Journal of Agricultural and Resource Economics 31(1):27–39

Rating Crop Insurance Policies with Efficient Nonparametric Estimators that Admit Mixed Data Types

Jeff Racine and Alan Ker

The identification of improved methods for characterizing crop yield densities has experienced a recent surge in activity due in part to the central role played by crop insurance in the Agricultural Risk Protection Act of 2000 (estimates of yield densities are required for the determination of insurance premium rates). Nonparametric kernel methods have been successfully used to model yield densities; however, traditional kernel methods do not handle the presence of categorical data in a satisfactory manner and have therefore tended to be applied on a county-by-county basis. By utilizing recently developed kernel methods that admit mixed data types, we are able to model the yield density *jointly* across counties, leading to substantial finite-sample efficiency gains. Findings show that when we allow insurance companies to strategically reinsure with the government based on this novel approach they accrue significant rents.

Key words: discrete data, insurance rating, kernel estimation, yield distributions

Introduction

Political forces have recently fashioned crop insurance as the cornerstone of U.S. agricultural policy. In 2000, Congress approved the Agricultural Risk Protection Act (ARPA). The additional cost of this legislation was estimated to be $8.2 billion over a five-year period, thereby *doubling* the federal budget on crop insurance programs to $16.1 billion. The program, only available to traditional field crops as recently as 1990, currently covers over 150 different crops including such nontraditional products as cut flowers, trees and shrubs, and most specialty crops such as avocados, blackberries, etc. ARPA has mandated the expansion of crop insurance in three important dimensions: (*a*) expanded product coverage including, for example, livestock products; (*b*) expanded geographical availability for existing crops; and (*c*) increasing producer demand by doubling subsidies from approximately 30% to 60% of the premium rate. Recent legislative actions indicate that crop insurance may become the policy instrument of choice to funnel resources to agricultural producers. Given the pivotal role played by crop insurance in U.S. agricultural policy and the substantial resources directed toward the support of agricultural producers, the accurate pricing of crop insurance policies, along with precise risk assessment, is more important than ever.

Jeff Racine is professor, Department of Economics, McMaster University, Hamilton, Ontario, Canada; Alan Ker is associate professor, Department of Agricultural and Resource Economics, University of Arizona, Tucson. The authors are indebted to two anonymous referees whose input led to a greatly improved version of this paper. The authors would like to thank but not implicate Qi Li, Thomas Knight, and Jack Rader for their thoughtful comments and suggestions. Racine gratefully acknowledges funding from the NSF, Award No. 3535905. Ker gratefully acknowledges support from the Cardon Endowment.

Review coordinated by T. F. Glover.

In this study we investigate a nonparametric method, recently introduced by Hall, Racine, and Li (2004), having the potential to substantially improve the accuracy of premium rate estimates. To this end, we compare this to the standard nonparametric kernel estimator used by Goodwin and Ker (1998) as well as the current rating procedures for the chosen plan of insurance. The distinguishing feature of the Hall, Racine, and Li (2004) estimator is that it can *appropriately* utilize data from extraneous sources to improve estimator efficiency. Specifically, we will be able to use yield data from comparable areas in estimating the yield density for a given area.

Because premium rates are derived from the lower tail of the estimated densities, we are necessarily comparing the competing methodologies with respect to their ability to accurately model this lower tail. This therefore requires a substantial amount of data in order to ensure sufficient statistical power to discriminate between the competing estimators in terms of their out-of-sample performance.[1] As a result, we are forced to focus on the less-purchased Group Risk Plan (GRP) of insurance allowing us to utilize the lengthy county-level yield series that was also used by Goodwin and Ker (1998). Ideally, we would prefer to focus on either the Multiple Peril Crop Insurance (MPCI) yield program or the Crop Revenue Coverage (CRC) revenue program, but both are based on the much shorter farm-level yield series. Recognizing this is a necessary limitation to any empirical analysis of this type, we provide simulation results in the appendix that compare, under a variety of situations, the performance of the Hall, Racine, and Li (2004) estimator with the traditional nonparametric kernel estimator used by Goodwin and Ker (1998).[2]

The U.S. crop insurance program is somewhat unique among insurance schemes in that three economic interests are served. The federal government through the United States Department of Agriculture's Risk Management Agency (RMA), the private insurance companies, and the farmers, all have vested interests. In 1980, the marketing of crop insurance policies, previously the domain of the RMA, was expanded to include private insurance companies in an attempt to increase farmer participation. While the pricing of the crop insurance policies remains the responsibility of the RMA, insurance companies receive compensation for administrative expenses and share, asymmetrically, the underwriting gains and losses of the policies.[3] The Standard Reinsurance Agreement (SRA) stipulates the terms of the sharing of these underwriting gains and losses. The structure of the SRA enables the private insurance companies to retain or cede—ex ante and subject to constraints—varying portions of the realized underwriting gains or losses of every federally subsidized crop insurance policy they sell.[4]

The unique arrangement among the producers, private insurance companies, and the RMA not only provides us with an ideal environment in which to evaluate proposed estimation methodologies, but also has important policy ramifications. Under the SRA,

---

[1] Unfortunately, there is a tradeoff between the estimator efficiency and the power of the out-of-sample test. The more of the sample we use to estimate the density, the less we have to conduct the out-of-sample test, thereby reducing the power of the test to distinguish between competing methodologies.

[2] In our concluding remarks, we also discuss the possibilities of using the Hall, Racine, and Li (2004) estimator with the more popular farm-level yield and revenue programs.

[3] Underwriting gain/loss for a set of policies is the total premium less the total indemnity payments.

[4] In practice, the insurance company can only retain or cede varying portions of the liability and associated premium rate for any insurance policy. See Ker and McGowan (2000) for a detailed discussion of the SRA. For the current analysis, it is sufficient to assume the insurance company can retain or cede 100% of the liability and premium for any given policy. This greatly reduces the complexity of the analysis without loss of generality.

an insurance company must decide which policies to retain and which to cede, thereby requiring the company to construct its own premium rate schedule. For example, consider the case where a farmer chooses to buy crop insurance from a private insurance company at the government-mandated price of, say, $100. The insurance company selling that policy must decide whether to retain or cede the premium and associated liability of the policy. Suppose the insurance company estimates the premium rate for that same policy to be $90. In this case, a risk-neutral insurance company will retain the policy because it expects a profit of $10. Suppose, instead, the insurance company estimates the premium rate for that policy to be $105. In this case, a risk-neutral insurance company will cede the policy.

Consequently, a risk-neutral insurance company will act according to the following decision rule: retain the subset of policies for which it expects a profit (the insurance company premium rate is less than the RMA premium rate) and cede the subset of policies for which it expects a loss (the insurance company premium rate is greater than the RMA premium rate).[5] As a result, the SRA represents an incentive for the RMA to employ the rating methodology which makes the most efficient use of the available data, thereby reducing adverse selection activities by the insurance companies. An appropriate context in which to evaluate a proposed methodology for rating crop insurance policies is to assume the role of an insurance company and determine if significant excess rents can be garnered from using the proposed methodology to identify which policies to retain and which to cede.

In the section below, we briefly review the U.S. crop insurance program and the SRA, outline the construction of premium rates, and discuss the yield data used in our analysis. Next, the RMA rating methodology is outlined, as well as the Hall, Racine, and Li (2004) (henceforth denoted as HRL) estimator. We then undertake an out-of-sample analysis designed to determine whether or not economically and statistically significant excess rents can be garnered using the HRL estimator. Policy implications and concluding remarks are offered in the last section. Finally, the appendix examines the finite-sample efficiency of the univariate kernel estimator which has previously been used to model crop yield densities (Goodwin and Ker, 1998), relative to that for the recently developed HRL estimator used in the current analysis.

Premium Rate Preliminaries

The U.S. Crop Insurance Program

Federally regulated crop insurance programs have been a prominent part of U.S. agricultural policy since the 1930s. In 2004, the number of crop insurance policies exceeded 1.23 million with total liabilities exceeding $48.6 billion. GRP is an area yield insurance program where both premiums and indemnities are calculated using county rather than farm yields. Group Risk Income Protection (GRIP) is the revenue analog to GRP. With respect to GRP and GRIP, the number of crop insurance policies in 2004 exceeded 48,000 with total liabilities exceeding $385 million.

[5] While this may not be economically inefficient, as it represents a simple transfer to insurance companies rather than agricultural producers, in a political economy framework this outcome may be undesirable. Political rents recovered from the agricultural production sector are likely to be significantly greater than those recovered from the private insurance companies involved in agricultural crop insurance.

Section II.A.2 of the 1998 SRA states that an insurance company "... must offer all approved plans of insurance for all approved crops in any State in which it writes an eligible crop insurance contract, and must accept and approve all applications from, all eligible producers." An eligible farmer will not be denied access to an available, federally subsidized, crop insurance product. Therefore, an insurance company wishing to conduct business in a state cannot discriminate among farmers, crops, or insurance products in that state. An unusual situation arises, however; the responsibility for pricing the crop policies lies with the RMA, but the insurance company must accept some liability for each policy it writes and cannot choose which policy it will or will not write.

Clearly, in the absence of additional incentive mechanisms, insurance companies are unlikely to become involved in such a risk-sharing arrangement. Therefore, to elicit their participation, two mechanisms are required that, necessarily, emulate a private market from the company's perspective. First, given that insurance companies do not set premium rates, there needs to be a mechanism by which they can cede the liability, or the majority thereof, of an undesirable policy (in a private market, the insurance company would simply refuse to write any policy deemed undesirable). Second, a mechanism providing an adequate return to the insurance company's capital and a level of protection against ruin (bankruptcy) is needed. Premium rates in a private market reflect a return to capital and a loading factor guarding against ruin. Premium rates set by the RMA do not reflect a return to capital but do include a loading factor. The SRA provides two such mechanisms which, in effect, emulate a private market from the perspective of the insurance company. In so doing, the SRA is the vehicle by which an insurance company can either retain or cede most of the premium and accompanying liability of policies of its choosing. By assuming the role of an insurance company, we can create an ideal environment in which to evaluate competing estimation methodologies.

Determining Premium Rates for Crop Contracts

Accurate pricing of crop insurance policies requires accurate estimation of yield densities. We define the premium rate as the probability of a loss multiplied by the expected loss given that a loss has occurred. Formally, the actuarially fair premium rate for a yield insurance contract that guarantees a percentage (say λ) of the expected yield (say y^e) is given as:

$$(1) \qquad \begin{aligned} \textit{Premium Rate} &= P(Y < \lambda y^e)(\lambda y^e - E(Y \mid y < \lambda y^e)) \\ &= \int_0^{\lambda y^e} (\lambda y^e - y) f_Y(y \mid I_t)\, dy, \end{aligned}$$

where $0 \le \lambda \le 1$, the expectation operator and probability measure are taken with respect to the conditional yield density $f_Y(y \mid I_t)$, and I_t is the information set known at the time of rating.[6] In the analysis that follows, the information set contains past yields and the county in which they were recorded. The RMA premium rate is taken with respect to its predicted yield and its estimate of the conditional yield density. Conversely, the insurance company determines its premium rate for the policy by integrating its estimate of the conditional yield density over the same space, $[0, \lambda y^e]$.

[6] The premium rate defined in (1) is in terms of expected loss with units equal to bushels per acre.

Yield Data

As noted earlier, our need to successfully discriminate between competing estimators in terms of their out-of-sample performance forces us to restrict attention to those crops and practice types for which a sufficiently long time series is available. Ideally, farm-level yield data would be used, but these data are available only for 20 years at most, and generally much less. While we could apply this estimator to farm-level yield series that are smaller in length, we would be unable to meaningfully discriminate between estimators because there would be far too few observations to permit us to differentiate between them in terms of their out-of-sample performance. Therefore, we use county-level yield data and the accompanying GRP plan of insurance for 87 Illinois counties with a complete yield series from 1956 to 2001 for all-practice corn. Historically, demand for GRP has been relatively high for this region-crop combination.

Estimating Yield Densities

The methods which have been used to model yield distributions fall into two camps, parametric and nonparametric. In the parametric camp, one common specification is the Beta distribution (see, for example, Hennessy, Babcock, and Hayes, 1997; Babcock and Hennessy, 1996; Coble et al., 1996; Borges and Thurman, 1994; Kenkel, Busby, and Skees, 1991; Nelson, 1990; and Nelson and Preckel, 1989). These authors found sufficient evidence of skewness and/or kurtosis in their yield data and opted to use the Beta distribution in lieu of the Normal distribution. Interestingly, none of these authors tested the appropriateness of the Beta distribution. Just and Weninger (1999) attempt to renew support for the Normal distribution by calling into question the use of aggregate yield data, inflexible trend modeling, and the interpretation of the normality test results. In contrast, Atwood, Shaik, and Watts (2000) attempt, using more diverse crop-region combinations, to reduce support for the Normal distribution, while Ker and Coble (2003) found empirical evidence rejecting the use of both the Normal and Beta distributions for modeling county corn yields in Illinois. In the nonparametric camp, Goodwin and Ker (1998) and Ker and Coble (2003) employ univariate nonparametric and semi-parametric kernel methods, respectively, to estimate yield densities and rate crop policies. Ker and Goodwin (2000) use empirical Bayes methods pointwise across the support to shrink the univariate nonparametric kernel estimates toward the mean.

GRP Rating Methodology

To model the temporal process of yields, the RMA employs a one-knot linear spline with once-iterated least squares while windsorizing outliers (determined based on residual estimates from the first iteration) in the second iteration to estimate the temporal process of yields (see Skees, Black, and Barnett, 1997).[7] The temporal model is the following:

[7] Windsorizing involves truncating the yield such that the absolute value of the residual is bounded below some determined level.

$$y_t = (\alpha_1 + \beta_1 t)I_{(0,T]}(t) + (\alpha_2 + \beta_2 t)I_{[T,TT)}(t), \tag{2}$$

where t is time, T is the knot point to be estimated, TT is the end of the time series, $I(\cdot)$ is an indicator function, and α_1, α_2, β_1, and β_2 are the parameters to be estimated subject to the constraint $\alpha_1 + \beta_1 T = \alpha_2 + \beta_2 T$. After correcting for heteroskedasticity, Skees, Black, and Barnett (1997) estimate a normal distribution and inflate the tails. The premium rate is the higher of the empirical rate or the rate derived from the Normal with inflated tails. The interested reader is referred to Skees, Black, and Barnett (1997) and the references contained therein for more detailed information on the GRP rating methodology.[8]

Nonparametric Methodology

Following the RMA, we first model the temporal process of yields employing the one-knot linear spline with once-iterated least squares while windsorizing outliers in the second iteration. Then, rather than presuming a parametric yield distribution or employing univariate kernel methods as has been done in the literature, we instead elect to use recently developed conditional nonparametric methods that are ideally suited to this setting. In the discussion below, we briefly outline the estimator for the interested reader, while the appendix presents some Monte Carlo results that examine the finite-sample performance of the estimator relative to the competing univariate kernel estimator.

For what follows, let $\mathbf{X}$ denote a vector of explanatory variables, and for a given value $\mathbf{X} = x$, we wish to estimate the conditional density of the response Y. In our case, $\mathbf{X}$ will denote the county in which yields are recorded, and Y will denote crop yields. We outline the estimator with a vector of mixed explanatory variables for completeness, and also to highlight the fact that additional information can be readily incorporated within the conditioning set $\mathbf{X}$ without modification.

We can always consistently estimate the density of crop yields conditional on county by simply computing each county's yield density separately, i.e., by taking a "frequency" approach. Specifically, we could first condition on county, then compute the density of yields for that county only. This is, of course, a consistent estimator, and is exactly that used by Goodwin and Ker (1998). However, this approach will be inefficient for a number of reasons leading to finite-sample efficiency losses. These efficiency losses will arise since the frequency approach does not use all available sample information (i.e., the county subsamples are much smaller than the overall sample size), it ignores any structure common to yield densities across counties, and it ignores any potential correlation structure among variables. Furthermore, if the yield density for all counties and time periods was in fact identical, then clearly one ought to pool all of the data in order to obtain the most efficient estimate. The frequency approach simply cannot take advantage of such possibilities. For what follows, we elect to use the kernel approach of HRL, which has improved finite-sample properties relative to the case where one simply computes yield densities separately by county. This approach is briefly described below.

[8] We thank Jerry Skees for providing the actual RMA GRP rating code and data.

Let $\hat{f}$ denote an estimator of the joint density, f, of $\mathbf{X}$, Y, and let $\hat{m}$ be an estimator of the marginal density, m, of $\mathbf{X}$. We estimate $g(y|x) = f(x,y)/m(x)$, the density of the response Y conditional on $\mathbf{X} = x$, by $\hat{g}(y|x) = \hat{f}(x,y)/\hat{m}(x)$. Our estimators of f and m are kernel estimators of the form:

$$\hat{f}(x,y) = \frac{1}{n}\sum_{i=1}^{n} K(x, X_i)L(y, Y_i) \quad \text{and} \tag{3}$$

$$\hat{m}(x) = \frac{1}{n}\sum_{i=1}^{n} K(x, X_i),$$

where K is a nonnegative "generalized kernel" that allows for the mix of categorical and continuous data. In particular, let $\mathbf{X} = (X^c, X^d)$ represent a division of $\mathbf{X}$ into continuous and discrete components. Reflecting these divisions, write $X_i = (X_i^c, X_i^d)$, where $X_i^d = (X_{i1}^d, \ldots, X_{iq}^d)$ and $X_i^p = (X_{i1}^c, \ldots, X_{ip}^c)$ denote the discrete and continuous components of X_i. We assume X_{ij}^d takes the values $0, 1, \ldots, r_j - 1$. Let $x_c = (x_1^c, \ldots, x_p^c)$ and $x_d = (x_1^d, \ldots, x_p^d)$, and define the continuous product kernel by:

$$K(x^c, X_i^c) = \prod_{j=1}^{p} \frac{1}{h_j} K\left(\frac{x_j^c - X_{ij}^c}{h_j}\right), \tag{4}$$

where $K((x_j^c - X_{ij}^c)/h_j)$ is a traditional kernel function. Next, define the categorical product kernel by:

$$K(x^d, X_i^d) = \prod_{j=1}^{p} \left\{\frac{\lambda_j}{r_j - 1}\right\}^{N_{ij}(x^d)} (1 - \lambda_j)^{1-N_{ij}(x^d)}, \tag{5}$$

where $N_{ij}(x^d) = I(X_{ij}^d \neq x_j^d)$, depending on x_j^d alone, and $I(\cdot)$ is the usual indicator function. The $h_1, \ldots, h_p$ are bandwidths for the continuous components of $\mathbf{X}$ satisfying $0 < h_j < \infty$, while $\lambda_1, \ldots, \lambda_q$ are smoothing parameters for the discrete components constrained by $0 \leq \lambda_j \leq (r_j - 1)/r_j$. The generalized kernel $K(x, X_i)$ and the kernel $L(y, Y_i)$ are given by:

$$K(x, X_i) = K^c(x^c, X_i^c)K^d(x^d, X_i^d) \quad \text{and} \tag{6}$$

$$L(y, Y_i) = \frac{1}{h} L\left(\frac{y - Y_i}{h}\right),$$

where $L((y - Y_i)/h)$ is another univariate kernel, typically identical to the univariate kernel $K((x_j^c - X_{ij}^c)/h_j)$ used in (4). Bandwidths are chosen by the cross-validation method outlined in HRL. For the estimation of unconditional distributions with mixed data types, see Li and Racine (2003), while for local constant and local polynomial regression with mixed data types, see Racine and Li (2004) and Li and Racine (2004) and the references therein. To best appreciate the benefits of this approach in the crop insurance setting, we undertook a modest simulation that underscores the performance of this approach in finite-sample settings. The main point underscored by the simulation results presented in the appendix is simply that we can never do worse than the frequency estimator employed by Goodwin and Ker (1998), since the HRL estimator collapses to the frequency estimator in the limit (i.e., as $\lambda \to 0$).

Analysis

As discussed above, an appropriate context in which to evaluate any proposed methodology for rating crop insurance policies is to assume the role of an insurance company. We can determine whether or not significant excess rents can be garnered when using a particular methodology by estimating the premium rate schedule and then determining which policies to retain and which to cede. In this section, we undertake a simulation designed to gauge the relative performance of the HRL kernel estimator, the standard cell-based univariate kernel estimator, and the RMA inflated Normal parametric estimator. Our simulation has the following salient features:[9]

- The RMA estimates its premium rate, denoted $\hat{\pi}_{RMA}$, using the GRP rating methodology. That is, the temporal models are estimated using a robust one-knot linear spline [see equation (2)], with premium rates being based on the maximum of the associated empirical rate or a rate derived from a Normal with inflated tails.
- The private insurance company estimates its premium rate, denoted $\hat{\pi}_{IC}$, by using the GRP methodology to estimate the temporal models [also equation (2)], and the HRL estimator to estimate the conditional yield density.
- The private insurance company, a profit maximizer, cedes a contract if $\hat{\pi}_{RMA} < \hat{\pi}_{IC}$ because the company believes the contract to be underpriced and expects a loss. Conversely, the private insurance company will retain a contract if $\hat{\pi}_{RMA} > \hat{\pi}_{IC}$.
- One-step-ahead premium rates are estimated for each of the 18 years, indexed from 1984 to 2001, based on yield data up to and including the preceding year—i.e., when constructing the 1989 estimated premium rates, only yield data from 1956 to 1988 are used.[10]
- The actual out-of-sample yield realizations are used to calculate the loss ratios for the set of contracts that the insurance company retains, the set of contracts the insurance company cedes (the set of contracts thereby held by the RMA), and the "program" or entire set.

If either nonparametric estimator better describes the yield distribution than the estimator used by the RMA, then the loss ratio for the contracts retained by the private insurance company would be expected to be lower than the overall loss ratio, and consequently the loss ratio for the government. Approximate randomization tests, which simulate the distribution of a desired statistic under the null, are used to ascertain statistical significance (see Kennedy, 1995). When evaluating the performance of each nonparametric method, our null is that the insurance company recovers rents by strategically reinsuring with the government, i.e., the insurance company's loss ratio is equal to the overall loss ratio. Under the null, the insurance company estimates every

[9] For the kernel estimators, bandwidths must be recomputed for each successive one-step forecast as the estimation sample size increases. Therefore, bandwidths are not reported here, but are available from the authors upon request.

[10] RMA must forecast two periods ahead because of a lag in the available yield data. We choose to forecast only one period ahead to conserve degrees of freedom. The sole effect of this is to maximize the power of our test, and the approach is valid for l-step forecasts ($l > 0$).

policy to have zero expected gain, and thus it is indifferent to retaining or ceding each and every policy. Having gauged each estimator's performance relative to the RMA, we are now in a position to assess their relative performance.

To obtain a realization from the null distribution, the insurance company randomly retains a policy with probability ρ, where ρ equals the fraction of policies retained in the original simulation (see table 1). We randomize over which policies are retained, not over the number of policies retained. We compare the insurance company's loss ratio from the analysis (denoted τ^*) to 1,000 simulated loss ratios under the null $\{\tau_1, \tau_2, \ldots, \tau_{1,000}\}$. The p-value for the test equals the fraction of $\{\tau_1, \tau_2, \ldots, \tau_{1,000}\}$ for which $\tau_i \leq \tau^*$.

Denote Ξ as the universe consisting of 1,566 policies (87 counties × 18 years), F the set of policies the insurance company retains, and F^c the set of policies the insurance company cedes. The loss ratio for a set, say F, is:

$$\text{(7)} \qquad Loss\ Ratio_F = \frac{\sum_{j \in F} \max\left(0, \lambda y_j^e - y_j\right)}{\sum_{j \in F} \hat{\pi}_{RMA,j}},$$

where j is the policy, y_j is the realized yield associated with policy j, λ is the coverage level, y_j^e is the RMA expected yield associated with policy j, and $\hat{\pi}_{RMA,j}$ is the RMA premium rate for policy j. We calculate the loss ratio for the program, the insurance company, and the RMA by summing over Ξ, F, and F^c, respectively.

Table 1 summarizes the program, RMA, and insurance company loss ratios for all simulations at both the 75% and 85% coverage levels based on a comparison of the nonparametric and RMA premium rates.

At the 75% coverage level, the insurance company using the HRL estimator would retain approximately 26% of the policies while ceding 74% of the policies to the RMA, suggesting the current RMA rating methodology may underestimate premium rates. More importantly, the insurance company's loss ratio based on the 26% of contracts it retains is reduced to 0.72, while the RMA's loss ratio increases to 1.18. The p-value of 0.036 strongly indicates the company is doing better than if it were to randomly select policies. By way of comparison, using the standard cell-based univariate kernel estimator applied to each county individually, the insurance company loss ratio actually increased to 1.297. Not surprisingly, the percentage of contracts retained is very small.[11] Note that the univariate kernel estimator performs quite poorly. For example, if one just randomly chooses 4.5% of the contracts, the resulting loss ratio will tend to be less than 1.297 with roughly 80% probability.

At the 85% coverage level, the insurance company using the HRL estimator retains approximately 18% of the policies while ceding 82% of the policies to the RMA, suggesting the current RMA rating methodology appears to underestimate the rates more significantly than at the 75% coverage level. This finding is consistent with the overall loss ratio being greater than 1. Interestingly, the insurance company's loss ratio based on the 18% of contracts it retains using the HRL estimator is reduced to 0.75, while the RMA loss ratio rises to 1.21. As suggested by the p-value of 0.017, the company is doing significantly better than if it were to randomly select policies. By way of comparison,

[11] The premium rate based on the univariate kernel will necessarily be higher than the empirical rate. As such, the univariate kernel rate will tend to be higher than the RMA premium rate, resulting in very few policies retained.

Table 1. GRP Simulation Results for All Counties in Illinois for All-Practice Corn: Nonparametric versus RMA Rates

	75% Coverage Level		85% Coverage Level	
Description	Univariate Kernel Estimator vs. RMA Rating Methodology	HRL Kernel Estimator vs. RMA Rating Methodology	Univariate Kernel Estimator vs. RMA Rating Methodology	HRL Kernel Estimator vs. RMA Rating Methodology
Program Loss Ratio	0.935	0.935	1.097	1.097
Insurance Company Loss Ratio	1.297	0.719	0.000	0.751
RMA Loss Ratio	0.925	1.182	1.102	1.213
Percent of Policies Retained	4.5%	26.4%	0.4%	18.0%
p-Value	0.799	0.036	0.287	0.017

using the standard cell-based univariate kernel estimator applied to each county individually, the insurance company only retains seven out of the 1,566 policies. Although the loss ratio for those seven policies is zero, if we randomly chose seven policies to retain, the insurance company would realize zero loss ratio with roughly 29% probability. Therefore, while the loss ratio has decreased, it does not differ statistically from the overall program loss ratio.

Based on the results reported in table 1, it is apparent that, by using a more efficient nonparametric estimator of yield densities than the univariate cell-based estimator, a private insurer could successfully adverse select against the government via its choice of which policies to retain.

Conclusions and Policy Implications

Given the increasing interest in crop insurance and agricultural risk arising in part due to the Agricultural Risk Protection Act, there has been a recent surge in interest in the identification of improved methods for characterizing yield distributions. There is mounting evidence indicating that common parametric yield distribution models may be inappropriate for characterizing the underlying data-generating process. Hence, some have turned instead to nonparametric estimation methods. In this study we continue this trend and investigate the application of a new nonparametric conditional distribution estimator proposed by Hall, Racine, and Li (HRL, 2004). This estimator has the potential to improve the accuracy of yield density estimates and their attendant insurance rates through the joint modeling of *continuous* data (yield) and *discrete* data (county in which the yield was recorded) using generalized product kernels.

We investigate the behavior of the HRL estimator by focusing on economic implications of estimator efficiency. Competing nonparametric estimators are used to estimate a set of yield densities and to derive the associated premium rates. The competing estimators are evaluated by calculating out-of-sample loss ratios based on decision rules for retaining or ceding GRP crop insurance contracts. This simulation is of interest from an economic and policy perspective because the SRA enables the private insurance companies to retain or cede, ex ante and subject to constraints, varying portions of the

realized underwriting gains or losses of every federally subsidized crop insurance contract they sell. Of the policies retained, the loss ratio suggests that the HRL estimator was successful at significantly increasing rents to insurance companies. This finding, along with the fact that the univariate estimator employed by Goodwin and Ker (1998) is not successful in this setting, simply confirms the new estimator is more efficient than the univariate kernel estimator which has been previously used to model yield densities.

Although our data requirements forced us to focus on a relatively small plan of insurance (GRP), we feel our results may be more generally applicable, particularly as more data become available. First, the RMA is shifting from separate rating methodologies to a consistent methodology for its farm-level yield (MPCI) and revenue-based [CRC and Revenue Assurance (RA)] products so that the revenue products are simply an additional premium to RMA's yield products. Currently, MPCI uses a modified loss cost approach and empirical rate relativities, while RA requires estimates of the density. It is likely the RMA will adopt a density-based approach, in which case the HRL estimator can be used to estimate the yield density component (or just the lower tail) for the county base rates across the coverage levels. Second, the HRL estimator may be applied not only for GRP but also GRIP. Finally, yield density estimates are used quite often in the literature in settings other than crop insurance, in which case the HRL estimator may be employed.[12]

[*Received September 2004; final revision received March 2006.*]

References

Atwood, J., S. Shaik, and M. Watts. "Normality of Crop Yields." Work. paper, Dept. of Agr. Econ. and Econ., Montana State University, Bozeman, 2000.

Babcock, B., and D. Hennessy. "Input Demand Under Yield and Revenue Insurance." *Amer. J. Agr. Econ.* 78(1996):416–427.

Borges, R., and W. Thurman. "Marketing Quotas and Subsidies in Peanuts." *Amer. J. Agr. Econ.* 76(1994):801–809.

Coble, K., T. Knight, R. Pope, and J. Williams. "Modeling Farm-Level Crop Insurance Demand with Panel Data." *Amer. J. Agr. Econ.* 78(1996):439–447.

Goodwin, B., and A. Ker. "Nonparametric Estimation of Crop Yield Distributions: Implications for Rating Group-Risk (GRP) Crop Insurance Contracts." *Amer. J. Agr. Econ.* 80(1998):139–153.

Hall, P., J. Racine, and Q. Li. "Cross-Validation and the Estimation of Conditional Probability Densities." *J. Amer. Statis. Assoc.* 99,2(2004):1015–1026.

Hennessy, D., B. Babcock, and D. Hayes. "Budgetary and Producer Welfare Effects of Revenue Insurance." *Amer. J. Agr. Econ.* 79(1997):1024–1034.

Just, R., and Q. Weninger. "Are Crop Yields Normally Distributed?" *Amer. J. Agr. Econ.* 81(1999): 287–304.

Kenkel, P., J. Busby, and J. Skees. "A Comparison of Candidate Probability Distributions for Historical Yield Distributions." Paper presented at annual SAEA meetings, Fort Worth, TX, 1991.

Kennedy, P. E. "Randomization Tests in Econometrics." *J. Bus. and Econ. Statis.* 13(1995):85–94.

Ker, A., and K. Coble. "Modeling Conditional Yield Distributions." *Amer. J. Agr. Econ.* 85(2003): 291–304.

Ker, A., and B. Goodwin. "Nonparametric Estimation of Crop Insurance Rates Revisited." *Amer. J. Agr. Econ.* 83(2000):463–478.

[12] We would caution against using kernel methods with data-driven bandwidth selectors for relatively small samples.

Ker, A., and P. McGowan. "Weather Based Adverse Selection: The Private Insurance Company Perspective." *J. Agr. and Resour. Econ.* 25,2(2000):386–410.

Li, Q., and J. Racine. "Nonparametric Estimation of Distributions with Categorical and Continuous Data." *J. Multivariate Analysis* 86(2003):266–292.

———. "Cross-Validated Local Linear Nonparametric Regression." *Statistica Sinica* 14,2(2004):485–512.

Nelson, C. "The Influence of Distribution Assumptions on the Calculation of Crop Insurance Premia." *N. Cent. J. Agr. Econ.* 12(1990):71–78.

Nelson, C., and P. Preckel. "The Conditional Beta Distribution as a Stochastic Production Function." *Amer. J. Agr. Econ.* 71(1989):370–378.

Racine, J., and Q. Li. "Nonparametric Estimation of Regression Functions with Both Categorical and Continuous Data." *J. Econometrics* 119,1(2004):99–130.

Skees, J., J. Black, and B. Barnett. "Designing and Rating an Area Yield Crop Insurance Contract." *Amer. J. Agr. Econ.* 79(1997):430–438.

Appendix: A Monte Carlo Examination of Finite-Sample Efficiency

We consider a modest Monte Carlo experiment to underscore the efficiency gains associated with the Hall, Racine, and Li (HRL, 2004) kernel estimator relative to the traditional frequency kernel estimator (the so-called "frequency" kernel estimator). We create a setting which parallels the crop yield application, and gently remind the reader that our sole intention is to demonstrate our approach can be expected to dominate the univariate (frequency) kernel density estimator such as that employed by Goodwin and Ker (1998). Furthermore, it is well suited to time-series panels as the estimator remains consistent for weakly dependent processes.

For what follows, let X denote a random scalar drawn from a binomial distribution having n_{trials} trials with probability of success on each trial equal to 0.5, and let Y denote a random scalar drawn from a χ^2 distribution with degrees of freedom $df_{\min} + X$, where $X \in \{0, 1, \ldots, n_{trials}\}$. As $df_{\min}$ increases, the conditional distribution of Y converges to $N(\mu, \sigma^2)$, where $\mu = df_{min} + X$ and $\sigma^2 = 2(df_{min} + X)$, while as n_{trials} increases, the number of cells increases. The smaller is $df_{\min}$, the more dissimilar the distribution of $Y \mid X$. Figure A1 presents two cases, one in which the *shape* of the distribution of Y differs dramatically with X, and one in which it is invariant with respect to X (though the mean and variance may, of course, vary).

In addition to allowing the shape of the distribution to vary with X, note there is also a trend present in Y *regardless* of the shape of its distribution—i.e., the mean of Y increases with X.[13] Hence, this simulation captures a number of features potentially present in the crop yield application, in particular, the fact that the distribution of yields may differ in both shape and level across counties, while county in which yield is recorded is a categorical explanatory variable. Finally, note that the sample size will be $\leq n_{subset} \times n_{trials}$.[14]

For the results that follow, we consider random samples $\{Y_i, X_i\}_{i=1}^n$ drawn from $f(y \mid x) = \chi^2_{(df_{\min}+X)}$. For each sample, we use least squares cross-validation to compute $\hat{h}, \hat{\lambda}$; compute $\hat{f}(y \mid x)$ using the method of HRL; and then compute the resulting mean squared error (MSE) given by

$$n^{-1} \sum_{i=1}^{n} \left(\hat{f}(y \mid x) - f(y \mid x) \right)^2.$$

We do the same for the frequency estimator, i.e., the univariate kernel estimator applied to those Y lying in each cell (each realization of X). We repeat this 1,000 times, and report the relative MSE generated as the median MSE of the HRL estimator divided by that for the univariate kernel estimator.

[13] Recall that the mean of a random variable drawn from a χ^2 distribution is its degrees of freedom, in this case $df_{\min} + X$.

[14] Some draws may not contain all possible realizations for X.

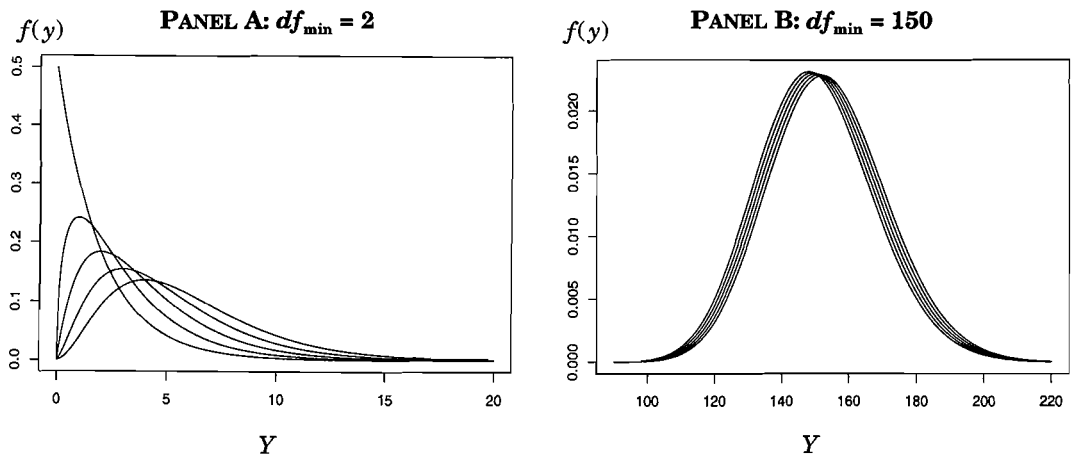

Figure A1. Simulated conditional distribution of $Y \sim \chi^2_{(df_{\min}+X)}$ for $X \in \{0, \ldots, 4\}$

Table A1. Summary of the Relative Efficiency of the HRL Estimator versus the Univariate Kernel (Frequency) Estimator

Description	$df_{\min}$	n_{trials}	n_{subset}	Efficiency
Effect of increasing similarity ($df_{\min}$):	2	4	25	1.00
	150	4	25	0.48
Effect of increasing number of cells (n_{trials}):	2	16	20	0.89
	2	32	20	0.86
Effect of increasing cell size (n_{subset}):	50	4	25	0.54
	50	4	50	0.56

Numbers < 1 indicate better performance of the HRL estimator. We consider a range of simulations beginning with a scenario in which the two approaches would be expected to deliver identical results (distributions differ dramatically with X) to the more typical case in which the HRL estimator is far more efficient. This latter case will occur when the number of cells (realizations of X) rises and/or the shape of the distributions becomes more similar. Representative results are reported in table A1.

Table A1 reveals a number of interesting features associated with the HRL estimator. First, we can never do worse than the frequency estimator since the HRL estimator collapses to the frequency estimator in the limit with cross-validation choosing $\hat{\lambda} = 0$ in such cases, i.e., when the shapes of the distributions differ dramatically across cells. This is the case for row 1 in table A1 corresponding to panel A in figure A1. Second, as the number of cells increases, other things equal (i.e., as n_{trials} increases), relative performance improves (rows 3 and 4). Third, as the shape of the distributions becomes more similar (i.e., $df_{\min}$ rises), relative performance improves (rows 1 and 2). Fourth, as the cell size increases, other things equal, relative performance converges (rows 5 and 6).

Note that, relative to these simulations, the number of cells in the crop yield application is very large (87) while the shape of the crop densities is naturally expected to be similar across counties. Hence, we would expect the efficiency gains on the HRL over the univariate kernel estimator would be large. Results reported in the text analysis section indicate this is indeed the case.