

RESEARCH ARTICLE

# When Less Is More: On the Use of Historical Yield Data with Application to Rating Area Crop Insurance Contracts

Yong Liu and Alan P. Ker\* 

Department of Food, Agricultural and Resource Economics, University of Guelph, Guelph, Ontario, Canada

\*Corresponding author. Email: [aker@uoguelph.ca](mailto:aker@uoguelph.ca)

## Abstract

Crop insurance is the cornerstone program of domestic farm policy in most developed countries. Although most countries' rating methodology corrects for time-varying movements in the first two moments, it is unclear whether using the entire yield series remains appropriate. We use distributional tests and an out-of-sample retain-cede rating game to answer whether governments/insurers should historically trim yields in estimating their premium rates. Despite small sample sizes and the need to estimate tail probabilities, the historical data appear to be sufficiently different such that trimming is justified.

**Keywords:** Premium rates; trimming; yield data

**JEL Classifications:** Q18; Q14

## 1. Introduction

Much of production agriculture in developed countries is produced under heavily subsidized insurance and has been for the past 20–25 years. Most crop insurance programs use historical yield data to set guarantees, estimate premium rates, and calculate indemnities. Moreover, many studies in the literature use historical yield data to consider a variety of issues related to rating crop insurance contracts (Goodwin and Hungerford, 2015; Goodwin and Ker, 1998; Miranda and Glauber, 1997; Zhang, 2017). In many cases, and certainly with respect to rating methodologies, yield data are detrended and adjusted for possible heteroscedasticity and then assumed to be independent and identically distributed (Zhu, Goodwin, and Ghosh [2011] denote this as the “two-stage method”). In many countries, county or provincial yield data exist from the 1950s onward and reflect very significant innovations in both seed and farm management technologies—innovations that have likely moved mass all around the support of the yield distribution. This begs the question, to what extent, if any, do yield losses in the 1950s and 1960s inform us about yield losses in 2019 even after accounting for time-varying movements in the first two moments? We ask whether yield data should be historically trimmed when estimating premium rates. Note that the issue of trimming is exacerbated in rating insurance contracts by the need to estimate tail probabilities.

Changes in seed and farm management technologies and their effects on yields have been well documented in the agronomy literature. Notable examples include the introduction of biotech seeds and precision farming. Many studies have shown that corn, soybean, and wheat yields in the United States have more than doubled from 1950 to the mid-1990s (Assefa et al., 2017;

Duvick, 2005; Egli, 2008, 2017; Fernandez-Cornejo, 2004; Fernandez-Cornejo et al., 2014; Reilly and Fuglie, 1998). They tend to suggest that roughly half of the yield gain is attributed to genetic seed improvements while the other half is attributed to improved agronomic practices. Although the agronomy literature has focused on changes in average yields, some studies have also documented increasing volatility in yields (Challinor et al., 2014; Kucharik and Ramankutty, 2005; Leng, 2017; Naylor, Falcon, and Zavaleta, 1997). Conversely, there has been a relatively large body of work on the changes in yield volatility by agricultural economists, primarily driven by issues related to crop insurance (Claassen and Just, 2011; Harri et al., 2011; Zhang, 2017). With respect to changes in the higher moments ( $>2$ ) of the yield distribution, there has been markedly less work. Zhu, Goodwin, and Ghosh (2011), using U.S. Department of Agriculture, National Agricultural Statistics Service (USDA-NASS) county-level yield data for corn, soybeans, and cotton, found changes in higher moments through time. Tack, Harri, and Coble (2012), using county-level cotton data in Arkansas, Mississippi, and Texas, found that the third moment, or skewness, changed with time for Mississippi and Texas. Note that changes in higher moments indicate that the common approach of correcting for changes in the first two moments is not sufficient for the identically distributed assumption in most of the literature as well as the rating methodologies for many government programs. However, given the need to estimate tail probabilities, the aforementioned results (which are very region-crop specific) do not necessarily suggest that historically trimming yield data will lead to more accurate premium rates; the loss function for each is over very different subsets of the density space.

The objective of this article is to answer the question of whether governments should or should not trim their historical yield data in estimating area crop insurance contracts. We focus our attention on the U.S. crop insurance program (administered by the Risk Management Agency) to be of greatest relevance to the existing literature (most of which is focused on the U.S. crop insurance program). Using county-level NASS yield data for corn, soybeans, and winter wheat, we first, for completeness, consider nonparametric distributional tests to assess if the adjusted yield data may result from different data generating processes. Second, we use an out-of-sample retain-cede rating game—commonly employed in the literature—to compare premium rates from the full versus historically trimmed yield data. Specifically, we focus our trimming at 1991 to reflect the distinction between the Federal Crop Insurance Corporation's (FCIC) area-based insurance programs and its newer shallow loss programs. In its area programs, all historical yield data are used in the rating process, whereas in its newer shallow loss programs only yield data from 1991 onward are used. In a somewhat related article, Shen, Odening, and Okhrin (2018) argue for trimming based on time-varying changes in the first two moments of the yield dgp (data generating process). However, the literature has generally corrected for changes in the lower moments using deterministic or stochastic trends (first moment) and procedures to accommodate heteroscedasticity (see Harri et al., 2011). The important question for trimming relates to the higher moments.

The remainder of this manuscript proceeds as follows. Section 2 details the NASS yield data, the FCIC detrending methodology, and the FCIC heteroscedasticity treatment. Section 3 presents the statistical results from testing the identically distributed assumption. Section 4 presents the economic results using an out-of-sample retain-cede rating game. The final section summarizes our findings.

## 2. NASS yield data, detrending methodology, and heteroscedasticity treatment

NASS provides 49 categories of field crops including beans, cotton, corn, grain, hay, peanuts, mint, rice, soybeans, and wheat. The data generally date back to the 1950s. We use county-level yield

data for corn, soybeans, and winter wheat for the period 1951–2017 (67 years). Our corn and soybean analysis focuses on states that account for the majority of national corn and soybean production. We removed counties with one or more missing yield observations, as well as any state that does not have 25 or more counties. We also removed all states that reported more than 10% of their acreage as irrigated in the 2012 Census of Agriculture. After doing so, we were left with seven states for corn: Illinois (IL), Indiana (IN), Iowa (IA), Minnesota (MN), Ohio (OH), South Dakota (SD), and Wisconsin (WI). These states accounted for 57.8% of harvested acreage and 61.8% of national production in 2017. All corn states except South Dakota met the inclusion criteria for soybeans. These six states accounted for 50.5% and 53.9% of national harvested acreage and production, respectively, in 2017. For winter wheat, we considered the top 15 states that had less than 10% of their acreage irrigated in 2012 Census of Agriculture, only two of which met the inclusion criteria: Kansas (KS) and Michigan (MI). These two states accounted for 29.2% and 28.9% of national harvested acreage and production, respectively, in 2017. In total, our data comprise 414 corn, 373 soybean, and 64 winter wheat counties.

Premium rates are estimated using a two-step process in which a trend is first estimated and then residuals are adjusted for possible heteroscedasticity. A two-step process is by far the most common in the literature as noted by Zhu, Goodwin, and Ghosh (2011). FCIC estimates the temporal process of yields, denoted  $y_t = (y_1, \dots, y_T)$ , for each crop-county combination using a robust two-knot linear spline:

$$y_t = \theta_1 + \theta_2 t + \delta_1 d_1(t - k_1) + \delta_2 d_2(t - k_2) + \varepsilon_t, \tag{1}$$

with  $d_1 = 1$  if  $t \geq k_1$  and  $d_2 = 1$  if  $t \geq k_2$  for knots  $k_1, k_2 \in (1 + \bar{k}, \dots, T - \bar{k})$  and  $k_2 - k_1 \geq \underline{k}$ . The  $\underline{k}, \bar{k} \geq 10$  are imposed, which prevent the knots from locating too close together ( $\underline{k}$ ) or too close to either end point ( $\bar{k}$ ). Knot locations  $k_i$  are selected using a grid search (least-squares criterion). The model is run with zero, one, and two knots and then the number of knots used is selected using the Akaike information criterion.<sup>1</sup> Given the number of knots, two robustness procedures are performed; the spline is iterated to convergence with Huber weights and then twice through a bisquare function. Specifically, let  $\hat{\varepsilon}_t$  be the estimated residuals from the robust spline with the chosen number of knots and  $\tilde{\eta}_t = \hat{\varepsilon}_t / \sqrt{T^{-1} \sum \hat{\varepsilon}_t^2}$ . The Huber function assigns weight one to observations if  $|\tilde{\eta}_t| < c$  and weight  $c/|\tilde{\eta}_t|$  otherwise with a default  $c = 1.345$ . Similarly, the bisquare function weights observations  $(1 - (\tilde{\eta}_t/c)^2)^2$  if  $|\tilde{\eta}_t| < c$  and zero otherwise with default  $c = 4.685$ .<sup>2</sup>

Denote the residuals from the aforementioned detrending process as  $\hat{\varepsilon}_t$  and the fitted values as  $\hat{g}(t) = \hat{y}_t$ . The heteroscedasticity adjustment via the Harri et al. (2011) estimates

$$\ln(\hat{\varepsilon}_t^2) = \alpha + \gamma \ln(\hat{y}_t) + v_t. \tag{2}$$

Note, constant and proportional variance in the underlying yield data correspond to  $\gamma = 0$  and  $\gamma = 2$ , respectively. Yields are adjusted based on a one-step ahead forecast ( $\hat{y}_{T+1}$ ) and the heteroscedasticity coefficient ( $\hat{\gamma}$ ).<sup>3</sup>

$$\hat{y}_t^* = \hat{y}_{T+1} + \hat{\varepsilon}_t \left( \frac{\hat{y}_{T+1}}{\hat{y}_t} \right)^{\frac{\hat{\gamma}}{2}}. \tag{3}$$

<sup>1</sup>We do not impose the spatial and temporal priors on knots used by the FCIC.

<sup>2</sup>Given any results are dependent on the choice of detrending method, we considered three alternative methodologies to ensure robustness of our results: (1) a linear model estimated by  $L_2$ , (2) a linear model estimated by  $L_1$ , and (3) nonparametric local lines using out-of-sample cross validation for the smoothing parameter. Our results are robust to any of these alternative detrending methodologies.

<sup>3</sup>FCIC uses a two-step ahead forecast because of data availability/timing issues. We choose a one-step ahead forecast for our analysis simply to gain an additional degree of freedom given we are truncating an already short time series.

The adjusted yields are then used to generate the empirical premium rate for period  $T + 1$ :

$$\pi_{T+1}^* = \frac{1}{T} \sum_{t=1}^T \max\{0, \lambda \hat{y}_{T+1}^* - \hat{y}_t^*\}, \tag{4}$$

where  $\lambda$  is the coverage level such that  $\lambda \hat{y}_{T+1}^*$  is the yield guarantee.

### 3. Testing the identically distributed assumption

When testing for structural change, generally a Chow-type test is used; the sample is split into different subpopulations, and residuals from regression equations within the subpopulations are combined with the residuals from a regression equation spanning the two samples to form a Wald-type test statistic. The Bai-Perron test is a sup-type test of the Chow test in that it does not assume the break point is known or the number of break points. The Wilcoxon rank sum test is like a Chow test and primarily has power against changes in location. Overall, these tests have power only against changes in the conditional mean function (first moment) and thus, unsurprisingly, resulted in very few rejections on the adjusted yields across the crop-county combinations.<sup>4</sup> We are interested in structural changes in the higher moments of the data generating process, beyond the conditional mean or variance. A common choice is the Kolmogorov-Smirnov (KS) test, which considers the maximum difference between two empirical distribution functions and thus has power against differences in all moments. Note that the test is nonparametric in that the test statistic is a function of the two empirical distribution functions. Also, the KS test has been shown to have relatively low power in comparison with the Chow or Bai-Perron tests as these tests have an infinitely smaller space of alternatives (Wilcox, 1997). Moreover, the difference between two empirical distribution functions is most pronounced for differences in the location, followed by differences in scale, and then higher moments in sequential order. Recall we will only be testing differences in the higher moments, and thus the power of the KS test is further weakened in that the two samples we are comparing have near identical first two moments. The KS test statistic is denoted  $D_{n,m}$  and is defined as

$$D_{n,m} = \sup_x |F_{1,n}(x) - F_{2,m}(x)|, \tag{5}$$

where  $F_{1,n}$  and  $F_{2,m}$  are the empirical distribution functions of the first and the second samples, respectively. Specifically, the entire yield series is detrended and corrected for heteroscedasticity and then split pre- and post-1991 corresponding to the different FCIC rating procedures; recall the area-based programs use all the yield data, whereas the newer supplemental loss programs only use yields from 1991 onward. The null of the KS test is rejected at level  $\alpha$  when

$$D_{n,m} > c(\alpha) \sqrt{\frac{n+m}{nm}}, \tag{6}$$

where  $c(\alpha)$  is calculated from the Kolmogorov distribution.

We also consider a second test forwarded by Li (1996) and further developed by Li, Maasoumi, and Racine (2009) (denoted the LMR test). This LMR test is similar to a Cramer-von Mises test in that rather than based on the supremum difference, it is based on the integrated squared difference. Specifically, the LMR test smooths the data using kernel

---

<sup>4</sup>Given we have corrected for the time-varying changes in the first moment in our detrending process, any rejections reflect the inappropriateness of the underlying functional form in the detrending process.

**Table 1.** Identically distributed test results

Crop-State	Number of Counties	Number of Rejections			
		KS Test		LMR Test	
		$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.10$
<b>Corn</b>					
Illinois	73	1	12	38	46
Indiana	60	2	4	13	20
Iowa	91	4	18	22	34
Minnesota	57	11	21	25	27
Ohio	58	1	5	5	12
Wisconsin	48	4	10	11	20
South Dakota	27	0	2	9	11
<b>Soybeans</b>					
Illinois	82	1	3	14	18
Indiana	59	1	1	6	12
Iowa	93	0	0	5	15
Minnesota	55	3	6	11	17
Ohio	51	0	2	9	14
Wisconsin	33	0	0	9	14
<b>Winter wheat</b>					
Kansas	35	0	0	1	3
Michigan	29	0	1	3	6

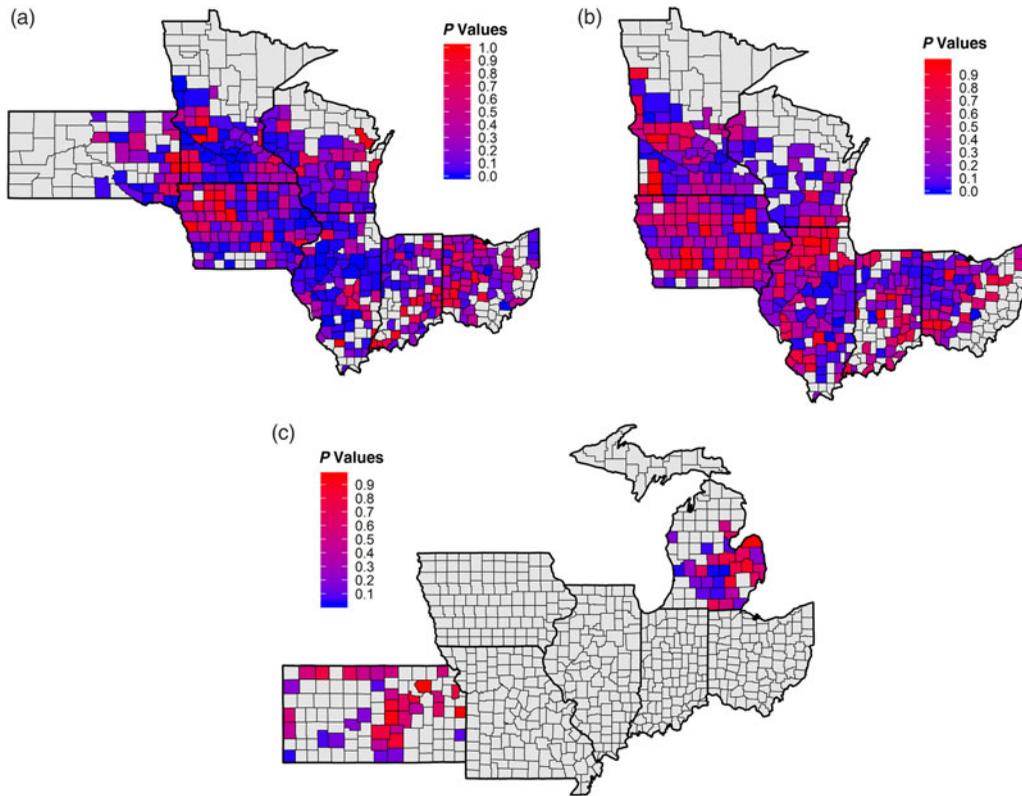
methods and calculates the integrated squared difference. Moreover, Li, Maasoumi, and Racine (2009) find that power is increased if one bootstraps the null, using randomization methods, rather than use an asymptotic expansion for the distribution of the test statistic. Specifically, the entire yield series is detrended, corrected for heteroscedasticity, and then divided into two subsets pre- and post-1991. The test statistic is defined as follows:

$$LMR = \int_{-\infty}^{\infty} (\hat{f}_1(x) - \hat{f}_2(x))^2 dx, \quad (7)$$

where  $\hat{f}_1$  and  $\hat{f}_2$  are kernel estimates based on the two subsets of data. Li, Maasoumi, and Racine (2009) suggest using least-squares cross validation for bandwidth selection. Moreover, the kernel estimates under the bootstrap samples to recover the distribution of  $LMR$  under the null use the same two bandwidths in each bootstrap. In our application, 500 bootstrap samples were used to construct the null.

The test results are presented in Table 1. As expected, the LMR test has noticeably more power than the KS test given the null for the LMR test is calculated using randomization methods.<sup>5</sup> Second, the results reject that the data pre- and post-1991 come from the same distribution in many of the crop-state combinations despite the small number of observations. In corn, 30% of the counties reject at the 5% significance level, while 41% reject at the 10% significance level.

<sup>5</sup>A similar randomization method can be used for the KS test, although this is rarely done in practice.



**Figure 1.** Maps of  $P$  values: LMR (Li, Maasoumi, and Racine) test. (a) Corn counties. (b) Soybean counties. (c) Wheat counties.

The results are similar across the seven states. Soybeans exhibit somewhat less significance as compared with corn; 14% of the counties reject at the 5% significance level, while 24% reject at the 10% significance level. Winter wheat exhibits very little statistical significance, just 6% of the counties reject at the 5% significance level, while 14% reject at the 10% significance level, barely above the size of the test. Interestingly, these results correspond to the level of research expenditures in the three crops over the past half-century. Corn has seen the most innovation, whereas wheat has seen very little. The LMR test results ( $P$  value) for corn, soybeans, and winter wheat are graphically illustrated by county-crop combinations in Figure 1. There does appear to be geographical clustering. For example, with respect to corn the majority of the central counties—the high production counties—reject the null of identically distributed. With respect to soybeans, the clustering rejections are in the more eastern counties. With respect to winter wheat, the clustering in Michigan is to the west, while in Kansas it is to the southwest.

#### 4. Trimming and estimating crop insurance rates

Results from the previous section call into question the identically distributed assumption from a statistical perspective but provide little information regarding economic importance. As previously mentioned, the loss functions for estimating a distribution versus a premium rate are over different subsets of the density space. We consider the effect of trimming in rating crop insurance contracts by using an out-of-sample retain-cede rating game consistent with the literature. Specifically, the game allows two players using different methodologies to estimate premium rates and adversely select against one another. The game was first proposed by Ker and McGowan

(2000) and has since been employed by Racine and Ker (2006), Harri et al. (2011), Annan et al. (2013), Tack and Ubilava (2015), Zhang (2017), and Shen, Odening, and Okhrin (2018) to justify alternative rating methodologies. The game was modified with an additional test of rating efficiency by Ker, Tolhurst, and Liu (2016). Park, Brorsen, and Harri (2019) utilized both tests in proposing an alternative rating methodology that exploits spatial closeness. The game was inspired by the retain-cede decision of private insurers in regard to the crop insurance contracts they sell. Some salient features of the U.S. crop insurance program are relevant to the game. First, FCIC rather than private insurers sets the rates for all policies. Second, the private insurer must sell all policies in a state in which it operates (even if it deems the policy to be underpriced). Third, the private insurer shares, asymmetrically, in the underwriting gains and losses of all policies it sells. Fourth, there is a mechanism by which private insurers can significantly reduce their exposure on policies they deem unwanted.<sup>6</sup> Given these salient features, private insurers determine which policies to retain and which to cede. That is, private insurers retain policies that they believe are overpriced and for which they expect an underwriting gain and cede policies that they believe are underpriced and for which they expect an underwriting loss. As a result, private insurers necessarily develop their own rates in the attempt to strategically adversely select against FCIC and recover excess rents. Mimicking this allows one to hypothetically compare two sets of premium rates: one based on the full yield series and one based on the trimmed yield series. This is in contrast to the past literature that employs the retain-cede game to evaluate alternative rating methodologies using the same data.

Specifically, we assume FCIC uses the full historical yield data from 1951 to 1997 on a county-crop basis to estimate the FCIC premium rates for 1998. Conversely, private insurers estimate their rates using a 25-year trimmed data set (i.e., yields from 1973 to 1997).<sup>7</sup> Both the FCIC and the private insurers use the FCIC rating methodology outlined previously, and as such, the only difference in the two sets of rates is the result of trimming the historical yield data. Based on the two sets of rates, the private insurer identifies which contracts to retain and which to cede. The underwriting gains or losses for the set of retained and ceded contracts are calculated using the actual yields in 1998. This process is repeated for 20 years, and the loss ratios (defined as the ratio of total underwriting losses to total premiums) for both the retained and ceded sets of contracts are calculated. We conduct the game for each crop-county combination at the 90% coverage level where the very large majority of area-based contracts are purchased.

As in the previously cited literature, we undertake two hypothesis tests. The first tests whether the loss ratio from the retained contracts is less than the loss ratio from retaining contracts randomly (choosing which contracts to retain randomly is equivalent to the private insurer being indifferent between the two sets of competing rates). Like Li, Maasoumi, and Racine (2009), randomization methods are used to recover the *P* value. Game 1 mimics the current reality of the U.S. crop insurance program. However, private insurers have an advantage because they react to the FCIC premium rates. As such, whichever of the two competing rates the private insurers use has an inherent competitive advantage in game 1. This advantage is nullified in game 2 by contrasting the changes in loss ratios under both sets of the competing rates (for details, see Ker, Tolhurst, and Liu, 2016). The number of contracts considered is the number of counties multiplied by 20 years: 8,280 contracts for corn, 7,460 contracts for soybeans, and 1,280 contracts for winter wheat. The results, which include percent retained by the private insurer, the government or ceded contracts loss ratio, the insurer or retained loss ratio, *P* value of game 1, and *P* value of game 2, are presented in Table 2.

Under a 25-year trimming decision rule, we find the private insurer's loss ratio is less than the FCIC loss ratio for all 14 of the 15 state-crop combinations (Michigan wheat is higher only in the

<sup>6</sup>Specific details are outlined in the USDA-FCIC Standard Reinsurance Agreement with approved private insurers.

<sup>7</sup>The 25-year trimming roughly corresponds to the 1991 cutoff. We also did trimming at 30 years, and our results were qualitatively identical and quantitatively very similar.

**Table 2.** Out-of-sample retain-cede rating game

Crop-State	Number of Counties	Retained by Private (%)	Loss Ratio Government	Loss Ratio Private	Game 1 <i>P</i> Value	Game 2 <i>P</i> Value
Corn						
Illinois	73	66.1	0.581	0.519	0.0207	0.1316
Indiana	60	72.5	0.676	0.560	0.0013	0.0577
Iowa	91	59.1	0.411	0.270	0.0013	0.0577
Minnesota	57	75.9	0.228	0.195	0.0000	0.0577
Ohio	58	68.1	0.713	0.585	0.0013	0.2517
Wisconsin	48	77.0	0.543	0.378	0.0013	0.7483
South Dakota	27	81.3	0.810	0.545	0.0000	0.0002
Soybeans						
Illinois	82	51.4	0.874	0.467	0.0059	0.2517
Indiana	59	70.5	1.019	0.538	0.0002	0.0207
Iowa	93	52.4	0.840	0.531	0.0059	0.1316
Minnesota	55	76.2	0.787	0.491	0.0013	0.0059
Ohio	51	76.1	0.734	0.640	0.0059	0.1316
Wisconsin	33	85.5	1.054	0.738	0.0013	0.2517
Winter wheat						
Kansas	35	32.9	1.525	0.814	0.0059	0.0577
Michigan	29	49.0	0.346	0.349	0.2517	0.1316

third decimal place). For corn, the private insurer loss ratio ranges from 66% to 89% of the FCIC loss ratio. For soybeans, the ratio ranges from 53% to 87%. Given only two states for winter wheat, the ratio is 53% for Kansas and 101% for Michigan. With respect to the first game, *P* value 1 is significant in all state-crop combinations but Michigan wheat, suggesting that economically and statistically significant rents can be recovered by private insurers by trimming the yield data. With respect to the second game, *P* value 2 is significant at the 10% level for 7 of the 15 state-crop combinations, suggesting that trimming leads to statistically significantly more accurate premium rates. In no cases did not trimming lead to statistically significantly more accurate premium rates. Specifically, with respect to corn, four of the seven states were significant. With respect to soybeans, two of the six states were significant. Finally, with respect to winter wheat, one of the two states was significant. In summary, the out-of-sample retain-cede rating game provides strong evidence for trimming, consistent with the results from the last section.

## 5. Conclusions

Historical yield data have been utilized in many empirical applications in literature, most notably in applications related to crop insurance. In general, all available historical yield data were used. Most applications account for time-varying lower moments but assume time-constant upper moments. In addition, FCIC does the same in estimating the premium rates of its area products. However, there have been significant innovations in farm management and seed technologies in the past half century such that mass has likely moved all around the yield distribution. Not surprisingly, a few studies have found changes in the upper moments of the yield data generating

processes, thus questioning the standard approach of correcting the first two moments only. Our distributional test results find strong evidence of the inappropriateness of the identically distributed assumption for corn and soybeans and markedly less so for winter wheat. These results are surprisingly strong in that the sample sizes are relatively small, and thus the power of the tests against economically relevant alternatives are weakened. Our out-of-sample retain-cede rating game, which represents a different loss function over only a subset of the density space, is consistent with our distributional tests. That is, trimming does not increase estimation error in the rating process and is shown in approximately half of the crop-state combinations considered to decrease estimation error. This result is quite noteworthy suggesting that despite small sample sizes and the need to estimate tail probabilities, the historical data appear to be sufficiently different such that trimming is justified. Our results across crops are fairly consistent with the research expenditures across crops in that we find the biggest efficiency gains from trimming in corn, which has experienced the most innovation. Although we caution against extrapolation, our results should give cause for consideration when using historical yield data in rating crop insurance contracts in all countries, as well as using historical yield data in other applications. Finally, although our results suggest trimming is preferred to nontrimming, alternative temporal weighting schemes may be more optimal but of course will be dependent on the empirical application, econometric methodology, and loss function (see, e.g., Liu and Ker [2019], in which cross-validation smoothing methods are used to simultaneously estimate temporal and spatial weights).

**Financial support.** We would like to thank the Institute for the Advanced Study in Food and Agricultural Policy, University of Guelph for financial support.

## References

- Annan, F., J. Tack, A. Harri, and K. Coble. "Spatial Pattern of Yield Distributions: Implications for Crop Insurance." *American Journal of Agricultural Economics* **96**, 1(2013):253–68.
- Assefa, Y., P. Prasad, P. Carter, M. Hinds, G. Bhalla, R. Schon, M. Jeschke, S. Paszkiewicz, and I.A. Ciampitti. "A New Insight into Corn Yield: Trends from 1987 through 2015." *Crop Science* **57**, 5(2017):2799–811.
- Challinor, A.J., J. Watson, D. Lobell, S. Howden, D. Smith, and N. Chhetri. "A Meta-analysis of Crop Yield under Climate Change and Adaptation." *Nature Climate Change* **4**(March 2014):287–91.
- Claassen, R., and R.E. Just. "Heterogeneity and Distributional Form of Farm-Level Yields." *American Journal of Agricultural Economics* **93**, 1(2011):144–60.
- Duvick, D.N. "The Contribution of Breeding to Yield Advances in Maize (*Zea mays* L.)." *Advances in Agronomy* **86**(2005): 83–145.
- Egli, D. "Comparison of Corn and Soybean Yields in the United States: Historical Trends and Future Prospects." *Agronomy Journal* **100**, S3(2008):S-79–88.
- Egli, D.B. *Seed Biology and Yield of Grain Crops*. Wallingford, UK: CABI, 2017.
- Fernandez-Cornejo, J. "The Seed Industry in US Agriculture: An Exploration of Data and Information on Crop Seed Markets, Regulation, Industry Structure, and Research and Development." Working paper, Washington, DC: U.S. Department of Agriculture, Economic Research Service, 2004.
- Fernandez-Cornejo, J., S. Wechsler, M. Livingston, and L. Mitchell. *Genetically Engineered Crops in the United States*. Washington, DC: U.S. Department of Agriculture, Economic Research Service, Economic Research Report No. 162, 2014.
- Goodwin, B.K., and A. Hungerford. "Copula-Based Models of Systemic Risk in US Agriculture: Implications for Crop Insurance and Reinsurance Contracts." *American Journal of Agricultural Economics* **97**, 3(2015):879–96.
- Goodwin, B.K., and A.P. Ker. "Nonparametric Estimation of Crop Yield Distributions: Implications for Rating Group-Risk Crop Insurance Contracts." *American Journal of Agricultural Economics* **80**, 1(1998):139–53.
- Harri, A., K.H. Coble, A.P. Ker, and B.J. Goodwin. "Relaxing Heteroscedasticity Assumptions in Area-Yield Crop Insurance Rating." *American Journal of Agricultural Economics* **93**, 3(2011):707–17.
- Ker, A.P., and P. McGowan. "Weather-Based Adverse Selection and the US Crop Insurance Program: The Private Insurance Company Perspective." *Journal of Agricultural and Resource Economics* **25**, 2(2000):386–410.
- Ker, A.P., T.N. Tolhurst, and Y. Liu. "Bayesian Estimation of Possibly Similar Yield Densities: Implications for Rating Crop Insurance Contracts." *American Journal of Agricultural Economics* **98**, 2(2016):360–82.
- Kucharik, C.J., and N. Ramankutty. "Trends and Variability in U.S. Corn Yields over the Twentieth Century." *Earth Interactions* **9**, 1(2005):1–29.
- Leng, G. "Recent Changes in County-Level Corn Yield Variability in the United States from Observations and Crop Models." *Science of the Total Environment* **607–608**(December 2017):683–90.

- Li, Q.** “Nonparametric Testing of Closeness between Two Unknown Distribution Functions.” *Econometric Reviews* **15**, 3(1996):261–74.
- Li, Q., E. Maasoumi, and J.S. Racine.** “A Nonparametric Test for Equality of Distributions with Mixed Categorical and Continuous Data.” *Journal of Econometrics* **148**, 2(2009):186–200.
- Liu, Y., and A.P. Ker.** “Simultaneous Borrowing of Extraneous Information across Space and Time for Estimating Crop Insurance Premium Rates.” Working Paper 2019.9, Institute for the Advanced Study of Food and Agricultural Policy, University of Guelph, Guelph, Ontario, Canada, 2019.
- Miranda, M.J., and J.W. Glauber.** “Systemic Risk, Reinsurance, and the Failure of Crop Insurance Markets.” *American Journal of Agricultural Economics* **79**, 1(1997):206–15.
- Naylor, R., W. Falcon, and E. Zavaleta.** “Variability and Growth in Grain Yields, 1950–94: Does the Record Point to Greater Instability?” *Population and Development Review* **23**, 1(1997):41–58.
- Park, E., B.W. Brorsen, and A. Harri.** “Using Bayesian Kriging for Spatial Smoothing in Crop Insurance Rating.” *American Journal of Agricultural Economics* **101**, 1(2019):330–51.
- Racine, J., and A.P. Ker.** “Rating Crop Insurance Policies with Efficient Nonparametric Estimators That Admit Mixed Data Types.” *Journal of Agricultural and Resource Economics* **31**, 1(2006):27–39.
- Reilly, J.M., and K.O. Fuglie.** “Future Yield Growth in Field Crops: What Evidence Exists?” *Soil and Tillage Research* **47**, 3–4(1998):275–90.
- Shen, Z., M. Odening, and O. Okhrin.** “Adaptive Local Parametric Estimation of Crop Yields: Implications for Crop Insurance Rate Making.” *European Review of Agricultural Economics* **45**, 2(2018):173–203.
- Tack, J., A. Harri, and K. Coble.** “More Than Mean Effects: Modeling the Effect of Climate on the Higher Order Moments of Crop Yields.” *American Journal of Agricultural Economics* **94**, 5(2012):1037–54.
- Tack, J.B., and D. Ubilava.** “Climate and Agricultural Risk: Measuring the Effect of ENSO on U.S. Crop Insurance.” *Agricultural Economics* **46**, 2(2015):245–57.
- Wilcox, R.R.** “Some Practical Reasons for Reconsidering the Kolmogorov-Smirnov Test.” *British Journal of Mathematical and Statistical Psychology* **50**, 1(1997):9–20.
- Zhang, Y.Y.** “A Density-Ratio Model of Crop Yield Distributions.” *American Journal of Agricultural Economics* **99**, 5(2017):1327–43.
- Zhu, Y., B.K. Goodwin, and S.K. Ghosh.** “Modeling Yield Risk under Technological Change: Dynamic Yield Distributions and the U.S. Crop Insurance Program.” *Journal of Agricultural and Resource Economics* **36**, 1(2011):192–210.

---

**Cite this article:** Liu Y and Ker AP (2020). When Less Is More: On the Use of Historical Yield Data with Application to Rating Area Crop Insurance Contracts. *Journal of Agricultural and Applied Economics* **52**, 194–203. <https://doi.org/10.1017/aae.2019.40>