



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
<http://ageconsearch.umn.edu>
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

**An Application of Kernel Density Estimation via Diffusion to Group
Yield Insurance**

Ford Ramsey, North Carolina State University, aframsey@ncsu.edu

*Paper prepared for presentation at the Agricultural & Applied
Economics Association's 2014 AAEA Annual Meeting, Minneapolis,
MN, July 27-29, 2014*

*Copyright 2014 by Austin F. Ramsey. All rights reserved. Readers may make
verbatim copies of this document for non-commercial purposes by any means,
provided that this copyright notice appears on all such copies.*

An Application of Kernel Density Estimation via Diffusion to Group Yield Insurance

FORD RAMSEY

North Carolina State University
aframsey@ncsu.edu

Abstract

The recent priority given to Federal Crop Insurance as an agricultural policy instrument has increased the importance of rate making procedures. Actuarial soundness requires rates that are actuarially fair: the premium is set equal to expected loss. Formation of this expectation depends, in the case of group or area yield insurance, on precise estimation of the probability density of the crop yield in question. This paper applies kernel density estimation via diffusion to the estimation of crop yield probability densities and determines ensuing premium rates. The diffusion estimator improves on existing methods by providing a cogent answer to some of the issues that plague both parametric and nonparametric techniques. Application shows that premium rates can vary significantly depending on underlying distributional assumptions; from a practical point of view there is value to be had in proper specification.

I. INTRODUCTION

Growth of the Federal Crop Insurance program has continued unabated for the past two decades. Created with the intent of protecting producers from a variety of risks, the program has also become a powerful tool for subsidy. With the passage of the Agricultural Act of 2014, crop insurance is poised to further solidify its position as the second largest spending item considered in the Farm Bill. The growing importance of crop insurance as an agricultural policy instrument has amplified the consequence of the rate setting policies and procedures of the Federal Crop Insurance Corporation (FCIC).

Actuarial soundness of the crop insurance program is one of the primary goals of the FCIC. A key component of any actuarially sound insurance program is an accurate premium rate. Actuarially fair insurance sets the policy premium equal to expected loss. The true probability of loss is often not known and is inferred from a given set of data and information. Error arises in estimation and must be constrained if ideal premium rates are to be

approached.

The number of policy types available from the FCIC is considerable. This assortment is slated to grow even larger with the introduction of additional policies for peanuts and specialty crops. Stacked Income Protection (STAX) for cotton producers will replace other subsidy measures with revenue insurance. Determining premium rates for these policies usually requires estimation of the population mean of one or more random variables and a distribution about this mean. In the case of yield insurance the random variable is some measure of crop yield; revenue insurance considers the randomness of yield and price jointly.

When modeling risk there are a number of statistical issues to take into account. Yield insurance is one of the simpler paradigms. Given the nature of observed yield data, there are two main concerns: elimination of the trend and estimation of the probability density function. With respect to the first, trends and autoregressive effects must be eliminated prior to the modeling of yield distributions. Direct use of observed yields is clearly inappropriate. As to the latter, premium rates are constructed us-

ing conditional yield distributions and mean yields. It is crucial that these distributions are estimated accurately.

This paper is largely concerned with the second of these two points of interest. Methods proposed for the estimation of conditional yield distributions have typically tracked advances in theoretical statistics. As in the statistics literature, density estimation procedures may be broadly grouped as either parametric or nonparametric. The suitability of either class of techniques to a given inferential problem depends on the nature of the problem itself. Naturally, certain approaches are to be preferred depending on the situation.

Appeal to parametric specification as a solution to the yield density problem is not uncommon and there is a long history of research in this vein. Botts and Boles, Just and Weninger, and Ozaki et al. utilized a normal distribution. Gallagher used a gamma distribution, Chen and Miranda a weibull distribution, and Sherrick et al. a logistic distribution. As with any parametric approach, these methods require prior specification of the distribution function. If the distribution is improperly specified it is likely that biased inference will result. It is disconcerting, given this possibility of specification bias, that goodness of fit tests typically do not validate any one underlying distributional form for all crops and aggregation levels.

Visual inspection of yield data, and experience with crop insurance over the past eighty years or so, has led to two conclusions about the nature of yields. If both of these conclusions are believed to be true, then a number of popular models for yield densities are almost surely misspecified. Acceptance of these results allows certain estimation techniques to be eliminated from consideration.

Gallagher and Nelson both explain that yields are often negatively skewed. Without weather, pests, and other outside factors affecting yields, it would be expected that yields always approach a given capacity constraint. The stochastic factors cause yields to be less than their ideal. Recognition of skewness invalidates use of a normal distribution for mod-

eling yields. If this negative skewness was the only feature of the data that contradicted normality, practitioners might be content with utilizing beta or gamma distributions to model yields. The weibull distribution is also able to accommodate skewness. However, there is another feature of the data that suggests that such distributional assumptions are inappropriate.

Ker and Goodwin suggest the idea that yields may fall under two regimes: catastrophic or normal. Under normal conditions yields vary about the capacity constraint as a result of typical variation in stochastic factors. Catastrophic conditions, like flooding or severe drought, severely decrease the yield and in some cases cause yields to approach the zero bound. The existence of underlying regimes of this nature generates bimodal or bitangential yield distributions.

When the data do not appear to fit any one parametric distribution, and when there is very little guidance from theory as to what distribution to utilize, the problem should be approached through nonparametric methods. Such methods are far better equipped to deal with the presence of multiple modes, skewness, and bitangentiality. In fact none of the parametric distributions mentioned thus far can take multiple modes. As Silverman notes, it is quite easy to smooth data with the eye. It is difficult to undo this process. Assuming a parametric distribution masks key features of the data that cannot be recovered.

Nonparametric techniques also protect against specification error. Where the evidence for any one parametric distribution is tenuous, as with crop yields, the possibility of introducing such error in estimation is a foremost concern. Failure to take this into account invariably leads to inaccurate premium rates.

A plethora of approaches to nonparametric density estimation is detailed in Silverman's now ubiquitous monograph. Techniques include kernel smoothing, orthogonal series estimators, and penalized likelihood approaches. Kernel smoothing has proven to be the most popular of the three, perhaps due to its comparative simplicity and relatively low compu-

tational cost. Goodwin and Ker and Ker and Goodwin utilized kernel density estimators in their examination of rate making for the Group Risk Plan.

Nonparametric methods are not without their drawbacks. Kernel density estimation suffers from a slow rate of convergence, requires a bandwidth choice, does not deal well with boundary bias, and oftentimes over inflates the importance of outliers. Nonetheless, there have been significant – and largely disparate – advances in the literature that correct for many of these problems. Kernel density estimation via diffusion brings many of these solutions under a single umbrella. Application of this technique represents a significant advance in the proper specification of a statistical approach to the estimation of conditional yield distributions.

II. THE GROUP RISK PLAN

Federal crop insurance has existed in one form or another since its implementation under the Agricultural Adjustment Act of 1938. Area yield crop insurance, or the Group Risk Plan (GRP) specifically, was first piloted in 1993. This policy is significantly different from farm level yield insurance products. Miranda and Skees et al. provide excellent reviews of both the theoretical and practical concerns of offering crop insurance contracts based on area yields. Their extended examinations of the benefits of area yield based policies provide further evidence of the pragmatic importance of rate making procedures.

Like most insurance products, crop insurance is not immune to problems of adverse selection and moral hazard. Many authors recognize that products based on area yields mitigate these complications. Additionally, there is typically more data available at aggregate levels than at the farm level. Lack of historical data can seriously impede rate making. The GRP is available for a large number of crops and is provided at the county level. It is the county yield that is used in determining indemnity payments and coverage levels.

From the perspective of administrative agencies there are supplementary benefits to using area yield policies. Compared to farm level plans, area yield policies require less paperwork and man hours. Instead of having to verify losses at each individual farm, the insurer must only compare a single realized county yield with its expected average.

If the realized county yield falls below a percentage of the expected average yield for the county, a payout is triggered. The percentage of expected county yield – more commonly termed the coverage level – may be 70, 75, 80, 85, or 90 percent of the expected yield. The indemnity payment made to farmers is the shortfall between the realized yield and the coverage level multiplied by a price protection level. Expected prices are determined exogenously by the Federal Crop Insurance Corporation. As Goodwin and Ker note, price election is not relevant for the premium rate in this case.

To be more concise, consider the area yield as a random variable Y . The Group Risk Plan pays an indemnity if the realized area yield falls below some percentage γ , of the population mean of such yields μ . The trigger yield that actually causes the insurance policy to pay out is then given by $\gamma\mu$. It should be clear that the probability of loss is calculated from the probability distribution of Y . Estimation of this distribution is of primary concern.

III. SOME CONSIDERATIONS

Crop yields trend upwards over time. Changes in technology, from new types of plants to more efficient machinery, can significantly alter the distribution of yields. Institutional change can also affect the distribution. To link this to previous notions of the processes affecting yields, the relevant capacity constraint can be viewed as gradually shifting over time. Before any direct estimation of the yield distribution can be attempted it is necessary to first remove the trend from the time series. It is usually the case that the residuals from the trend line will show evidence of heteroskedasticity.

A common approach to detrending is to

specify the overall problem as one of two stages. The trend model may be viewed as

$$Y_t = h(X_t) + \epsilon_t \quad (1)$$

where Y is the aggregate yield, $h(\cdot)$ is an unspecified regression function, and X is a time index to capture trend. ϵ is a simple error term that is independently distributed with mean zero. Estimation of the function $h(\cdot)$ is the goal of the first stage and this process can be completed in a number of different ways. Examples of different approaches are given in Miranda and Glauber and Atwood et al. Zhu, Goodwin, and Ghosh also review common applications of the two stage framework.

Residuals from the first stage are given as $\hat{\epsilon}_t = Y_t - \hat{h}(X_t)$. As shown in Goodwin and Ker, the residuals tend to be proportional to the level of yields. An admittedly ad-hoc approach to correct for this heteroskedasticity is to use a rescaled version of the residuals. Each error is divided by its yield forecast and residuals are scaled to the equivalent predicted yield of the last year in the series. It should be noted that there is estimation error inherent in the two stage approach. Such error arises in the initial estimation of yield forecasts.

Provided the first stage is properly dealt with, the result is a series of observations that are independent and identically distributed. Nearly the full gamut of density estimation methods is then available for the second stage problem: estimation of the conditional yield density.

As already mentioned, parametric methods are not as flexible in accommodating skewness and bimodality. Both of these features have been observed empirically. Ker and Goodwin note that Central Limit Theorems for dependent processes provide theoretical justification for bimodal behavior. Given that these features should be accommodated, and that the first stage estimation is specified correctly, the class of suitable estimation methods for the second stage has been considerably reduced. Nonparametric methods are certainly applicable and kernel density estimation possesses a number of advantages within this reduced group.

IV. KERNEL DENSITY ESTIMATION AND CONCERNS

Kernel density estimators can typically be used in situations where the data is independent and identically distributed. Fixed bandwidth kernel density estimators are given by

$$\hat{g}(Y) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{y - Y_i}{h}\right) \quad (2)$$

where K is a kernel function satisfying a number of conditions that essentially ensure that the kernel is a valid probability density function. h is the bandwidth or window width and Y now represents the yield data conditional on temporal effects having been removed. The basic idea is to place a weighted kernel on top of each observation. These individual kernels are summed vertically to obtain an estimate of the density.

An immediate concern with any kernel density estimation is both the nature (fixed or variable) and size of the bandwidth chosen. Fixed bandwidth methods smooth each observation equally. Larger bandwidths increase smoothing while smaller bandwidths decrease smoothing. Intuitively it makes sense for observations in the tails of the density to be smoothed more while observations near the mode of the density are smoothed less. This notion cannot be entertained within the fixed bandwidth framework. The bandwidth will be either too large for observations in the tails or too small for those near the mode. Premium rates for yield insurance are heavily dependent on tail estimates and the inadequacy of fixed bandwidth estimators in this area is cause for concern.

Variable bandwidth methods adjust the degree of smoothing in a way that gives observations in sparse areas of the data less emphasis. Instead of having a single bandwidth parameter h , each observation is assigned its own bandwidth which is inversely proportional to the density of the data about the point. If the data are dense around observation i then h_i will be small and less smoothing will occur. For the purpose of estimating yield densities, this variable bandwidth approach is preferable

because it ignores spurious features in the tails of the densities.

While variable bandwidth methods are preferred to fixed bandwidth as far as tail probabilities are concerned, nothing has been said thus far about the methods actually used to select the bandwidth. The bandwidth is typically selected to minimize either the mean integrated squared error of the density estimator or the asymptotic approximation to this error. Details of the form of the mean integrated squared error (MISE) and minimization procedures can be found in Silverman, Marron and Wand, and Li and Racine. Without reproducing these calculations, it should be noted that the optimal bandwidth in terms of the AMISE depends on a functional of the true density $f(Y)$. (Specifically it depends on $\int f''(Y)^2 dY$)

Of course the true density is unknown – else there would not be a density estimation problem at all. Popular methods of bandwidth selection deal with this obstacle in a variety of ways. Silverman’s Rule of Thumb uses a normal reference rule where the underlying true density is assumed to be normal for bandwidth calculation. Least squares cross validation and likelihood cross validation have also received attention as the former possesses appealing asymptotic properties. Sheather and Jones suggest a plug-in method that assumes normality in a way, but the assumption is so deeply embedded that it is of very little consequence.

A sterling survey of the drawbacks and advantages of various bandwidth selection rules can be found in Jones, Marron, and Sheather. Silverman’s Rule of Thumb has been shown to oversmooth the data. Least squares cross validation does perform well, but only when the sample size is large and there are few outliers. The Sheather Jones method is shown to be optimal based on a number of criterion and details of this approach can be found predictably in Sheather and Jones. Of all the bandwidth methods considered, Sheather Jones performs optimally in the Marron and Wand test suite.

One aspect of kernel density estimation that has not been addressed in the estimation of yield densities is boundary bias. Standard ker-

nel methods take as given that the support of the distribution is the whole real line. A problem presents itself whenever the density must be estimated on some subspace of the real line. Physical reality prevents crop yields from being negative so there is a natural boundary for the support at zero. Solutions to this problem have taken a number of forms including boundary kernels, reflection methods, and data transformations. While such techniques are capable of dealing with the issue, they typically are not able to accommodate variable bandwidth kernels or do not lead to true probability densities.

Ostensibly, one might think that boundary bias is of no consequence as the majority of the time yields lie away from the boundary. Indeed if a variable bandwidth kernel estimator is used the problem is further alleviated. While it may be comforting to assume that the error from boundary bias is negligible, there is some evidence that the bias may affect yield distributions. Analysis of state level yields by Goodwin and Ker, using a fixed bandwidth kernel estimator, seems to imply that boundary bias can come into play. Results of this paper indicate a similar possibility.

The degree to which boundary bias is present will vary with both crop and location. Non-irrigated crops are in general more susceptible to catastrophic yield behavior. The same may be said of crops grown in developing nations or areas where modern farming practices are not employed. The trend in American agricultural policy toward new insurance policies for specialty crops, cotton, and peanuts could mean an increase in the consequence of this type of bias.

Goodwin and Ker, Ker and Goodwin, and Ker and Coble all note the slow rate of convergence of kernel density estimators. It is perhaps the main drawback of many nonparametric methods. The standard fixed bandwidth kernel estimator has a best possible mean integrated squared error order of magnitude of $N^{-4/5}$. Convergence rates are quite slow when compared with a properly specified parametric model and are calculated assuming that the

bandwidth is chosen optimally.

National Agricultural Statistics Service (NASS) data on county level yields is generally available for the past fifty years or so. The choice then is between using a parametric method that is misspecified or a nonparametric method with a slow rate of convergence. The latter should be favored as, at the very least, it is theoretically consistent. There have been several techniques proposed to artificially increase the sample size. Goodwin and Ker utilized information from neighboring counties. Ker and Goodwin considered an estimator that used Bayesian routines to increase efficiency. This issue may also present an opportunity for the use of semiparametric methods should a trade off for efficiency be desired.

V. KERNEL DENSITY ESTIMATION VIA DIFFUSION

Botev et al. offer kernel density estimation via diffusion as a comprehensive solution to what are some of the leading drawbacks of the kernel density approach. Proof of the following results and further analysis can be found in Botev and Botev et al.

The underlying model is based on the information mixing properties of the linear diffusion process governed by

$$\frac{\partial}{\partial t} \hat{f}(Y, t) = L[\hat{f}(Y, t)] \quad (3)$$

where $t > 0$ and $x \in \psi$. For this partial differential equation, the linear differential operator is of the form $L[\cdot] = \frac{1}{2} \frac{d}{dY} (a(Y) \frac{d}{dY} (\frac{\cdot}{p(Y)}))$. The function $a(Y)$ is arbitrary but positive on ψ . The only initial condition required for a solution is that $g(Y, 0) = \Delta(Y)$ where the term on the right is the empirical density of the data. Note that mixing occurs between the observed data Y and the, until now, unknown function $p(Y)$.

The solution to this diffusion process $\hat{f}(Y, t)$ is a type of kernel density estimator sharing many of the properties of other estimators in this class. t – the mixing time for the partial differential equation – takes the role of the band-

width parameter. At time 0, the solution is exactly equal to the empirical density $\Delta(Y)$ as specified in the initial condition. This is analogous to the convergence of the standard kernel density estimator to a sum of dirac delta functions as the bandwidth tends to zero. If $p(Y)$ is a probability density function on ψ , then the limit of the solution as t goes to infinity is the specified probability density. Thus $p(Y)$ can be viewed as the limiting distribution of the process.

Botev shows that the solution to this process can be written in the form

$$\hat{f}(Y, t) = \frac{1}{N} \sum_{i=1}^N K(y, Y_i, t) \quad (4)$$

where the kernel is a diffusion kernel satisfying certain conditions. Though there is no analytical form for the diffusion kernel, it can be written as a Fourier series when ψ is bounded. The expression of the diffusion estimator as a sum of individual kernels makes the relationship with the kernel density estimator evident.

If the set ψ has boundaries, then the Neumann conditions given by $\frac{\partial}{\partial Y} (\frac{\hat{f}(Y, t)}{p(Y)}) = 0$ for $\partial\psi$ may also be added. These conditions are sufficient to ensure that the density estimate always integrates to one and account for boundary bias in a way that is similar to the reflection method. All that is required is to solve the partial differential equation over the specified domain with the Neumann conditions. Unlike standard kernel estimators, the diffusion estimator is consistent at these boundaries.

The asymptotic mean integrated squared error of the diffusion estimator and the standard fixed bandwidth kernel estimator are given as:

$$AMISE(\hat{f}) = \frac{1}{4} t^2 ||(a(f/p)')'||^2 + \frac{E[\sigma[Y]]^{-1}}{2N\sqrt{\pi t}} \quad (5a)$$

$$AMISE(\hat{g}) = \frac{1}{4} t^2 ||f''||^2 + \frac{1}{2N\sqrt{\pi t}} \quad (5b)$$

where $\sigma^2(Y) = \frac{a(Y)}{p(Y)}$

Careful inspection of this form reveals further information about the model. In both

cases, choice of bandwidth is crucial and it is possible to find the optimal bandwidth under AMISE criterion. The rate of convergence is the same provided that $p(Y)$ is not chosen as the true f – i.e. $O(N^{4/5})$. Suitable manipulation also shows consistency as the AMISE of the diffusion estimator approaches zero in the limit as N goes to infinity.

In fact the diffusion estimator is capable of nesting the fixed bandwidth kernel estimator and variable bandwidth estimators. If $a(Y) = p(Y)$ and if these are proportional to one, then the differential equation of interest is the Fourier heat equation. In this particular case the solution to the heat equation happens to be the fixed bandwidth gaussian kernel estimator. If $a(Y)$ is proportional to one, but $p(Y)$ is a pilot estimate of the density, the result is the variable bandwidth kernel estimator of Abramson. Optimal bandwidths for these nested estimators are given by their respective literature.

To depart briefly from technical aspects, Botev gives the following broad interpretation of the mechanics of this method:

If we think of each empirical observation as a point source of heat, then $\Delta(x)$ is an initial heat profile and the pde models the dissipation of this heat into a medium with nonuniform diffusivity. The nonuniform diffusivity depends on the prior $p(x)$ in such a way that in regions where we expect a lot of observations (i.e. high prior density), the empirical data is diffused (is smoothed away) at a slow rate. In regions where we expect few observations (low prior density) or features, the empirical data is diffused at a fast rate.

This interpretation may help in understanding the following approach to estimating densities where there is no prior information assumed. In such situations, as in the case of yields, $p(Y)$ is initially assumed proportional to one.

Botev et al. call this Algorithm 2 and it is

an extension of what is termed the Improved Sheather Jones Method. In the first step a pilot density is constructed by taking t^* as given by the Sheather Jones method. As no prior information is assumed in constructing the pilot density, both $a(Y)$ and $p(Y)$ are proportional to one and the estimator reduces to the standard fixed bandwidth kernel density estimator. In the second stage, $p(Y)$ is replaced by the pilot estimate with $a(Y) \propto 1$. Estimation is accomplished using the diffusion estimator $\hat{f}(Y, \hat{t})$ where $\hat{t} = t^* \times E[\sigma^{-1}Y]$. The bandwidth \hat{t} is chosen such that the asymptotic variance of the pilot is equal to the asymptotic variance in the second stage. Computational details can be found in Botev et al.

The algorithm might then be described in the following way. The first stage of the process models the dissipation of the heat into a space with uniform diffusivity. Uniformity is a result of a lack of prior information about the density and $a(Y) = p(Y) \propto 1$. The second stage models dissipation into the same space, but now the diffusivity is nonuniform. It depends on the nature of the density estimate from the first stage. Where the pilot estimate has little mass, the heat is diffused quickly. This quickened diffusion leads to more smoothing in these areas. In this way the amount of smoothing adapts to local features of the data.

VI. RESULTS

For this application, historical yield data was obtained from the NASS database. Data was generally available from 1962 until 2013. Exceptions are Castro, TX where data was available from 1968 and Coahoma, MS where data was available from 1972. In the first stage a quadratic trend was estimated and a correction for conditional heteroskedasticity was implemented using techniques mentioned previously. Conditional yields were generated about the yield forecast for 2013.

In the second stage, maximum likelihood was used to fit both normal and weibull distributions to the conditional yields. Nonparametric estimation was conducted using a fixed

bandwidth kernel estimator and the diffusion estimator of Botev et al. In the former case, the bandwidth was selected using Silverman's Rule of Thumb and in the latter the bandwidth was selected using Botev's Algorithm 2. A normal kernel was assumed.

Given a trigger yield of $\gamma\mu$, the probability of loss is given by $\int_0^{\gamma\mu} \hat{m}(Y)$ where $\hat{m}(Y)$ is some estimate of the probability density of the yield. Expected loss is then given by $E[LOSS] = Prob(Y < \gamma\mu)[\gamma\mu - E(Y|Y < \gamma\mu)]$. It is often not possible to calculate these integrals analytically. Loss probabilities for the parametric distributions were calculated using Monte Carlo simulation. Probabilities from the kernel estimators were approximated using the trapezoid rule.

While the GRP is a policy available at the county level, several state level distributions are included. This allows comparison with the state level distributions estimated in Goodwin and Ker and it also facilitates scrutiny of the possibly differing nature of state and county yields. The panels of figure 1 contain graphs of density estimates for Georgia peanuts, Indiana corn, Iowa corn, Kansas sorghum, Kansas wheat, Mississippi cotton, Mississippi wheat, and Texas cotton. All crops are all-practice.

Many of the state level crops exhibit negative skewness, but Texas cotton and Georgia peanuts show positive skew. This feature may be significant when designing new policies for these crops. There is also evidence of bitangentiality and bimodality in some crops, such as Indiana and Iowa corn. The densities are usually quite smooth which is a reflection of the area of aggregation. To sum up, even at a high level of aggregation, parametric methods are incapable of capturing significant features of the data.

Figure 2 displays graphs of density estimates for eight different combinations of counties and crops. Table 1 gives premium rates at the 75 and 90 percent coverage levels for these counties and crops which include Adair, IA corn, Adams, IN corn, Atchison, KS wheat, Barton, KS sorghum, Bolivar, MS wheat, Boone, NE corn, Castro, TX cotton, and Coahoma, MS

cotton. Again, all crops are all-practice except Coahoma cotton which is non-irrigated only.

As expected, many of the densities are negatively skewed. A majority of the kernel estimates also reveal bitangential features or multiple modes. Rates constructed using the normal distribution and the weibull distribution were surprisingly similar. Rates under the weibull should be larger than those calculated under the normal if negative skewness was present, but this is not the case. This similarity suggests that the weibull distribution may not sufficiently capture skewness in practice.

In all of the cases, rates obtained through nonparametric methods are larger than those given by the two parametric distributions. The difference in these approaches is especially clear in the case of Castro, TX cotton. West Texas cotton is subject to particularly volatile growing conditions. At the 75% coverage level, nonparametric rates are almost double those obtained under the normal.

The rates from the diffusion estimator exceed those of the standard kernel estimator in every case. This is most likely a result of the adaptive smoothing of the diffusion estimator. Consider again the case of Castro, TX cotton. The diffusion estimator, which is consistent at the lower bound, properly captures the increased risk associated with exceptionally low yields for this crop. The standard kernel estimator suffers from boundary bias and does not capture this feature.

In sum, results show that the method used in modeling yield distributions can have a significant effect on policy parameters (premium rates in this case). With new insurance policies being proposed for peanuts and cotton, the impact of these parameters will be significant. Based on Goodwin and Ker, it can be conjectured that existing rates are likely less than those estimated using nonparametric methods. Be that as it may, no attempt is made to compare these estimates with existing rates as of yet.

VII. CONCLUSION

Varying the specification of yield distributions can have a large effect on premium rates. The problem that rate makers must address is in some ways different from – but in many ways similar to – classic cases where a density estimate is desired. As in those classic cases, prior information and a clear understanding of the problem allows the range of tolerable estimation methods to be tightened. The desire for flexibility in capturing skewness and bimodality forces the exclusion of parametric methods from the range of tolerable choices. There is no theoretical basis for many of the parametric methods commonly used.

Further tightening of the tolerable set, based on considerations of computational ease, the bounded domain of yield densities, and the importance of tail behavior, establishes kernel density estimation via diffusion as a preferred method. By taking advantage of the unique properties of this estimator, rate makers may be able to adjust Group Risk Plan premium rates to bring them in line with rates that are actuarially fair.

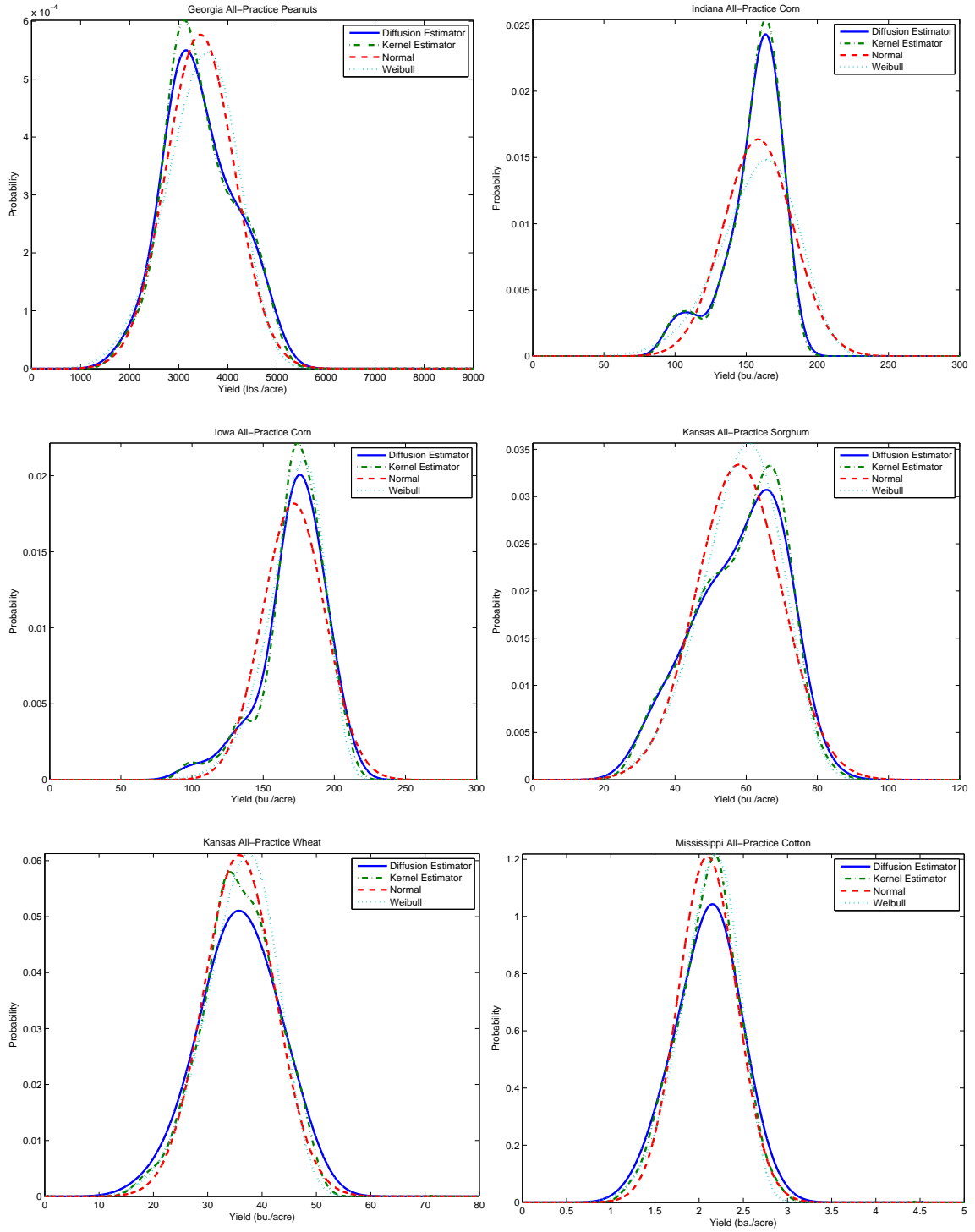
Further research on this topic will address a number of issues. Noting concerns regarding slow rates of convergence, it would be agreeable to incorporate methods that increase the sample size and utilize spatial information in some way. Detrending procedures should also be more closely examined to try and correct for error in first stage estimation. By bringing the estimation techniques employed here in line with current GRP rating procedures, and by considering a larger number of counties, it will also be possible to pursue policy simulation. The monetary impact of varying policy parameters could then be considered.

REFERENCES

- [Abramson, 1982] Abramson, I. S. (1982) On Bandwidth Variation in Kernel Estimates-A Square Root Law. *The Annals of Statistics*, 10:4:1217–1223.
- [Atwood, Shaik, and Watts, 2003] Atwood, J. A., Shaik, S., and Watts, M. J. (2003) Are Crop Yields Normally Distributed? A Re-examination. *American Journal of Agricultural Economics*, 85:4:888–901.
- [Botev, 2007] Botev, Z. I. (2007) Nonparametric density estimation via diffusion mixing. Technical Report, Department of Mathematics, University of Queensland.
- [Botev, Grotowski, and Kroese, 2010] Botev, Z. I., Grotowski, J. F. and Kroese, D. P. (2010) Kernel Density Estimation via Diffusion. *The Annals of Statistics*, 38:5:2916–2957.
- [Botts and Boles, 1958] Botts, R. R. and Boles, J. N. (1958) Use of Normal-Curve Theory in Crop Insurance Rate Making. *Journal of Farm Economics*, 40:2:733–740.
- [Chen and Miranda, 2004] Chen, S. L. and Miranda, M. (2004) Modeling Multivariate Crop Yield Densities with Frequent Extreme Events. Paper for AAEA Annual Meeting, Denver, CO, August 2004.
- [Gallagher, 1987] Gallagher, P. (1987) U.S. Soybean Yields: Estimation and Forecasting with Nonsymmetric Disturbances. *American Journal of Agricultural Economics*, 69:4:798–803.
- [Glauber, 2013] Glauber, J. P. (2013). The growth of the Federal Crop Insurance Program, 1990-2011. *American Journal of Agricultural Economics*, 95:2:482–488.
- [Goodwin and Ker, 1998] Goodwin, B. K. and Ker, A. P. (2000). Nonparametric Estimation of Crop Yield Distributions: Implications for Rating Group-Risk Crop Insurance Contracts. *American Journal of Agricultural Economics*, 80:1:139–153.
- [Hall, Marron, and Park, 1992] Hall, P., Marron, J. S. and Park, B. U. (1992). Smoothed cross-validation. *Probability Theory and Related Fields*, 92:1–20.

-
- [Jones, Marron, and Sheather, 1996] Jones, M. C., Marron, J. S. and Sheather, S. J. (1996). A Brief Survey of Bandwidth Selection for Density Estimation. *Journal of the American Statistical Association*, 91:433:401–407.
- [Just and Weninger, 1999] Just, R. E. and Weninger, Q. (1991). Are Crop Yields Normally Distributed? *American Journal of Agricultural Economics*, 81:2:287–304.
- [Ker and Coble, 2003] Ker, A. P. and Coble, K. (2003). Modeling Conditional Yield Densities. *American Journal of Agricultural Economics*, 85:2:291–304.
- [Ker and Goodwin, 2000] Ker, A. P. and Goodwin, B. K. (2000). Nonparametric Estimation of Crop Insurance Rates Revisited. *American Journal of Agricultural Economics*, 82:2:463–478.
- [Li and Racine, 2007] Li, L. and Racine, J. S. (2007). *Nonparametric Econometrics*. Princeton: Princeton University Press.
- [Marron and Wand, 1992] Marron, J. S. and Wand, M. P. (1992). Exact Mean Integrated Squared Error. *The Annals of Statistics*, 20:2:712–736.
- [Miranda, 1991] Miranda, M. J. (1991). Area Yield Crop Insurance Reconsidered. *American Journal of Agricultural Economics*, 73:2:233–242.
- [Miranda and Glauber, 1997] Miranda, M. J. and Galuber, J. W. (1997) Systemic Risk, Reinsurance, and the Failure of Crop Insurance Markets. *American Journal of Agricultural Economics*, 79:1:206–215.
- [Nelson, 1990] Nelson, C. H. (1990) The Influence of Distributional Assumptions on the Calculation of Crop Insurance Premia. *North Central Journal of Agricultural Economics*, 12:1:71–78.
- [Ozaki, Ghosh, Goodwin, and Shirota, 2008] Ozaki, V. A., Ghosh, S. K., Goodwin, B. K. and Shirota, R. (2010) Spatio-Temporal Modeling of Agricultural Yield Data with an Application to Pricing Crop Insurance Contracts. *American Journal of Agricultural Economics*, 90:4:951–961.
- [Sheather and Jones, 1991] Sheather, S. J. and Jones, M. C. (1991). A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation. *Journal of the Royal Statistical Society. Series B*, 53:3:683–690.
- [Sherrick, Zanani, Schnitkey, and Irwin, 2004] Sherrick, B., Zanani, F., Schnitkey, G., and Irwin, S. (2004) Crop Insurance Valuation Under Alternative Yield Distributions. *American Journal of Agricultural Economics*, 86:2:406–419.
- [Silverman, 1986] Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall.
- [Skees, Black, and Barnett, 1997] Skees, J. R., Black, J. R., and Barnett, B. J. (1997) Designing and Rating and Area Yield Crop Insurance Contract. *American Journal of Agricultural Economics*, 79:2:430–438.
- [Zhu, Goodwin, and Ghosh, 2011] Zhu, Y., Goodwin, B. K. and Ghosh, S. J. (2011) Modeling Yield Risk Under Technological Change: Dynamic Yield Distributions and the U.S. Crop Insurance Program. *Journal of Agricultural and Resource Economics*, 36:1:192–210.

Figure 1: State Level Yield Distributions



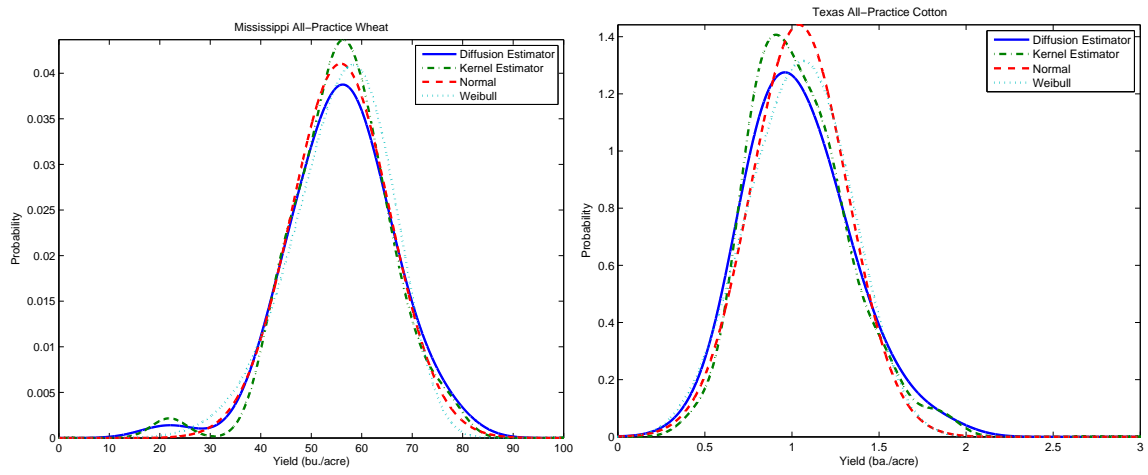
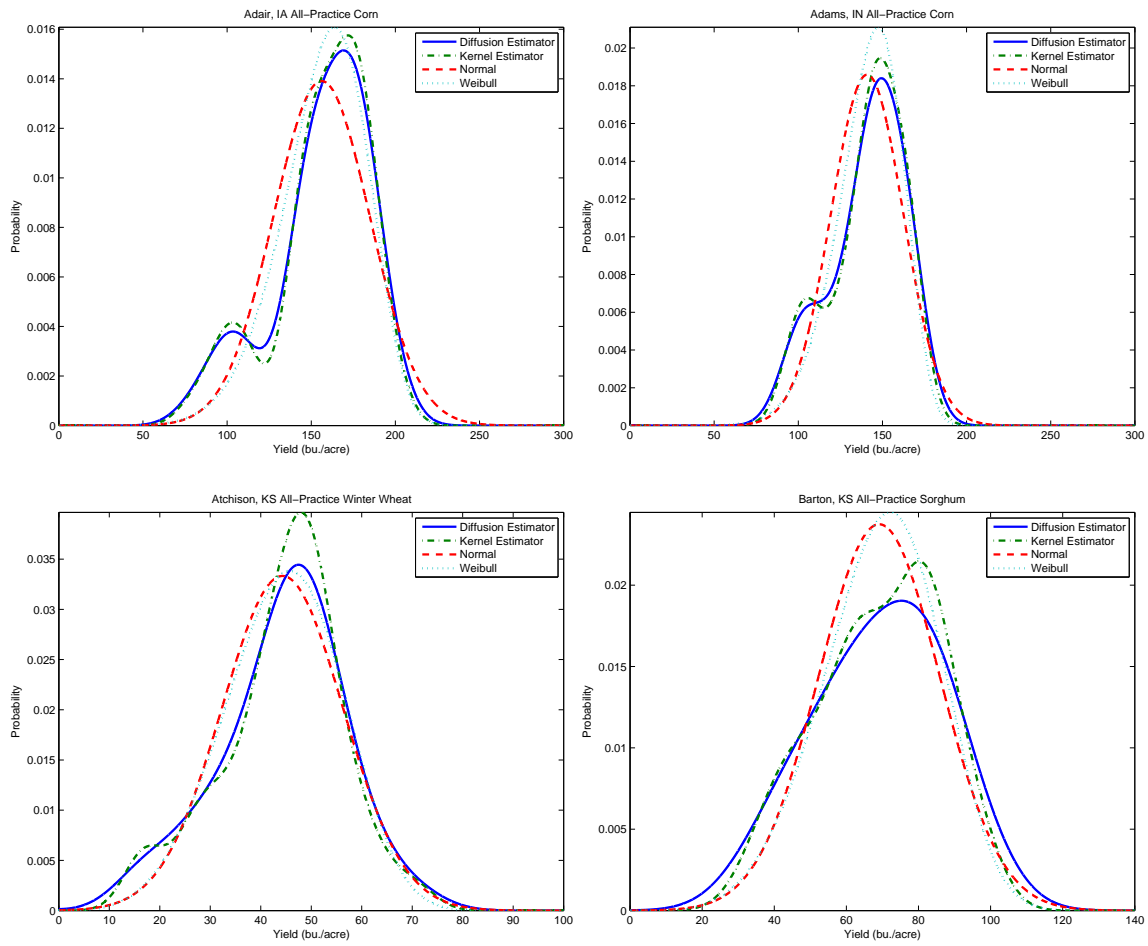


Figure 2: County Level Yield Distributions



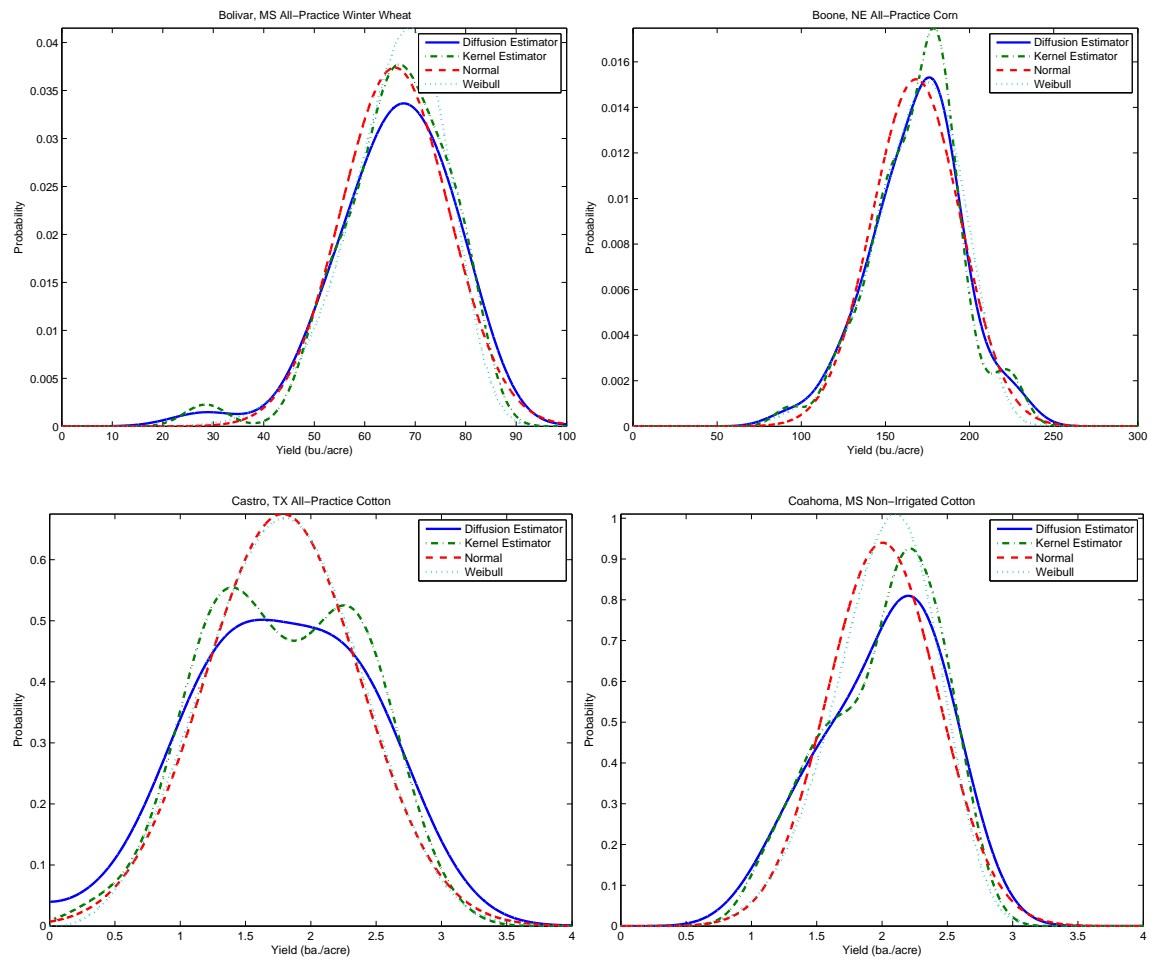


Table 1: County Level Premium Rates

County and Coverage Level	Normal	Weibull	Kernel	Diffusion
Adair, IA Corn				
75%	1.00	0.92	2.29	2.35
90%	3.80	3.32	4.80	4.95
Adams, IN Corn				
75%	0.41	0.55	0.94	1.06
90%	2.59	2.45	3.68	3.82
Atchison, KS Wheat				
75%	3.50	3.31	4.88	5.34
90%	7.18	7.01	8.05	8.71
Barton, KS Sorghum				
75%	2.58	2.54	3.66	4.54
90%	6.15	5.83	7.43	8.41
Bolivar, MS Wheat				
75%	0.57	0.70	1.26	1.55
90%	2.94	2.84	3.48	4.06
Boone, NE Corn				
75%	0.49	0.93	0.86	1.00
90%	2.68	3.25	3.15	3.45
Castro, TX Cotton				
75%	4.14	5.36	6.78	8.47
90%	8.10	9.51	11.34	12.80
Coahoma, MS Cotton				
75%	1.65	1.69	2.67	3.22
90%	4.92	4.51	6.22	6.82
