Is There Too Much History in Historical Yield Data

Yong Liu

Ph.D. Candidate

Department of Food, Agricultural and Resource Economics

University of Guelph, Canada

Alan P. Ker[1]

Professor, Department of Food, Agricultural and Resource Economics

OAC Research Chair in Agricultural Risk and Policy

Director, Institute for the Advanced Study of Food and Agricultural Policy

University of Guelph, Canada

Institute for the Advanced Study of Food and Agricultural Policy

Working Paper 2019.1

January, 2019

[1]Corresponding author. Mailing address: University of Guelph, J.D. MacLachlan Building, Guelph, Ontario, Canada N1G 2W1. Phone: 519-824-4120. E-mail: aker@uoguelph.ca.

Is There Too Much History in Historical Yield Data

Abstract

County crop yield data from United States Department of Agriculture - National Agricultural Statistics Service has and continues to be extensively used in the literature as well as practice. The most notable practical example is crop insurance, as the Risk Management Agency uses the data to set guarantees, estimate premium rates, and calculate indemnities for their area programs. In many applications including crop insurance, yield data are detrended and adjusted for possible heteroscedasticity and then assumed to be independent and identically distributed. For most major crop-region combinations, county yield data exist from the 1950s onwards 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. Despite correcting for movements in the first two moments of the yield data generating process (dgp), these innovations raise doubt regarding the identically distributed assumption. This manuscript considers the question of how much historical yield data should be used in empirical analyses. The answer is obviously dependent on the empirical application, crop-region combination, econometric methodology, and chosen loss function. Nonetheless, we hope to provide some guidance by tackling this question in three ways. First, we use distributional tests to assess if and when the adjusted yield data may result from different dgps. Second, we consider the application to crop insurance by using an out-of-sample rating game -- commonly employed in the literature -- to compare rates from the full versus historically restricted data sets. Third, we estimate flexible time-varying dgps and then simulate to quantify the additional error when the identically distributed assumption is erroneously imposed. Our findings suggest that despite accounting for time-varying movements in the first two moments, using yield data more than 30 years past increases estimation error. Given that discarding historical data is unappetizing, particularly so in applications with relatively small T, we investigate three methodologies that re-incorporate the discarded data while explicitly acknowledging: (i) the retained and discarded data are from different dgps; and (ii) the extent and form of those differences is unknown. Our results suggest gains in efficiency may be realized by using these more flexible methodologies. While our results are most applicable to the crop insurance literature, we suggest proceeding with caution when using historical yield data in other applications as well.

Some key words: yield data, changing technology, borrowing information, crop insurance

Introduction

United States Department of Agriculture - National Agricultural Statistics Service (USDA-NASS) county crop yield data continues to be extensively used in the literature as well as practice. The Risk Management Agency (RMA) uses this data to set guarantees, estimate premium rates, and calculate indemnities for their area programs. Moreover, these county level rates are used in estimating farm level rates. Examples from the literature include investigation of rating methodologies (Ramirez, Misra, and Field, 2003; Goodwin and Hungerford, 2015; Ker, Tolhurst, and Liu, 2016; Yvette Zhang, 2017), issues related to reinsurance (Miranda and Glauber, 1997; Ker and Ergun, 2007), issues related to land use (Wu et al., 2004; Claassen, Hellerstein, and Kim, 2013), modeling the climate-yield relationship (Roberts, Schlenker, and Eyer, 2013; Tolhurst and Ker, 2015; Miao, Khanna, and Huang, 2016; Cooper, Nam Tran, and Wallander, 2017), and supply analysis (Hendricks, Smith, and Sumner, 2014). Since 1995, over 100 *American Journal of Agricultural Economics* articles have made reference to NASS yield data and at least 30 have used the data in their empirical analyses. In many cases, and certainly with respect to RMA and the crop insurance literature, yield data are often detrended and adjusted for possible heteroscedasticity and then assumed to be independent and identically distributed. Zhu, Goodwin, and Ghosh (2011) denote this the "two-stage method". For most major crop-region combinations, annual NASS county yield data exist from the 1950s onwards and reflect very significant innovations in both seed and farm management technologies, which have likely moved mass all around the support of the yield distribution. Despite correcting for movements in the first two moments of the yield data generating process (dgp), the identically distributed assumption may not hold. This assumption is employed in both the literature and RMA's rating methodology for their area yield and revenue programs. However, it is unlikely that yield losses in the 1950s and 1960s inform much about yield losses in 2019, let alone be given equal weight to yield losses from recent years. Quite interestingly, and in contrast to their area programs, RMA uses yield data starting in 1991 to rate the newer area-based shallow loss products.[2] This highlights an underlying, but empirically undocumented notion, that the yield dgp has significantly changed over the past half-century possibly rendering the more historical data useless or even harmful in empirical analyses. Similarly, and presumably for the same reason, some in the recent literature have used a markedly shorter historical yield series in their empirical analyses (Claassen, Hellerstein, and Kim, 2013; Hendricks, Smith, and Sumner, 2014; Claassen, Langpap, and Wu, 2017).

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 have shown that corn, soybean and wheat yields in the United States have more than doubled from 1950 to mid-1990s (Reilly and Fuglie, 1998; Fernandez-Cornejo, 2004; Duvick, 2005; Egli, 2008; Fernandez-Cornejo et al., 2014; Assefa et al., 2017; Egli, 2017). 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

[2] Our understanding is RMA aggregates their individual farm data to the county level to construct a county yield series rather than use NASS yield data.

practices. Although the agronomy literature has focused on changes in average yields, some have also documented increasing volatility in yields (Naylor, Falcon, and Zavaleta, 1997; Kucharik and Ramankutty, 2005; Challinor et al., 2014; Leng, 2017). 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 (Ramirez, 1997; Harri et al., 2011; Claassen and Just, 2011; Tolhurst and Ker, 2015; Yvette 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 NASS county level yield data for corn, soybean and cotton, find 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, was changing with time for Mississippi and Texas. Tolhurst and Ker (2015) using county level corn, soybean, winter wheat, and cotton data, found that skewness and kurtosis was changing over time for the vast majority of county-crop combinations. Finally, Ker and Tolhurst (2019) propose correcting heteroscedasticity differentially between the upper and lower tails because of time-varying skewness. Note that changes in higher moments indicate the common approach of correcting for changes in the first two moments is not sufficient for the identically distributed assumption. However, while this assumption may not be *statistically* valid, this does not provide guidance as to if, and to what extent, yield data should be truncated, and, whether it only matters *statistically* and is of little *economic* consequence.

The purpose of this manuscript is to answer the question of how much historical yield data -- after correcting for changes in the first two moments as per much of the literature and RMA -- should be used in empirical analyses. The answer is dependent on the empirical application, crop-region combination, econometric methodology, and chosen loss function. Using county-level NASS yield data for corn, soybean, and winter wheat we take a three-prong approach to answer this question. First, we use nonparametric distributional tests to assess if and when the adjusted yield data may result from different dgps; this highlights the statistical significance. Second, we consider the application to crop insurance by using an out-of-sample retain-cede rating game -- commonly employed in the literature -- to compare premium rates from the full versus restricted data sets; this highlights the economic significance. Third, to get a sense of the magnitude of the error when the identically distributed assumption is imposed, we undertake simulations using estimated mixture models allowing for time-varying moments.

Given we find economic and statistical justification for restricting the historical yield data, a subsequent purpose of this manuscript is to consider methodologies that can make use of the discarded data to increase estimation efficiency while acknowledging the data are not from the same dgp. Specifically, we investigate three approaches that explicitly acknowledge: (i) the unknown dgps are different; and (ii) the extent and form of those differences are unknown. The first is a Bayesian Model Averaging (BMA) approach forwarded by Ker and Liu (2017) and smooths between multiple dgps. The second approach is generally used for smoothing across categories of independent variables (in our case different time periods) forwarded by Li and Racine (2003), and Racine and Li (2004). The final approach forwarded by Ker (2016) reduces bias by combining data from multiple dgps to form a start estimate and then corrects that start according to the dgp of interest. All three methodologies are purposely nonparametric and thus our findings are not conditional on any parametric assumptions. We replicate our economic game and simulation analyses with these more flexible methodologies and find gains in efficiency may be realized by incorporating the discarded historical yield data back into the estimation process.

The remainder of this manuscript proceeds as follows. The next section details the NASS yield data, the detrending methodologies, and the heteroscedasticity treatment. The third section presents the statistical results from testing the identically distributed assumption. The fourth presents the economic results using an out-of-sample retain-cede rating game. The fifth section summarizes the results with the simulated yield data and known dgp. The sixth section outlines the methodologies that borrow information from other dgps. The seventh section presents the economic game and statistical simulation results using these more flexible methodologies. The final section summarizes our findings.

NASS Yield Data, Detrending Methodologies, and Heteroscedasticity Treatment

NASS Yield Data

The mission of NASS is to serve those "working in and depending upon U.S. agriculture." To that end, NASS collects, summarizes, analyzes, and publishes agricultural production and marketing data covering various areas. NASS collects information mainly from farmers, ranchers, livestock feeders, slaughterhouse managers, grain elevator operators, and other agribusinesses. 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, soybean, 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 ten percent of their acreage as irrigated in the 2012 Census of Agriculture. After doing so, we are 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 percent of harvested acreage and 61.8 percent of national production in 2017. All corn states except South Dakota met the inclusion criteria for soybean. These six states accounted for 50.5 and 53.9 percent of national harvested acreage and production, respectively, in 2017. For winter wheat, we considered the top 15 states that had less than ten percent 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 percent and 28.9 percent of national harvested acreage and production, respectively, in 2017. In total, our data comprises 414 corn, 373 soybean, and 64 winter wheat counties.

Detrending Methodologies

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). Given any results are dependent on the choice of detrending method, we consider four different methodologies to ensure robustness of our results. The four detrending methods are: (i) the current RMA methodology (given the relevance of our results to crop insurance); (ii) linear model estimated by L_2; (iii) linear model estimated by L_1; and (iv) nonparametric local lines using out-of-sample cross validation for the smoothing parameter. The four methods span most of the functional forms and error metrics used in the literature.

RMA estimates the temporal process of yields, denoted $y_t = (y_1, ..., y_T)$, for each crop-county combination using a robust two-knot linear spline:

$$y_t = \theta_1 + \theta_2 t + \delta_1 d_1(t - k_1) + \delta_2 d_2(t - k_2) + \varepsilon_t \quad (1)$$

with $d_1 = 1$ if $t \geq k_1$ and $d_2 = 1$ if $t \geq k_2$ for knots $k_1, k_2 \in (1 + \bar{k}, \ldots, T - \bar{k})$ and $k_2 - k_1 \geq \underline{k}$. The $\underline{k}, \bar{k} \geq 10$ are *a prior* imposed bounds which prevent the knots from locating too close together ($\underline{k}$) or too close to either endpoint ($\bar{k}$). 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 AIC.[3] 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.[4]

Our second method assumes a linear model estimated by minimizing least squares. That is, $y_t = \theta_1 + \theta_2 t + \varepsilon_t$ where $(\hat{\theta_1}, \hat{\theta_2}) = argmin_{(\theta_1, \theta_2)}(\varepsilon' \varepsilon)$. Our third method assumes the same linear model estimated via minimizing least absolute deviations (median regression). That is, $y_t = \theta_1 + \theta_2 t + \varepsilon_t$ where $(\tilde{\theta_1}, \tilde{\theta_2}) = argmin_{(\theta_1, \theta_2)}(\sum(|\varepsilon_i|)$.

Finally, our fourth method estimates $g(t)$ by nonparametric local lines. The underlying notion in nonparametric regression is continuity; local data is relatively more informative than non-local data. The fitted yield at some point of interest (t_0), denoted $\hat{g}(t_0)$, is derived from a weighted average of neighboring yields where the weights are defined by the distance in the time space. Specifically, let $z_i = (t_i - t_0)/h$ denote the scaled, signed distance between the t_i and the point t_0. The scale factor h, called *bandwidth*, controls the relative weights. The general choice of $K(z)$ is the *Gaussian* or *normal* kernel. Define a $T \times T$ diagonal matrix W_0 where $W_0[i, i] = K((t_i - t_0)/h)$. The estimated coefficients are derived from the locally weighted least squares regression

$$\hat{\beta_0} = (\mathbf{X}'\mathbf{W}^0\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}^0 y \quad (2)$$

[3] We do not impose the spatial and temporal priors on knots used by the RMA.

[4] Let $\tilde{\varepsilon}_t$ be the estimated residuals from the robust spline with the chosen number of knots and $\tilde{\eta}_t = \tilde{\varepsilon}_t / \sqrt{T^{-1} \sum \tilde{\varepsilon}_t^2}$. 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 $c = 4.685$.

where y is the vector of yields, and X is a $T \times 2$ matrix with the first column being a vector of ones and the second column a sequence from 1 to T. The fitted value, denoted $\hat{y}_0 = X_0\hat{\beta}_0$ where $X_0 = (1, t_0)$, can be calculated for any point $t_0 \in R$. As is commonly done, we undertake at points $t = 1, ..., T$ to recover $\hat{y}_1, ..., \hat{y}_T$, a set of fitted annual yield values. The bandwidth h is chosen by minimizing an out-of-sample cross-validation one-step ahead prediction error squared metric.

Heteroscedasticity Treatment

Denote the residuals from any of the above 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[5]

$$\ln(\hat{\varepsilon}_t^2) = \alpha + \gamma \ln(\hat{y}_t) + v_t. \tag{3}$$

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 ($\hat{\gamma}$):[6]

$$\hat{y}_t^* = \hat{y}_{T+1} + \hat{\varepsilon}_t \left(\frac{\hat{y}_{T+1}}{\hat{y}_t}\right)^{\frac{\hat{\gamma}}{2}} \tag{4}$$

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

$$\pi_{T+1}^* = \frac{1}{T}\sum_{t=1}^{T} \max\left\{0,\ \lambda\hat{y}_{T+1}^* - \hat{y}_t^*\right\} \tag{5}$$

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

Testing the Identically Distributed Assumption

When testing for structural change, generally a Chow-type test is used where the sample is split into different sub-populations and residuals from regression equations within the sub-populations 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 breakpoint is known or the number of breakpoints. These tests have power only against changes in the conditional mean function (first moment) and thus, unsurprisingly, resulted in very little rejections on the adjusted yields across the crop-county combinations.[7] We are interested in structural changes in the higher moments of the dgp, beyond the conditional mean or variance. Therefore, we necessarily use a Kolmogorov-Smirnov (KS) test which considers the maximum difference between two empirical distribution functions and thus has power against differences in all moments. Note, the test is nonparametric in that the test statistic is a function of the empirical distribution functions. Also, the KS test has been shown to have relatively low power

[5] See Ker and Tolhurst (2019) for an alternative methodology. Our findings are robust to either heteroscedasticity treatment.

[6] RMA 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 very short time series.

[7] 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 that detrending process.

in comparison to 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 defined as:

$$D_{n,m} = \sup_x \mid F_{1,n}(x) - F_{2,m}(x) \mid, \tag{6}$$

where $F_{1,n}$ and $F_{2,m}$ are the empirical distribution functions of the first and the second sample respectively. Specifically, the raw yield data are detrended (using one of the four methods) and corrected for heteroscedasticity. The adjusted yield data are subsequently divided into two samples. Given the breakpoint is unknown and we have no prior we consider every fifth year between 1968 and 2003 as possible breakpoints (8 possible breakpoints).[8] Unfortunately, there is not an analogous Bai-Perron generalization of the Chow test for the KS test; multiple testing upward biases the level of our tests. The null of the KS test is rejected at level α when

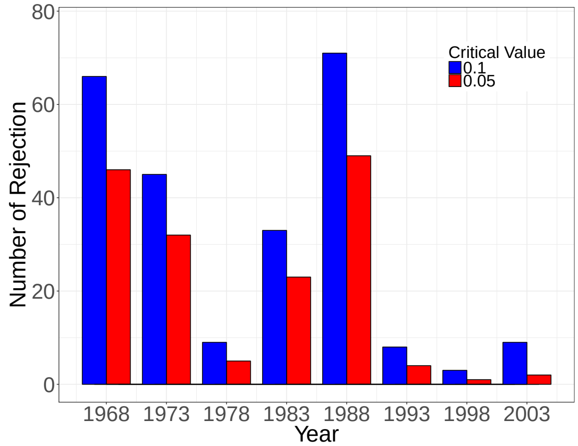
$$D_{n,m} > c(\alpha)\sqrt{\frac{n+m}{nm}}, \tag{7}$$

where $c(\alpha)$ is calculated from the Kolmogorov distribution. All tests were conducted at the 10% and 5% significance level given our relatively small sample and low power of the test.

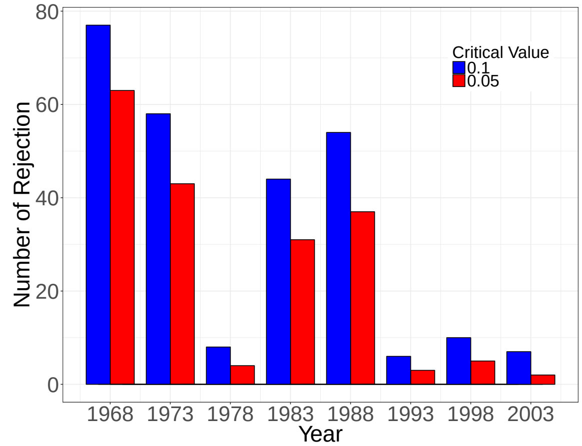
A histogram of the test results from different detrending methodologies are shown in Figures 1-3 for each of the three crops. The results are fairly consistent across the different detrending methodologies. For example, when the critical value of the test is 10% (blue), in corn with 414 counties, the number of rejections tend to be just over 50% clustering around late 1960s and late 1980s. With respect to soybean with 373 counties, the number of rejections tend to be around 30% clustering toward the late 1960s and early 1970s. With respect to winter wheat with just 64 counties, the number of rejections tend to be 70% with again clustering around the late 1960s. When the size of the test is reduced to 0.05 approximately one third of the rejections disappear for corn, one half for soybean, and one quarter for wheat. In response to the multiple testing problem, if we use the excessively conservative Holm-Bonferroni method (assumes independence) for multiple testing (Holm, 1979), all rejections for all three crops disappear. Nonetheless, given the downward bias in the size of the test in small samples, we find the results bring into question the identically distributed assumption.[9]

[8]As is customary, we do not test close to the boundaries of the sample.

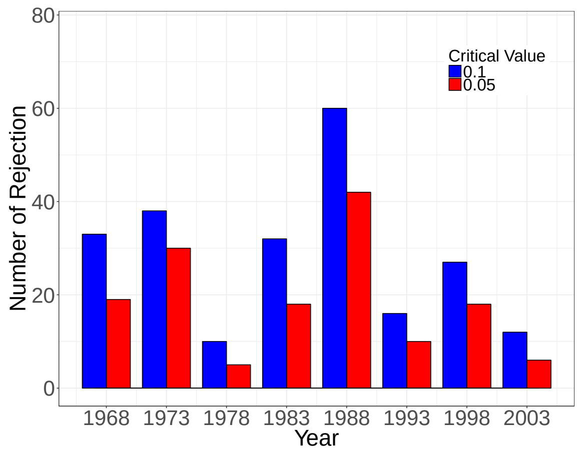
[9]Maps of the test results per county are available from the authors. The results did not illustrate any significant spatial clustering in the rejections.



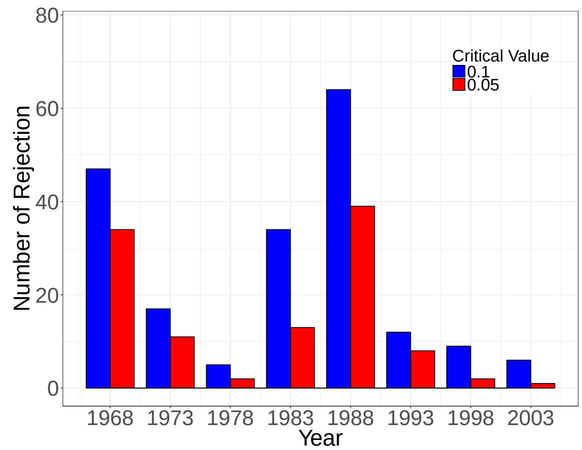
(A) RMA Detrending



(B) L_2 Detrending



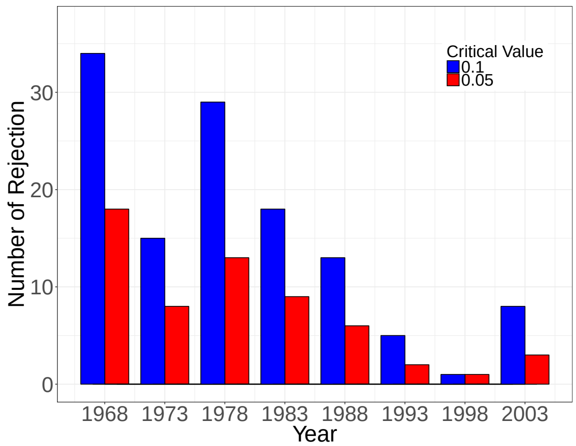
(C) L_1 Detrending



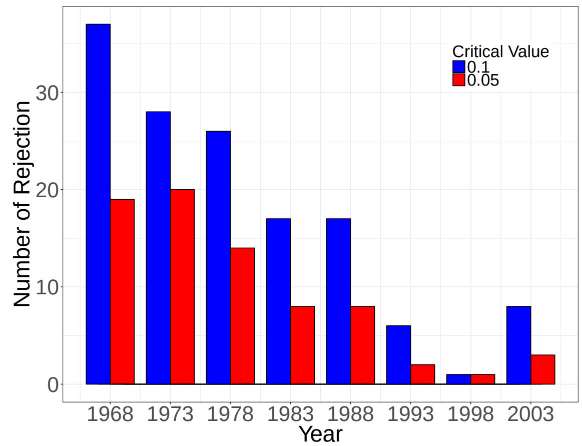
(D) NP Detrending

FIGURE 1. KS Test Rejections for Corn by Detrending Method.

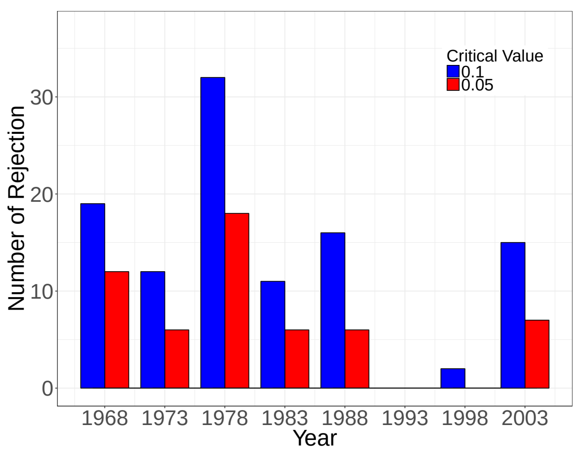
Note: For counties with multiple rejections, only the year with maximum KS test statistic is shown. Given 414 total counties, the number of rejections with size of the test 0.10(0.05) is 244(162) using RMA detrending, is 264(188) using L_2, is 228(148) using L_1, and is 194(110) using nonparametric is 194.



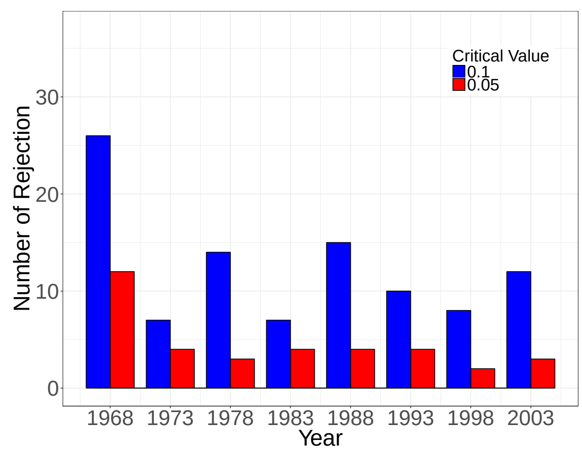
(A) RMA Detrending



(B) L_2 Detrending



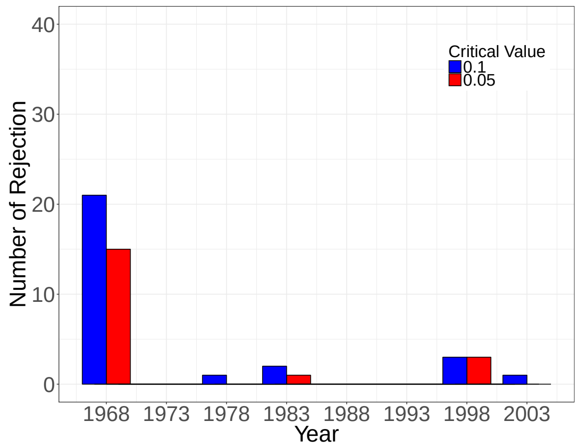
(C) L_1 Detrending



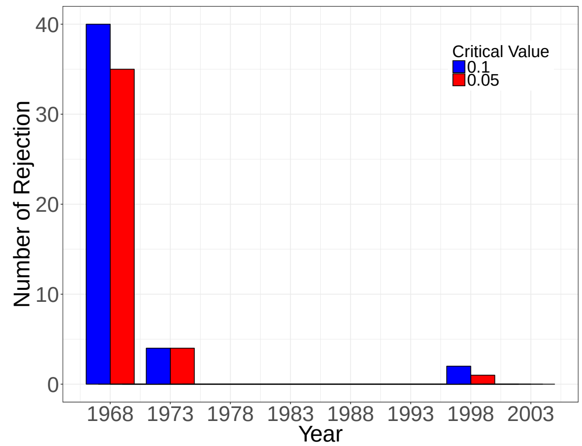
(D) NP Detrending

Figure 2. KS Test Rejections for Soybean by Detrending Method.

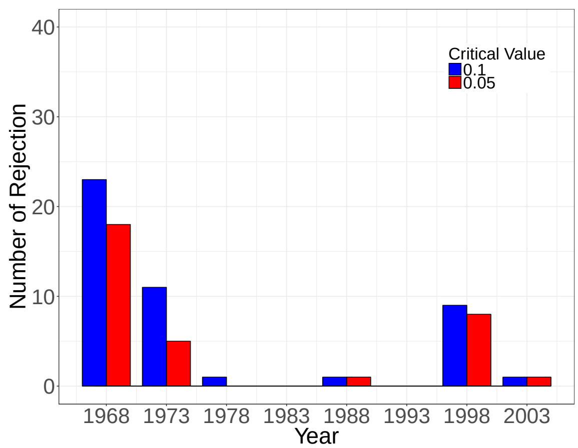
Note: For counties with multiple rejections, only the year with maximum KS test statistic is shown. Given 373 total counties, the number of rejections with size of the test 0.10(0.05) is 123(60) using RMA detrending, is 140(75) using L_2, is 107(55) using L_1, and is 99(36) using nonparametric is 194.



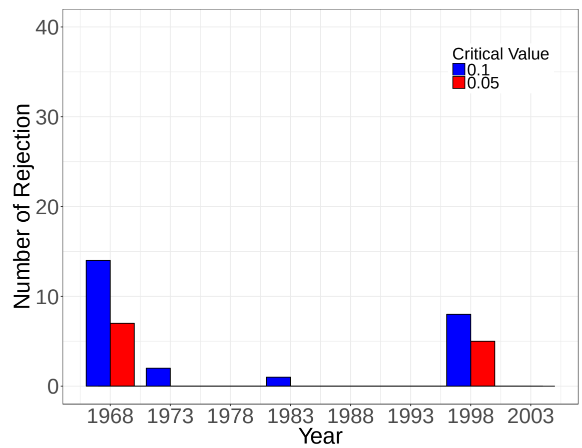
(A) RMA Detrending



(B) L_2 Detrending



(C) L_1 Detrending



(D) NP Detrending

FIGURE 3. KS Test Rejections for Winter Wheat by Detrending Method.

Note:
For counties with multiple rejections, only the year with maximum KS test statistic is shown. Given 64 total counties, the number of rejections with size of the test 0.10(0.05) is 28(19) using RMA detrending, is 46(40) using L_2, is 46(33) using L_1, and is 25(12) using nonparametric is 194.

Crop Insurance Application

Results from the previous section call into question the identically distributed assumption from a statistical perspective, but provide no information regarding economic importance. We consider an application to rating crop insurance contracts consistent with the literature. Specifically, we conduct an out-of-sample retain-cede rating game, where two players using different methodologies estimate premium rates and averse select against one another. The game was first proposed by Ker and McGowan (2000) and has since been employed by Ker and Coble (2003), Racine and Ker (2006), Harri et al. (2011), Annan et al. (2013), Tack and Ubilava (2015), Tolhurst and Ker (2015), Yvette Zhang (2017), and Shen, Odening, and Okhrin (2018) to justify alternative rating methodologies. The game was modified with an additional test of rating efficiency in Ker, Tolhurst, and Liu (2016). Park, Brorsen, and Harri (2018) utilized both tests in proposing an alternative rating methodology which exploits spatial closeness. The game was inspired by the retain-cede decision of private insurance companies in regards to the crop insurance contracts they sell. Some salient features of the U.S. crop insurance program are relevant to the game. First, RMA rather than private insurers set the rates for all policies. Second, the private insurer must sell all policies in a state that it operates in (even if it deems the policy to be under-priced). Third, the private insurer shares, asymmetrically, in the underwriting gains and losses of all policies it sells. Fourth, there is a mechanism in which private insurers can significantly reduce their exposure on policies they deem unwanted.[10] Given these salient features, private insurers determine which policies to retain and which to cede; that is, policies insurers believe over-priced and expect an underwriting gain (retain) versus policies insurers believe under-priced and expect an underwriting loss (cede). As a result, private insurers necessarily develop their own rates in attempts to strategically averse select against RMA and recover excess rents. Mimicking this allows one to hypothetically compare two sets of premium rates. In contrast to the past literature that employs the retain-cede game to evaluate alternative rating methodologies using the same data, we evaluate alternative data sets -- full versus historically restricted -- using the same rating methodology.

More specifically, we assume RMA uses the full historical yield data from 1951-2002 on a county-crop basis to estimate the RMA premium rates for 2003. Conversely, the private insurer estimates their rates using a restricted data set (truncation rules are defined below). Both the RMA and the private insurer use the RMA rating methodology and as such the only difference in the two sets of rates is the result of truncating 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 2003. This process is repeated for 15 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 at the 70%, 80%, and 90% coverage levels. We consider two truncation rules for the private insurer: (i) a fixed 30-year cut-off; and (ii) calculate the KS test statistic at five-year intervals and truncate at the maximum test statistic.

[10] Specific details are outlined in the USDA-RMA Standard Reinsurance Agreement with approved private insurers.

As in the above 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). Randomization methods are used to recover the p-value. Game 1 mimics the current reality of the US crop insurance program. However, the private insurer has an advantage because they react to the RMA premium rates. As such, whichever of the two competing rates the private insurer uses 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 (see Ker, Tolhurst, and Liu, 2016, for details). The number of contracts considered is the number of counties multiplied by 15 years; 6,210 contracts for corn, 5,595 contracts for soybean, and 960 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 1.

Under a 30-year fixed cut-off decision rule, we find the private insurer's loss ratio is less than the RMA loss ratio for all nine cases; that is, for all three crops at all three coverage levels. The private insurer loss ratio ranges from 38% to 77% of the RMA loss ratio. We find the differences statistically significant in six of the nine cases at the 5% level; all three coverage levels for corn, the 70% coverage level for soybean, and the 70% and 90% coverage levels for wheat. Even under the excessively conservative Holm-Bonferroni method for multiple testing (by crop across coverage levels at 5% significance), all six remain significant. These results strongly suggest that private insurers can recover statistically and economically significant rents by averse selecting against the RMA via simply restricting the length of the historical yield data to 30 years in the rating process. With respect to game 2, the historically restricted rates are more accurate than the unrestricted rates in all nine cases but the difference is only significant for soybeans at the 70% coverage level.

Under a more sophisticated decision rule using the KS test, the results are very similar. We again find the loss ratio for the retained contracts is less than the loss ratio for the ceded contracts in all nine cases. The private insurer loss ratio ranges from 44% to 89% of the RMA loss ratio; slightly above that with the fixed truncation rule. We find the differences statistically significant in seven of the nine cases at the 5% level; all three coverage levels for corn, the 70% and 90% coverage level for soybean, and the 70% and 80% coverage levels for wheat. If we consider the Holm-Bonferroni method for multiple testing, all seven remain significant. These results also suggest that private insurers can recover statistically and economically significant rents by averse selecting against the RMA via simply restricting the length of the historical yield data. With respect to game 2, the historically restricted rates are as or more accurate than the unrestricted rates in seven of the nine cases and those differences are significant for both soybean and wheat at the 70% coverage level; slightly greater significance than with the 30 year fixed truncation rule. In no cases are the unrestricted rates shown to be statistically significantly more accurate than the unrestricted rates.

TABLE 1. Out-Of-Sample Retain-Cede Rating Game

Crop-Coverage	Retained by Private (%)	Loss Ratio RMA	Loss Ratio Private	Game 1 p-value	Game 2 p-value
Truncated by 30-Year Cutoff:					
Corn					
70% Coverage	54.0	0.398	0.272	0.0037	0.1509
80% Coverage	58.8	0.521	0.300	0.0176	0.1509
90% Coverage	66.0	0.672	0.386	0.0037	0.3036
Soybean					
70% Coverage	62.5	0.538	0.272	0.0037	0.0176
80% Coverage	61.2	0.918	0.347	0.1509	0.5000
90% Coverage	64.9	1.122	0.476	0.1509	0.1509
Winter Wheat					
70% Coverage	29.2	1.135	0.721	0.0176	0.3036
80% Coverage	28.4	1.008	0.773	0.0592	0.1509
90% Coverage	39.5	1.088	0.760	0.0037	0.1509
Truncated by Maximum KS Test Statistic:					
Corn					
70% Coverage	56.9	0.338	0.301	0.0037	0.3036
80% Coverage	60.7	0.436	0.338	0.0176	0.6964
90% Coverage	66.3	0.525	0.436	0.0176	0.3036
Soybean					
70% Coverage	64.9	0.478	0.295	0.0176	0.0037
80% Coverage	64.5	0.862	0.375	0.1509	0.5000
90% Coverage	67.8	1.043	0.508	0.0592	0.3036
Winter Wheat					
70% Coverage	39.6	1.351	0.614	0.0037	0.0176
80% Coverage	41.8	1.131	0.698	0.0176	0.5000
90% Coverage	49.2	1.023	0.844	0.1509	0.6964

Note: Games conducted over 15 years (2003–2017). Game 1 is the adverse selection game of Ker and McGowan (2000) and game 2 is the efficiency game of Ker, Tolhurst, and Liu (2016). Reported p-values for both are binomial and based on annual Bernoulli trial with number of trials equal to 15 and binomial parameter equal to 0.5 (under null).

Estimating and Simulating Yield DGPs

Ideally, we would know the dgp of yields and be able to quantify the error from using too much historical yield data in the estimation process. Our third of the three-prong approach attempts to quantify the unknown error by assuming a known dgp. For each county-crop combination, we estimate the yield dgp as a mixture of normals with time-varying mixing, location, and variance parameters; an exceedingly flexible form meant to capture movement in the higher moments of the yield dgp. Second, assuming these estimates to be true, we sample various lengths of historical yields and quantify the estimation error from correcting only the first two moments. Specifically, we estimate

$$y_t \sim \sum_{k=1}^{K} \lambda_{tk} N(\alpha_k + \beta_k t,\ \gamma_k + \delta_k t), \tag{8}$$

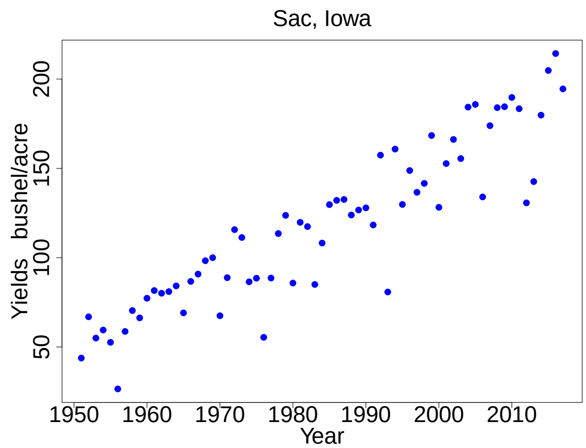
where y_t is the yield in year t, K is the number of mixtures, λ_{tk} is weight given to the kth mixture in year t satisfying $\sum_{k=1}^{K} \lambda_{tk} = 1\ \forall\ t$; $\alpha_k + \beta_k t$ are the mean parameters for mixture k; and, $\gamma_k + \delta_k t$ are the time-varying variance parameters for mixture k. We estimate the mixture model using the Expectation-Maximization (EM) algorithm with multiple starting values and BIC to choose the number of mixtures. For illustrative purposes, the original yield data and a sample from the corresponding estimated time-varying mixture model are shown in Figure 4. The original yield data and the sampled yield data *appear* drawn from *near*-identical dgps.

To quantify the estimation error associated with using various lengths of historical yield data, we draw 500 time-dependent samples of size 67 years (1951-2017) from each estimated time-varying mixture model. We do this for each county-crop combination; 414 corn counties; 373 soybean counties; and 64 wheat counties. We truncate the sampled historical yield data to the most recent 10, 15, 20, ..., 55, and 67 years. As done in the literature, we detrend and correct for heteroscedasticity to recover 500 sets of adjusted yield data per county-crop combination. With the adjusted yield data, we estimate the conditional yield density at year 2018 and the associated premium rates at 70%, 80%, and 90% coverage levels. We detrend the data using the four previously described methodologies: RMA two-knot robust linear spline; linear line estimated by L_2; linear line estimated by L_1; and nonparameteric local lines. We correct for heteroscedasticity using method from Harri et al. (2011). Given we know the true dgp, we can calculate the estimation error according to any metric. With respect to the conditional density estimates, we necessarily use a function space metric; we use mean integrated squared error (MISE) which is the function space analogy to mean squared error. With respect to the estimated premium rates, we choose root mean squared error (RMSE). Figure 5 presents the average (across counties) of the MISE and RMSE for the 80% coverage level premium rates by crop and sample size. The results for the 70% and 90% are presented in the appendix.

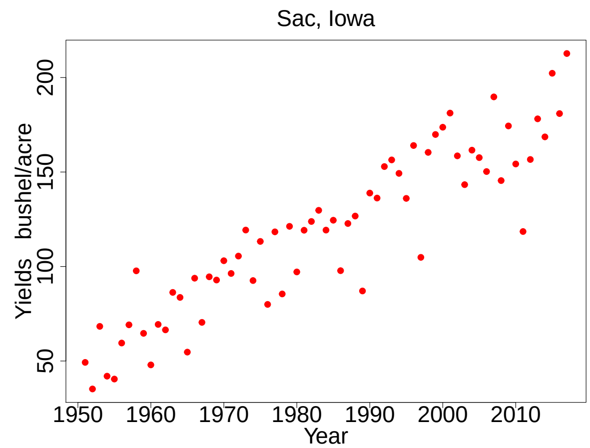
The results are consistent with the findings from the distributional tests and the out-of-sample retain-cede rating games. As one uses more historical yield data, estimation error decreases to a point and then begins to increase. Estimation error for both the conditional yield density and premium rates increases after roughly 30 years of historical yield data. The pattern emerges for all crops and is not dependent on detrending methodology.

Incorporating Data from Like dgps

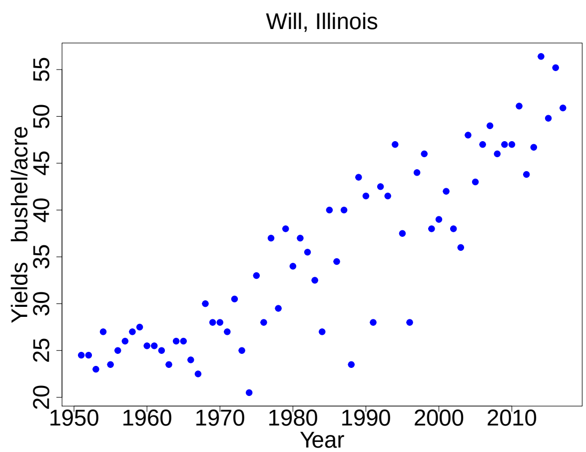
Collectively, the above findings suggest that historical yield data beyond 30 years should be discarded. While this is not surprising given the significant advances in technology and farm management, discarding



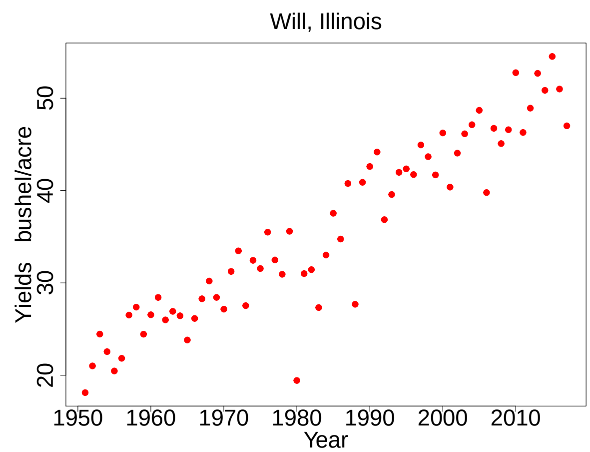
(A) Corn Yields



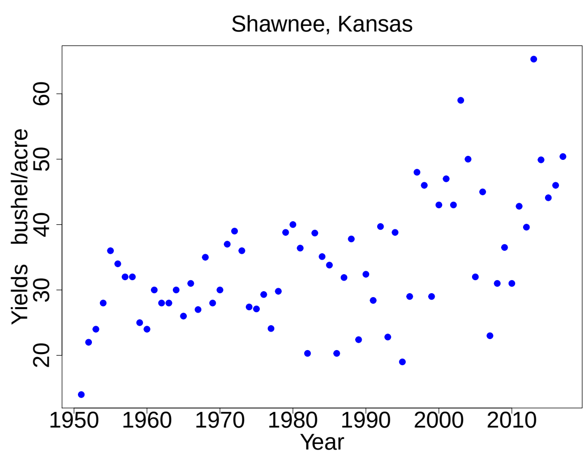
(B) Corn Yields(Simulated)



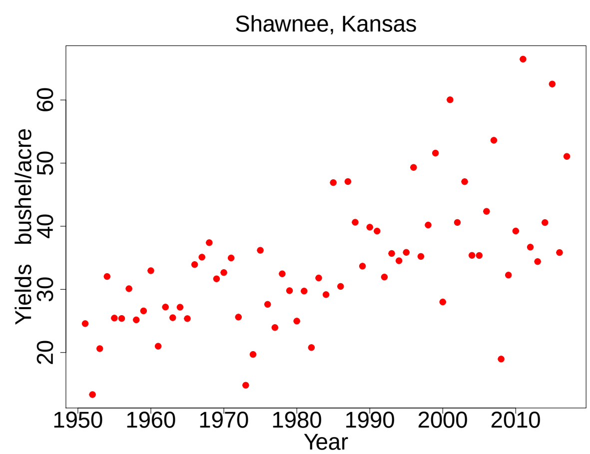
(C) Soybean Yields



(D) Soybean Yields(Simulated)

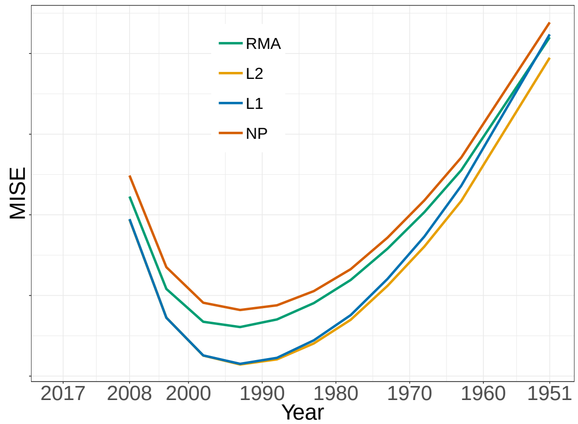


(E) Winter Wheat Yields

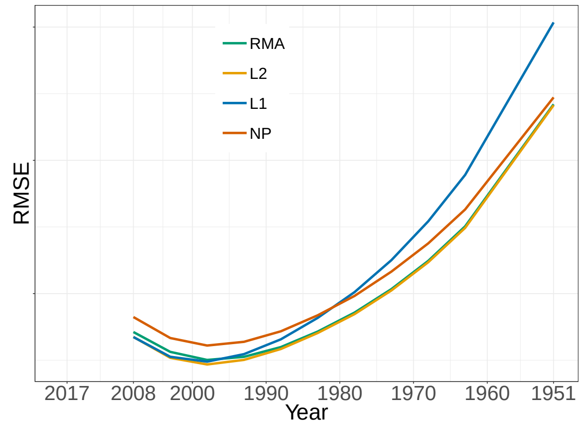


(F) Winter Wheat Yields(Simulated)

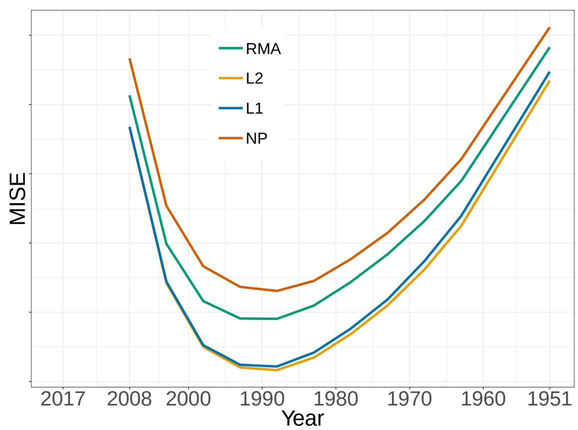
Figure 4. Actual & Simulated Crop Yields of Randomly Selected Counties



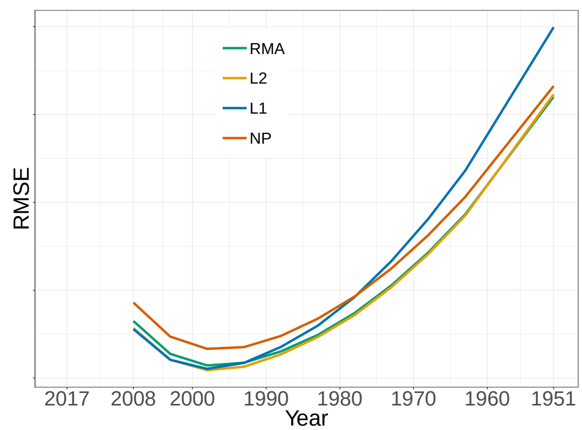
(A) Corn, MISE of Kernel Density



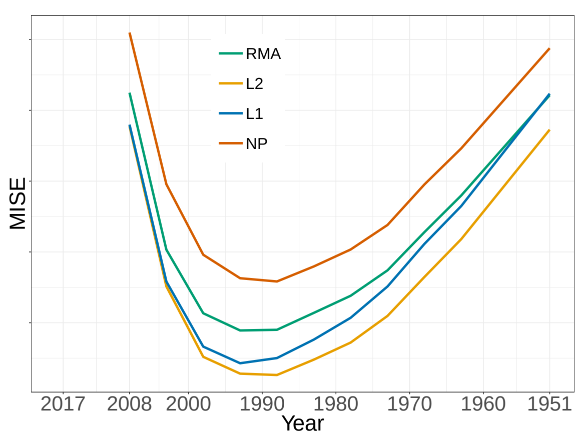
(B) Corn, RMSE of 80% Empirical Rate



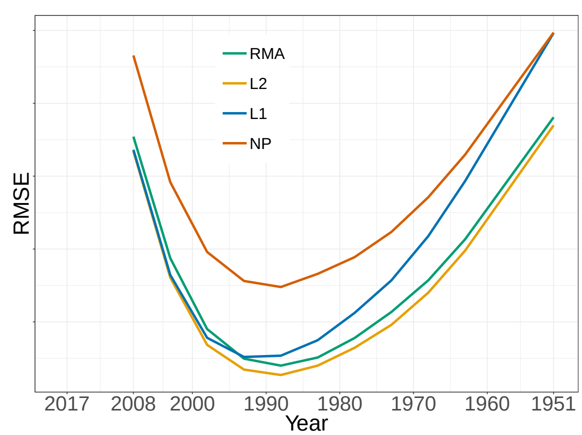
(C) Soybean, MISE of Kernel Density



(D) Soybean, RMSE of 80% Empirical Rate



(E) Winter Wheat, MISE of Kernel Density



(F) Winter Wheat, RMSE of 80% Empirical Rate

Figure 5. MISE & RMSE by Crop under Different Detrending Methods

50% or more of the historical yield data when $T = 67$ is already small is unappealing. Therefore, we consider three approaches that can incorporate the discarded data back into the estimation process while recognizing that the discarded data come from a different dgp. The first is a Bayesian Model Averaging (BMA) approach forwarded by Ker and Liu (2017) and smooths between multiple dgps. The second approach is generally used for smoothing across categories of independent variables forwarded by Li and Racine (2003), and Racine and Li (2004). The final approach forwarded by Ker (2016) reduces bias by combining data from multiple dgps to form a start estimate and then corrects that start according to the dgp of interest. All three methodologies are purposely nonparametric and thus our findings are not conditional on any parametric assumptions. We replicate our economic game and simulation analyses with these more flexible methodologies to determine if gains in efficiency may be realized by incorporating the discarded historical yield data in alternative ways.

For all three methods define the set of yields as $(y_1, \ldots, y_T)$, the truncation point as t_1, the discarded historical yields as $(y_1, \ldots, y_{t_1-1})$, and the recent yields as $(y_{t_1}, \ldots, y_T)$. Note, the discarded and recent yield sets partition the overall set of yields. The Nadaraya-Watson (NW) or the standard kernel estimator is the basis of all three estimators presented below. It is defined as:

$$\hat{f}(y) = (1/Th) \sum_{t=1}^{T} K((y - y_t)/h), \tag{9}$$

where K is kernel function assumed to be a square integrable symmetric probability density function with a finite second moment, h is the bandwidth or smoothing parameter chosen by likelihood cross-validation, and the summation is over set of interest.[11]

Bayesian Model Averaging Estimator (BMA)

Bayesian Model Averaging (BMA) has been widely used in the literature to deal with model uncertainty. For a comprehensive introduction, we refer to Hoeting et al. (1999). In this manuscript, we follow the estimator proposed in Ker and Liu (2017). In our application, the BMA estimator has the form

$$\tilde{f}(y) = \omega_r \hat{f}_r(y) + (1 - \omega_c)\hat{f}_d(y), \tag{10}$$

where $\hat{f}_r$ and $\hat{f}_d$ are density estimates from recent and discarded yield data sets respectively and ω_r is weight on recent yields. This weight is defined as

$$\omega_c = \frac{\exp\Big\{-\frac{1}{2}\text{BIC}_r\Big\}}{\exp\Big\{-\frac{1}{2}\text{BIC}_r\Big\} + \exp\Big\{-\frac{1}{2}\text{BIC}_d\Big\}}, \tag{11}$$

where BIC_r is the Bayesian information criterion (BIC) of model $\hat{f}_r$ with the recent yield data and BIC_d is the BIC of model $\hat{f}_d$ also with the recent yield data. Note, $\hat{f}_d$ is estimated with the historical yield data

[11] Kernel estimators essentially take a given point and spread the mass, via the kernel, around that point thereby adding variation. As a result, kernel estimators that use second order kernels (standard normal) have variance greater than the unbiased sample variance with that bias being of $O(T^{7/5})$ (Ker and Goodwin, 2000). This is non-trivial in samples of our size and can greatly increase the estimated premium rates. We rectify this bias by using a fourth-order kernel that yields an unbiased second moment in the estimated density.

but for the BIC the estimate is evaluated using the recent yield data. When uniform priors are assigned to both models $\hat{f}_r$ and $\hat{f}_d$, the weight ω_r assigned to the recent yields is determined by likelihood that the current data $(y_{t_1}, \ldots, y_T)$ comes from the estimated model $\hat{f}_r$ and $\hat{f}_d$ respectively, where the likelihood is represented by the BIC_r and BIC_d. We obtain BIC values by interpreting the standard kernel estimator as a mixture of normals with means equal to the sample points, variances equal to the smoothing parameter squared, and equal mixing weights. For a specific normal mixture model M_i, the likelihood that it generates the yield data D with sample size T $(d_1, \ldots, d_T)$ can be expressed as

$$\text{pr}(D \mid M_i) = L_i = \prod_{t=1}^{T} (1/Th) \sum \phi((d_t - y_i)/h), \tag{12}$$

which is the likelihood function of data D under the model assumption M_i that is specified by sample y_i and corresponding bandwidth h. Note, i takes on only two values in our case; the kernel estimate based on the discarded yields and the kernel estimate based on the recent yields. The advantage of this estimator is that it converges to the kernel estimate based on the recent yields as T grows, and, it does not require assumptions about the form or extent of similarity between the recent and discarded yield dgps. That is, the estimator does not require any assumptions on how similar the discarded yield data is to the more recent yield data.

Li-Racine Estimator (LR)

The Li-Racine (LR) estimator smooths across mixed data-types; both continuous and discrete. The continuous component is the random variable of interest, in our case yields, and the discrete component is the two different time periods (recent and discarded). The LR estimator essentially controls for the relative weights between the data from within a unit (recent yield data) versus outside the unit (discarded historical yield data) via a discrete smoothing parameter. The LR estimator has the following form

$$\hat{f}(y) = (Th)^{-1} \left[\underbrace{\sum_{t=1}^{t_1-1} (1-\lambda) K\left(\frac{y - y_t}{h}\right)}_{\text{discarded data}} + \underbrace{\sum_{t=t_1}^{T} \lambda K\left(\frac{y - y_t}{h}\right)}_{\text{recent data}} \right]. \tag{13}$$

Asymptotic normality of the estimator was established by Racine and Li (2004). They also showed that the convergence rate of the estimator is the same as the case when there are only continuous random variables. They show it outperforms the conventional frequency or bin estimator that was previously used to handle mixed data types. The frequency or bin estimator would correspond to the standard kernel using only the recent yield data in our case. The key to the LR estimator is that it uses cross-validation to choose the amount of smoothing between bins. As a result, if the data are very different between bins, LR smooths very little between bins, and the LR estimator behaves like the bin estimator. Conversely, if the data are similar between bins, LR smooths between bins, and gains in efficiency are realized. The estimator has been widely employed in the literature, including empirical analyses of firm distributions (Huynh et al., 2011), business formation (Vliamos and Tzeremes, 2012), economic growth (Anderson and Hachem, 2013), risk analysis (Hernandez and Torero, 2014), health care (Koch and Racine, 2016), and rating crop insurance contracts (Racine and Ker, 2006). The LR estimator also requires no assumptions about the form or extent of similarity between dgps. More details about Li-Racine estimator can be obtained from Li and Racine (2003); Racine and Li (2004); Li, Simar, and Zelenyuk (2016).

Possibly Similar Estimator (PS)

The PS estimator introduced by Ker (2016) is based on the standard kernel estimator but uses extraneous data to reduce bias. Bias is a function of global curvature and the estimator attempts to reduce the curvature of the underlying density to be estimated. This is done by first estimating a start density using all available data, in our case both the discarded and recent yield data, and then nonparametrically correct that start by only using the recent yield data. The PS estimator has the form:

$$\tilde{f}(y) = \hat{g}(y)\hat{r}(y) = \sum_{t=t_1}^{T} \frac{1}{h(T-t_1+1)} K\left(\frac{y-y_t}{h}\right)\frac{\hat{g}(y)}{\hat{g}(y_t)}, \tag{14}$$

where $\hat{g}$ is the start estimate with smoothing parameter h_p. Note that the summation is only over the recent yield series although as previously stated $\hat{g}$ is the kernel estimate based on the entire yield series. Ker (2016) found the PS estimator to perform very well in small samples. The asymptotic mean integrated squared error (AMISE) of the PS estimator is less than the standard kernel estimator within the neighborhood is:

$$\begin{aligned} 1/4h^4(R(g(f/g)'') - R(f'')) + 2(Nh_p)^{-1}\left[\int (K^2(u) - K(u)(K*K(u))du\right] \\ +(Nh_p)^{-1}\left[\int (K(u) - K*K(u))^2 du\right] < 0, \end{aligned} \tag{15}$$

where $u_2(K) = \int u^2K(u)du$ and $R(K) = \int K(u)^2du$.

Games and Simulations Revisited

Retain-Cede Rating Game

We consider possible efficiency gains from using BMA, LR, and PS. We repeat both the retain-cede rating games and the simulations from known dgps. The basis for comparison is the historically restricted estimates (densities and premium rates) as our earlier findings suggest truncating the data versus not is more efficient. That is, for the rating games we assume the RMA uses the premium rates based on the truncated yield data while the private insurer used BMA, LR, and PS estimators to derive their rates. The game results are shown in Tables 2 to 5. We maintained both 30-year fixed and maximum KS test statistic truncation rules.

Using BMA, we find the private insurer loss ratio is less than the RMA loss ratio in all nine cases with the fixed truncation rule and seven of the nine cases with the maximum KS test statistic truncation rule (see Table 2). The private insurer loss ratio ranges from 62% to 90% under fixed truncation, and 44% to 106% under KS truncation, of the RMA loss ratio. We find the differences statistically significant in eight of the 18 cases at the 5% significance level. Under the Holm-Bonferroni method for multiple testing (by crop across coverage levels at 5% significance), two of five remain significant for corn, three of three remain significant for soybean, and two of three remain significant for wheat. These results strongly suggest that private insurers using BMA can recover statistically and economically significant rents by averse selecting against the RMA. With respect to game 2, BMA rates are more accurate than the restricted rates in 16 of 18 cases of which eight are statistically significant. Interestingly, of the two cases that are not more accurate using the BMA, one is statistically significant. In Table 5, the mean and median weights assigned by the BMA estimation method to the recent yield data are given. In general, most of the weight is given to the recent yield data and very little to the discarded yield data. In fact, for corn and soybean the median weight is almost zero on the discarded yield data.

Using LR, we find the private insurer loss ratio is less than the RMA loss ratio in six of the nine cases with both the fixed and flexible truncation rules (see Table 3). The private insurer loss ratio ranges from 53% to 113% under fixed truncation, and 44% to 126% under KS truncation, of the RMA loss ratio. We find the differences statistically significant in 11 of the 18 cases at the 5% significance level. Under the Holm-Bonferroni method for multiple testing, all remain significant save one with wheat. These results, consistent with the BMA results, suggest private insurers using LR can recover statistically and economically significant rents by averse selecting against the RMA. With respect to game 2, LR rates are more accurate than the restricted rates in 15 of 18 cases of which eight are statistically significant. None of the three cases where rates are less accurate are statistically significant. In Table 5, the mean and median weights assigned by the LR estimation method to the recent yield data are also given. In general, most of the weight is given to the recent yield data but markedly less so than with the BMA.

Finally, using PS, we find the private insurer loss ratio is less than the RMA loss ratio in six of the nine cases with the fixed truncation rule and eight of the nine cases with the KS truncation rule. The private insurer loss ratio ranges from 54% to 112% under fixed truncation, and, 31% to 127% under KS truncation, relative to the RMA loss ratio. We find the differences statistically significant in eight of the 18 cases at the 5% significance level. Under the Holm-Bonferroni method for multiple testing, all remain significant for corn and soybean, none for wheat. These results, consistent with both the BMA and LR results, suggest private insurers using PS can recover statistically and economically significant rents by averse selecting against the RMA. With respect to game 2, PS rates are more accurate than the restricted rates in 12 of 18 cases of which seven are statistically significant. None of the six cases where PS rates are less accurate are statistically significant. Given the construction of the PS estimator, the weight put on the recent versus discarded yield data cannot be recovered.

Table 2. Out-Of-Sample Retain-Cede Rating Game, BMA Method

Crop-Coverage	Retained by Private (%)	Loss Ratio RMA	Loss Ratio Private	Game 1 p-value	Game 2 p-value
Truncated by 30-Year Cutoff:					
Corn					
70% Coverage	79.8	0.549	0.346	0.0000	0.0037
80% Coverage	74.3	0.495	0.405	0.1509	0.1509
90% Coverage	31.4	0.548	0.493	0.0176	0.1509
Soybean					
70% Coverage	66.0	0.670	0.419	0.0000	0.0000
80% Coverage	70.1	0.712	0.571	0.0037	0.0037
90% Coverage	39.2	0.746	0.741	0.0592	0.0592
Winter Wheat					
70% Coverage	63.9	1.141	0.752	0.0005	0.0037
80% Coverage	35.9	0.964	0.656	0.0176	0.0592
90% Coverage	8.6	0.937	0.721	0.3036	0.9408
Truncated by Maximum KS Test Statistic:					
Corn					
70% Coverage	79.0	0.792	0.348	0.0037	0.0000
80% Coverage	73.1	0.601	0.414	0.0176	0.0176
90% Coverage	33.2	0.579	0.539	0.0176	0.1509
Soybean					
70% Coverage	69.0	0.776	0.436	0.0000	0.0000
80% Coverage	69.1	0.573	0.679	0.0592	0.0592
90% Coverage	38.4	0.793	0.752	0.1509	0.3036
Winter Wheat					
70% Coverage	60.3	1.005	0.917	0.0037	0.0037
80% Coverage	39.1	0.971	0.795	0.0592	0.1509
90% Coverage	10.3	0.976	1.035	0.5000	0.9824

Note: Games conducted over 15 years (2003–2017). Game 1 is the adverse selection game of Ker and McGowan (2000) and game 2 is the efficiency game of Ker, Tolhurst, and Liu (2016). Reported p-values for both are binomial and based on annual Bernoulli trial with number of trials equal to 15 and binomial parameter equal to 0.5 (under null).

Simulations

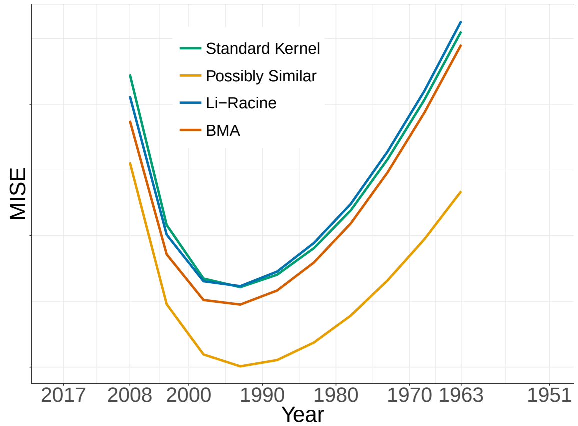
Recall, we previously estimated the county-crop yield dgps using a time-varying mixture model outlined in equation 8. In this subsection, we again recover 500 simulated yield series for each county-crop combination and quantify the estimation error (MISE and RMSE) by sample size (or truncation point). We compare the bin estimator (standard kernel based on the recent yield only), the BMA estimator, the LR estimator, and the PS estimator. The results are presented in Figure 6 (70% and 90% coverage levels are in the appendix). As expected, the errors (MISE and RMSE at all three coverage levels) for all three methods, BMA, LR,

Table 3. Out-Of-Sample Retain-Cede Rating Game, Li-Racine Method

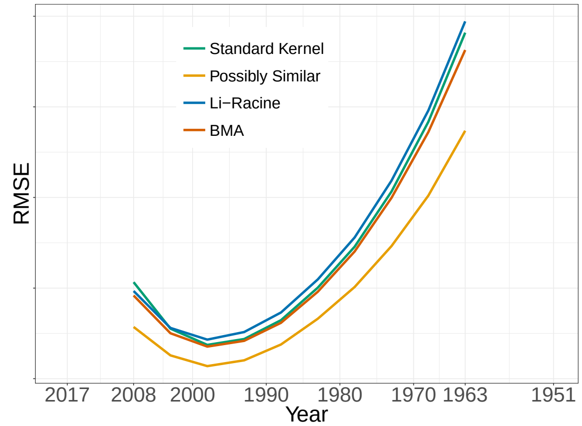
Crop-Coverage	Retained by Private (%)	Loss Ratio RMA	Loss Ratio Private	Game 1 p-value	Game 2 p-value
Truncated by 30-Year Cutoff:					
Corn					
70% Coverage	81.6	0.529	0.354	0.0000	0.0037
80% Coverage	71.0	0.421	0.432	0.5000	0.1509
90% Coverage	27.1	0.512	0.580	0.3036	0.5000
Soybean					
70% Coverage	67.4	0.756	0.403	0.0000	0.0000
80% Coverage	69.3	0.667	0.587	0.0037	0.0005
90% Coverage	35.6	0.732	0.769	0.0592	0.1509
Winter Wheat					
70% Coverage	68.9	1.336	0.709	0.0005	0.0005
80% Coverage	46.8	0.942	0.778	0.0176	0.1509
90% Coverage	15.7	0.937	0.831	0.0592	0.8491
Truncated by Maximum KS Test Statistic:					
Corn					
70% Coverage	78.3	0.730	0.355	0.0000	0.0000
80% Coverage	67.7	0.574	0.412	0.1509	0.1509
90% Coverage	26.1	0.580	0.530	0.5000	0.6964
Soybean					
70% Coverage	69.7	0.784	0.433	0.0000	0.0000
80% Coverage	67.3	0.551	0.694	0.0592	0.0037
90% Coverage	33.0	0.772	0.794	0.1509	0.1509
Winter Wheat					
70% Coverage	63.2	1.034	0.902	0.0005	0.0005
80% Coverage	43.9	0.990	0.813	0.0592	0.0592
90% Coverage	15.2	0.981	0.982	0.3036	0.9408

Note: Games conducted over 15 years (2003–2017). Game 1 is the adverse selection game of Ker and McGowan (2000) and game 2 is the efficiency game of Ker, Tolhurst, and Liu (2016). Reported p-values for both are binomial and based on annual Bernoulli trial with number of trials equal to 15 and binomial parameter equal to 0.5 (under null).

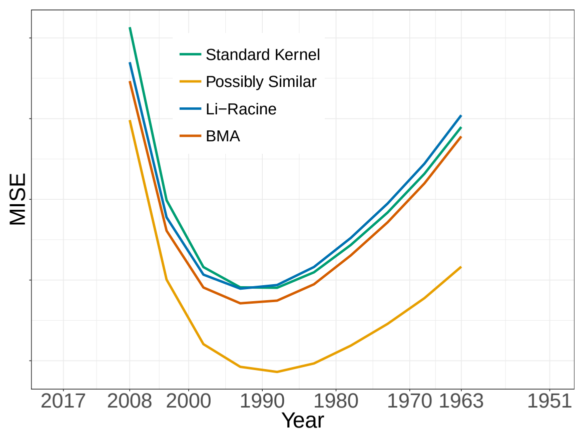
and PS, are bounded above by the bin estimator. Unlike the retain-cede rating game results which suggest that all three methods perform somewhat similarly, PS appears to perform markedly better than BMA and LR under this *constructed* example. Note, the PS estimator is a bias reduction method while the LR and BMA are variance reduction methods. Also quite interestingly, the minimum appears to be around 1990 or 30 years for the truncation point across all three crops for all three estimators.



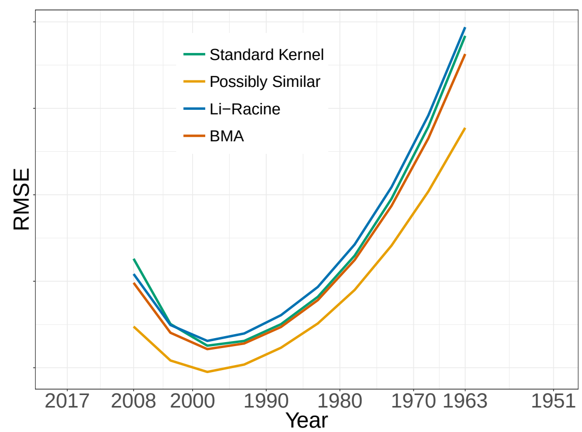
(A) Corn, MISE



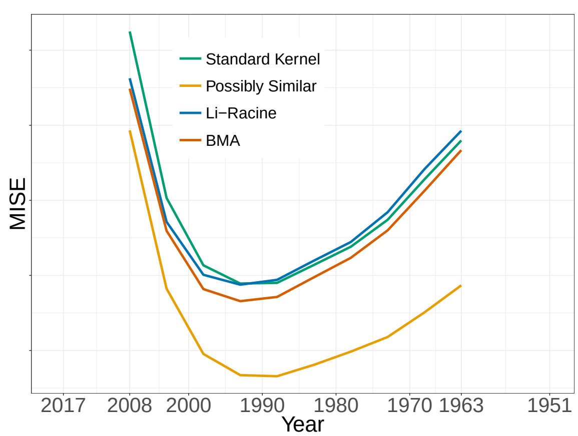
(B) Corn, RMSE of 80% Rate



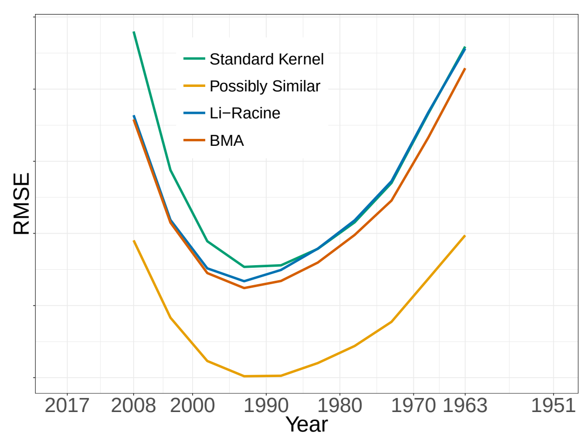
(C) Soybean, MISE



(D) Soybean, RMSE of 80% Rate



(E) Winter Wheat, MISE



(F) Winter Wheat, RMSE of 80% Rate

FIGURE 6. MISE & RMSE by Crop and Smoothing Methodology

Table 4. Out-Of-Sample Retain-Cede Rating Game, Possibly Similar Method

Crop-Coverage	Retained by Private (%)	Loss Ratio RMA	Loss Ratio Private	Game 1 p-value	Game 2 p-value
Truncated by 30-Year Cutoff:					
Corn					
70% Coverage	81.5	0.626	0.341	0.0000	0.0037
80% Coverage	70.1	0.456	0.417	0.5000	0.8491
90% Coverage	29.2	0.540	0.509	0.5000	0.9408
Soybean					
70% Coverage	64.6	0.737	0.415	0.0000	0.0000
80% Coverage	69.6	0.562	0.631	0.0592	0.0037
90% Coverage	40.5	0.727	0.769	0.6964	0.9408
Winter Wheat					
70% Coverage	71.8	1.239	0.778	0.0176	0.0005
80% Coverage	51.1	0.894	0.839	0.0592	0.0592
90% Coverage	20.1	0.897	0.995	0.5000	0.9408
Truncated by Maximum KS Test Statistic:					
Corn					
70% Coverage	79.2	1.063	0.325	0.0000	0.0000
80% Coverage	70.1	0.560	0.425	0.0037	0.3036
90% Coverage	31.9	0.588	0.520	0.5000	0.3036
Soybean					
70% Coverage	67.1	0.813	0.437	0.0000	0.0000
80% Coverage	67.6	0.546	0.691	0.0592	0.0592
90% Coverage	41.5	0.793	0.758	0.3036	0.8491
Winter Wheat					
70% Coverage	70.5	1.092	0.895	0.0176	0.0037
80% Coverage	55.1	1.012	0.838	0.0037	0.0592
90% Coverage	24.1	1.039	0.835	0.5000	0.9408

Note: Games conducted over 15 years (2003–2017). Game 1 is the adverse selection game of Ker and McGowan (2000) and game 2 is the efficiency game of Ker, Tolhurst, and Liu (2016). Reported p-values for both are binomial and based on annual Bernoulli trial with number of trials equal to 15 and binomial parameter equal to 0.5 (under null).

Conclusions

Historical yield data has been utilized in many empirical applications, most notably, applications related to crop insurance. In general, yield data is adjusted for time-varying lower moments and assumed to have time-constant upper moments. However, there has been significant innovations in farm management and seed technologies such that mass has likely moved all around the yield distribution. Not surprisingly, many have found changes in the upper moments of the yield dgps thus nullifying the standard two-step procedure which simply corrects for changes in the first two moments. In this manuscript, we attempt to answer the question

Table 5. Weight Table of BMA and LR Methods

	BMA (ω_r)		LR (λ)	
Crop	Mean	Median	Mean	Median
Truncated by 30-Year Cutoff				
Corn	0.8376	0.9970	0.8020	0.8959
Soybean	0.8661	0.9932	0.7577	0.8071
Winter Wheat	0.8273	0.9661	0.6805	0.5000
Truncated by Maximum KS Test Statistic:				
Corn	0.8867	0.9995	0.8677	0.9744
Soybean	0.8924	0.9982	0.8242	0.9239
Winter Wheat	0.9092	0.9961	0.7617	0.8051

how much historical yield data should be used in empirical analysis. While the answer is dependent on the crop-region combination, empirical question at hand, loss function, and econometric methodology, we used a three-prong approach to provide some insight. First, we found statistical evidence, albeit weak, for non-identical dgps using KS tests; not overly surprising given the test is known to under reject in small samples. Second, we undertook an out-of-sample retain-cede rating game commonly employed in the literature. We compared the same rating methodology using the full versus historically truncated yield data. We find strong evidence the private insurers can gain statistically significant excess rents averse selecting against RMA by simply restricting the amount of historical yield data in the rating process. While we find less statistical evidence that truncating produces more accurate rates, we find no evidence that truncating produces less accurate rates. Third, we simulated yield data from time-varying dgps and recovered the error in estimating premium rates as well as the conditional yield densities. Recognizing this is a constructed exercise, we find that error begins to increase after approximately 30 years of historical yield data.

The notion of discarding data in any situation is unappealing. As a result, we considered three flexible methodologies (BMA, LR, and PS) that allow one to borrow information from like dgps without assuming the extent or form to which those dgps are alike. We compare these estimators to the results using the truncated yield series. We find fairly strong evidence the private insurers can gain statistically significant excess rents averse selecting against RMA by incorporating the historical yield data back into the rating process in alternative ways. We also find statistical evidence that incorporating the historical yield data produces rates that are more accurate. Quite interestingly, the estimators significantly downweight the historical yield data relative to the current yield data. Our simulations from known dgps are also consistent in that we find efficiency gains may be realized using the BMA, LR, or PS estimators.

Our findings may help clarify for RMA the inconsistency in their rating methodologies. Recall, RMA uses the yields dating back to the 1950s to rate their area-yield and revenue products but uses only yields dating back to 1991 to rate their newer shallow loss products. Our results suggest that the latter rating methodology appears to be the preferred approach. Finally, our findings support the more recent literature that chooses to restrict the amount of historical yield data used in their empirical analyses.

References

Anderson, G., and K. Hachem. 2013. "Institutions and economic outcomes: A dominance-based analysis." *Econometric Reviews* 32:164–182.

Annan, F., J. Tack, A. Harri, and K. Coble. 2013. "Spatial Pattern of Yield Distributions: Implications for Crop Insurance." *American Journal of Agricultural Economics*, pp. 253–268.

Assefa, Y., P. Prasad, P. Carter, M. Hinds, G. Bhalla, R. Schon, M. Jeschke, S. Paszkiewicz, and I.A. Ciampitti. 2017. "A New Insight into Corn Yield: Trends from 1987 through 2015." *Crop Science*, pp. 1–10.

Challinor, A.J., J. Watson, D. Lobell, S. Howden, D. Smith, and N. Chhetri. 2014. "A meta-analysis of crop yield under climate change and adaptation." *Nature Climate Change* 4:287–291.

Claassen, R., D. Hellerstein, and S.G. Kim. 2013. "Using mixed logit in land use models: can expectation-maximization (EM) algorithms facilitate estimation?" *American Journal of Agricultural Economics* 95:419–425.

Claassen, R., and R.E. Just. 2011. "Heterogeneity and distributional form of farm-level yields." *American Journal of Agricultural Economics* 93:144–160.

Claassen, R., C. Langpap, and J. Wu. 2017. "Impacts of Federal Crop Insurance on Land Use and Environmental Quality." *American Journal of Agricultural Economics* 99:592–613.

Cooper, J., A. Nam Tran, and S. Wallander. 2017. "Testing for Specification Bias with a Flexible Fourier Transform Model for Crop Yields." *American Journal of Agricultural Economics* 99:800–817.

Duvick, D.N. 2005. "The contribution of breeding to yield advances in maize (Zea mays L.)." *Advances in Agronomy* 86:83–145.

Egli, D. 2008. "Comparison of corn and soybean yields in the United States: historical trends and future prospects." *Agronomy Journal* 100:1–79.

Egli, D.B. 2017. *Seed biology and yield of grain crops*. CABI.

Fernandez-Cornejo, J. 2004. "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, United States Department of Agriculture, Economic Research Service.

Fernandez-Cornejo, J., S. Wechsler, M. Livingston, and L. Mitchell. 2014. "Genetically engineered crops in the United States.", pp. .

Goodwin, B.K., and A. Hungerford. 2015. "Copula-based models of systemic risk in US agriculture: implications for crop insurance and reinsurance contracts." *American Journal of Agricultural Economics* 97:879–896.

Harri, A., K.H. Coble, A.P. Ker, and B.J. Goodwin. 2011. "Relaxing heteroscedasticity assumptions in area-yield crop insurance rating." *American Journal of Agricultural Economics*, pp. aar009.

Hendricks, N.P., A. Smith, and D.A. Sumner. 2014. "Crop supply dynamics and the illusion of partial adjustment." *American Journal of Agricultural Economics*, pp. 1469–1491.

Hernandez, M.A., and M. Torero. 2014. "Parametric versus nonparametric methods in risk scoring: an application to microcredit." *Empirical Economics* 46:1057–1079.

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

Holm, S. 1979. "A simple sequentially rejective multiple test procedure." *Scandinavian Journal of Statistics*, pp. 65–70.

Huynh, K.P., D.T. Jacho-Chávez, R.J. Petrunia, and M. Voia. 2011. "Functional principal component analysis of density families with categorical and continuous data on Canadian entrant manufacturing firms." *Journal of the American Statistical Association* 106:858–878.

Ker, A., and T.A. Ergun. 2007. "On the revelation of private information in the US crop insurance program." *Journal of Risk and Insurance* 74:761–776.

Ker, A.P. 2016. "Nonparametric estimation of possibly similar densities." *Statistics & Probability Letters* 117:23–30.

Ker, A.P., and K. Coble. 2003. "Modeling conditional yield densities." *American Journal of Agricultural Economics* 85:291–304.

Ker, A.P., and B.K. Goodwin. 2000. "Nonparametric estimation of crop insurance rates revisited." *American Journal of Agricultural Economics* 82(2):463–478.

Ker, A.P., and Y. Liu. 2017. "Bayesian model averaging of possibly similar nonparametric densities." *Computational Statistics* 32:349–365.

Ker, A.P., and P. McGowan. 2000. "Weather-based adverse selection and the US crop insurance program: The private insurance company perspective." *Journal of Agricultural and Resource Economics*, pp. 386–410.

Ker, A.P., and T. Tolhurst. 2019. *American Journal of Agricultural Economics*, pp. forthcoming.

Ker, A.P., T.N. Tolhurst, and Y. Liu. 2016. "Bayesian estimation of possibly similar yield densities: implications for rating crop insurance contracts." *American Journal of Agricultural Economics* 98:360–382.

Koch, S.F., and J.S. Racine. 2016. "Healthcare facility choice and user fee abolition: regression discontinuity in a multinomial choice setting." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 179:927–950.

Kucharik, C.J., and N. Ramankutty. 2005. "Trends and variability in US corn yields over the twentieth century." *Earth Interactions* 9:1–29.

Leng, G. 2017. "Recent changes in county-level corn yield variability in the United States from observations and crop models." *Science of The Total Environment* 607:683–690.

Li, D., L. Simar, and V. Zelenyuk. 2016. "Generalized nonparametric smoothing with mixed discrete and continuous data." *Computational Statistics & Data Analysis* 100:424–444.

Li, Q., and J. Racine. 2003. "Nonparametric estimation of distributions with categorical and continuous data." *Journal of Multivariate Analysis* 86:266–292.

Miao, R., M. Khanna, and H. Huang. 2016. "Responsiveness of crop yield and acreage to prices and climate." *American Journal of Agricultural Economics* 98:191–211.

Miranda, M.J., and J.W. Glauber. 1997. "Systemic risk, reinsurance, and the failure of crop insurance markets." *American Journal of Agricultural Economics* 79:206–215.

Naylor, R., W. Falcon, and E. Zavaleta. 1997. "Variability and growth in grain yields, 1950-94: does the record point to greater instability?" *Population and Development Review*, pp. 41–58.

Park, E., B.W. Brorsen, and A. Harri. 2018. "Using Bayesian Kriging for Spatial Smoothing in Crop Insurance Rating." *American Journal of Agricultural Economics* 101:330–351.

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.

Racine, J., and Q. Li. 2004. "Nonparametric estimation of regression functions with both categorical and continuous data." *Journal of Econometrics* 119:99–130.

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.

Ramirez, O.A., S. Misra, and J. Field. 2003. "Crop-yield distributions revisited." *American Journal of Agricultural Economics* 85:108–120.

Reilly, J.M., and K.O. Fuglie. 1998. "Future yield growth in field crops: what evidence exists?" *Soil and Tillage Research* 47:275–290.

Roberts, M.J., W. Schlenker, and J. Eyer. 2013. "Agronomic weather measures in econometric models of crop yield with implications for climate change." *American Journal of Agricultural Economics* 95:236–243.

Shen, Z., M. Odening, and O. Okhrin. 2018. "Adaptive local parametric estimation of crop yields: implications for crop insurance rate making." *European Review of Agricultural Economics* 45:173–203.

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* 94:1037–1054.

Tack, J.B., and D. Ubilava. 2015. "Climate and agricultural risk: measuring the effect of ENSO on US crop insurance." *Agricultural Economics* 46:245–257.

Tolhurst, T.N., and A.P. Ker. 2015. "On technological change in crop yields." *American Journal of Agricultural Economics* 97:137–158.

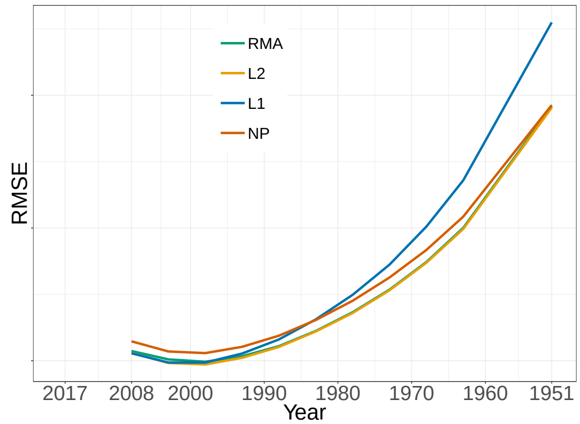
Vliamos, S.J., and N.G. Tzeremes. 2012. "Factors influencing entrepreneurial process and firm start-ups: evidence from central Greece." *Journal of the Knowledge Economy* 3:250–264.

Wilcox, R.R. 1997. "Some practical reasons for reconsidering the Kolmogorov-Smirnov test." *British Journal of Mathematical and Statistical Psychology* 50:9–20.

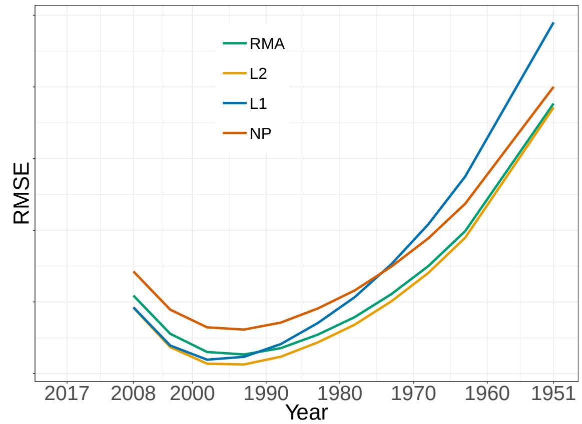
Wu, J., R.M. Adams, C.L. Kling, and K. Tanaka. 2004. "From microlevel decisions to landscape changes: an assessment of agricultural conservation policies." *American Journal of Agricultural Economics* 86:26–41.

Yvette Zhang, Y. 2017. “A Density-Ratio Model of Crop Yield Distributions.” *American Journal of Agricultural Economics* 99:1327–1343.

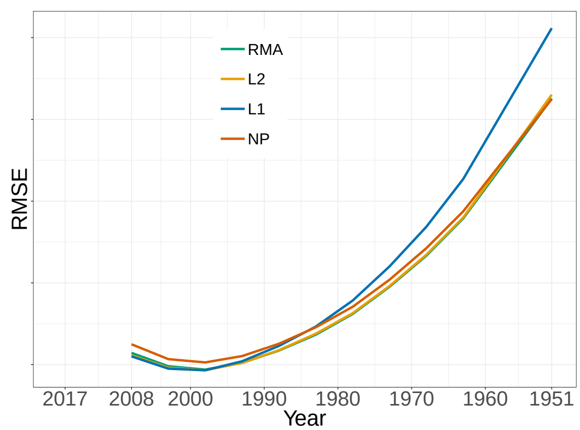
Zhu, Y., B.K. Goodwin, and S.K. Ghosh. 2011. “Modeling yield risk under technological change: Dynamic yield distributions and the US crop insurance program.” *Journal of Agricultural and Resource Economics*, pp. 192–210.



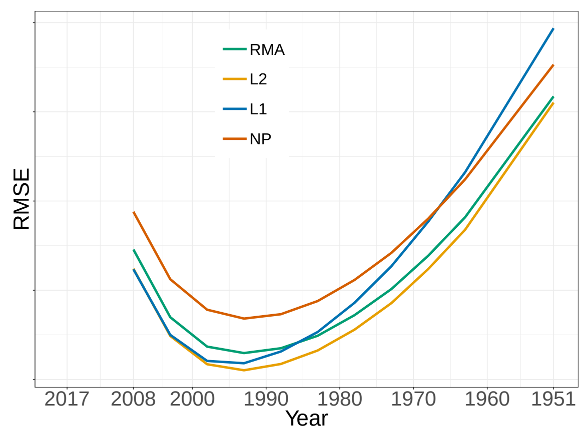
(A) Corn, RMSE of 70% Empirical Rate



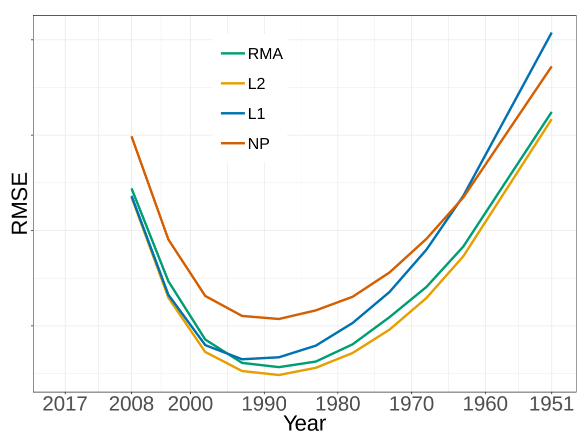
(B) Corn, RMSE of 90% Empirical Rate



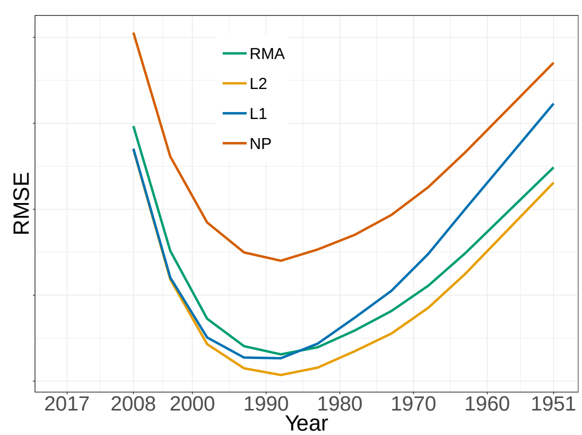
(C) Soybean, RMSE of 70% Empirical Rate



(D) Soybean, RMSE of 90% Empirical Rate

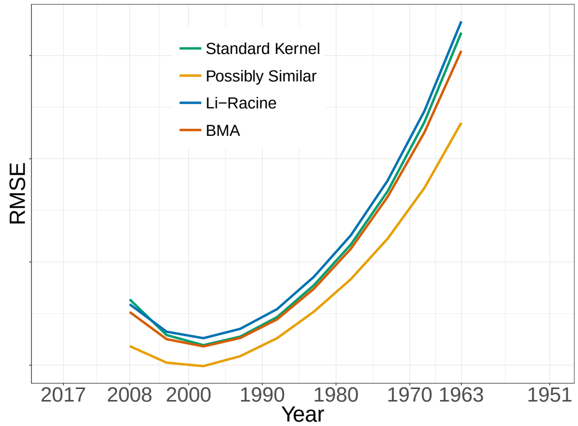


(E) Winter Wheat, RMSE of 70% Empirical Rate

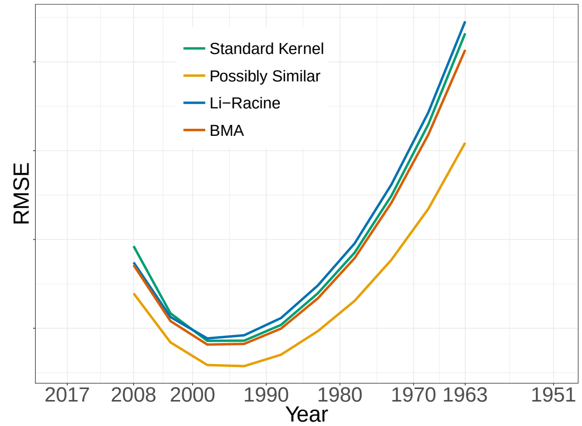


(F) Winter Wheat, RMSE of 90% Empirical Rate

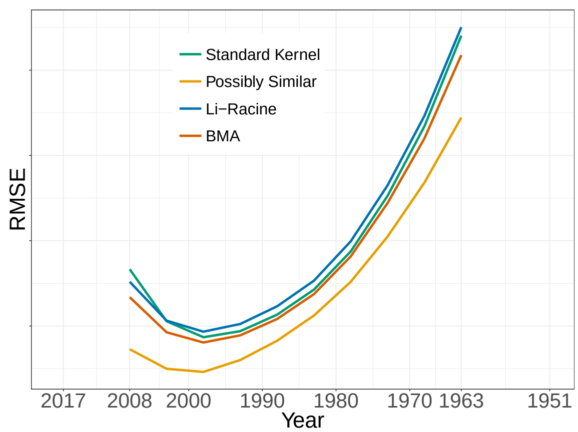
Figure 7. RMSE of Empirical Rate by Crop and Detrending Methodology



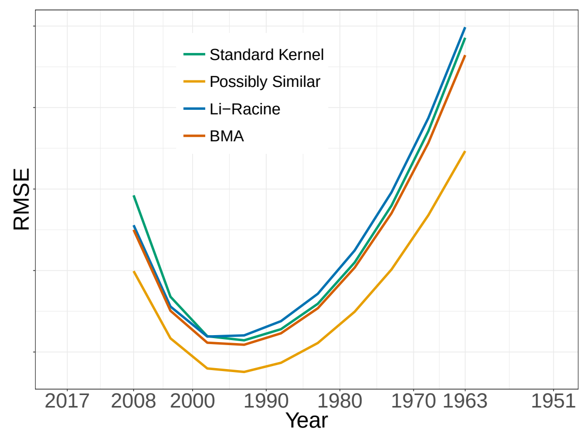
(A) Corn, RMSE of 70% Rate



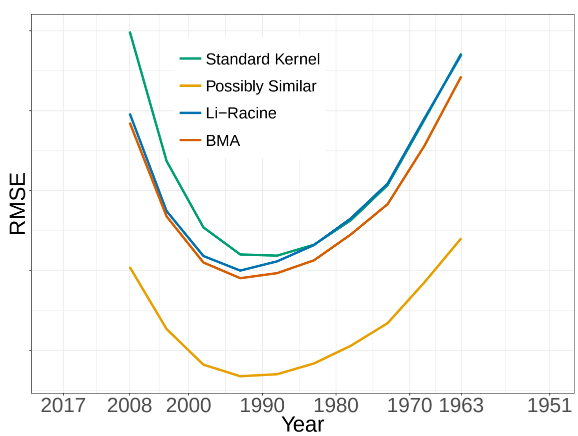
(B) Corn, RMSE of 90% Rate



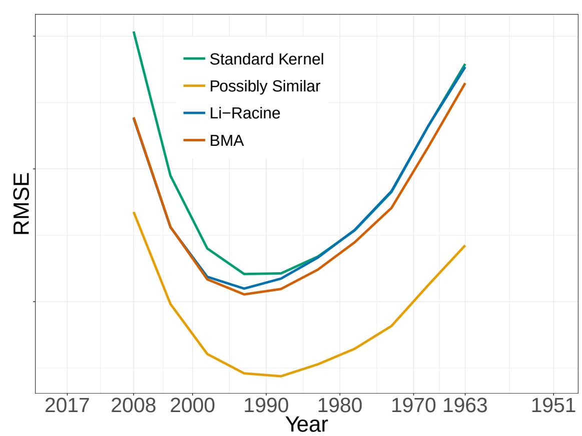
(C) Soybean, RMSE of 70% Rate



(D) Soybean, RMSE of 90% Rate



(E) Winter Wheat, RMSE of 70% Rate



(F) Winter Wheat, RMSE of 90% Rate

Figure 8. RMSE by Crop, Coverage Level, and Smoothing Methodology