	0.0302**
Tennessee	0.1024	0.0026***	0.3918	0.4810
Winter Wheat				
Illinois	0.6406	0.2842	0.0012***	0.0008***
Indiana	0.0016***	0.0888*	0.0080***	0.0000***
Kansas	0.0000***	0.0000***	0.0000***	0.0000***
Maryland	0.2250	0.2106	0.1244	0.1496
Michigan	0.0666*	0.0004***	0.0000***	0.0000***
Missouri	0.0156**	0.0000***	0.0028***	0.0018***
Ohio	0.1510	0.0810*	0.0258**	0.0026***
Oklahoma	0.0002***	0.0000***	0.0048***	0.1376
Tennessee	0.2002	0.6680	0.4812	0.0464**

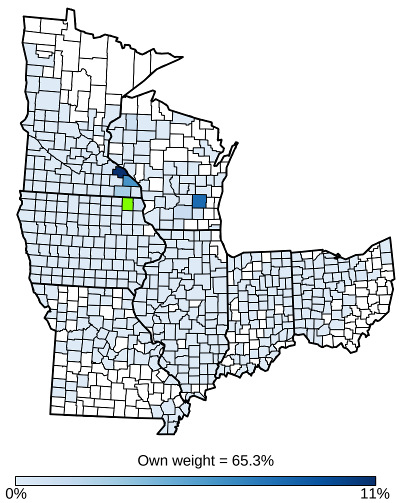
Note: Statistical significance of lower private (government) loss ratios indicated by *(†)-10%, **(††)-5%, and ***(†††)-1%.

Table 9. Efficacy Test Results: p-values

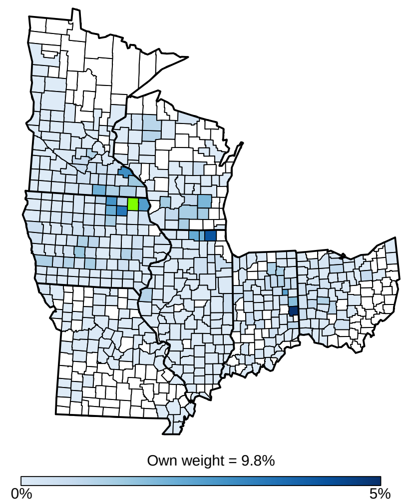
	BMA versus Standard Approach			
	Full Sample	25 Years	20 Years	15 Years
Corn				
Illinois	0.0000***	0.0577	0.1316	0.0013***
Indiana	0.0002***	0.0059**	0.0002***	0.0000***
Iowa	0.0207**	0.2517	0.1316	0.1316
Minnesota	0.0207**	0.0577	0.0577	0.0059**
Missouri	0.0577	0.0577	0.0577	0.1316
Ohio	0.0059**	0.0059**	0.0059**	0.0059**
Wisconsin	0.1316	0.0059**	0.0059**	0.0013***
Soybean				
Illinois	0.0059**	0.0059**	0.0207**	0.7483
Indiana	0.0207**	0.0002***	0.0013***	0.0059**
Iowa	0.0577	0.0059**	0.0013***	0.0013***
Minnesota	0.0577	0.0000***	0.0059**	0.0000***
Missouri	0.1316	0.0207**	0.0059**	0.1316
Ohio	0.0000***	0.0013***	0.0059**	0.1316
Cotton				
Wisconsin	0.0577	0.0059**	0.0013***	0.0002***
Arkansas	0.0059**	0.0002***	0.0207**	0.0207**
Georgia	0.0013***	0.0013***	0.0002***	0.0000***
Louisiana	0.0577	0.0013***	0.0002***	0.0002***
Mississippi	0.0000***	0.0013***	0.0002***	0.0013***
Tennessee	0.0013***	0.0207**	0.7483	0.4119
Winter Wheat				
Illinois	0.0207**	0.0013***	0.0013***	0.0000***
Indiana	0.0013***	0.0577	0.0577	0.1316
Kansas	0.0207**	0.2517	0.0059**	0.0207**
Maryland	0.2517	0.0207**	0.0207**	0.0059**
Michigan	0.4119	0.0013***	0.0059**	0.0577
Missouri	0.1316	0.0059**	0.0577	0.0059**
Ohio	0.0207**	0.0013***	0.0059**	0.0002***

Note: Statistical significance of lower private (government) loss ratios indicated by $^{*}(^{\dagger})$-10%, $^{**}(^{\dagger\dagger})$-5%, and $^{***}(^{\dagger\dagger\dagger})$-1%.

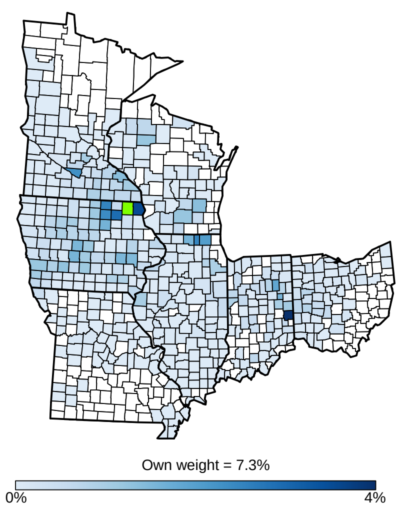
Appendix B. BMA weights in Out-of-Sample Game



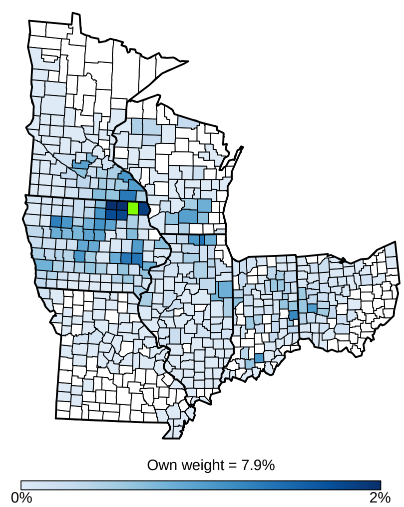
(A) Full sample (39+ Years)



(B) 25 Year Sample

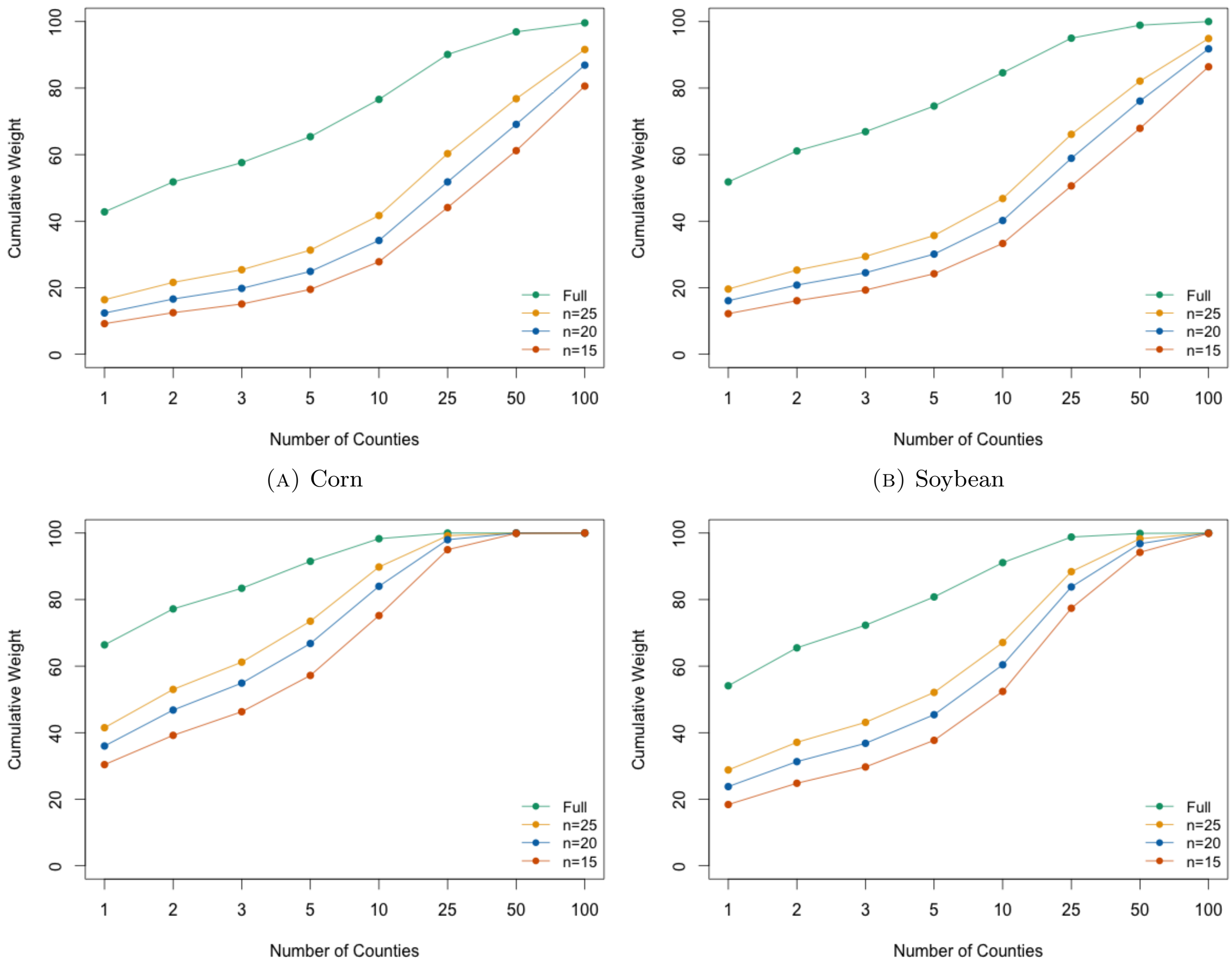


(C) 20 Year Sample



(D) 15 Year Sample

FIGURE 1. Spatial Distribution of BMA Weights for Winneshiek, Iowa Soybean (Bright Green)



(A) Corn

(B) Soybean

(C) Cotton

(D) Winter Wheat

FIGURE 2. Crop-Level Average of Cumulative BMA Weight on n Counties

Bayesian Estimation of Possibly Similar Yield Densities: Implications for Rating Crop Insurance Contracts

Online Supplement

1 Detailed Out-of-Sample Game Results

1.1 Two Trend BMA versus RMA

Table 9: Out-of-Sample Rating Game: Full Sample, Two Trend BMA versus RMA

Crop-State	Number of Counties	Payouts (%)	Retained by Private (%)	Loss Ratio Government	Loss Ratio Private	p-value
Corn						
Illinois	82	24.5	87.1	1.014	0.649	0.0332**
Indiana	69	25.6	83.1	1.080	0.758	0.0538*
Iowa	99	16.1	73.8	0.614	0.347	0.0038***
Minnesota	57	16.1	59.6	0.181	0.145	0.1858
Missouri	48	36.8	82.7	0.861	0.765	0.2694
Ohio	63	21.5	83.7	1.075	0.676	0.0262**
Wisconsin	54	20.6	64.5	0.891	0.476	0.0034***
Soybean						
Illinois	91	20.4	44.0	1.068	0.594	0.0000***
Indiana	66	27.0	51.5	1.475	0.613	0.0000***
Iowa	98	21.6	40.2	0.955	0.786	0.0882*
Minnesota	63	21.7	51.0	0.721	0.415	0.0024***
Missouri	57	33.1	39.6	1.432	0.894	0.0012***
Ohio	52	23.3	67.3	1.173	0.719	0.0050***
Wisconsin	42	34.4	74.4	1.485	0.731	0.0000***
Cotton						
Arkansas	7	25.0	9.3	0.486	0.000	0.0000***
Georgia	20	37.8	43.0	0.809	0.436	0.0012***
Louisiana	6	38.3	26.7	2.303	1.194	0.0934*
Mississippi	15	29.0	8.7	1.435	0.552	0.0588*
Tennessee	10	24.5	0.0	0.758	–	–
Winter Wheat						
Illinois	10	26.0	31.5	1.201	0.232	0.0000***
Indiana	28	27.9	45.5	1.101	0.725	0.0296**
Kansas	71	33.9	55.4	1.607	1.194	0.0048***
Maryland	7	32.9	5.7	2.725	0.336	0.1470
Michigan	32	15.2	15.5	0.466	0.505	0.6000
Missouri	23	25.4	48.3	0.893	0.728	0.2474
Ohio	19	21.8	41.8	1.198	0.910	0.2510
Oklahoma	22	37.0	80.2	3.377	1.531	0.0002***
Tennessee	13	19.6	67.7	0.734	0.367	0.0590*

Note: Statistical significance of lower private (government) loss ratios indicated by *(†)-10%, **(††)-5%, and ***(†††)-1%.

Table 10: Out-of-Sample Rating Game: 25 Year Sample, Two Trend BMA versus RMA

Crop-State	Number of Counties	Payouts (%)	Retained by Private (%)	Loss Ratio Government	Loss Ratio Private	*p*-value
Corn						
Illinois	82	28.7	45.9	1.868	0.929	0.0002***
Indiana	69	27.8	59.2	1.298	1.100	0.1512
Iowa	99	17.5	32.0	0.651	0.712	0.7146
Minnesota	57	16.7	41.1	0.368	0.268	0.0530*
Missouri	48	33.0	61.1	1.883	1.067	0.0006***
Ohio	63	26.3	46.0	1.746	0.880	0.0000***
Wisconsin	54	20.3	47.6	1.666	0.401	0.0000***
Soybean						
Illinois	91	24.4	27.3	1.694	0.833	0.0000***
Indiana	66	29.0	40.9	2.747	0.795	0.0000***
Iowa	98	25.2	30.0	1.377	0.974	0.0082***
Minnesota	63	21.4	51.5	1.862	0.271	0.0000***
Missouri	57	35.3	39.5	2.810	0.981	0.0000***
Ohio	52	31.5	37.9	2.730	0.668	0.0000***
Wisconsin	42	31.8	54.4	2.803	0.863	0.0000***
Cotton						
Arkansas	7	25.7	48.6	2.326	0.326	0.0000***
Georgia	20	32.2	59.2	1.086	0.648	0.0342**
Louisiana	6	21.7	53.3	2.782	0.968	0.0252**
Mississippi	15	30.7	53.3	2.352	0.745	0.0000***
Tennessee	10	19.5	50.5	2.443	0.277	0.0000***
Winter Wheat						
Illinois	10	23.5	61.5	1.095	0.453	0.0060***
Indiana	28	28.0	45.5	1.597	0.771	0.0002***
Kansas	71	36.3	38.6	1.928	0.818	0.0000***
Maryland	7	24.3	41.4	1.908	1.583	0.3640
Michigan	32	23.8	34.4	1.131	0.787	0.0422**
Missouri	23	26.7	47.4	1.566	0.736	0.0012***
Ohio	19	27.4	37.9	1.658	1.569	0.4244
Oklahoma	22	36.8	44.5	2.867	0.873	0.0000***
Tennessee	13	28.1	31.2	0.812	1.213	0.8652

Note: Statistical significance of lower private (government) loss ratios indicated by $^*(^\dagger)$-10%, $^{**}(^{\dagger\dagger})$-5%, and $^{***}(^{\dagger\dagger\dagger})$-1%.

Table 11: Out-of-Sample Rating Game: 20 Year Sample, Two Trend BMA versus RMA

Crop-State	Number of Counties	Payouts (%)	Retained by Private (%)	Loss Ratio Government	Loss Ratio Private	*p*-value
Corn						
Illinois	82	28.5	46.2	1.962	0.996	0.0000***
Indiana	69	28.5	49.3	1.833	0.918	0.0000***
Iowa	99	19.1	40.1	0.690	0.765	0.7602
Minnesota	57	19.6	40.1	0.624	0.271	0.0000***
Missouri	48	35.8	53.5	2.537	0.982	0.0000***
Ohio	63	29.3	35.8	1.982	0.740	0.0000***
Wisconsin	54	25.0	37.5	1.991	0.421	0.0000***
Soybean						
Illinois	91	24.9	27.1	1.842	0.895	0.0000***
Indiana	66	27.7	36.9	2.809	0.744	0.0000***
Iowa	98	25.8	34.8	1.399	0.930	0.0022***
Minnesota	63	25.1	40.3	1.734	0.375	0.0000***
Missouri	57	34.1	39.9	3.126	0.853	0.0000***
Ohio	52	31.2	35.9	2.945	0.910	0.0000***
Wisconsin	42	31.0	48.1	2.547	0.881	0.0000***
Cotton						
Arkansas	7	25.0	47.1	1.766	0.403	0.0016***
Georgia	20	34.5	48.5	1.525	0.659	0.0008***
Louisiana	6	20.8	46.7	3.267	0.595	0.0010***
Mississippi	15	31.3	46.3	2.684	0.619	0.0000***
Tennessee	10	26.0	44.0	2.718	0.300	0.0000***
Winter Wheat						
Illinois	10	25.5	53.5	1.365	0.349	0.0000***
Indiana	28	26.4	50.0	1.704	0.791	0.0000***
Kansas	71	38.0	42.2	2.151	0.889	0.0000***
Maryland	7	27.1	40.7	2.890	0.896	0.0084***
Michigan	32	21.6	39.5	1.423	0.384	0.0000***
Missouri	23	28.7	40.4	1.742	0.621	0.0000***
Ohio	19	23.4	40.5	2.078	0.764	0.0020***
Oklahoma	22	42.5	39.8	2.950	1.115	0.0000***
Tennessee	13	28.5	29.6	1.098	1.088	0.5030

Note: Statistical significance of lower private (government) loss ratios indicated by *(†)-10%, **(††)-5%, and ***(†††)-1%.

Table 12: Out-of-Sample Rating Game: 15 Year Sample, Two Trend BMA versus RMA

Crop-State	Number of Counties	Payouts (%)	Retained by Private (%)	Loss Ratio Government	Loss Ratio Private	p-value
Corn						
Illinois	82	28.4	43.2	2.189	1.230	0.0000***
Indiana	69	28.1	48.6	2.084	1.037	0.0000***
Iowa	99	20.3	37.9	1.105	0.561	0.0000***
Minnesota	57	21.2	34.6	0.880	0.178	0.0000***
Missouri	48	37.0	50.3	2.826	1.092	0.0000***
Ohio	63	30.1	37.1	1.959	0.912	0.0000***
Wisconsin	54	25.5	29.7	2.003	0.642	0.0000***
Soybean						
Illinois	91	25.4	29.4	2.158	0.774	0.0000***
Indiana	66	27.2	38.8	3.464	0.507	0.0000***
Iowa	98	24.0	40.6	1.178	1.077	0.2498
Minnesota	63	25.8	37.9	1.584	0.630	0.0000***
Missouri	57	36.1	40.9	2.955	1.201	0.0000***
Ohio	52	29.6	34.0	3.047	0.764	0.0000***
Wisconsin	42	30.1	46.7	2.173	1.008	0.0000***
Cotton						
Arkansas	7	27.1	43.6	1.768	0.664	0.0154**
Georgia	20	32.5	40.2	1.701	0.680	0.0000***
Louisiana	6	23.3	36.7	3.550	0.256	0.0002***
Mississippi	15	32.0	38.7	2.446	0.647	0.0000***
Tennessee	10	28.5	32.0	2.201	0.314	0.0002***
Winter Wheat						
Illinois	10	25.5	50.5	1.896	0.167	0.0000***
Indiana	28	26.4	53.6	2.113	0.695	0.0000***
Kansas	71	37.9	40.4	2.169	0.920	0.0000***
Maryland	7	26.4	36.4	2.782	0.586	0.0034***
Michigan	32	21.1	33.9	1.496	0.405	0.0000***
Missouri	23	25.4	36.1	1.638	0.761	0.0062***
Ohio	19	24.5	34.7	2.155	0.638	0.0004***
Oklahoma	22	40.9	40.7	2.608	1.283	0.0016***
Tennessee	13	31.2	27.3	1.553	0.409	0.0008***

Note: Statistical significance of lower private (government) loss ratios indicated by $^{*}(^{\dagger})$-10%, $^{**}(^{\dagger\dagger})$-5%, and $^{***}(^{\dagger\dagger\dagger})$-1%.

1.2 Two Trend BMA versus Standard Approach

Table 13: Out-of-Sample Rating Game: Full Sample, Two Trend BMA versus Standard Density Estimate

Crop-State	Number of Counties	Payouts (%)	Retained by Private (%)	Loss Ratio Government	Loss Ratio Private	p-value
Corn						
Illinois	82	24.5	37.5	1.413	0.708	0.0000***
Indiana	69	25.6	41.0	1.640	0.764	0.0002***
Iowa	99	16.1	36.8	0.716	0.347	0.0002***
Minnesota	57	16.1	34.0	0.237	0.140	0.0224**
Missouri	48	36.8	49.0	1.278	0.926	0.0306**
Ohio	63	21.5	47.7	1.475	0.676	0.0000***
Wisconsin	54	20.6	31.5	1.049	0.475	0.0008***
Soybean						
Illinois	91	20.4	33.3	1.007	0.612	0.0036***
Indiana	66	27.0	30.1	1.144	0.773	0.0168**
Iowa	98	21.6	43.4	1.212	0.606	0.0000***
Minnesota	63	21.7	40.6	0.863	0.338	0.0000***
Missouri	57	33.1	46.6	1.223	0.845	0.0054***
Ohio	52	23.3	42.3	1.558	0.626	0.0000***
Wisconsin	42	34.4	38.1	1.412	1.047	0.0406**
Cotton						
Arkansas	7	25.0	58.6	0.288	0.185	0.1662
Georgia	20	37.8	46.8	0.696	0.476	0.0558*
Louisiana	6	38.3	38.3	1.673	0.835	0.0340**
Mississippi	15	29.0	45.7	0.940	0.406	0.0038***
Tennessee	10	24.5	56.5	0.245	0.149	0.1024
Winter Wheat						
Illinois	10	26.0	30.0	0.614	0.696	0.6406
Indiana	28	27.9	46.1	1.136	0.587	0.0016***
Kansas	71	33.9	45.4	1.806	1.062	0.0000***
Maryland	7	32.9	25.7	0.990	0.610	0.2250
Michigan	32	15.2	46.9	0.413	0.236	0.0666*
Missouri	23	25.4	39.3	0.984	0.523	0.0156**
Ohio	19	21.8	26.3	1.140	0.756	0.1510
Oklahoma	22	37.0	47.0	3.755	1.965	0.0002***
Tennessee	13	19.6	33.1	0.684	0.445	0.2002

Note: Statistical significance of lower private (government) loss ratios indicated by *(†)-10%, **(††)-5%, and ***(†††)-1%.

Table 14: Out-of-Sample Rating Game: 25 Year Sample, Two Trend BMA versus Standard Density Estimate

Crop-State	Number of Counties	Payouts (%)	Retained by Private (%)	Loss Ratio Government	Loss Ratio Private	p-value
Corn						
Illinois	82	28.7	51.9	1.516	1.002	0.0020^{***}
Indiana	69	27.8	59.3	1.316	1.068	0.0996^{*}
Iowa	99	17.5	47.6	0.385	0.626	$0.9962^{\dagger\dagger\dagger}$
Minnesota	57	16.7	66.3	0.380	0.148	0.0002^{***}
Missouri	48	33.0	79.1	1.969	0.848	0.0000^{***}
Ohio	63	26.3	42.9	1.608	0.881	0.0000^{***}
Wisconsin	54	20.3	48.1	1.536	0.409	0.0000^{***}
Soybean						
Illinois	91	24.4	43.4	1.097	0.810	0.0172^{**}
Indiana	66	29.0	52.0	1.915	0.773	0.0000^{***}
Iowa	98	25.2	41.2	1.137	0.789	0.0098^{***}
Minnesota	63	21.4	60.6	1.727	0.281	0.0000^{***}
Missouri	57	35.3	70.2	1.512	0.919	0.0000^{***}
Ohio	52	31.5	43.3	2.243	0.711	0.0000^{***}
Wisconsin	42	31.8	60.6	2.383	0.741	0.0000^{***}
Cotton						
Arkansas	7	25.7	93.6	0.870	0.212	0.0160^{**}
Georgia	20	32.2	71.2	0.616	0.366	0.0398^{**}
Louisiana	6	21.7	69.2	2.330	0.480	0.0008^{***}
Mississippi	15	30.7	71.7	1.271	0.383	0.0004^{***}
Tennessee	10	19.5	96.5	1.223	0.148	0.0026^{***}
Winter Wheat						
Illinois	10	23.5	77.0	0.674	0.545	0.2842
Indiana	28	28.0	67.3	1.010	0.757	0.0888^{*}
Kansas	71	36.3	33.3	1.948	0.785	0.0000^{***}
Maryland	7	24.3	68.6	1.062	0.683	0.2106
Michigan	32	23.8	62.3	0.709	0.335	0.0004^{***}
Missouri	23	26.7	58.0	1.429	0.484	0.0000^{***}
Ohio	19	27.4	48.2	1.235	0.839	0.0810^{*}
Oklahoma	22	36.8	40.5	3.222	1.236	0.0000^{***}
Tennessee	13	28.1	34.2	0.728	0.841	0.6680

Note: Statistical significance of lower private (government) loss ratios indicated by $^{*}(^{\dagger})$-10%, $^{**}(^{\dagger\dagger})$-5%, and $^{***}(^{\dagger\dagger\dagger})$-1%.

Table 15: Out-of-Sample Rating Game: 20 Year Sample, Two Trend BMA versus Standard Density Estimate

Crop-State	Number of Counties	Payouts (%)	Retained by Private (%)	Loss Ratio Government	Loss Ratio Private	p-value
Corn						
Illinois	82	28.5	56.8	1.559	0.999	0.0008***
Indiana	69	28.5	60.1	1.603	0.924	0.0000***
Iowa	99	19.1	53.4	0.495	0.568	0.8144
Minnesota	57	19.6	66.8	0.482	0.182	0.0000***
Missouri	48	35.8	72.5	2.011	0.851	0.0000***
Ohio	63	29.3	38.5	1.716	0.824	0.0000***
Wisconsin	54	25.0	44.3	1.538	0.498	0.0000***
Soybean						
Illinois	91	24.9	45.2	1.141	0.756	0.0020***
Indiana	66	27.7	51.2	1.808	0.621	0.0000***
Iowa	98	25.8	43.1	1.111	0.771	0.0028***
Minnesota	63	25.1	55.2	1.281	0.415	0.0000***
Missouri	57	34.1	69.4	1.676	0.818	0.0000***
Ohio	52	31.2	41.2	2.230	0.726	0.0000***
Wisconsin	42	31.0	55.7	2.196	0.634	0.0000***
Cotton						
Arkansas	7	25.0	90.0	0.260	0.225	0.3894
Georgia	20	34.5	67.0	0.762	0.385	0.0056***
Louisiana	6	20.8	65.0	2.632	0.407	0.0002***
Mississippi	15	31.3	71.7	1.399	0.431	0.0000***
Tennessee	10	26.0	95.5	0.235	0.216	0.3918
Winter Wheat						
Illinois	10	25.5	69.5	1.083	0.415	0.0012***
Indiana	28	26.4	65.0	1.091	0.652	0.0080***
Kansas	71	38.0	40.4	2.130	0.887	0.0000***
Maryland	7	27.1	70.0	1.396	0.733	0.1244
Michigan	32	21.6	62.7	0.795	0.288	0.0000***
Missouri	23	28.7	54.8	1.147	0.588	0.0028***
Ohio	19	23.4	51.1	1.205	0.657	0.0258**
Oklahoma	22	42.5	40.7	2.914	1.726	0.0048***
Tennessee	13	28.5	37.3	0.838	0.811	0.4812

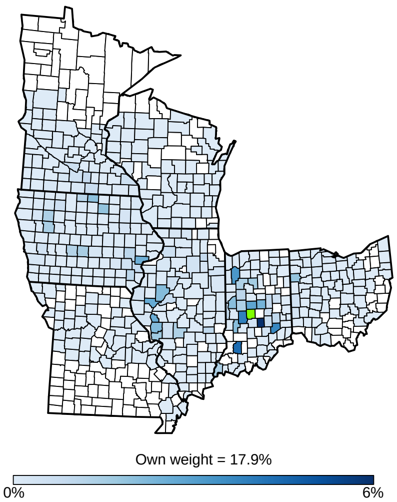
Note: Statistical significance of lower private (government) loss ratios indicated by *(†)-10%, **(††)-5%, and ***(†††)-1%.

Table 16: Out-of-Sample Rating Game: 15 Year Sample, Two Trend BMA versus Standard Density Estimate

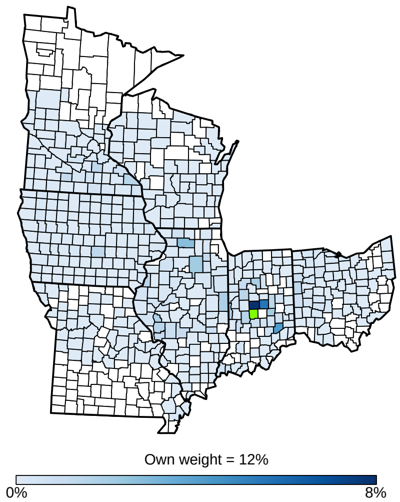
Crop-State	Number of Counties	Payouts (%)	Retained by Private (%)	Loss Ratio Government	Loss Ratio Private	p-value
Corn						
Illinois	82	28.4	61.1	1.816	0.950	0.0000***
Indiana	69	28.1	62.0	1.569	0.997	0.0016***
Iowa	99	20.3	62.1	0.642	0.530	0.1064
Minnesota	57	21.2	73.5	0.625	0.184	0.0000***
Missouri	48	37.0	73.4	2.123	0.860	0.0000***
Ohio	63	30.1	39.9	1.731	0.908	0.0000***
Wisconsin	54	25.5	45.6	1.462	0.597	0.0000***
Soybean						
Illinois	91	25.4	49.1	1.173	0.797	0.0024***
Indiana	66	27.2	51.1	1.963	0.596	0.0000***
Iowa	98	24.0	46.5	0.875	0.858	0.4546
Minnesota	63	25.8	56.9	1.224	0.477	0.0000***
Missouri	57	36.1	70.7	1.716	0.856	0.0000***
Ohio	52	29.6	43.2	2.146	0.811	0.0000***
Wisconsin	42	30.1	57.4	2.047	0.690	0.0000***
Cotton						
Arkansas	7	27.1	89.3	0.194	0.243	0.5562
Georgia	20	32.5	70.0	0.690	0.392	0.0258**
Louisiana	6	23.3	56.7	2.287	0.444	0.0004***
Mississippi	15	32.0	72.0	1.015	0.544	0.0302**
Tennessee	10	28.5	98.5	0.000	0.238	0.4810
Winter Wheat						
Illinois	10	25.5	70.0	1.144	0.426	0.0008***
Indiana	28	26.4	68.6	1.465	0.525	0.0000***
Kansas	71	37.9	41.1	2.231	0.892	0.0000***
Maryland	7	26.4	65.7	1.279	0.731	0.1496
Michigan	32	21.1	63.6	0.976	0.256	0.0000***
Missouri	23	25.4	58.5	1.070	0.518	0.0018***
Ohio	19	24.5	51.1	1.420	0.592	0.0026***
Oklahoma	22	40.9	38.9	2.443	1.970	0.1376
Tennessee	13	31.2	40.8	1.117	0.588	0.0464**

Note: Statistical significance of lower private (government) loss ratios indicated by $^{*}(^{\dagger})$-10%, $^{**}(^{\dagger\dagger})$-5%, and $^{***}(^{\dagger\dagger\dagger})$-1%.

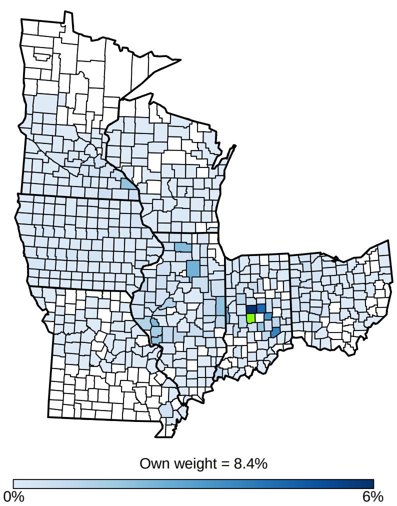
2 Spatial Distribution of BMA Weights



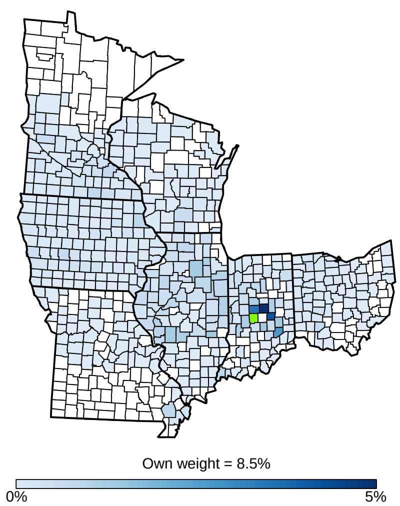
(a) Full sample (39+ Years)



(b) 25 Year Sample

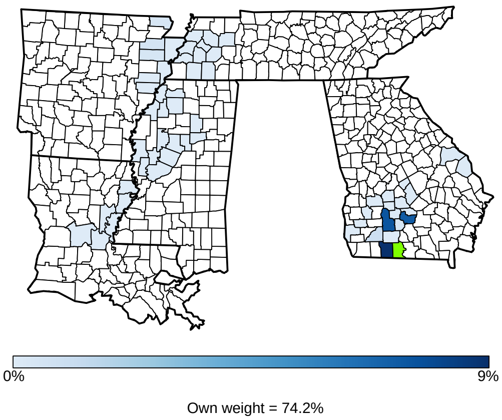


(c) 20 Year Sample

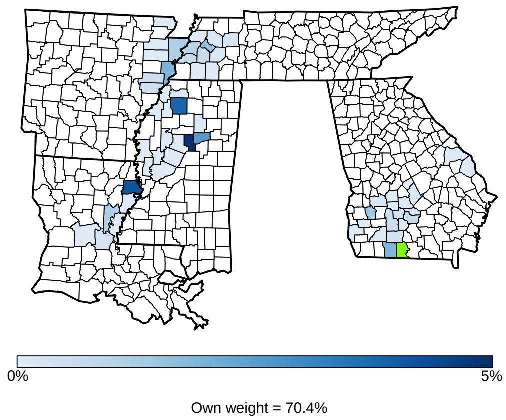


(d) 15 Year Sample

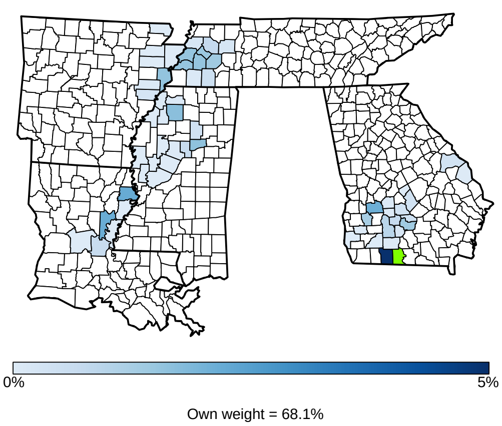
Figure 3: Hendricks, Indiana Corn (Green) Average BMA Weight



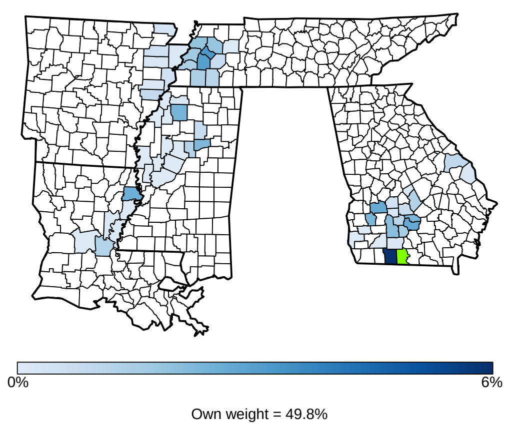
(a) Full sample (39+ Years)



(b) 25 Year Sample

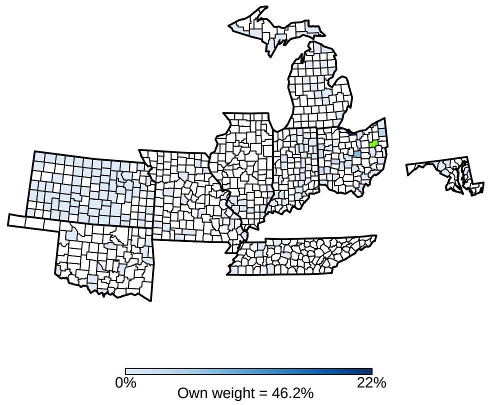


(c) 20 Year Sample

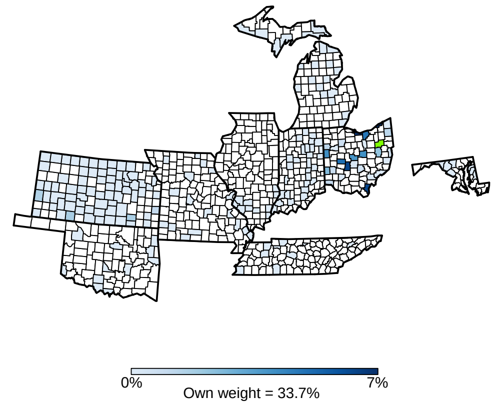


(d) 15 Year Sample

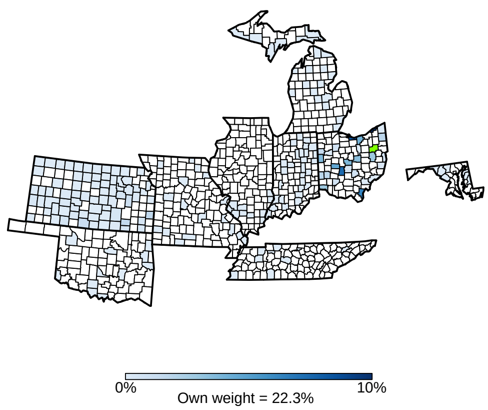
Figure 4: Brooks, Georgia Cotton (Green) Average BMA Weight



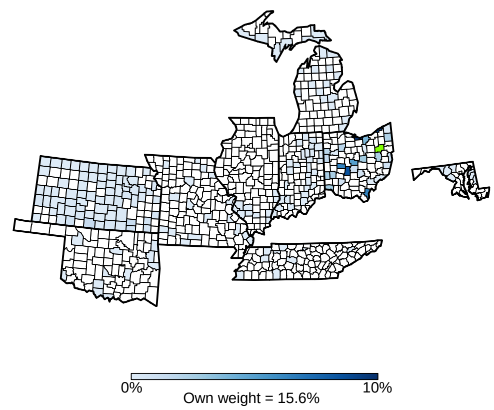
(a) Full sample (39+ Years)



(b) 25 Year Sample



(c) 20 Year Sample



(d) 15 Year Sample

Figure 5: Tillman, Oklahoma Winter Wheat (Green) Average BMA Weight

3 Efficacy Test Statistic

We consider the effect of the relative variance and bias between the RMA rating methodology and the insurance company rating methodology on D. We assume yields are normal with mean 180 and standard deviation 30. Subsequently, the true rate is 0.041782. We assume the RMA methodology yields rates that are uniformly distributed with mean equal to the true rate and bounds ± 0.005. We first assume that the insurance company rating methodology yields rates that are uniformly distributed with mean equal to the true rate and bounds $\pm\alpha$ where α $[0.000, 0.0001, ..., 0.015]$. Secondly, we assume the insurance company rating methodology yields rates that are uniformly distributed with mean equal to the true rate plus β and bounds ± 0.005 where $\beta \in [-0.007, -0.006, ..., 0.008]$. We take $n = \{50, 100, 500, 1000\}$ iid draws from the yield and rate distributions 10,000 times. Based on the realized rates, the insurance company chooses which policies to retain and which to cede but it must retain 50% of the policies.[1] We then calculate the loss ratios for the program overall, government, and insurance company. We then reverse the roles and let the government choose which policies to retain and which to cede and then calculate the *reverse* loss ratios. Based on the standard and reverse loss ratios we calculate D. The median of D is reported in the table of results. Note, the following: (i) bias does not effect the median of D; (ii) if the insurance company has a more efficient rating methodology than RMA, in terms of variation, then the median of D is greater than 1; (iii) if the insurance company and RMA have equally efficient rating methodologies then the median of D is 1; and (iv) if the insurance company has a less efficient rating methodology than RMA (in terms of variation) then the median of D is less than 1. We note that these results hold for all levels of n, when the rates are normally distributed rather than uniformly distributed, when yields are distributed as a mixture of normals versus a normal, and when bias and variance are interacted.

[1] If neither the insurance company or RMA methodology is biased, then the insurance company naturally retains 50% of the contracts. This is not true when there is bias and so we restrict the insurance company to retain 50%, of what it believes to be, the most profitable contracts.

Table 17: Effect of Relative Variance and Bias on D-Statistic

Bias		Std. Dev.		Median of D-Statistic			
Government	Ins. Comp.	Government	Ins. Comp.	n=50	n=100	n=500	n=1000
0	0.000000	0.028868	0.000000	1.166329	1.138577	1.122865	1.130395
0	0.000000	0.028868	0.012910	1.118314	1.129131	1.133062	1.118651
0	0.000000	0.028868	0.018257	1.083488	1.087681	1.111727	1.100811
0	0.000000	0.028868	0.022361	1.074511	1.071410	1.081611	1.082506
0	0.000000	0.028868	0.025820	1.036057	1.055539	1.047695	1.038083
0	0.000000	0.028868	0.028868	0.955884	1.001314	0.998295	1.000864
0	0.000000	0.028868	0.031623	0.980893	0.969439	0.960685	0.955322
0	0.000000	0.028868	0.034157	0.930025	0.924326	0.918558	0.919616
0	0.000000	0.028868	0.036515	0.861471	0.892770	0.889293	0.888774
0	0.000000	0.028868	0.038730	0.879252	0.866166	0.858438	0.858242
0	0.000000	0.028868	0.040825	0.833146	0.840363	0.839476	0.836791
0	0.000000	0.028868	0.042817	0.812507	0.808627	0.812007	0.808775
0	0.000000	0.028868	0.044721	0.789803	0.782338	0.783352	0.784801
0	0.000000	0.028868	0.046547	0.773162	0.772444	0.766452	0.768062
0	0.000000	0.028868	0.048305	0.753966	0.753070	0.745987	0.744013
0	0.000000	0.028868	0.050000	0.747547	0.728813	0.727806	0.724605
0	-0.007000	0.028868	0.028868	1.018019	0.958969	0.987931	0.988592
0	-0.006000	0.028868	0.028868	0.968814	0.990977	0.984899	0.985135
0	-0.005000	0.028868	0.028868	0.993887	1.004893	0.982229	0.985441
0	-0.004000	0.028868	0.028868	0.992630	0.988634	0.991309	0.995869
0	-0.003000	0.028868	0.028868	1.013007	0.976353	1.005506	0.994504
0	-0.002000	0.028868	0.028868	1.004892	1.008515	0.991976	0.992910
0	-0.001000	0.028868	0.028868	1.000708	1.003491	0.999741	1.003012
0	0.000000	0.028868	0.028868	1.000462	1.004408	1.000318	0.992451
0	0.001000	0.028868	0.028868	1.025930	1.012099	1.007393	0.998255
0	0.002000	0.028868	0.028868	1.002509	0.997201	1.003592	1.003909
0	0.003000	0.028868	0.028868	1.027996	0.993578	1.010359	1.005075
0	0.004000	0.028868	0.028868	1.008680	1.020211	1.008870	1.008146
0	0.005000	0.028868	0.028868	1.002681	1.007549	1.008014	1.011023
0	0.006000	0.028868	0.028868	1.004485	1.018828	1.018640	1.009963
0	0.007000	0.028868	0.028868	0.996161	1.004829	1.011970	1.013115
0	0.008000	0.028868	0.028868	1.014466	1.001738	1.013047	1.011943